



UNIVERSIDAD  
POLITECNICA  
DE VALENCIA



# TESIS DOCTORAL

Mejora de la Arquitectura de Acceso Radio UMTS  
mediante Multinodos B

Universidad Politécnica de Valencia

Departamento de Comunicaciones

Valencia, Septiembre de 2007

Presentada por:

*Jose Francisco Monserrat del Río*

Dirigida por:

*Dr. Narcís Cardona Marcet*

*Dr. Lorenzo Rubio Arjona*



*A los que siempre han estado ahí, mis padres y mi mujer.*



## *Agradecimientos*

*En primer lugar, quisiera tener unas sinceras palabras de agradecimiento a mis dos directores de Tesis, al Dr. Narcís Cardona y al Dr. Lorenzo Rubio, por su constante dedicación, orientación y ejemplo y por la gran confianza que han depositado en mí.*

*También tengo que agradecer la oportunidad que he tenido de trabajar con tan grandes compañeros como Daniel Calabuig, Rubén Fraile, David Gómez, Paolo García, Vicent Soler, Javier Gozávez y Oscar Lázaro. En esta Tesis Doctoral hay un poco de todos vosotros, gracias.*

*Finalmente, quiero hacer una mención especial a mis padres y hermanas por su apoyo y ayuda en todos los sentidos, y a la mujer de mi vida, Lorena.*



# Resumen

Dada la creciente demanda de servicios multimedia en movilidad, las operadoras de telefonía móvil necesitan aumentar constantemente la capacidad de sus sistemas. En ocasiones la única solución posible para mejorar las prestaciones del sistema es incrementar el número de transmisores, siendo ésta una solución costosa. Además, en UMTS utilizar más emplazamientos macro-celulares no siempre garantiza una mayor capacidad del sistema, por los grandes problemas de interferencia mutua que sufren dos macro-celdas si se encuentran demasiado próximas. La única solución que se puede adoptar en estos casos es hacer uso de redes jerárquicas celulares, añadiendo micro-celdas que utilicen una frecuencia de portadora distinta.

Esta Tesis Doctoral aporta una visión complementaria al despliegue de las redes jerárquicas y propone aprovechar dicho despliegue para modificar las redes de acceso radio UMTS con nueva arquitectura de nodos B múltiples o Multinodo B.

El sistema a desarrollar utiliza la tecnología de nodos B UMTS a los que se conectan un número de antenas distribuidas en diferentes puntos del área de servicio, conformando el equivalente a un *array* distribuido espacialmente. La incorporación de técnicas de procesado en *array* junto con mecanismos de gestión dinámica de recursos y la tecnología de enlaces ópticos proporcionarán a este sistema una estructura de acceso radio óptima en entornos micro-celulares y en interiores, con gran flexibilidad y adaptabilidad.

Las ventajas de utilizar esta configuración, que se detallan a lo largo de la Tesis Doctoral, son fundamentalmente la mejora de la capacidad de la red de acceso, minimización del impacto radioeléctrico y óptima gestión de recursos.

En primer lugar, en esta Tesis Doctoral se demuestra que la nueva arquitectura radio es capaz de dotar al sistema UMTS de una diversidad espacial adicional, diversidad que mejorará los requerimientos de potencia recibida en los terminales, sobre todo en los escenarios micro-celulares.

Respecto a la gestión dinámica de recursos, esta Tesis Doctoral desarrolla una formulación matemática basada en redes neuronales de Hopfield capaz de resolver el problema de optimización planteado en pocos microsegundos. La nueva formulación descrita maximiza el uso de los recursos a la vez que minimiza el *delay* en la transmisión de los paquetes. El mejor comportamiento respecto a otros algoritmos propuestos en la literatura se debe, sobre todo, a la gran ventaja que aportan las redes neuronales puesto que son capaces de considerar simultáneamente la satisfacción de múltiples objetivos.

Además, la red Hopfield distribuye de manera óptima la señal transmitida al usuario entre las antenas que conforman el Multinodo B. Gracias a la diversidad espacial y esta gestión óptima de recursos, la potencia consumida se puede dividir hasta por un factor cinco. Además de esta reducción en la potencia consumida, la calidad de servicio también se ve mejorada, reduciéndose así el número de usuarios en cola y el tiempo medio de espera en la transmisión de los paquetes.

Por último, en esta Tesis Doctoral se prueba que el sistema Multinodo B es realizable a medio plazo y con un coste de implementación relativamente bajo, en comparación con otras alternativas más complejas.



# Resum

Donada la creixent demanda de servicis multimèdia en mobilitat, les operadores de telefonia mòbil necessiten augmentar constantment la capacitat dels seus sistemes. En ocasions l'única solució possible per a millorar les prestacions del sistema és incrementar el nombre de transmissors, sent aquesta una solució costosa. A més, en UMTS utilitzar més emplaçaments macro cel·lulars no sempre garanteix una major capacitat del sistema, pels grans problemes d'interferència mútua que sofreixen dos macro cel·les si es troben massa pròximes. L'única solució que es pot adoptar en estos casos és fer ús de xàrcies jeràrquiques cel·lulars, afegint micro cel·les que utilitzen una freqüència de portadora distinta.

Esta Tesi Doctoral aporta una visió complementària al desplegament de les xàrcies jeràrquiques i proposa aprofitar aquest desplegament per a modificar les xàrcies d'accés ràdio UMTS amb una nova arquitectura de nodes B múltiples o Multinodo B.

El sistema a desenrotllar utilitza la tecnologia de nodes B UMTS als que es connecten un nombre d'antenes distribuïdes en diferents punts de l'àrea de servicis, conformant l'equivalent a un *array* distribuït espacialment. La incorporació de tècniques de processat en *array* junt amb mecanismes de gestió dinàmica de recursos i la tecnologia d'enllaços òptics proporcionaran a aquest sistema una estructura d'accés ràdio òptima en entorns micro cel·lulars i en interiors, amb gran flexibilitat i adaptabilitat.

Els avantatges d'utilitzar esta configuració, que es detallen al llarg de la Tesi Doctoral, són fonamentalment la millora de la capacitat de la xàrcia d'accés, minimització de l'impacte radioelèctric i òptima gestió de recursos.

En primer lloc, en esta Tesi Doctoral es demostra que la nova arquitectura ràdio és capaç de dotar al sistema UMTS d'una diversitat espacial addicional, diversitat que millorarà els requeriments de potència rebuda en els terminals, sobretot als escenaris micro cel·lulars.

Respecte a la gestió dinàmica de recursos, esta Tesi Doctoral desenrotlla una formulació matemàtica basada en xàrcies neuronals d'Hopfield capaç de resoldre el problema d'optimització plantejat en pocs microsegons. La nova formulació descrita maximitza l'ús dels recursos al mateix temps que minimitza el *delay* en la transmissió dels paquets. El millor comportament respecte a altres algoritmes proposats en la literatura es deu, sobretot, a el gran avantatge que aporten les xàrcies neuronals ja que són capaces de considerar simultàniament la satisfacció de múltiples objectius.

A més, la xàrcia Hopfield distribueix de manera òptima la senyal transmesa a l'usuari entre les antenes que conformen el Multinodo B. Gràcies a la diversitat espacial i esta gestió òptima de recursos, la potència consumida es pot dividir fins per un factor cinc. A més d'esta reducció en la potència consumida, la qualitat de servici també es veu millorada, reduint-se així el nombre d'usuaris en cua i el temps mitjà d'espera en la transmissió dels paquets.

Finalment, en aquesta Tesi Doctoral es prova que el sistema Multinodo B es realitzable a mig termini i amb un cost d'implementació relativament baix, en comparació amb altres alternatives més complexes.

# Abstract

Given the increasing demand of services on the move, the mobile operators need to increase the capability of 3G systems constantly. Sometimes, the only possible solution for improving the system performance is to increase the number of transmitters, being this an expensive approach. Besides, in UMTS utilizing more macro-cellular sites does not always entail an increasing system capability if they are too near, since mutual interferences can harm the global operation. In these cases the only solution is to make use of hierarchical cell structures (HCS), adding micro-cells transmitting at different carriers.

This Thesis provides a complementary point of view of the HCS deployment and proposes to take advantage of such deployment to improve the UMTS radio access network with the introduction of a new architecture of B-nodes or Multinodes B.

The system to be developed uses UMTS B nodes where several spatially distributed antennas are connected. The Multinode antennas cover the service area and form an equivalent to a spatially distributed array. An optimum radio access structure for indoor and micro-cellular environments, with high adaptability and flexibility, is achieved when array processing techniques, optical links and dynamic resource allocation algorithms are incorporated to the Multinode B.

Advantages of using this configuration, as described along this Thesis, are mainly on improving the capacity of the radio access network, minimizing the radioelectric impact and optimizing the resource allocation.

Firstly, in this Thesis it is demonstrated that the new radio access architecture is able to provide an additional spatial diversity to the UMTS system, diversity that reduces the power required by the terminals, most of all at micro-cellular scenarios.

Regarding dynamic resource allocation (DRA), an efficient scheduler based on Hopfield neural networks (HNN) has been proposed to support Quality of Service (QoS) for different services. This new scheme provides an optimum user bit rate allocation together with a controlled delay and a high utilization of system resources.

The proposed algorithm improves all analyzed performance indicator over other conventional algorithms. Moreover, another benefit of using a HNN-based DRA algorithm is precisely that its hardware implementation makes possible a real-time running of the algorithm.

On the other hand, the HNN distributes in an optimal manner the signal transmitted to the user among the antennas included in the Multinode B. Thanks to the spatial and the optimal DRA, the power consumption can be divided by five. In addition to this reduction in the power, QoS is also improved, decreasing the number of queued users and the average service response time.

Finally, in this Thesis it has been proven that the Multinode B system could be implemented in the mean term and the implementations costs are lower than other more complex alternatives found in the literature.

# ÍNDICE

<b>ÍNDICE</b> .....	<b>13</b>
<b>LISTA DE ACRÓNIMOS</b> .....	<b>17</b>
<b>LISTA DE PUBLICACIONES DEL AUTOR</b> .....	<b>21</b>
<b>1 INTRODUCCIÓN Y OBJETIVOS</b> .....	<b>29</b>
1.1 MOTIVACIÓN DE LA TESIS DOCTORAL.....	29
1.2 ESTADO DEL ARTE.....	32
1.2.1 <i>Arquitectura de redes UMTS</i> .....	32
1.2.2 <i>Diversidad en transmisión en sistemas 3G</i> .....	36
1.2.3 <i>Estimación de canal</i> .....	40
1.3 OBJETIVOS PLANTEADOS.....	42
1.4 PRINCIPALES APORTACIONES DE LA TESIS .....	44
1.5 ESTRUCTURA DE LA TESIS .....	47
REFERENCIAS.....	50
<b>2 DESCRIPCIÓN DE LA ARQUITECTURA MULTINODO B</b> .....	<b>55</b>
2.1 INTRODUCCIÓN.....	55
2.2 ARQUITECTURA DE PROTOCOLOS DE LA INTERFAZ RADIO MNB .....	57
2.2.1 <i>Introducción</i> .....	57
2.2.2 <i>Arquitectura de protocolos UMTS</i> .....	58
2.2.3 <i>Transmisión mediante DCH</i> .....	68
2.2.4 <i>Transmisión mediante DSCH</i> .....	74
2.2.5 <i>Obtención de parámetros de la red</i> .....	83
2.3 ENTIDADES FUNCIONALES DEL MULTINODO B.....	86
2.3.1 <i>Unidad de control</i> .....	87
2.3.2 <i>Procesado en array</i> .....	88
2.3.3 <i>Unidad Tx/Rx</i> .....	89
2.3.4 <i>Convertidores de radio - fibra</i> .....	91
2.3.5 <i>Amplificador bidireccional</i> .....	93
2.3.6 <i>Antena</i> .....	96
REFERENCIAS.....	99
<b>3 EVALUACIÓN DE LAS PRESTACIONES DE LA ARQ. MNB</b> .....	<b>101</b>
3.1 INTRODUCCIÓN.....	101
3.2 SIMULACIÓN DE ENLACE DE LA RED MNB.....	103
3.2.1 <i>Estructura del simulador de enlace</i> .....	108
3.2.2 <i>Suposiciones y simplificaciones</i> .....	110
3.2.3 <i>Validación del simulador</i> .....	114
3.2.4 <i>Mejoras en la estimación de canal</i> .....	127
3.3 PRESTACIONES DE LA ARQUITECTURA MNB .....	133
3.3.1 <i>Efecto del número de fingers considerado</i> .....	135
3.3.2 <i>Efecto de la diferencia relativa de potencia</i> .....	138
3.3.3 <i>Efecto del retardo relativo</i> .....	140

3.3.4	<i>Procesado en array y combinación coherente de la señal en recepción...</i>	142
3.3.5	<i>Estimación de canal en el MNB</i> .....	146
3.4	<i>LOOK-UP TABLES CONSIDERADAS</i> .....	154
3.4.1	<i>Servicio de voz a 12.2 kbps</i> .....	155
3.4.2	<i>Servicios de datos</i> .....	158
3.5	CONCLUSIONES.....	166
	REFERENCIAS .....	168
<b>4</b>	<b>GESTIÓN DE RECURSOS EN REDES UMTS.....</b>	<b>173</b>
4.1	INTRODUCCIÓN .....	173
4.2	MODELO DEL SISTEMA.....	178
4.2.1	<i>Métricas de calidad de servicio</i> .....	180
4.3	ALGORITMOS DE GESTIÓN DE RECURSOS .....	182
4.3.1	<i>Control de potencia</i> .....	182
4.3.2	<i>Selección inicial de celda</i> .....	184
4.3.3	<i>Algoritmo de traspaso</i> .....	184
4.3.4	<i>Control de admisión</i> .....	185
4.3.5	<i>Control de congestión</i> .....	185
4.3.6	<i>Cronoejecución o scheduling</i> .....	188
4.3.7	<i>Gestión de códigos</i> .....	197
4.4	FUNCIONAMIENTO DE LAS REDES NEURONALES.....	200
4.4.1	<i>Introducción</i> .....	200
4.4.2	<i>Redes neuronales</i> .....	202
4.4.3	<i>Redes neuronales Hopfield</i> .....	209
4.5	UTILIZACIÓN DE REDES HOPFIELD PARA ASIGNAR RECURSOS EN UMTS.....	224
4.5.1	<i>Requisitos del algoritmo de gestión de recursos</i> .....	225
4.5.2	<i>Optimización del problema DRA mediante redes Hopfield</i> .....	229
4.5.3	<i>Evaluación de resultados y discusión</i> .....	246
4.6	CONCLUSIONES.....	270
	REFERENCIAS .....	272
<b>5</b>	<b>GESTIÓN DE RECURSOS EN REDES MNB.....</b>	<b>277</b>
5.1	INTRODUCCIÓN .....	277
5.2	GESTIÓN DE RECURSOS EN EL MNB .....	280
5.2.1	<i>Control de potencia</i> .....	281
5.2.2	<i>Selección de celda</i> .....	286
5.3	CRONOEJECUCIÓN MEDIANTE REDES HOPFIELD .....	288
5.3.1	<i>Formulación del problema</i> .....	289
5.3.2	<i>Dinámica del sistema y parámetros hardware</i> .....	294
5.3.3	<i>Cálculo de las constantes</i> .....	295
5.4	ANÁLISIS DE RESULTADOS Y DISCUSIÓN .....	298
5.4.1	<i>Algoritmos de distribución de recursos utilizados</i> .....	298
5.4.2	<i>Escenarios de simulación</i> .....	301
5.4.3	<i>Discusión de los resultados de simulación</i> .....	327
5.5	CONCLUSIONES.....	349
	REFERENCIAS .....	351

<b>6</b>	<b>MODELADO DEL <i>SHADOWING</i> Y SU INFLUENCIA SOBRE LA EVALUACIÓN DE SISTEMAS.....</b>	<b>353</b>
6.1	INTRODUCCIÓN.....	353
6.2	MODELOS DE <i>SHADOWING</i> .....	354
6.2.1	<i>Modelo 1: Unidimensional Incorrelado</i> .....	356
6.2.2	<i>Modelo 2: Unidimensional con Autocorrelación Espacial</i> .....	357
6.2.3	<i>Modelo 3: Bidimensional con Autocorrelación Espacial</i> .....	359
6.2.4	<i>Modelo 4: Bidimensional con Correlación Cruzada Constante</i> .....	367
6.2.5	<i>Modelo 5: Bidimensional con Correlación Cruzada Basada en Medidas</i> .....	370
6.2.6	<i>Modelo 6: Bidimensional con Correlación Cruzada Completa</i> .....	376
6.3	INFLUENCIA DE LOS DISTINTOS MODELOS.....	385
6.3.1	<i>Introducción</i> .....	385
6.3.2	<i>Análisis de los resultados y discusión</i> .....	387
6.4	CONCLUSIONES.....	396
	REFERENCIAS.....	398
<b>7</b>	<b>CONCLUSIONES Y LÍNEAS FUTURAS DE INVESTIGACIÓN.....</b>	<b>401</b>
7.1	CONCLUSIONES.....	401
7.2	LÍNEAS FUTURAS DE INVESTIGACIÓN.....	411
<b>A</b>	<b>DESCRIPCIÓN DEL SIMULADOR DE ENLACE.....</b>	<b>413</b>
A.1	BLOQUES FUNCIONALES DEL SIMULADOR.....	414
A.1.1	<i>Generación aleatoria de bits</i> .....	414
A.1.2	<i>CRC</i> .....	414
A.1.3	<i>Concatenación de bloques de transporte</i> .....	415
A.1.4	<i>Codificación de canal</i> .....	416
A.1.5	<i>Ajuste de tasa de transmisión</i> .....	430
A.1.6	<i>Entrelazado</i> .....	432
A.1.7	<i>Segmentación del canal físico</i> .....	433
A.1.8	<i>Multiplexación del canal de control</i> .....	435
A.1.9	<i>Ensanchado de la señal y modulación</i> .....	435
A.1.10	<i>Filtrado Nyquist</i> .....	437
A.1.11	<i>Canal de propagación</i> .....	438
A.1.12	<i>Inserción de ruido</i> .....	442
A.1.13	<i>Receptor Rake</i> .....	443
A.2	VALIDACIÓN DE LOS BLOQUES FUNCIONALES.....	448
A.2.1	<i>Validación del modelo de canal</i> .....	448
A.2.2	<i>Validación de la codificación de canal</i> .....	450
	REFERENCIAS.....	458
<b>B</b>	<b>DESCRIPCIÓN DEL SIMULADOR DE SISTEMA.....</b>	<b>459</b>
B.1	MODELO DEL SISTEMA.....	459
B.1.1	<i>Entidades funcionales</i> .....	460
B.1.2	<i>Estructura celular</i> .....	463
B.1.3	<i>Estructura del MNB</i> .....	464
B.1.4	<i>Estación base</i> .....	467
B.1.5	<i>Modelo de propagación</i> .....	467
B.1.6	<i>Modelo de interferencias</i> .....	470

<i>B.1.7</i>	<i>Modelo de movilidad</i> .....	481
<i>B.1.8</i>	<i>Modelos de tráfico</i> .....	483
B.2	INTERFAZ ENTRE SIMULACIÓN DE ENLACE Y SIMULACIÓN DE SISTEMA .....	509
REFERENCIAS	.....	512



# LISTA DE ACRÓNIMOS

3G	Tercera Generación
3GPP	3rd Generation Partnership Project
AAD	Angle-of-Arrival Difference
ACORP	Asignación Combinada de Recursos y Potencias
ADP	Algoritmo de Distribución de Potencia
AM	Acknowledged Mode
AP	Alternating Projection
ARMAD	Asignación de Recursos por Márgenes de Diversidad
ARMIP	Asignación de Recursos por Mínima Potencia
ARQ	Automatic Repeat reQuest
AS	Access Stratum
ASN.1	Abstract Syntax Notation 1
AWGN	Additive White Gaussian Noise
BCCH	Broadcast Control Channel
BCH	Broadcast Channel
BDA	Bidirectional Amplifier
BER	Bit Error Rate
BLER	Block Error Rate
BMC	Broadcast/Multicast Control Protocol
CAC	Call Admission Control
CB	Code Block
CBS	Cell Broadcast Service
CCCH	Common Control Channel
CDF	Cumulative Distribution Function
CDMA	Code Division Multiple Access
CIR	Carrier to Interference Ratio
CLMTD	Closed Loop Mode Transmit Diversity
CMT	Comisión del Mercado de las Telecomunicaciones
CN	Core Network
CPCH	Common Packet Channel
CPICH	Common Pilot Channel
CRC	Cyclic Redundance Check
CRC	Cyclic Redundancy Code
CRNC	Controlling Radio Network Controller
CS	Circuit Switched
DBR	Descent Bit Rate
DCCH	Dedicated Control Channel
DCH	Dedicated Channel
DPCCH	Dedicated Physical Control Channel
DPCH	Dedicated Physical Channel
DRA	Dynamic Resource Allocation
DRNC	Drift Radio Network Controller
DS-CDMA	Direct Sequence CDMA

DSCH	Downlink Shared Channel
ETSI	European Telecommunications Standards Institute
FACH	Forward Access Channel
FBI	Feedback Information
FDD	Frequency Division Duplex
FDMA	Frequency Division Multiple Access
FEC	Forward Error Correction
FER	Frame Error Rate
FIFO	First in First Out
FIR	Finite Impulse Response
FPGA	Field Programmable Gate <i>Array</i>
FRA	Fixed Resource Allocation
FTP	File Transfer Protocol
HNN	Hopfield Neural Networks
HSDPA	High Speed Downlink Packet Access
ISI	Inter Symbol Interference
KPI	Key Performance Indicator
LA	Location Area
LLR	Log Likelihood Ratio
LOG-MAP	Logarithmic Maximum A Posteriori
LOS	Line of Sight
LUT	Look-Up Table
MAC	Medium Access Control
MAP	Maximum a Posteriori
ME	Mobile Equipment
MHA	Mast Head Amplifier
ML	Maximum Likelihood
MNB	Multinodo B
MNR	Minimum Noise Rise
MRC	Maximum Ratio Combining
NAS	Non-Access Stratum
NBAP	Node B Application Protocol
NLOS	Non Line of Sight
NP	Non-deterministic Polynomial time
NRT	Non Real Time
OBR	Optimum Bit Rate
OVSF	Orthogonal Variable Spreading Factor
PCCH	Paging Control Channel
PCH	Paging Channel
PDCP	Packet Data Convergence Protocol
PDF	Probability Distribution Function
PDP	Power <i>Delay</i> Profile
PDP	Packet Data Protocol
PDSCH	Physical Downlink Shared Channel
PDU	Packet Data Unit
PEDF	Priorized Earliest Delay First

PS	Packet Switched
QoS	Quality of Service
RA	Routing Area
RAB	Radio Access Bearer
RACH	Random Access Channel
RANAP	Radio Access Network Application Protocol
RAT	Radio Access Technology
RB	Radio Bearer
RCS	Recursive Systematic Convolutional
RF	Radio Frecuencia
RLC	Radio Link Control
RNC	Radio Network Controller
RNS	Radio Network Subsystem
RNSAP	Radio Network System Application Part
RR	Round Robin
RRC	Radio Resource Control
RRM	Radio Resource Management
RSC	Recursive Systematic Convolutional
RT	Real Time
SDU	Service Data Unit
SF	Spreading Factor
SIR	Signal to Interference Ratio
SMS	Short Message Service
SOVA	Soft-Output Viterbi Algorithm
SRB	Signalling Radio Bearer
SRNC	Serving Radio Network Controller
SS	Spread Spectrum
STTD	Space Time Transmit Diversity
TB	Transport Block
TDMA	Time Division Multiple Access
TF	Transport Format
TFCS	Transport Format Combination Sets
TM	Transparent Mode
TTI	Transmission Time Interval
UE	User Equipment
UIT	Unión Internacional de Telecomunicaciones
UM	Unacknowledged Mode
UMTS	Universal Mobile Telecommunications System
USIM	User Services Identity Module
UTRAN	UMTS Terrestrial Radio Access Network
VBR	Variable Bit-Rate
WARC	World Administrative Radio Conference
WCDMA	Wideband Code Division Multiple Access
WMSA	Weighted Multi- <i>Slot</i> Averaging
WRR	Weighted Round Robin



# LISTA DE PUBLICACIONES DEL AUTOR

## PUBLICACIONES DIRECTAMENTE RELACIONADAS CON LA TESIS DOCTORAL

### Revistas Internacionales

- J. F. Monserrat, R. Fraile, L. Rubio, “Application of the Alternating Projection Method to Ensure the Feasibility of *Shadowing* Cross-Correlation Models,” IEE Electronic Letters, vol. 43, no. 13, pp. 724-725, June 2007.
- D. Calabuig, J. F. Monserrat, D. Gómez-Barquero, O. Lázaro, “User Bandwidth Usage – Driven HNN Neuron Excitation Method for Maximum Resource Utilization within Packet – Switched Communication Networks,” IEEE Communication Letters, vol. 10, no. 11, pp. 766-768, November 2006.
- R. Fraile, J. F. Monserrat, J. Gozávez , N. Cardona, “Mobile Radio Bi-dimensional Large-Scale Fading Modelling with Site-to-Site Cross-Correlation,” European Transactions on Telecommunications, Article in press, 2006.
- J. F. Monserrat, D. Calabuig, D. Gómez-Barquero, N. Cardona, “Cronoejecución y Gestión Adaptativa de Paquetes en el Sistema de Comunicaciones Móviles UMTS,” Revista Gerencia Tecnología Informática, vol. 5, no. 11, pp. 11-20, April 2006.

### Capítulo de Libro de Investigación

- N. Cardona, R. Verdone, T. Kürner, P. Grazioso, J.F. Monserrat, R. Patelli, A. Zanella, Mobile Broadband Multimedia Networks. Chapter: 9. UMTS Radio Networks, Ed.Academic Press Publications Elsevier.

### Revistas Sometidas a Evaluación

- J.F. Monserrat, D. Calabuig, L. Rubio, N. Cardona, “Hopfield Neural Networks-Based Dynamic Resource Allocation Scheme for Non Real Time

Traffic in Wireless Networks”, European Transactions on Telecommunications.

- D. Calabuig, J.F. Monserrat, D. Gómez, O. Lázaro, “An Efficient Dynamic Resource Allocation Algorithm for Packet – Switched Communication Networks Based on Hopfield Neural Excitation Method”, Elsevier Neurocomputing.
- D. Calabuig, J.F. Monserrat, V. Torres, N. Cardona “On the Practical Use of Recurrent Neural Networks. Discrete Hardware Implementation and Convergence Analysis”, IEEE Transactions on Neural Networks.
- D. Calabuid, J. F. Monserrat, D. Gómez, N. Cardona, “Delay-Centric Dynamic Resource Allocation Algorithm for Wireless Communication Systems based on HNN”, IEEE Transactions on Vehicular Technology.

### **Congresos Internacionales con Actas**

- R. Fraile, J. Gozalvez, O. Lazaro, J.F. Monserrat, N. Cardona, “Effect of a Two Dimensional Shadowing Model on System Level Performance Evaluation”, WPMC, Abano Terme, 2004.
- J. F. Monserrat, R. Patelli, N. Cardona, “Performance Comparison of Two Dynamic System Simulators Over COST273-MORANS Reference Scenarios”, WPMC, Aalborg, 2005.
- J. F. Monserrat, N. Cardona, D. Gómez-Barquero, L. Rubio, “Radio Performance Evaluation of an Enhanced Multi Node-B UTRAN Architecture”, WPMC, Aalborg, 2005.
- N. Cardona, S. Ruiz, J. Monserrat, “Mobile Radio Access network Reference Scenarios (MORANS)”, COST 273 Final Workshop, Leuven, 2005.
- J.F. Monserrat, L. Rubio, O. Lázaro, “Simplified Model of the Channel Estimation Error in UMTS Link Level Simulations”, IEEE ISWCS, Siena, 2005.

- J.F. Monserrat, J. Gozávez, R. Fraile, N. Cardona, “Effect of Shadowing Correlation Modeling on the System Level Performance of Adaptive Radio Resource Management Techniques”, IEEE ISWCS, Siena, 2005.
- D. Calabuig, J. F. Monserrat, D. Gómez-Barquero, N. Cardona, “Hopfield Neural Network Algorithm for Dynamic Resource Allocation in WCDMA Systems,” IEEE ISWCS, Valencia, 2006.
- J. F. Monserrat, D. Gómez-Barquero, D. Calabuig, L. Rubio, N. Cardona, “Evaluation of Soft Handover Micro Diversity Gain on the UMTS System Capacity and QoS,” IEEE Personal, Indoor and Mobile Radio Communication (PIMRC), Helsinki, 2006.
- D. Gómez-Barquero, D. Calabuig, J. F. Monserrat, N. Cardona, N. García, J. Pérez-Romero, “Hopfield Neural Network Algorithm for Joint Dynamic Resource Allocation in Heterogeneous Wireless Networks,” NEWCOM and ACORN Joint Workshop, Vienna, 2006.
- D. Gómez-Barquero, D. Calabuig, J. F. Monserrat, N. García, J. Pérez-Romero, “Hopfield Neural Network – based Approach for Joint Dynamic Resource Allocation in Heterogeneous Wireless Networks,” IEEE Vehicular Technology Conference (VTC) Fall, Montreal, 2006.

### **Congresos Internacionales sin Actas**

- R. Fraile, J.F. Monserrat, N. Cardona, L. Rubio, “MORANS Layer 1: Geographic Data Synthetic Scenario”, COST 273 TD (04)070, Gothenburg, 2004.
- J.F. Monserrat, L. González, R. Fraile, N. Cardona, “MORANS Layer 2: Traffic Information Synthetic Scenario”, COST 273 TD (04)071, Gothenburg, 2004.
- J.F. Monserrat, L. González, R. Fraile, O. Lázaro, “MORANS Layer 5: Site Locations Synthetic Scenario”, COST 273 TD (04)072, Gothenburg, 2004.

- R. Fraile, J.F. Monserrat, N. Cardona, L. Rubio, “MORANS Layer 6: Propagation Data Synthetic Scenario”, COST 273 TD (04)073, Gothenburg, 2004.
- R. Fraile, J.F. Monserrat, L. González, O. Lázaro, “MORANS Layer 7: Node B Parameters Synthetic Scenario”, COST 273 TD (04)074, Gothenburg, 2004.
- A. Munna, J.F. Monserrat, “MORANS White Paper - after Bologna”, COST 273 TD (04)093, Gothenburg, 2004.
- J. Monserrat, L.M. González, R. Fraile, O. Lázaro, N.Cardona, J.Gozálvez, “WCDMA Downlink Simulator for Capacity Evaluation”, COST 273 TD (04)186, Duisburg, 2004.
- R. Fraile, J. Gozalvez, O. Lazaro, J.F. Monserrat, “Effect of a Two Dimensional Shadowing Model on System Level Performance Evaluation”, COST 273 TD (04)190, Duisburg, 2004.
- R. Patelli, J.F. Monserrat, “Report of the Short Term Mission on MORANS”, COST 273 TD (05)180, Leuven, 2004.
- J. F. Monserrat, D. Martín-Sacristán, D. Gozávez, N. Cardona and J. Gozávez, “SPHERE – A Simulation Tool for CRRM Investigations”, COST2100, February 2007.

### **Congresos Nacionales**

- L. González, J.F. Monserrat, R. Fraile, O. Lázaro, N. Cardona, S. Ruiz, “Simulador Dinámico Avanzado WCDMA.Una Propuesta de Simulación en el Entorno Sintético de Referencia Europeo MORANS”, Telecom I+D '04, Madrid, 2004.
- J.F. Monserrat, D. Argilés, L. Rubio, “Mejora de la arquitectura UTRAN y Evaluación de sus Prestaciones”, URSI' 05, Gandía, 2005.
- J.F. Monserrat, R. Patelli, N. Cardona, “Simulación en el Entorno Sintético de Referencia Europeo MORANS”, URSI' 05, Gandía, 2005.



- J.F. Monserrat, R. Fraile, L.Rubio, N. Cardona, “Modelo Completo de Shadowing y su Influencia sobre la Evaluación de Sistemas”, URSI’07, Tenerife, 2007.

## **OTRAS PUBLICACIONES DEL AUTOR**

### **Revistas Internacionales**

- D. Gozávez, J. F. Monserrat, D. Calabuig and J. Gozávez, “Policy-Based Channel Access Mechanism Selection for QoS Provision in IEEE 802.11e,” IEEE Vehicular Technology Magazine, Article in press, 2007.
- J. Gozávez, A. Rodríguez-Mayol, J. Sánchez-Soriano, J.F. Monserrat, “Game Theoretic and Coordinated Interference-based Channel Allocation Schemes for Packet Mobile Communication Systems,” International Journal of Mobile Network Design and Innovation, vol. 1, no. 2, 2006.

### **Congresos Internacionales con Actas**

- R. Fraile, J.F. Monserrat, N. Cardona, J. Nasreddine, “Impact of Shadowing Modelling on TD-CDMA System-level Simulations,” IEEE ISWCS, Valencia, 2006.
- J. Díaz, D. Argilés, J.F. Monserrat, L. Rubio, “Characterization of the UWB Mobile Radio Channel Time Dispersion at 0.3 - 3GHz Band”, IEEE ISWCS, Valencia, 2006.
- D. Gómez-Barquero, A. Bria, J.F. Monserrat, N. Cardona, “Minimal Cost Planning of DVB-H Networks on Existing Wireless Infrastructure”, IEEE Personal, Indoor and Mobile Radio Communication (PIMRC), Helsinki, 2006.
- D. Martín-Sacristán, J. F. Monserrat, J. Gozávez, Narcís Cardona, “Effect of Channel-Quality Indicator Delay on HSDPA Performance”, IEEE Vehicular Technology Conference (VTC) Spring, Dublin, 2007.

- D. Gozávez, J. F. Monserrat, J. Gozávez, L. Rubio, “An Efficient HCF Scheduling Mechanism in Mixed Traffic Scenarios”, IEEE Vehicular Technology Conference (VTC) Spring, Dublin, 2007.
- J. Gozávez, D. Martín-Sacristán, M. Lucas-Estañ, J.F. Monserrat, J.J. Gonzalez-Delicado, D. Gozávez and M. Marhuenda, “SPHERE – A Simulation Platform for Heterogeneous Wireless Systems”, IEEE Conference on Testbeds and Research Infrastructures for the Development of Networks and Communities (Tridentcom), Orlando, 2007.
- D. Martín-Sacristán, J.F. Monserrat, D. Calabuig, N. Cardona, “HSDPA Link Adaptation Improvement Based on Node-B CQI Processing”, IEEE International Symposium on Wireless Communications Systems (ISWCS), Norway, 2007.

### **Congresos Nacionales**

- D. Gómez Barquero, J.F. Monserrat del Río, N. Cardona Marcet, “Planificación y Despliegue de Red Eficiente de DVB-H”, Telecom I+D '05, Madrid, 2005.
- D. Argilés, V. Soler, J. Fernández, J.F. Monserrat, L.Rubio “Localización de terminales móviles en redes UMTS mediante parametros de la UTRAN”, URSI'05, Gandía, 2005.
- J. Díaz, D.Argilés, J.F.Monserrat, L.Rubio, “Caracterizacion del canal movil indoor de UWB en la banda 0.3 a 3 GHz”, URSI'06, Oviedo, 2006.
- D. Gómez-Barquero, A. Salieta, A.P. García, J.F. Monserrat, N. Cardona, “Planificación de una Red DVB-H en la Ciudad de Valencia”, Telecom I+D'06, Madrid, 2006.





# Capítulo 1

## *Introducción y Objetivos*

*Lo último que uno sabe es por dónde empezar.*

*Blaise Pascal*

### **1.1 Motivación de la Tesis Doctoral**

En la actualidad los servicios de comunicación móviles han alcanzado una enorme penetración en el mercado. Estudios actuales señalan que la penetración de la telefonía móvil sobre la población es de más del 103% alcanzando los 48 millones de terminales móviles en España a día de hoy, lo que supone que cada español tiene de media un teléfono móvil, según la última memoria trimestral de la Comisión del Mercado de las Telecomunicaciones (CMT) [1], que pone de manifiesto que los ingresos de los operadores móviles crecieron un 10,5% en 2006, hasta los 13.394 M€, con lo que han superado el negocio generado por la telefonía fija. El importante incremento del número de usuarios y la creciente demanda de una nueva variedad de servicios ha hecho surgir una serie de necesidades que no podían ser cubiertas por los sistemas de comunicaciones móviles de segunda generación (2G). Ante tales expectativas se acometió, ya en el año 1999 y a nivel mundial, el desarrollo de una nueva generación de sistemas de comunicaciones móviles, capaces de superar las limitaciones de sus predecesores. Los sistemas de tercera generación (3G), cuya versión europea recibe el nombre de UMTS (*Universal Mobile Telecommunications System*), soportan una mayor capacidad y mayores velocidades de transmisión.

Este aumento en la velocidad de transmisión ha permitido el desarrollo de nuevos servicios, como la videotelefonía, el acceso a Internet o la transmisión de video bajo demanda. Estos nuevos servicios multimedia se han convertido en un producto clave para las operadoras de telecomunicaciones, no sólo por la ventaja competitiva que puede representar un buen servicio respecto a la competencia, sino sobre todo por los elevados ingresos repercutidos. Por ejemplo, según el informe de la CMT, el peso de los ingresos por servicios de descarga, navegación y transmisión de datos sobre el total de ingresos de datos creció hasta el 16% en 2006. Son precisamente los servicios de datos no SMS (*Short Message Service*) los que se están convirtiendo en el factor clave de este crecimiento. Los ingresos por estos servicios alcanzaron un crecimiento interanual de casi un 45% con respecto al 2005 y supusieron en el cuarto trimestre de 2006 casi el 42% de los ingresos de datos de las compañías.

Sin embargo, el crecimiento y masificación del uso de los servicios multimedia lleva implícito una necesidad creciente en la capacidad de transmisión y en la calidad de servicio que los sistemas de comunicaciones móviles deben ser capaces de proporcionar. Los sistemas de tercera generación son capaces de cubrir estas necesidades siempre que el despliegue de estaciones base, también denominados Nodos B, sea el suficiente. Sin embargo, aumentar el número de estaciones base resulta muy costoso, no sólo por la inversión en equipamiento necesaria, sino también por el arrendamiento de las azoteas y la creciente reticencia de las comunidades de vecinos ante el hipotético peligro para la salud pública de las radiaciones electromagnéticas.

Ante la creciente dificultad de situar más antenas en las azoteas y el imparable incremento en la demanda de servicios móviles de mayor capacidad, parece razonable que la única solución es recurrir a emplazamientos a nivel de calle, utilizando mucha menor potencia de transmisión y reduciendo las dimensiones de las antenas y, por tanto, el impacto visual. Es lo que se denomina despliegue micro-celular. Una micro-celda tiene un radio de cobertura de entre 100 y 300 m., ya que las limitaciones en la potencia máxima de transmisión y la situación no dominante en altura provocan que la señal se vea fuertemente atenuada en su propagación, reduciéndose por tanto el radio de cobertura respecto a una macro-celda.

Por otro lado, el elevado ancho de banda utilizado en WCDMA (*Wideband Code Division Multiple Access*), la tecnología de acceso radio de UMTS, permite una alta resolución temporal en el receptor *Rake*, por lo que en recepción se pueden separar las distintas componentes multicamino aprovechando al máximo la potencia transmitida. Esta diversidad en tiempo es también conocida como micro-diversidad, diversidad multicamino o diversidad contra desvanecimientos rápidos. Se denomina así porque el desvanecimiento rápido, o *fast fading*, está incorrelado en las distintas ramas del receptor *Rake* de manera que, si una rama está altamente atenuada, las otras pueden compensar este desvanecimiento. La diversidad multicamino es un elemento básico en el buen funcionamiento de UMTS, por lo que un gran problema en el despliegue de las redes micro-celulares, agravado aun más en escenarios pico-celulares, es la pérdida de esta diversidad en tiempo y la consiguiente reducción de la ganancia por diversidad conseguida en el receptor *Rake*. Este hecho fue ratificado en [2] donde, a partir de medidas experimentales, se demostró que en un escenario UMTS micro-celular el receptor *Rake* dispone de la información proveniente de un sólo camino de propagación al menos durante el 40% del tiempo.

Otro problema de las redes jerárquicas son los usuarios de alta velocidad. A estos usuarios no se les puede dar servicio desde las micro-celdas porque los procedimientos de trasposos son demasiado lentos como para poder garantizar la continuidad en el servicio. La solución planteada a este problema suele pasar siempre por asignar los usuarios móviles a las redes macro-celulares, desaprovechando la gran capacidad de las micro-celdas.

Estos problemas han retrasado el despliegue de las estructuras jerárquicas en las redes UMTS, estructuras hoy por hoy imprescindibles para la provisión de servicios en puntos de alta densidad de usuarios. En [3] se resalta la importancia de este despliegue micro-celular y, sobre todo, el necesario desarrollo de algoritmos de gestión de recursos adaptativos que permitan optimizar el uso de tales emplazamientos.

Esta preocupación por los aspectos relacionados con el despliegue de redes jerárquicas macro-micro celulares, al igual que la adaptación necesaria de la arquitectura de acceso radio UMTS a tales estructuras, ha abierto una incipiente línea

de investigación en la cual se encuadra esta Tesis Doctoral. En este sentido, en la presente Tesis Doctoral se abordan diversos aspectos relacionados con una nueva propuesta de arquitectura de acceso radio UMTS, denominada arquitectura Multinodo B (MNB), caracterizada por la utilización de antenas distribuidas. Además de analizar las prestaciones de esta arquitectura y la ganancia asociada a la diversidad espacial introducida, en esta Tesis Doctoral se define un nuevo algoritmo de gestión de recursos radio adaptativo basado en redes neuronales, necesario para el correcto funcionamiento de la nueva arquitectura.

## **1.2 Estado del Arte**

Antes de poder formular los objetivos de la Tesis Doctoral es necesario analizar el estado del arte. Esta expresión, heredada del inglés *State of the Art*, denota el nivel más alto de desarrollo de un dispositivo, técnica o campo científico, alcanzado en un tiempo determinado. El análisis del estado del arte es, sin duda, una de las primeras etapas que debe desarrollarse dentro de una investigación. Se deberá determinar cómo ha sido tratado el tema de investigación hasta la fecha, cómo se encuentra en el momento de realizar la investigación y cuáles son las tendencias. Este apartado ofrece una visión sucinta del estado del arte de los distintos aspectos tecnológicos involucrados en la presente Tesis Doctoral para, a partir de este estudio, poder plantear los objetivos de la misma, de manera que se garantice el avance de la ciencia.

### **1.2.1 Arquitectura de redes UMTS**

La principal idea que maneja esta Tesis Doctoral consiste en modificar la arquitectura radio UMTS añadiendo micro-celdas distribuidas mediante radio sobre fibra. A continuación se analiza el estado del arte al respecto de estos tres conceptos, la arquitectura UMTS, el desarrollo micro-celular y la distribución de la información mediante radio sobre fibra.



La tecnología UMTS superó su proceso de estandarización en el año 1999 y en marzo de 2000 el Gobierno adjudicó las doce licencias disponibles de frecuencia a cuatro operadoras en España. Tuvieron que pasar cuatro años de investigación, de desarrollo de las redes y de pruebas piloto para testar el producto hasta que en febrero de 2004 se anunciara el lanzamiento en España de los primeros servicios de UMTS. Sin embargo, no sería hasta finales de 2005 y principios de 2006 cuando los servicios UMTS llegaron al gran público.

A día de hoy, cualquier persona puede comprar en una tienda un terminal UMTS o una tarjeta de datos para que los ordenadores portátiles se conecten a Internet a velocidad de banda ancha y en movilidad. Aunque la cobertura es limitada y el despliegue no se ha dado por concluido, se puede afirmar que la red UMTS está disponible en las 52 capitales de provincia españolas. Asimismo, desde los últimos meses se dispone de datos reales sobre el comportamiento de las redes de acceso radio UMTS e igualmente se ha iniciado la fase de optimización en la que se introducirán mejoras en los métodos de planificación, implementación y gestión de las redes. A pesar de estos cambios no se deberían producir modificaciones sustanciales en el estándar, que en este momento cuenta ya con 5 revisiones publicadas [4].

Uno de los aspectos más novedosos que se ha introducido en la tecnología de estaciones base UMTS es la posibilidad de que los nodos B – estaciones base del sistema UMTS – se desarrollen con arquitectura compartida, pudiendo por ejemplo compartir elementos de procesamiento entre distintos sectores conectados en dicho nodo, así como la posibilidad de compartir amplificadores de potencia entre distintos sectores e incluso distintas portadoras del mismo nodo B. En el caso de macro-celdas, los nodos B más típicos pueden alojar hasta 6 tarjetas de radiofrecuencia (RF) en un mismo bastidor (en el argot se les llama "portadoras", aunque se corresponden con las celdas lógicas de UMTS), y tienen la posibilidad de configurar hasta 7 sectores en el mismo emplazamiento y estar dotados de hasta 120 elementos de canal por celda lógica, elementos que se pueden utilizar de forma compartida [5].

Esta alta flexibilidad *hardware* de los nodos B hace que sea muy sencillo generar múltiples replicas de la señal ya modulada e introducir las diferentes replicas en

equipos de radio sobre fibra que distribuyan esta información a equipos remotos encargados únicamente de la amplificación. Este concepto es muy similar al aplicado en la amplificación remota de nodos B [6], lo que corrobora la viabilidad de esta solución.

Básicamente, los sistemas radio sobre fibra se aprovechan de la fibra óptica para transportar señales de radio directamente hasta un punto remoto de radiación [7]. Dependiendo de la aplicación considerada, las señales radio pueden ser VHF, UHF, microondas o incluso ondas milimétricas. En cuanto a la fibra óptica, éste es el medio de transmisión ideal por varias razones. En primer lugar, sus bajas pérdidas (inferiores a 0,25 dB/km. para 1550 nm.), unido a la existencia de amplificadores ópticos de fibra dopada con erbio, permite alcanzar grandes distancias. En segundo lugar, su inmunidad frente a las interferencias electromagnéticas. Y en tercer lugar, su gran ancho de banda de transmisión únicamente limitado por la dispersión cromática.

Entregando las señales de radio directamente, el enlace de fibra evita así la necesidad de generar la portadora de alta frecuencia en la estación base, el cual es normalmente un emplazamiento de difícil acceso. Para cualquier sistema de transmisión, cuestiones prácticas como tamaño, peso, fiabilidad, coste y consumo del equipamiento remoto son de vital importancia. Así, la gran ventaja de los sistemas radio sobre fibra es su aptitud para concentrar la mayoría del equipamiento más costoso de alta frecuencia en un lugar centralizado (estación de control o lo que sería la unidad central del MNB), permitiendo que el resto de equipamiento del emplazamiento remoto sea simple, ligero, de pequeño tamaño y bajo consumo. Esto redundará en una instalación más rápida y sencilla, con baja necesidad de mantenimiento y con mayor simplicidad en la provisión de potencia eléctrica. Al mismo tiempo, la centralización del equipamiento de alta frecuencia aumenta la flexibilidad y el potencial de reutilización o distribución entre usuarios.



**Figura 1.1 Equipo de radio sobre fibra**

Existen en el mercado soluciones para la ubicación de cabezales de RF alejados del bastidor del Nodo B. Éstos se conectan mediante fibra óptica a la unidad principal y constan del transceptor radio, amplificador de potencia y unidad de acoplamiento a la antena. En la Figura 1.1 se muestra un ejemplo de una solución comercial de radio sobre fibra.

También están disponibles en la actualidad estaciones base UMTS de reducidas dimensiones aptas para un despliegue micro-celular y para su instalación en puntos de alta concentración de tráfico con el objetivo de reforzar la cobertura de la red. En la Figura 1.2 se muestra un ejemplo de caseta de reducidas dimensiones (60cm. x 70cm.) al igual que una antena omnidireccional para entornos micro-celulares (62cm.) que fácilmente puede ser camuflada dadas sus dimensiones.

Ya que tecnológicamente es posible tanto realizar un despliegue de micro-celdas o unidades remotas de reducidas dimensiones como distribuir la información entre ellas utilizando tecnología de radio sobre fibra, los objetivos que plantea esta Tesis Doctoral se corresponden con una solución tecnológicamente viable y de posible interés para los fabricantes de elementos de red UMTS.



**Figura 1.2 Equipos de UMTS para entornos micro-celulares**

## 1.2.2 Diversidad en transmisión en sistemas 3G

Las técnicas de diversidad en transmisión son un elemento clave en los sistemas 3G WCDMA. Dentro de las técnicas de diversidad en transmisión, la diversidad espacial se introduce en la señal transmitida mediante la utilización de múltiples antenas en transmisión. Dado que se pretende aprovechar el despliegue micro-celular del Multinodo B para garantizar una elevada diversidad en transmisión espacial y siendo un requisito de esta Tesis Doctoral el mantener una continuidad con el estándar de UMTS, es imprescindible analizar en profundidad las especificaciones del 3GPP (3rd *Generation Partnership Project*), organismo responsable de la estandarización de UMTS, al respecto.

El 3GPP no obliga a los terminales móviles a disponer de elementos de diversidad en recepción, sin embargo, el móvil debe ser capaz de aprovechar la diversidad en transmisión del Nodo B. Las mejores prestaciones en las técnicas de diversidad en transmisión se consiguen cuando la estación base tiene un conocimiento preciso del canal en enlace descendente. Sin embargo, al ser UMTS un estándar FDD, las

estimaciones en enlace ascendente no son necesariamente válidas para el canal en enlace descendente.

Existen dos tipos de técnicas de diversidad en transmisión:

- Técnicas en bucle cerrado, en las que el terminal realiza una estimación del canal en enlace descendente y envía dicha información en el enlace ascendente a la estación base.
- Técnicas en bucle abierto, más sencillas de implementar, más robustas frente a errores en la estimación de canal y en las que se utilizan técnicas de transmisión espacio-temporales.

Las principales técnicas de diversidad en transmisión son las siguientes:

- Diversidad de modo de transmisión en bucle cerrado, CLMTD (*Closed Loop Mode Transmit Diversity*).
- Diversidad de transmisión espacio-temporal, STTD (*Space-Time Transmit Diversity*).

### **Diversidad de Transmisión Espacio-Temporal**

Esta técnica es, sin duda, la técnica de diversidad en transmisión más extendida por su sencillez y por la importante ganancia que es capaz de conseguir. Esta técnica introduce diversidad en transmisión entre antenas de la misma celda. La misma información se transmite en antenas distintas (diversidad espacial) y en instantes de tiempo distintos (diversidad temporal). En concreto, cuando se utiliza esta forma de diversidad STTD, los símbolos QPSK que se transmitirían si sólo existiese una antena se envían por ambas antenas: por la primera se transmiten tal como son mientras que por la segunda se transmiten alternando el orden en parejas y, además, conjugados (considerando cada símbolo QPSK como un número complejo) y uno de los dos cambiado de signo. Al retransmitir los símbolos cambiados de signo y conjugados, se facilita la recuperación de los símbolos en recepción.

En la Figura 1.3 se resume el funcionamiento del transmisor STTD.

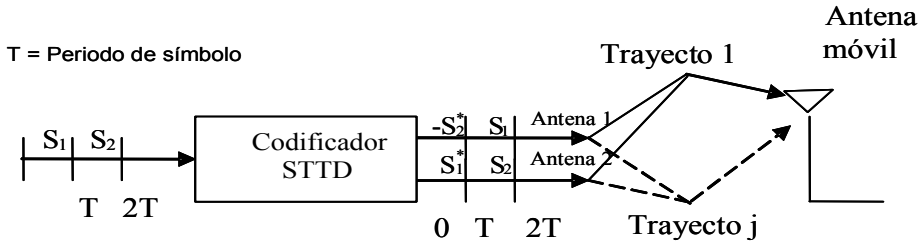


Figura 1.3 Funcionamiento del transmisor STTD [8]

En recepción, en el intervalo temporal  $[0, T]$  el terminal móvil recibirá:

$$r_1 = \alpha S_1 + \beta S_2 + n_1 \tag{1.1}$$

Mientras que en el intervalo  $[T, 2T]$  se recibirá la siguiente señal:

$$r_2 = -\alpha S_2^* + \beta S_1^* + n_2 \tag{1.2}$$

Para obtener la señal transmitida, en recepción se deberá operar de la siguiente manera con las señales recibidas:

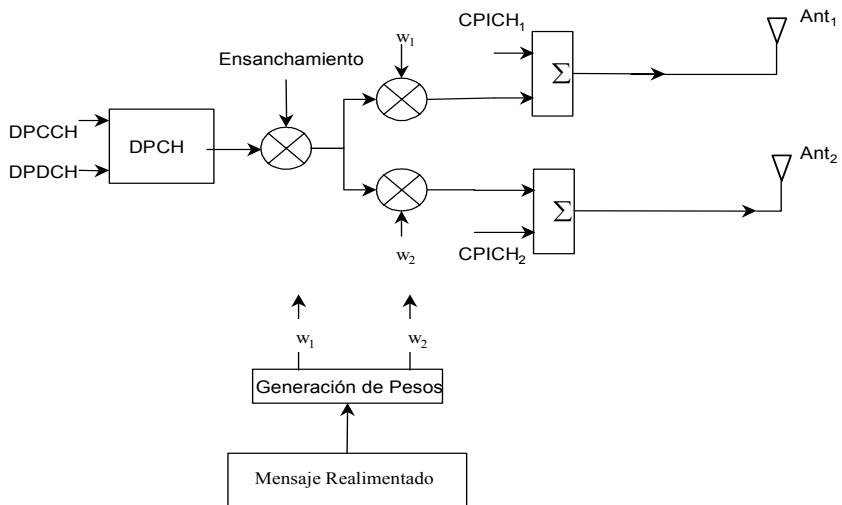
$$\begin{aligned} S_1 &\rightarrow \alpha^* r_1 + \beta r_2^* \\ S_2 &\rightarrow \beta^* r_1 - \alpha r_2^* \end{aligned} \tag{1.3}$$

La codificación STTD se puede aplicar a todos los símbolos salvo a los símbolos piloto, ya que cada antena transmite símbolos pilotos distintos.

### Diversidad de Transmisión en Bucle Cerrado

Esta técnica ofrece mejores prestaciones que la anterior, aunque a costa de una mayor complejidad en el transmisor y la necesidad de realimentar una serie de cálculos realizados por el terminal móvil.

En esta técnica de transmisión se envía distinta información piloto por cada una de las antenas, por lo que el equipo terminal puede realizar una estimación distinta del canal para ambas contribuciones a partir de la información de cada CPICH (*Common Pilot Channel*). A partir de aquí, el móvil realiza una serie de cálculos y envía hacia la red los parámetros de estima de canal que ha realizado.



**Figura 1.4 Funcionamiento del transmisor en bucle cerrado [8]**

De esta manera, tras analizar la realimentación del terminal, el Nodo B anticipa la corrección de la respuesta del canal de forma que se garantice una combinación coherente por parte del receptor. Gracias a esta técnica se puede llegar a conseguir una importante reducción de los requerimientos de  $E_b/N_0$  para una determinada calidad de servicio.

Son muchos los artículos encontrados en la literatura que abordan el tema de la diversidad espacial en transmisión. La mayor parte de ellos sin embargo se centran en un estudio para canales dispersivos planos en frecuencia (*flat fading*) [9]-[12]. En todos ellos se concluye que para canales poco dispersivos las ganancias ofrecidas por la diversidad espacial en transmisión son muy elevadas.

Otros trabajos sí que están centrados en el análisis de técnicas de diversidad espacial en WCDMA [13]-[19]. Las principales conclusiones alcanzadas en estos artículos son:

- Existe una enorme mejora en las prestaciones del canal radio al utilizar las distintas técnicas de diversidad en transmisión en WCDMA.

- Se puede conseguir una reducción en la  $E_b/N_0$  para un valor de BLER (*Block Error Rate*) fijo de entre 2 y 5 dB, dependiendo de la velocidad del móvil, el canal considerado y la potencia transmitida por las antenas.
- Las técnicas de diversidad en bucle cerrado son mucho más potentes que las de bucle abierto, sobre todo para velocidades del terminal móvil inferiores a 50km/h, por lo que son más interesantes en entornos urbanos.
- Se reconoce cierta reducción de ganancia en las técnicas de diversidad cuando se consideran entornos multicamino por el hecho de utilizar el mismo código de ensanchado para ambas antenas. Además, los errores cometidos en la estimación de canal es un aspecto no abordado en estos análisis que normalmente suponen una estimación perfecta.

Es importante destacar que, hasta la fecha, no se ha realizado un estudio riguroso de la ganancia obtenida gracias a las técnicas de diversidad estandarizadas para el conjunto de servicios “clásicos” definidos en el 3GPP [20]. Por otro lado, tampoco se ha realizado ningún estudio sobre la ganancia obtenida cuando las dos antenas que participan en la diversidad se separan una distancia mayor que  $10\lambda$  ( $\approx 1.5\text{m}$ ).

Una vez conocidas las técnicas de diversidad estandarizadas por el 3GPP, es importante destacar que en esta Tesis Doctoral se ha utilizado una variante de la técnica de diversidad en bucle cerrado (al ser la más potente y mejor adaptada para entornos urbanos) compatible con ésta. Además, se ha propuesto un nuevo mecanismo de estimación de canal específico para la transmisión con diversidad en bucle cerrado.

### 1.2.3 Estimación de canal

La transmisión de información sobre canales limitados en banda deterministas (fijos en el tiempo) o dinámicos (variables en el tiempo) es posible mediante el uso de técnicas de equalización que permiten compensar el efecto de dispersión temporal o distorsión en frecuencia introducido por el canal.



La estimación de canal en los sistemas de tercera generación WCDMA (UMTS) se basa en el conocimiento de una secuencia de bits piloto única para cada transmisor y que se repite en la transmisión de cada *slot*.

En un receptor *Rake* convencional, las estimaciones del canal se utilizan para combinar las distintas contribuciones de la señal recibida mediante la técnica de combinación MRC (*Maximum Ratio Combining*). El objetivo es estimar la fase y la amplitud de cada canal identificado. A continuación, esta información se utiliza para combinar cada contribución de la siguiente forma: si  $c_j$  es la estima de canal,  $d_j$  el retardo estimado de la contribución,  $J$  es el número de contribuciones consideradas,  $x_j$  es la secuencia de desensanchado en el camino  $j$ -ésimo y  $det(\bullet)$  es un simple detector o decisor, entonces la secuencia de símbolos obtenidos a la salida del receptor *Rake* es:

$$\hat{a} = det\left(\sum_{j=0}^{J-1} \hat{c}_j^* x_j(\hat{d}_j)\right) \quad (1.4)$$

En WCDMA la estimación de canal se puede realizar utilizando los bits piloto (de 1 a 8 símbolos) multiplexados dentro de cada *slot* del canal físico dedicado de control (DPCCH – *Dedicated Physical Control Channel*) [21]. La estimación de canal se realiza por comparación directa entre los bits piloto recibidos y los bits piloto esperados. Siendo  $I_n$  los símbolos piloto transmitidos en un *slot* y  $v_n$  los símbolos piloto recibidos, entonces la estimación del canal es:

$$\hat{c}_j \approx \frac{1}{N} \frac{\sum_{n=1}^N v_n I_n^*}{\sigma_j^2} \quad (1.5)$$

donde  $N$  es el número total de símbolos piloto. Además, tal y como se demuestra en [22], un promediado ponderado multi-*slot* (WMSA – *Weighted Multi-Slot Average*) de las estimas del canal reduce hasta en 0.8dB los requerimientos de  $E_b/N_0$  para un valor de BLER fijo. Este promediado multi-*slot* realiza un filtrado paso bajo capaz de reducir en gran medida el error cometido en la estimación de canal. Una vez se dispone de las muestras de canal en cada *slot*, la estimación del canal experimentado en los bits situados entre dos secuencias consecutivas de bits piloto (dos *slots*) se pueden obtener por mera interpolación. Cuando se utiliza interpolación lineal, la fase y la amplitud de

las estimas del canal se interpolan por separado para evitar así las rotaciones de fase de  $180^\circ$  producidas al cambiar de signo la magnitud.

Por otro lado en [23] se destaca la enorme utilidad de canal piloto común (CPICH) para la realización de la estimación de canal, pudiendo combinar el conocimiento de ambas señales piloto, la común y la del canal dedicado de control.

Para garantizar el buen funcionamiento de la nueva arquitectura radio propuesta, se deberá tener un conocimiento lo más amplio posible del canal para las distintas contribuciones recibidas, por lo que será necesario desarrollar la técnica de estimación de canal más adecuada para el Multinodo B.

## 1.3 Objetivos Planteados

Una vez analizada la motivación y el estado del arte en los aspectos que ocupan a esta Tesis Doctoral, es momento de plantear los objetivos de la misma.

Las soluciones que existen en la actualidad en tecnología de antenas distribuidas basadas en el transporte óptico de señales de radiofrecuencia son empleadas principalmente para dar cobertura al interior de edificios y grandes recintos, o bien para extender la cobertura celular a túneles y otros entornos de difícil acceso [25]. En estos casos las unidades remotas actúan a modo de repetidor o como simple extensión de la conexión entre la antena y el correspondiente equipo de radiofrecuencia. Esta configuración presenta ventajas adicionales en dichos entornos en cuanto a la mejora de capacidad intrínseca debida al incremento del número de puntos de acceso a la red, aunque los operadores de GSM despliegan esta tecnología con el único objetivo de incrementar o mejorar la cobertura.

Sin embargo, en UMTS cada celda del sistema transmite utilizando la misma portadora, aunque con códigos de celda distintos. La limitación en capacidad del enlace ascendente se debe principalmente a las interferencias que usuarios de la propia celda y de celdas vecinas producen en la estación base, mientras que en enlace descendente dicha limitación está provocada por las interferencias que recibe el móvil de otras

estaciones y de otros canales de la propia estación [26]. UMTS es por tanto un sistema limitado por interferencias. Cualquier estructura que reduzca el número de interferentes que capta un móvil o una estación base UMTS, garantiza una mejora en la capacidad.

La distribución de antenas formando celdas más pequeñas permite transmitir menor potencia, desde ubicaciones menos dominantes y con un nivel de interferencias externas inferior a la configuración tradicional basada en macro-celdas.

El objetivo principal de la Tesis Doctoral es, por tanto, umentar la capacidad, flexibilidad y calidad de las redes de acceso radio UMTS estandarizadas por el 3GPP [4], definiendo una nueva arquitectura radio, denominada Arquitectura Multinodo B (MNB), que modifica la configuración de las estaciones base, o nodos B, gracias a una estructura de antenas distribuidas espacialmente en el área de cobertura.

Desde el punto de vista de la red UMTS, el Multinodo B será percibido como un único emplazamiento, aunque esté conformado por un número de celdas mayor que el habitual. A tal efecto, se desarrollará la unidad de control necesaria para adaptar los procesos de gestión internos del nodo B a la interfaz estándar Iub [27].

Los enlaces entre el nodo B y las antenas que conforman cada una de las celdas del nodo se realizarán mediante tecnología óptica, aunque alternativamente podrían ser sustituidos por coaxial en determinadas aplicaciones en las que las distancias sean menores, como en el caso de pico-celdas *indoor*.

El esquema de Multinodo lo completa un bloque sencillo de procesado encargado de distribuir los recursos entre las antenas. En este caso, este algoritmo de gestión de recursos no es convencional, ya que las antenas están distribuidas por la celda y no agrupadas en el mismo panel, como en el caso de antenas adaptativas.

Teniendo en cuenta el objetivo principal mencionado, pueden definirse los siguientes objetivos parciales:

- Profundizar en los mecanismos más adecuados de transmisión que garanticen el máximo aprovechamiento de la diversidad multicamino.
- Desarrollar las técnicas de estimación de canal adecuadas a la estructura Multinodo.

- Caracterizar la ganancia conseguida por la diversidad en transmisión añadida.
- Definir los algoritmos de gestión adaptativa de recursos del Multinodo B, respetando al máximo las interfaces recogidas en el estándar para garantizar la integración de estos nuevos elementos en redes UMTS ya operativas.
- Evaluar el funcionamiento completo del sistema MNB, comparando sus prestaciones con la arquitectura UMTS convencional.

Gracias al trabajo desarrollado en esta Tesis Doctoral y la evaluación de la nueva arquitectura de acceso radio, se dispondrá de todos los elementos necesarios para evaluar la idoneidad de la propuesta y poder realizar la implementación de un prototipo del sistema.

## 1.4 Principales Aportaciones de la Tesis

Este apartado pretende analizar qué aporta el desarrollo de la presente Tesis Doctoral en la evolución de la tecnología de acceso radio UMTS. La nueva propuesta de arquitectura de acceso radio utiliza la estructura de nodos B convencional ya existente definida por el 3GPP [4] añadiendo la conexión de varias antenas físicamente distribuidas en el área de servicio, formando una estructura equivalente a una agrupación distribuida de antenas. La incorporación de enlaces ópticos y gestión de recursos dinámica garantiza una estructura flexible y adaptativa, una estructura óptima de acceso radio.

Además, contar con antenas distribuidas espacialmente incrementa la percepción de diversidad en ambos enlaces tanto descendente como ascendente. En el enlace ascendente, gracias a la estructura MNB, se dispone de un conjunto de antenas en recepción conectadas a la misma estructura *Rake* comportándose de manera similar a tener múltiples caminos de propagación recibidos por la misma antena. De esta manera es posible aprovechar al máximo toda la energía de la señal de interés disponible en el canal. Por otro lado en sentido descendente, mediante la aplicación de técnicas de

procesado, es posible una conformación inteligente del perfil de potencia (PDP - *Power Delay Profile*), es decir, ajustar los retardos y amplitudes de cada una de las contribuciones recibidas de la señal.

La configuración MNB permite disponer de canales incorrelados en recepción debido a la distribución espacial de las antenas. Para poder determinar el beneficio de esta distribución es fundamental evaluar las prestaciones, en términos de ganancia, del receptor *Rake* al contemplar dicha diversidad espacial. A primera vista puede parecer que tanto la diversidad por multiplicidad de caminos y antenas son análogas. No obstante, debe pensarse que el MNB está orientado a entornos micro o pico-celulares donde el *delay-spread* que caracteriza el canal es del orden de la duración de *chip* (0.26  $\mu$ s) - en entornos micro-celulares este valor puede ser superior-, por lo que, al no poder detectarse más de una contribución en el receptor *Rake*, se carece de diversidad por multicamino. De esta manera, la tecnología de acceso basada en Multinodos proporciona un incremento muy importante en la diversidad experimentada, pudiendo procesar y combinar diferentes señales en recepción.

Otros beneficios de la nueva arquitectura de acceso radio, y por tanto de la Tesis Doctoral son:

- La mejora en el despliegue de cobertura en redes UMTS, puesto que la estructura de células homogéneas no reporta ventajas en entornos donde el tráfico no es homogéneo en su distribución espacial ni en el tipo de servicios. Las unidades remotas, o antenas distribuidas, pueden ubicarse en posiciones óptimas para conformar el área de servicio o cobertura del nodo B que mejor se adapte a cada entorno. El Multinodo B podrá gestionar dinámicamente qué antenas forman parte de una misma célula en cada momento, convirtiéndose en una parte de la red de acceso con celdas reconfigurables, reduciendo el número de antenas de una celda -y por lo tanto su área de servicio- en condiciones de saturación o aumentar éste si el tráfico es suficientemente bajo.
- Se consigue una reducción del impacto ambiental, al transmitir menor potencia por antena para un mismo nivel de capacidad. Los emplazamientos son equiparables a los actuales repetidores de una o varias celdas del nodo B, lo que

permite trabajar con antenas en ubicaciones menos privilegiadas e instalar antenas con menor impacto visual.

- Mejora el comportamiento de las redes micro-celulares, tanto desde el punto de vista del diseño de la red de acceso como de la gestión de recursos. El *handover* entre antenas será siempre *softer* puesto que se producirá entre antenas del mismo nodo. Además la ganancia por combinación mejorará al tratarse de antenas en ubicaciones distintas.
- Tiene aplicación directa en redes *indoor* ya que en el interior de edificios o grandes recintos es habitual disponer de una mayor facilidad para realizar el cableado adecuado de múltiples antenas a un único nodo. Además, el comportamiento del receptor *Rake* en *indoor* puede mejorarse, puesto que será posible forzar una mayor dispersión del multitrayecto mediante la adición controlada de un retardo a las señales transmitidas desde las antenas del mismo nodo.
- Permite proveer macro-diversidad por *soft handover* en canales de transporte comunes y compartidos en los que, según el estándar del 3GPP, no está recogido el uso de esta técnica.
- Se puede aprovechar el despliegue de estructuras jerárquicas convencionales, donde las distintas capas tienen asignadas distintas frecuencias, desplegando también el MNB, de manera que se mejore la capacidad de cada portadora.
- Existen otras aplicaciones derivadas de esta configuración, como por ejemplo la localización de los equipos terminales móviles, puesto que el nodo B puede determinar distancias o pseudo-rangos a cada punto de acceso, con todas las consecuencias que se podrían derivar a servicios orientados a localización y a gestión de movilidad.

## 1.5 Estructura de la Tesis

El resto del documento, o memoria de la Tesis, se organiza de la siguiente forma:

### **Capítulo 2: Descripción de la Arquitectura Multinodo B**

En él se presentará la nueva arquitectura radio Multinodo B. Para ello se describirán, en primer lugar, dos arquitecturas de protocolos para del MNB. La primera de ellas, denominada MNB-pasivo, requerirá menos cambios en el Nodo-B, mientras que la segunda arquitectura, denominada MNB-activo, supondrá añadir más capacidades de procesado al Nodo-B. En cualquier caso, ambas configuraciones serán totalmente transparentes para la red UMTS *Release 99*, garantizado así la compatibilidad con el sistema actual UMTS. Finalmente, se realizará un análisis minucioso de todas las entidades funcionales que componen el Multinodo B.

### **Capítulo 3: Evaluación de las Prestaciones de la Arquitectura MNB**

Tras describir y validar el simulador de enlace utilizado, en este capítulo se explicará cómo se sincroniza la transmisión desde el MNB para obtener la máxima ganancia por diversidad. Además, se mostrará un mecanismo para estimar el canal a partir de la señal piloto recibida. Gracias a estas nuevas técnicas, el MNB dotará de una ganancia en transmisión adicional, ganancia que dependerá de diversos factores, tal y como se demostrará en este capítulo. Considerando una serie de servicios de referencia de UMTS, se calcularán sus *Look-up Tables* y se obtendrá la ganancia del MNB para cada escenario. Esta ganancia se deberá incluir en las simulaciones a nivel de sistema para evaluar el impacto final del MNB en UMTS.

### **Capítulo 4: Gestión de Recursos en Redes UMTS**

Una vez analizada en profundidad la ganancia por diversidad que la nueva arquitectura propuesta en esta Tesis Doctoral introduce en el sistema UMTS, aun queda

por resolver otro de los retos que plantea esta nueva arquitectura, es decir, el reparto óptimo de los recursos disponibles en la red, de manera que se pueda aprovechar al máximo esta ganancia. En este capítulo se estudiará el papel de los distintos algoritmos de gestión de recursos en el sistema convencional UMTS. Además, se abordará de forma genérica la optimización de la gestión de recursos en UMTS mediante el uso de redes neuronales de Hopfield. Los resultados obtenidos demostrarán el enorme potencial del nuevo gestor de recursos propuesto en esta Tesis Doctoral.

### **Capítulo 5: Gestión de Recursos en Redes MNB**

En él se extenderá la formulación anteriormente descrita para redes UMTS convencionales, incluyendo ahora también la distribución de potencia entre las distintas antenas del MNB. Para cada una de las arquitecturas MNB propuestas, se evaluará un algoritmo específico, comparando los resultados obtenidos. Además, en este capítulo se realizará el análisis final de la nueva arquitectura MNB, del que se podrá extraer hasta qué punto esta propuesta es capaz de mejorar la capacidad del sistema.

### **Capítulo 6: Modelado del *Shadowing* y su Influencia sobre la Evaluación de Sistemas**

Este capítulo se centrará en el estudio de los modelos de pérdidas a largo plazo, puesto que este aspecto del modelado del canal radio ha sido objeto de un meticuloso análisis en esta Tesis Doctoral. Además de proponer un conjunto de modelos de creciente realismo y complejidad, este capítulo analizará el efecto del uso de los distintos modelos sobre la evaluación realizada de las prestaciones de un sistema, particularizando este estudio para el escenario de simulación propuesto en esta Tesis Doctoral. A partir de este análisis se podrá concluir la utilidad de los modelos complejos de *shadowing* en función del tipo de estudio que se pretenda realizar.



**Capítulo 7: Conclusiones y Líneas Futuras de Investigación**

En este capítulo se presentarán las conclusiones derivadas del estudio realizado, al tiempo que se propondrán distintas líneas futuras de investigación como continuación del trabajo desarrollado.

**Anexos:**

Al final de la memoria se incluyen 2 anexos. El Anexo A describirá con detalle el simulador de enlace empleado, mientras que el Anexo B hará lo propio con el simulador de sistema. El por qué de incluir la descripción de los simuladores en la Tesis Doctoral es debido a que muchos de los modelos propuestos en ambos simuladores han sido utilizados con posterioridad en distintos foros de trabajo europeos, como COST 273 y NEWCOM, lo que demuestra la importancia que el trabajo de modelado ha supuesto en el desarrollo de esta Tesis Doctoral.

## Referencias

- [1] Comisión del Mercado de las Telecomunicaciones. “Estadísticas del Sector, IV Informe Trimestral de 2006”. [www.cmt.es](http://www.cmt.es).
- [2] E.T.S.I., “Selection Procedures for the Choice of Radio Transmission Technologies of the UMTS”, UMTS 30.03 v3.2.0, pp 42 - 43, 2004.
- [3] E. Berruto et al., “Research Activities on UMTS Radio Interface, Network Architectures and Planning”, IEEE Communications Magazine, February 1.998.
- [4] 3GPP specifications, [www.3gpp.org](http://www.3gpp.org).
- [5] H. Kaanen, “UMTS Networks, Architecture, Mobility and Services”, Ed. Wiley, 2001.
- [6] J. Laiho, A. Wacker and T. Novosad, “Radio Network Planning and Optimisation for UMTS”, Ed. Wiley, 2005.
- [7] H. Al-Raweshidy, “Radio Over Fiber Technologies for Mobile Communications Networks”, Ed. Artech House, 2002.
- [8] 3GPP TS 25.214 V5.9.0, “Physical Layer Procedures (FDD)”, 2004.
- [9] S.M. Alamouti, “A Simple Transmit Diversity Technique for Wireless Communications”, IEEE Journal on Selected Areas in Communications, Vol. 16, No. 8, pp. 1451 - 1458, 1998.
- [10] J. Wu, H. Horng, H. J. Zhang, C. Xiao, “Adaptive Transmit Diversity with Quadrant Phase Constraining Feedback”, in Proc. of the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, Vol. 3, pp. 2262 - 2266, 2004.
- [11] J. Hamalainen, R. Wichman, “Asymptotic Bit Error Probabilities of some Closed-Loop Transmit Diversity Schemes”, in Proc. of the IEEE Global Telecommunications Conference, Vol. 1, pp 360 - 364, 2002.
- [12] P. Malm, B. Lindoff, “Theoretical Analysis of a Simple Transmit Diversity Scheme”, in Proc. of the IEEE Workshop on Signal Processing Advances in Wireless Communications, pp 26 - 30, 2003.
- [13] D.A. Novakovic, M.J. Juntti, M.L. Dukic, “Generalized Closed Loop Transmit Diversity Algorithm for WCDMA”, in Proc. of the IEEE International Symposium on Spread Spectrum Techniques and Applications, Vol. 3, pp. 797 - 801, 2002.

- 
- [14] B. Yang, X. Zhang, X. Dazhuan, "A Novel Space-Time Transmit Diversity Scheme for DS-CDMA Downlink in Multipath Channel", in Proc. of the Radio Science Conference, pp. 131 - 134, 2004.
- [15] L. Jalloul, K. Rohani, K. Kuchi, J. Chen, "Performance Analysis of CDMA Transmit Diversity Methods", in Proc. of the IEEE *Vehicular* Technology Conference, Vol. 3, pp. 1326 - 1330, 1999.
- [16] M. Canales, A. Valdovinos, J.R. Gallego, F. Gutierrez, "Performance Analysis of Downlink Transmit Diversity System Applied to the UTRA FDD Mode", in Proc. of the IEEE International Symposium on Spread Spectrum Techniques and Applications, Vol. 2, pp. 410 - 414, 2002.
- [17] R. Vishwakarma, K.S. Shanmugan, "Performance Analysis of Transmit Antenna Diversity in 3G WCDMA System", in Proc. of the IEEE International Conference on Personal Wireless Communications, pp. 1 - 4, 2000.
- [18] A. Hottinen, R. Wichman, "Transmit Diversity by Antenna Selection in CDMA Downlink", in Proc. of the IEEE International Symposium on Spread Spectrum Techniques and Applications, Vol. 3, pp. 767 - 770, 1998.
- [19] Y. Yangxin, X. You, "Closed Loop Transmit Diversity Techniques for WCDMA Systems", in Proc. of the International Conference on Neural Networks and Signal Processing, Vol. 2, pp. 1460 - 1463, 2003.
- [20] 3GPP TS 34.108 V5.3.0, "Common Test Environments for User Equipment (UE); Conformance Testing", 2004.
- [21] 3GPP TS 25.211 V3.12.0, "Physical channels and mapping of transport channels onto physical channels (FDD)", 2002.
- [22] H. Andoh, M. Sawahashi, F. Adachi, "Channel Estimation Using Time Multiplexed Pilot Symbols for Coherent *Rake* Combining for DS-CDMA Mobile Radio", in Proc. of the IEEE Personal Indoor and Mobile Radio Communications Conference, Vol. 3, pp. 954 - 958, 1997.
- [23] K.A. Qaraqe, S. Roe, "Channel Estimation Algorithms for Third Generation W-CDMA Communication Systems", in Proc. of the IEEE *Vehicular* Technology Conference, Vol. 4, pp. 2675 - 2679, 2001.
- [24] 3GPP TS 23.107 V3.9.0, "QoS Concept and Architecture", 2002.

- [25] J. M. Hernando, C. Lluch, "Comunicaciones Móviles de Tercera Generación", Telefónica Móviles España, 2000.
- [26] H. Holma, A. Toskala, "WCDMA for UMTS", John Wiley & Sons Ltd., 2000.
- [27] 3GPP TS 25.430 V7.0.0, "UTRAN Iub Interface: General Aspects and Principles", 2006.

# Parte I

## *Arquitectura Radio Multinodo B*

---



# Capítulo 2

## *Descripción de la Arquitectura Multinodo B*

*La ciencia es el simple sentido común llevado al máximo:  
observación cuidadosa y rigor ante las falacias lógicas*

*Thomas Henry Huxley*

### **2.1 Introducción**

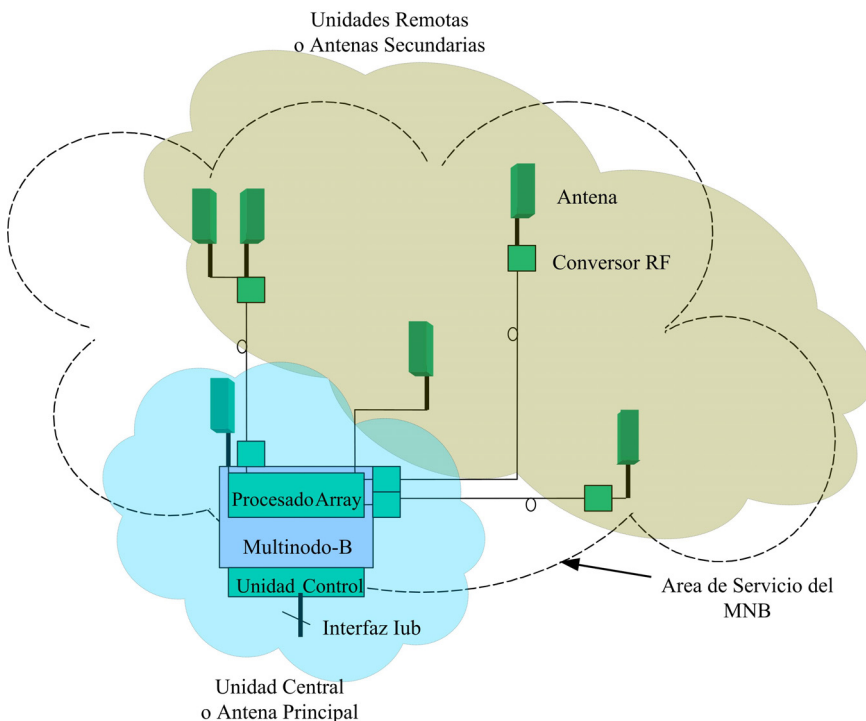
En la actualidad existe un gran número de grupos de investigación inmersos en la elaboración de nuevas tecnologías que permitan sacar el máximo partido a las capacidades de los sistemas de tercera generación. En este sentido, el Grupo de Comunicaciones Móviles (GCM) de la Universidad Politécnica de Valencia, en colaboración con otros grupos, llevó a cabo un proyecto de investigación, financiado parcialmente por la Comisión Interministerial de Ciencia y Tecnología (CICYT), denominado “Mejora de Arquitecturas de Acceso Radio UMTS Mediante Mutinodos B” [1]. Este proyecto de investigación, en cuyo seno se desarrolla la presente Tesis Doctoral, tiene como objetivo principal desarrollar una nueva configuración de la red de acceso, denominada Mutinodo B, que permita obtener mejores prestaciones que la arquitectura de red de acceso radio UMTS tradicional.

Tal y como se describió en el capítulo de introducción, la arquitectura Mutinodo B (MNB) utiliza la tecnología de nodos B ya existente y definida por el 3GPP [2] añadiendo la conexión de varias antenas físicamente distribuidas en el área de servicio, formando una estructura equivalente a una agrupación distribuida de antenas, como se puede ver en la Figura 2.1. La incorporación de enlaces ópticos, gestión de recursos

dinámica y procesado en *array* garantiza una estructura flexible y adaptativa, una estructura óptima de acceso radio.

La estructura Multinodo B incluye dos tipos de elementos: la unidad central y la unidad remota. La unidad central es la que concentra toda la inteligencia del sistema y además puede incluir una antena transmisora, la antena principal. La unidad remota incluye simplemente un conversor de radio sobre fibra, un amplificador bidireccional y la antena transmisora secundaria. Es muy importante para la implementación del Multinodo B que estas unidades remotas sean lo más compactas posible ya que su instalación está pensada a nivel micro-celular, es decir, en las propias calles, por lo que reducir su consumo de potencia y su tamaño es una cuestión de vital importancia.

Por cada unidad central puede haber varias unidades remotas. Tal y como se explicó en el Capítulo 1, para que el MNB sea compatible con las estructuras de amplificación y procesado actuales se ha considerado un máximo de hasta 7 celdas en el Multinodo por lo que podrá haber un máximo de 6 unidades remotas más la unidad central.



**Figura 2.1** Arquitectura Mutinodo B



Una vez planteada la nueva arquitectura, es necesario estudiar tres aspectos clave dentro de la Tesis Doctoral: (1) se ha de evaluar el impacto que la nueva configuración de la red tendrá sobre la arquitectura de protocolos y si realmente sería factible implementar una solución de estas características. (2) hay que determinar hasta qué punto se obtiene una ganancia en diversidad por la distribución de las antenas. (3) se tiene que idear un algoritmo de gestión de recursos radio que permita utilizar al máximo las capacidades de la nueva arquitectura radio.

Este capítulo abarca el primero de estos análisis y pretende describir la arquitectura Mutinodo B propuesta. Además, se evaluarán las modificaciones que se tendrían necesariamente que aplicar a la arquitectura de protocolos UMTS convencional para permitir el funcionamiento de la nueva configuración.

La sección 2.2 analiza dichos cambios en la arquitectura de protocolos de la interfaz radio, mientras que la sección 2.3 enumera las entidades que componen el MNB haciendo especial hincapié en cómo introducir en ellas las nuevas funcionalidades detectadas en la arquitectura de protocolos.

## **2.2 Arquitectura de Protocolos de la Interfaz Radio MNB**

### **2.2.1 Introducción**

Uno de los principales objetivos de la nueva arquitectura radio Multinodo B es modificar de manera mínima la arquitectura de red de acceso estandarizada, de manera que el coste de su implementación sea el menor posible y, además, sea compatible su despliegue con el uso de celdas convencionales.

Para salvaguardar esta compatibilidad, es fundamental estudiar la especificación llevada a cabo por el 3GPP respecto a la arquitectura radio y a la interfaz de protocolos. Una vez se conozca el proceso de toma de decisiones y el intercambio de datos entre las distintas entidades que conforman el subsistema de red radio RNS (*Radio Network*

*Subsystem*), será posible determinar qué entidades son objeto de modificación y de qué manera hay que abordar dichos cambios para optimizar el uso del MNB.

En general, se van a proponer dos nuevas arquitecturas radio MNB, una más sencilla que se ha denominado MNB-pasivo y otra más potente y compleja denominada MNB-activo. En la arquitectura MNB-pasivo, es el RNC (*Radio Network Controller*) quien realiza toda la gestión de recursos radio y el Nodo-B se encarga únicamente de distribuir los recursos entre las antenas que conforman el MNB, determinando la potencia transmitida desde cada antena. La arquitectura MNB-activo, más compleja, implementa un Nodo-B con muchas más capacidades que será capaz de determinar conjuntamente los recursos asignados y la distribución de potencias realizada.

Esta sección se compone de los siguientes apartados: el apartado 2.2.2 introduce los principales conceptos de la arquitectura de protocolos UMTS. Este análisis es fundamental ya que permitirá al lector entender qué cambios en la arquitectura radio implica las distintas alternativas propuestas en esta Tesis Doctoral. En el apartado 2.2.3 se analiza el mecanismo de transmisión de datos mediante conmutación de circuitos, distinguiendo los cambios necesarios en la arquitectura del sistema para introducir el uso de MNB dentro de la red UTRAN. El apartado 2.2.4 realiza un estudio similar al anterior pero centrándose en la transmisión por conmutación de paquetes mediante el uso del canal compartido DSCH. Finalmente, el apartado 2.2.5 analiza el procedimiento seguido en el Nodo-B para obtener toda la información, tanto de los usuarios como del sistema, necesaria para la toma de decisiones en la gestión de recursos radio.

## **2.2.2 Arquitectura de protocolos UMTS**

Los niveles superiores, o niveles de aplicación, que constituyen el estrato de no acceso NAS (*Non-Access Stratum*) hacen uso del conjunto de RABs (*Radio Access Bearers*) que la red de acceso radio pone a su disposición. Las RABs son los recursos que la red de acceso radio utiliza para la provisión de servicios. Las RABs se definen

de forma genérica, sin hacer relación a ninguna característica específica del acceso radio UMTS. Los parámetros de configuración incluyen:

- El tipo de tráfico.
- Las tasas binarias mínima y garantizada (en función de la clase de tráfico).
- El tratamiento que debe darse a los errores de transmisión, incluyendo la tasa de errores y tasa de errores residual (correspondiente a los errores no detectados) permitidas. Además debe especificarse si las SDUs (*Service Data Units*) erróneas se entregan o se descartan.
- La necesidad de entregar los paquetes en orden.
- Los formatos de SDUs permitidos, que se utiliza para la configuración de la interfaz Iu entre SRNC (*Serving RNC*) y red troncal.
- Finalmente, se especifican también las prioridades, tanto de admisión/retención (prioridad para el establecimiento de la comunicación, en relación con otros servicios) como de tráfico (prioridad para los datos en relación con los de otros servicios, una vez establecida la comunicación).

Una vez definidas las propiedades del servicio, es necesario ajustar los parámetros de la interfaz radio para garantizar dichas características. De esta manera, los protocolos de la interfaz radio son necesarios para configurar, modificar o liberar las portadoras radio, o lo que se conoce en inglés como *Radio Bearers*, RB.

La arquitectura de protocolos de la interfaz radio hereda una estructura similar a la pila de protocolos OSI [3]. El nivel más bajo, también denominado nivel 1, es el nivel físico, controlado por el nivel 2, el nivel de enlace de datos (*Data Link Layer*) y éste a su vez por el nivel 3, el nivel de red (*Network Layer*). En la interfaz radio de la UTRAN (*UMTS Terrestrial Radio Access Network*) con duplexado FDD (*Frequency Division Duplex*), la capa 2 se divide en varias subcapas. En el plano de control, el nivel 2 contiene dos subniveles, el protocolo de acceso al medio MAC (*Medium Access Control*) y el protocolo de control del radio enlace RLC (*Radio Link Control*). En el

plano de usuario, además de las capas MAC y RLC, existen dos protocolos adicionales que dependen del tipo de servicio: el protocolo de convergencia de paquetes PDCP (*Packet Data Convergence Protocol*) y el protocolo de control de las comunicaciones de difusión BMC (*Broadcast/Multicast Control Protocol*). A nivel 3 sólo se ha definido un protocolo de control, el protocolo de control de los recursos radio, el RRC (*Radio Resource Control*).

Las *Radio Bearers* sí que son específicas del acceso radio empleado. Básicamente, los parámetros de las RB corresponden a los parámetros de configuración de los distintos protocolos de la interfaz radio.

Para el protocolo PDCP, que se emplea únicamente para compresión de cabeceras en servicios de conmutación de paquetes PS (*Packet Switching*), la única configuración relevante es si se soporta el cambio de SRNC sin pérdidas (*lossless*) o no. En caso afirmativo debe garantizarse una posible transferencia de contexto PDP (*Packet Data Protocol*) al nuevo SRNC.

Para el protocolo RLC, el parámetro de configuración básico es la selección de uno de los modos de operación, en función de las características del servicio.

La configuración de la capa MAC incluye la asignación entre canales lógicos y de transporte en función de los requerimientos del servicio. Como se puede multiplexar varios canales lógicos sobre un mismo canal de transporte, será preciso asignar también prioridades.

Tanto la capa MAC como la capa física se configuran con los TFCSs (*Transport Format Combination Sets*) correspondientes a los distintos canales de transporte, que regulan el intercambio de información entre ambas capas: duración de la trama, número de bloques, tamaño del CRC (*Cyclic Redundancy Code*), tipo de codificación de canal, etc.

Finalmente, un parámetro muy importante de la capa física es el factor de ensanchado, SF (*Spreading Factor*), que influye directamente en la tasa binaria a nivel físico.

Una vez que el núcleo de red ha enviado los parámetros de configuración de los distintos RABs al SRNC, utilizando para ello el protocolo RANAP (*Radio Access Network Application Protocol*), el SRNC debe traducir los RABs a una configuración de los RB definidos entre el propio SRNC y el terminal. Para ello, el SRNC seleccionará los parámetros de configuración de los distintos protocolos radio, de modo que se satisfagan todos los requerimientos del RAB. Una vez determinada la configuración más apropiada, el SRNC la comunicará tanto al Nodo B, vía NBAP (*Node B Application Protocol*) a través de la interfaz Iub, como al terminal, vía RRC, a través de la interfaz radio.

La arquitectura general de protocolos de la interfaz radio se muestra en la Figura 2.2. La información proveniente de capas superiores o del estrato de no acceso, se transmite a través de la interfaz radio sin modificaciones. Esta información incluye el tráfico y la señalización NAS, que se intercambian entre los terminales y la red troncal. El protocolo RRC es el encargado de dar servicio a las capas superiores. De esta manera, toda la señalización proveniente de la red troncal se encapsula en mensajes RRC para su posterior transmisión sobre la interfaz radio.

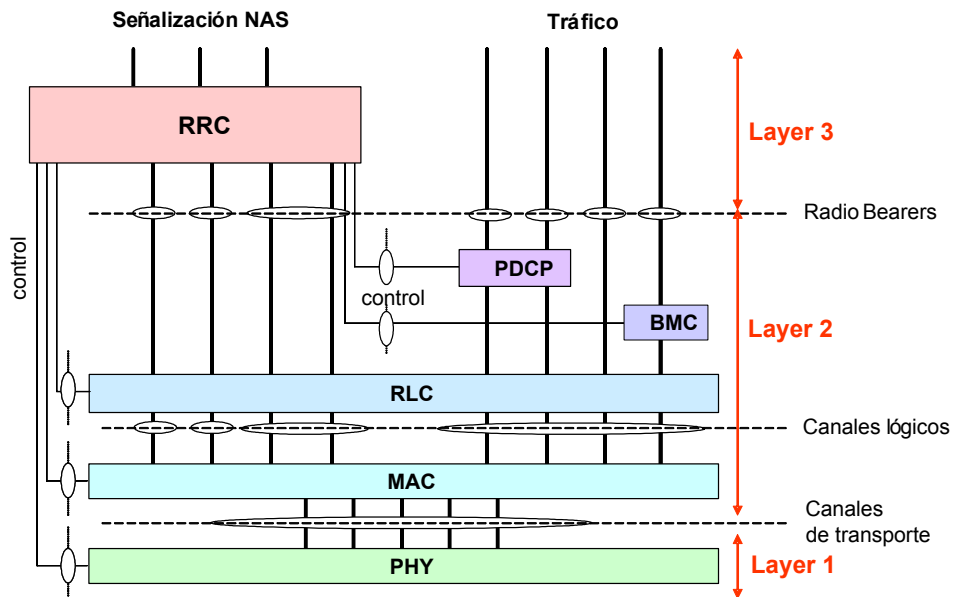


Figura 2.2 Arquitectura de protocolos de la red UTRAN FDD [3]

Un aspecto interesante de esta pila de protocolos es que permite definir diversos tipos de servicios portadores, definidos a distintos niveles dentro de la arquitectura. Así, además de los canales físicos (que corresponden a los canales radio empleados por la capa física), es posible hablar también de canales de transporte (definidos entre la capa física y la capa MAC), de canales lógicos (entre MAC y RLC) y de *Radio Bearers* (entre la capa 2 y las capas superiores).

Los RB son servicios de transferencia de información ofrecidos por la capa 2 a las capas superiores. Dependiendo del tipo de información que se transmita, un RB estará definido inmediatamente por encima de la capa RLC (por ejemplo para servicios CS o para señalización), por encima de la capa PDCP (para servicios PS) o por encima de la capa BMC (para servicios de difusión de mensajes). Cuando los RB transportan señalización (es decir, mensajes RRC) se denominan *Signalling Radio Bearers* (SRBs).

Los canales lógicos se definen entre MAC y RLC, y tienen una relación directa con los RB. Un canal lógico se caracteriza principalmente por el tipo de información que transporta: tráfico o datos, dedicados para un usuario o comunes para todos los usuarios de la celda.

Los canales de transporte se definen entre la capa física y la capa MAC. Un mismo canal de transporte puede emplearse para transmitir información perteneciente a diversos canales lógicos. Por tanto, la información que transmiten puede ser mixta. Este tipo de canales se caracterizan no tanto por el tipo de información que transportan como por el formato que se da a dicha información para el intercambio entre las capas física y MAC.

A continuación, se describen, brevemente, las principales características de los protocolos incluidos en la UTRAN.

## **El Protocolo MAC**

El protocolo MAC (*Medium Access Control*) [4] es un protocolo de acceso al medio, por lo que entre sus funciones se encuentra el mapeo entre canales lógicos y canales de

transporte, que incluye la multiplexación de distintos canales lógicos sobre un mismo canal de transporte, la gestión de prioridades entre usuarios y entre flujos de datos y los mecanismos para la identificación de usuarios en canales de transporte comunes.

Una función fundamental del protocolo MAC es la de realizar la cronoejecución o *scheduling* de las tramas radio para los distintos canales de transporte, es decir, determinar el momento exacto en que una trama debe ser transmitida por la capa física. Puesto que el *scheduling* lo realiza la capa MAC y la transmisión la realiza la capa física, el transporte de tramas radio pertenecientes a los distintos canales de transporte entre ambas capas tiene características de tiempo real.

Por último, el protocolo MAC también se encarga del cifrado necesario de los datos llevado a cabo mediante la operación XOR.

La capa MAC está subdividida en diversas subcapas, que son responsables del *scheduling* para distintos tipos de canales de transporte. Así, cabe destacar las subcapas MAC-d (para canales de transporte dedicados, además es la responsable del cifrado de los datos), MAC-c/sh (para canales de transporte comunes o compartidos), MAC-b (para el canal de *broadcast*) y, a partir del *Release 5*, MAC-hs (para HSDPA, *High Speed Downlink Packet Access*, un servicio de datos basado en el empleo de un canal compartido de alta velocidad en el enlace descendente).

Por último, el protocolo MAC también se encarga de seleccionar el formato de transporte TF (*Transport Format*) más adecuado para cada canal de transporte, dependiendo de la tasa binaria de la fuente. La definición de un formato de transporte incluye dos partes, la semi-estática y la dinámica. La parte semi-estática no cambia para los distintos formatos del TF, por lo que los parámetros correspondientes son fijos para un determinado canal de transporte una vez definido entre el SRNC y el usuario (salvo reconfiguración del canal por parte de las capas superiores). Los parámetros de la parte semi-estática son los siguientes:

- *Transmission Time Interval* (TTI), que es el intervalo de tiempo que transcurre entre dos intercambios consecutivos de bloques de transporte. El TTI toma típicamente un valor de 10, 20, 40 ó 80 ms. A menor TTI, más estrictos son los

requerimientos temporales para el intercambio de información entre capa física y MAC.

- Tipo de codificación del canal, que va a influir en la resistencia frente a los errores de transmisión. Para cada canal se puede seleccionar codificación convolucional o turbo. Además, se puede especificar un ratio de redundancia distinto ( $1/2$ ,  $1/3$  y  $2/3$ ) y se puede permitir o no el empleo de *puncturing*.
- El tamaño del código de redundancia cíclica (CRC), que se emplea para la detección de errores en la transmisión de un bloque de transporte. Dependiendo de la tasa de error residual que requiera el servicio (proporción de errores no detectados), se seleccionará un tamaño mayor o menor para el CRC (típicamente 16 ó 32 bits).

En cuanto a la parte dinámica, como bien indica su propio nombre, sí puede cambiar dinámicamente de TTI a TTI. Esta parte es mucho más sencilla que la anterior, ya que incluye únicamente el tamaño de los bloques de transporte y el número de bloques que se transmitirá en cada TTI.

## El Protocolo RLC

El RLC (*Radio Link Control*) [5] es un protocolo clásico del nivel de enlace de datos que realiza, en enlace descendente o transmisión, la segmentación de la información recibida de capas superiores en paquetes RLC o la concatenación de los bloques recibidos y el posterior rellenado con ceros (*padding*), en caso de que la información de los niveles superiores no complete el tamaño de un paquete RLC. En sentido ascendente o recepción, el RLC realizará la función inversa a la llevada a cabo en transmisión, es decir, o bien reensamblará los bloques RLC recibidos o bien extraerá los paquetes de capas superiores incluidos en cada paquete RLC. Además, el protocolo RLC se encarga de funciones de control de flujo y de corrección de errores, descartando o retransmitiendo las tramas recibidas erróneamente.



En función de las características del RAB se pueden definir tres modos de operación apropiados para distintos tipos de servicios:

- El modo AM (*Acknowledged Mode*), para la transmisión fiable de datos, está basado en la corrección de errores mediante la retransmisión de las tramas erróneas utilizando el algoritmo ARQ (*Automatic Repeat reQuest*). El número máximo de retransmisiones lo determina el RRC y una vez sobrepasado este máximo el paquete se descarta, informando de ello a los niveles superiores. Este modo de operación lo emplean la mayoría de servicios de datos tradicionales como FTP, e-mail, servicio Web, etc., ya que en los servicios de datos es fundamental garantizar la correcta entrega de todos los datos.
- El Modo UM (*Unacknowledged Mode*), se utiliza para servicios de datos en tiempo real que no pueden permitir la retransmisión de las tramas pero que, por el contrario, necesitan garantizar la entrega ordenada de los datos. Las tramas se numeran, permitiendo su correcta secuenciación en recepción, pero no se retransmiten en caso de error. Este modo de operación se emplea fundamentalmente para la prestación de servicios multimedia sobre IP, como *streaming* de audio y vídeo o Voz sobre IP (VoIP).
- El modo TM (*Transparent Mode*), que transmite la información de las capas superiores sin añadir ningún tipo de información de control. Se emplea para los servicios del dominio de conmutación de circuitos, sobre todo llamadas de voz. En este modo de operación el protocolo RLC sólo se encarga, en caso de que sea necesario, de la segmentación o concatenación de los paquetes recibidos de capas superiores.

La capa RLC también será la encargada de garantizar el envío de la información recibida a los niveles superiores de manera ordenada y sin duplicados.

Por último, cabe destacar que la capa RLC realiza la encriptación de la información destinada a un usuario concreto, aunque únicamente en los modos AM y UM, puesto que en modo TM no se altera la información de las capas superiores.

## El Protocolo PDCP

La principal función del protocolo PDCP (*Packet Data Convergence Protocol*) [6] es comprimir las cabeceras de los paquetes antes de la transmisión, para ahorrar ancho de banda en la interfaz radio. Por lo tanto, únicamente tiene sentido emplear este protocolo para los servicios del dominio PS. El porqué de la necesidad de compresión de las cabeceras se explica fácilmente si se considera que en IPv4 las cabeceras son de al menos 40 bytes, incrementándose este tamaño hasta 60 bytes para IPv6. Teniendo en cuenta que para un servicio de voz el campo de datos IP no supera los 20 bytes, se entiende perfectamente la necesidad de reducir el tamaño de las cabeceras.

Cada usuario concreto tendrá asociada una aplicación PDCP exclusiva, por lo que el protocolo PDCP está siempre situado en el SRNC asociado con el usuario. Los algoritmos de compresión empleados, muy eficientes, requieren que el protocolo PDCP almacene cierta información de contexto. Por este motivo, si se ejecuta el procedimiento de SRNS *Relocation*, mediante el cual se produce un cambio del SRNC asociado con el usuario, puede ser necesario transferir el contexto PDCP entre el antiguo SRNC y el nuevo para evitar la pérdida de datos.

## El Protocolo BMC

El protocolo BMC (*Broadcast/Multicast Control*) [7] únicamente se utiliza para un servicio denominado CBS (*Cell Broadcast Service*), que consiste en la difusión en una celda de mensajes de texto cortos (tipo SMS), utilizados habitualmente para proporcionar a los usuarios avisos o servicios de noticias breves. Generalmente, este tipo de mensajes se retransmite un cierto número de veces, con una sincronización determinada, para asegurar su recepción por los usuarios. El protocolo BMC se ocupa principalmente de la sincronización y retransmisión.

Existen dos modalidades del servicio, la modalidad de *broadcast* y la modalidad de *multicast*. En la modalidad de *broadcast*, los mensajes son recibidos por todos los

usuarios de la celda, mientras que en la de *multicast*, únicamente los usuarios suscritos a un servicio concreto recibirán los mensajes correspondientes.

Puesto que el servicio CBS siempre está asociado a una celda concreta, el protocolo BMC estará situado siempre en el CRNC (*Controlling RNC*) asociado con dicha celda.

## El Protocolo RRC

El protocolo RRC (*Radio Resource Control*) es el protocolo de señalización utilizado entre la UTRAN y los terminales de usuario [8]. Los mensajes RRC se definen en las especificaciones mediante la notación ASN.1 (*Abstract Syntax Notation 1*). El RRC se emplea para el intercambio de información de control relativa al nivel de acceso, AS (*Access Stratum*), aunque también puede transportar, de forma transparente, información de control relativa al nivel NAS.

Para la transmisión de mensajes RRC se emplean diversos canales lógicos de control, como los canales comunes BCCH (*Broadcast Control Channel*), PCCH (*Paging Control Channel*) y CCCH (*Common Control Channel*) y los canales de control dedicados DCCH (*Dedicated Control Channel*). Los canales comunes no precisan el establecimiento de la conexión para su uso. Sin embargo, un DCCH está asociado siempre a un terminal concreto y para poder emplearlo es necesario establecer previamente una conexión RRC.

Respecto al protocolo RRC, cada terminal puede encontrarse en dos modos de operación denominados RRC *idle* (inactivo) y RRC *connected* (conectado).

En el modo RRC *idle*, el terminal no dispone de conexión RRC, por lo que tampoco tiene un SRNC asociado ni ningún canal de control dedicado DCCH. La red UTRAN no almacena ningún tipo de información sobre los terminales que se encuentran en este modo de operación, por lo que la gestión de la movilidad se lleva a cabo únicamente en la red troncal. El terminal monitoriza los canales de transporte BCH (*Broadcast Channel*) y PCH (*Paging Channel*) de la celda en la que se encuentre. De esta forma, el terminal obtiene la información del sistema necesaria para operar en la red y detectar

cambios de celda, LA (*Location Area*) o RA (*Routing Area*), y puede ser avisado en caso de llamada entrante.

En el modo RRC *connected*, el terminal dispone de una conexión RRC y, por tanto, de un SRNC asociado y un canal de control dedicado DCCH. El SRNC almacena información de estado y de posición sobre el usuario y será responsable del control de la movilidad dentro de la red UTRAN.

### 2.2.3 Transmisión mediante DCH

Una vez descrita la arquitectura de protocolos convencional de UMTS, los siguientes apartados analizan en profundidad los canales de transporte considerados en esta Tesis Doctoral para la transmisión de datos. Los cambios introducidos por la nueva arquitectura radio MNB se reflejarán únicamente en los niveles MAC y físico, por lo que se deberá concretar, para cada modo de transporte, qué cambios deberían ser introducidos para maximizar las prestaciones del MNB.

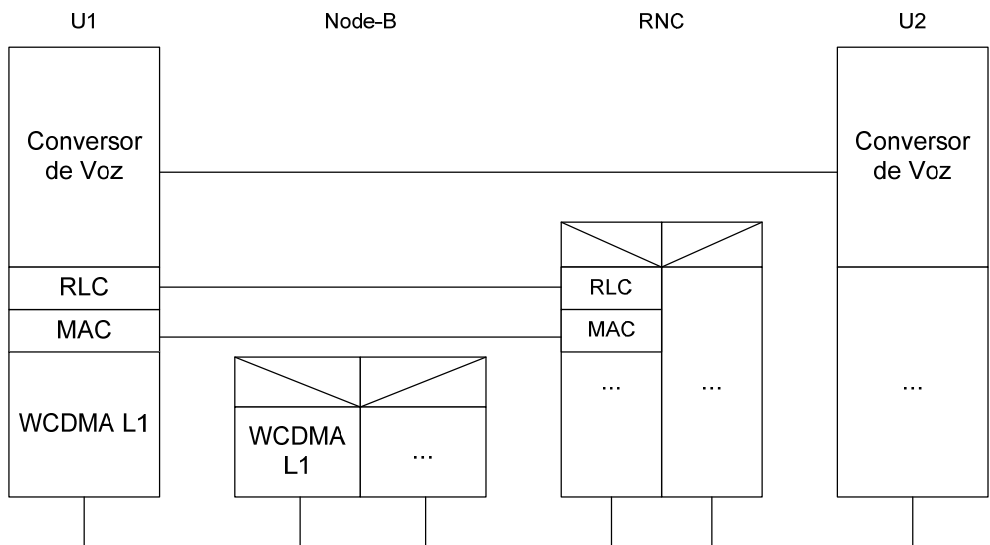
La transmisión mediante el canal de transporte DCH (*Dedicated Channel*) es la más sencilla y será la primera en ser analizada. Siempre que se trabaja con servicios orientados a la conmutación de circuitos se recomienda el uso del canal DCH. Por lo tanto, los principales servicios transmitidos mediante DCH serán los de voz y vídeo conferencia, ya que la tasa de transmisión es constante y los requerimientos temporales son estrictos. Sin embargo, en las simulaciones llevadas a cabo se ha optado por incluir el servicio de vídeo conferencia dentro de conmutación de paquetes, para así evaluar las prestaciones de los algoritmos de gestión de recursos desarrollados cuando las restricciones temporales son estrictas.

Dado que el canal DCH es muy sencillo de implementar, en las primeras versiones de UMTS solamente se incluyó este tipo de canal, incluso para los servicios de datos. Tuvieron que pasar un par de años hasta encontrar los primeros Nodos B comerciales que incluyeran también la transmisión de datos mediante canales compartidos.

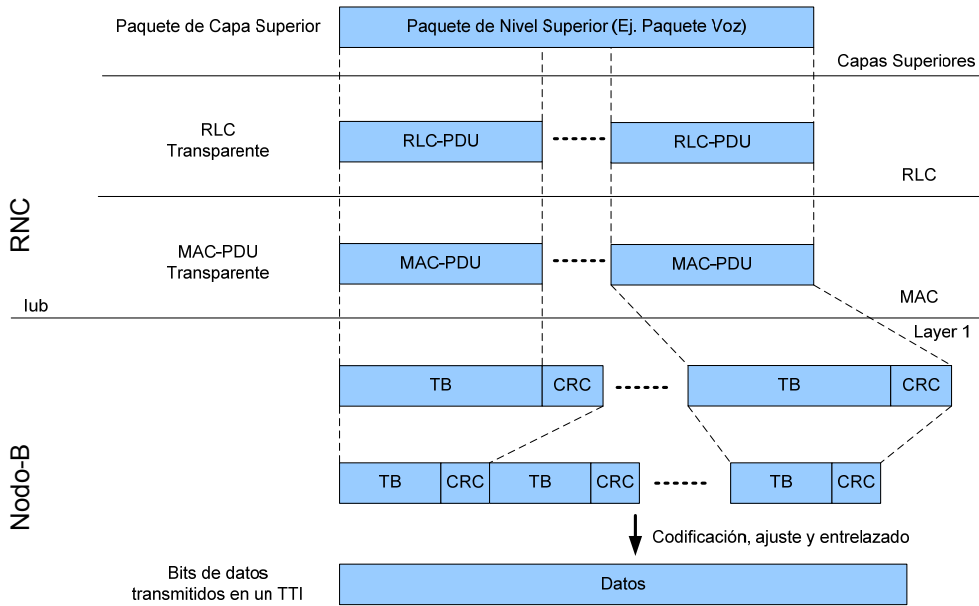
Respecto a la localización de los distintos protocolos, en la transmisión mediante DCH el Nodo-B es una entidad carente de inteligencia y sólo se encarga de codificar la información y de constituir las tramas físicas WCDMA que serán transmitidas por la interfaz aire.

En la Figura 2.3 se muestra la distribución de los protocolos en las distintas entidades para el servicio de voz. En este caso, el protocolo PDCP no se utiliza puesto que no se envían paquetes IP.

En este modo de transmisión el RLC sólo se encarga de fragmentar los datos ya que la transmisión es en modo transparente y no hay retransmisiones ni se garantiza la entrega ordenada de los datos. De esta manera, el nivel RLC no añade ningún tipo de cabecera a los datos segmentados. Por su parte, el protocolo MAC trabajará de manera también transparente por lo que no aplicará ninguna funcionalidad salvo la de selección del formato de transporte correspondiente al flujo de datos recibido.



**Figura 2.3 Estructura de protocolos convencional de la interfaz radio para DCH**



**Figura 2.4 Flujo de datos a través de la interfaz radio para transmisión DCH**

El flujo de datos a través de la estructura de protocolos se muestra en la Figura 2.4. Es importante destacar que en la transmisión DCH no se realiza ningún procedimiento de gestión de recursos, por lo que, una vez configurado un RB, éste se mantiene fijo durante toda la comunicación, salvo que a nivel RRC se decida realizar un *Radio Bearer Reconfiguration* de manera que se cambie la configuración deseada. Sin embargo, la modificación de la configuración del RB no se produce de manera inmediata, al contrario, desde que se inicia la reconfiguración hasta que el usuario se sincroniza puede transcurrir un tiempo aproximado de  $2.5s^1$ . Con estos tiempos de reconfiguración es inviable considerar un algoritmo de gestión de recursos eficaz para la transmisión de datos eficiente mediante el uso de canales dedicados.

<sup>1</sup> valor medio obtenido mediante la lectura de KPI (*Key Performance Indicator*) de una red real de Motorola

Higher layer	RAB/Signalling RB	RAB subflow #1	RAB subflow #2	RAB subflow #3	
RLC	Logical channel type	DTCH			
	RLC mode	TM	TM	TM	
	Payload sizes, bit	0, 39, 42, 55, 75, 81	53, 63, 84, 103	60	
	Max data rate, bps	12 200			
	TrD PDU header, bit	0			
MAC	MAC header, bit	0			
	MAC multiplexing	N/A			
Layer 1	TrCH type	DCH	DCH	DCH	
	TB sizes, bit	0, 39, 42, 55, 75, 81	53, 63, 84, 103	60	
	TFS	TF0, bits	1x0	0x103	0x60
		TF1, bits	1x39	1x53	1x60
		TF2, bits	1x42	1x63	N/A
		TF3, bits	1x55	1x84	N/A
		TF4, bits	1x75	1x103	N/A
		TF5, bits	1x81	N/A	N/A
	TTI, ms	20	20	20	
	Coding type	CC 1/3	CC 1/3	CC 1/2	
	CRC, bit	12	N/A	N/A	
	Max number of bits/TTI after channel coding	303	333	136	
	RM attribute	180-220	170-210	215-256	

**Tabla 2.1 Parámetros de la portadora radio del servicio de voz UMTS [9]**

Puesto que el único servicio que se va a prestar mediante el canal DCH es el servicio de voz, es conveniente comprender la estructura de datos del servicio para saber cuál es la capacidad de maniobra de la que puede hacer uso el MNB.

La voz en UMTS está compuesta por tres flujos de datos, que en la Tabla 2.1 se denominan RAB *subflow* #1, #2 y #3. Como el TTI es de 20 ms, se transmiten 50 tramas de voz por segundo. Trama a trama cada flujo de datos puede transmitir un bloque de datos de tamaño variable o bien no transmitir nada. Cuando se transmite la máxima información posible sobre todos los subflujos se configura el bloque de mayor tamaño, de manera que el total de bits de información transmitidos por trama resulta ser de 244 bits (81+103+60). Con esta configuración la tasa binaria de transmisión, teniendo en cuenta que se transmiten 50 tramas por segundo, sería de 12.2 kbps.

En UMTS se han definido distintas combinaciones de estos bloques para definir 4 tasa binarias distintas: 12.2, 7.95, 5.9 y 4.75 kbps. El flujo de datos #1 siempre se transmite ya que contiene los bits fundamentales para decodificar la voz, mientras que el flujo #2 y #3 permiten tener una mayor calidad de la señal.

Como se ve en la Tabla 2.1, ni el nivel RLC ni el nivel MAC introducen ningún bit adicional en las cabeceras, por lo que al Nodo-B le llegan directamente los bloques a transmitir en el siguiente TTI y el formato en que tienen que ser transmitidos. De esta manera, si en el nivel MAC del RNC se decide transmitir la voz a una determinada tasa binaria el Nodo-B no podrá nunca transmitir a mayor tasa porque carecerá de esos bits de información. Por el contrario, el Nodo-B sí que podrá ajustar la transmisión a una menor tasa sin más que eliminar los bits redundantes o directamente todo un flujo de datos. Respecto a la potencia transmitida a cada usuario por el Nodo-B, ésta es una decisión que debe tomar el propio Nodo-B a partir de los informes del control de potencia de los usuarios y de los objetivos de  $E_b/N_0$  marcados por el RNC.

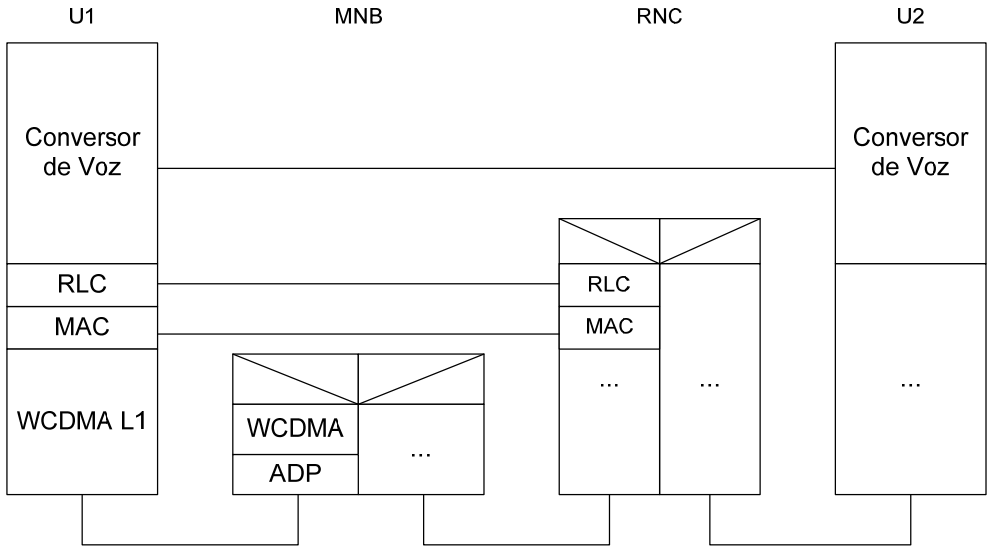
La conclusión que se puede extraer de este análisis es que sería muy sencillo modificar el Nodo-B para que éste pudiera ajustar la tasa binaria de transmisión de los usuarios de canales de voz DCH, así como distribuir la potencia de transmisión entre todas las antenas que conforman el MNB. Al protocolo de asignación de potencia de transmisión a cada antena del MNB se le denomina Algoritmo de Distribución de Potencia (ADP).

Respecto a los cambios necesarios en las dos nuevas arquitecturas radio MNB, en el caso del MNB-pasivo, el MNB asumiría la tasa binaria establecida en el nivel MAC del RNC y sólo se modificaría la capa física para distribuir las potencias transmitidas por las antenas del MNB, utilizando el protocolo ADP.

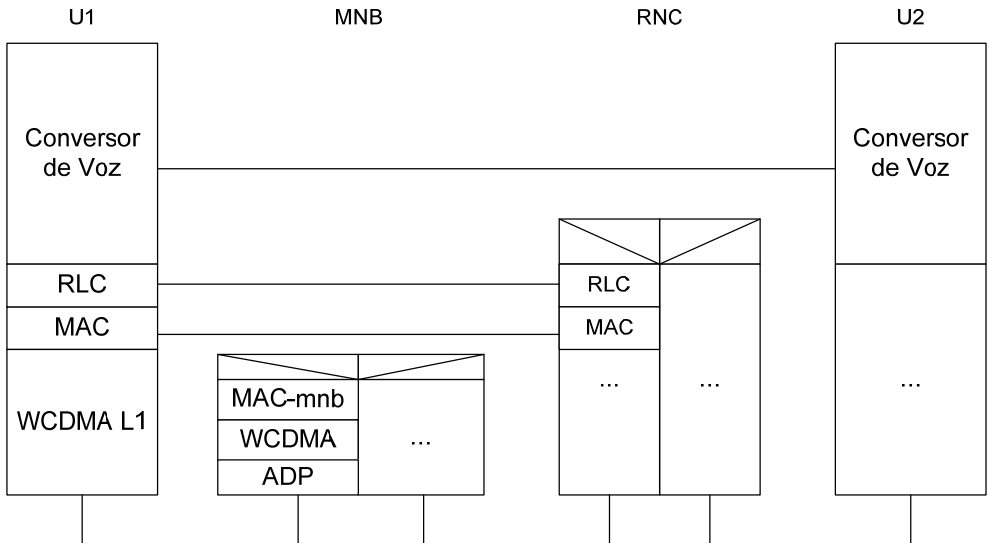
La arquitectura MNB-activo, requeriría de una capa MAC adicional (MAC-mnb) dentro del MNB, de manera que se pudiera modificar la tasa binaria asignada a los usuarios en función de las decisiones tomadas por el gestor de recursos. Como el modo de transmisión es transparente, no haría falta informar al RNC de los cambios realizados. Además, incluiría los mismos cambios en la capa física para la distribución de potencias.



La ventaja de ambas configuraciones es que no se precisa de ningún cambio ni en el terminal ni en el RNC. Las Figuras 2.5 y 2.6 representan la nueva arquitectura de protocolos para ambas configuraciones.



**Figura 2.5 Estructura de protocolos del MNB-pasivo para DCH**



**Figura 2.6 Estructura de protocolos del MNB-activo para DCH**

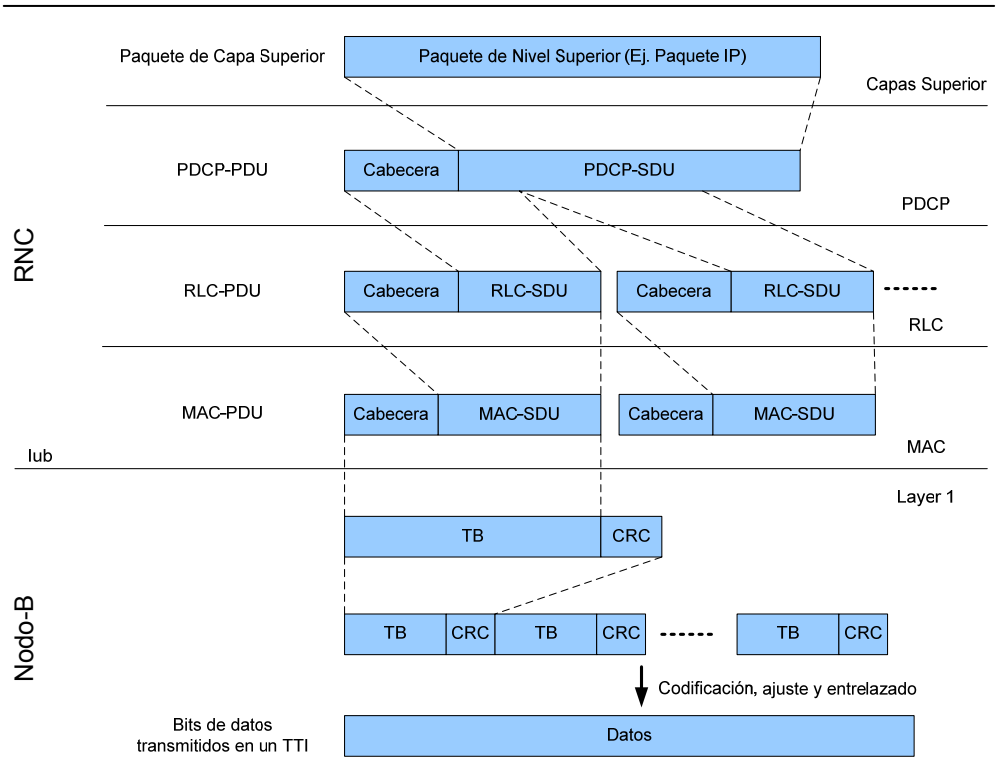
## 2.2.4 Transmisión mediante DSCH

El canal DSCH (*Downlink Shared Channel*) fue estandarizado para permitir la transmisión por conmutación de paquetes en UMTS. Posteriormente a la estandarización de este canal, en la Release 5 se estandarizó el uso del canal de transporte HS-DSCH, en principio mucho más orientado a la transmisión eficiente de paquetes. Hasta junio de 2005 ambos canales, el DSCH y el HS-DSCH coexistieron en las especificaciones como válidos para la transmisión de paquetes en el modo FDD. Sin embargo, el 30 de junio de 2005 el canal DSCH se descartó para la transmisión de paquetes y aunque se implementó en Nodos-B de algunos fabricantes (Motorola, Nortel y Siemens) el canal HSDPA parece haber relegado este canal de transmisión de paquetes a un segundo plano.

En el momento de iniciarse esta Tesis Doctoral el canal HS-DSCH ni siquiera estaba planteado en las especificaciones, por lo que el trabajo realizado siempre se centró en la utilización del canal DSCH. El canal HS-DSCH no es más que una evolución del canal DSCH por lo que los resultados obtenidos en esta Tesis Doctoral se pueden, sin duda, extrapolar a otros canales más recientes.

Cuando se realiza una transmisión mediante el canal DSCH nunca se utiliza el modo transparente del protocolo RLC, ya que es necesario garantizar la entrega sin errores y ordenada de los paquetes, por lo que se debe insertar numeración a cada paquete RLC-SDU. Además, tal y como se explica en el apartado 4.3.6, el canal DSCH se comparte entre varios usuarios, por lo que es imprescindible identificar el destinatario de cada uno de los paquetes MAC. De esta manera, tanto la capa RLC como la capa MAC introducirán una cabecera adicional.

Además, el flujo de datos en la transmisión mediante DSCH es distinto al seguido por los datos transmitidos mediante canales dedicados DCH. A continuación, y a modo de ejemplo, se muestra el flujo de datos que sigue un paquete IP a través de todos los protocolos de la interfaz radio hasta ser transmitido.



**Figura 2.7 Flujo de datos para transmisión mediante DSCH**

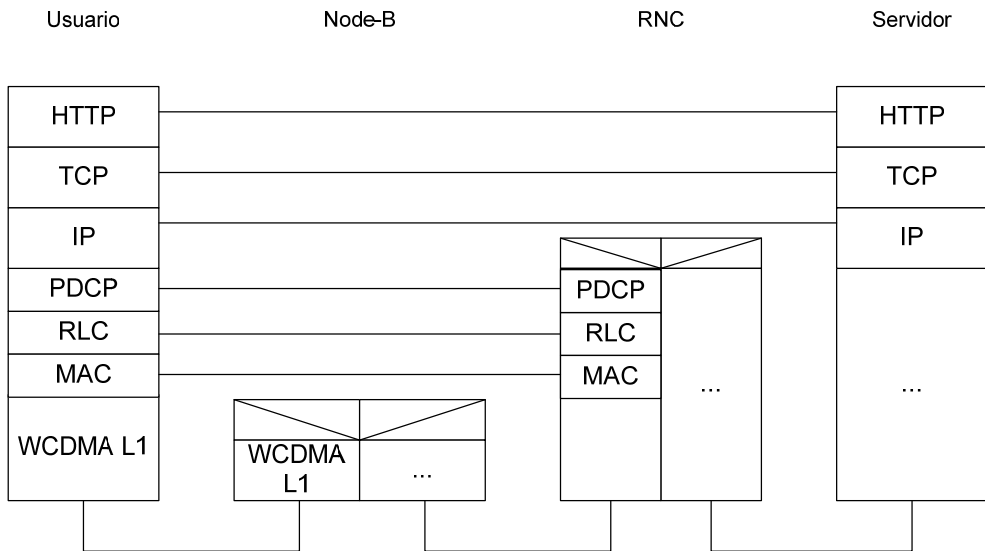
En primer lugar, el protocolo PDCP comprime la cabecera del paquete IP creando el paquete PDCP-SDU (PDCP *Service Data Unit*) al que añadirá su correspondiente cabecera conformando lo que se conoce como paquete PDCP-PDU (PDCP *Packet Data Unit*). Posteriormente, el protocolo RLC será el encargado de fragmentar el paquete PDCP-PDU y añadir las correspondientes cabeceras formando los bloques RLC-PDU, que serán los paquetes que administrará el protocolo ARQ encargado del control de flujo y control de errores. A continuación, el paquete RLC-PDU pasará por la entidad MAC que será la encargada de multiplexar los distintos canales de datos recibidos, seleccionando el formato de transporte, TF, más indicado y cifrando la información. El protocolo MAC enviará varios MAC-PDU al Nodo-B para su transmisión. El paquete MAC-PDU también se conoce como TB (*Transport Block*) y constituye la mínima cantidad de información que puede detectarse como errónea y descartarse. Ya en el Nodo-B se añaden unos determinados bits de redundancia cíclica

(CRC) a cada TB de manera que se pueda detectar su correcta recepción. Según el TFC el Nodo-B procederá a la concatenación de un determinado número de TB junto con su CRC y a su posterior codificación. Tras la codificación, y una vez ajustado el número de bits físicos a transmitir, se entrelazan los datos, se crea la trama física, se ensancha, se modula y finalmente se transmite. La Figura 2.7 resume todo este procedimiento.

Por otro lado, la estructura de protocolos y su localización en las distintas entidades del RNS se recoge en la Figura 2.8.

Algunos conceptos importantes de la transmisión DSCH se resumen a continuación:

- El canal DSCH sólo existe en sentido descendente.
- El canal DSCH sólo tiene impacto en los niveles 1 y 2 de la arquitectura de protocolos puesto que a nivel lógico no existe el concepto de canal compartido.



**Figura 2.8 Estructura de protocolos de la interfaz radio para transmisión de paquetes**

- Además, el canal DSCH no acepta macro-diversidad por *soft handover* por lo que en este caso la transmisión en diversidad ofrecida por el MNB resulta especialmente atractiva.
- En cada trama se puede configurar uno o varios canales físicos PDSCH. Esto quiere decir que el canal DSCH soporta multiplexación por código, es decir, en cada trama varios usuarios pueden recibir datos del mismo canal de transporte compartido sin más que asignar a cada usuario parte del árbol de código reservado para DSCH.
- Por otro lado, cada TTI un canal físico PDSCH sólo puede transmitirse a un único usuario, pudiendo cambiar el usuario destino TTI a TTI.
- El tamaño de los TB es constante, por lo que según el código asignado a cada usuario en el canal DSCH habrá un número determinado de TB a transmitir por el nivel físico.
- Por último, el TFCI transmitido en el canal físico dedicado lleva la información que indica qué código del DSCH debe ser utilizado por el usuario.

En la arquitectura MNB-pasivo el principal problema radica en que la gestión de recursos se realiza en el RNC mientras que la gestión de la transmisión sobre cada antena se lleva a cabo en el Multinodo B. Esta primera variante, mucho más conservadora, mantiene la estructura de protocolos establecida en las especificaciones [3] de manera que la gestión de recursos se mantiene en el RNC. Si, por el contrario, el Multinodo B fuera el responsable no sólo de la distribución de potencia entre las antenas sino también de la gestión de recursos, se podría optimizar el funcionamiento del sistema. Esta segunda variante requiere un cambio mínimo en las especificaciones y en el software del RNC, pero a cambio permite que toda la gestión de recursos la realice el propio Multinodo B, que es quien mejor conoce el estado de su celda. Esta segunda variante, que se ha denominado Arquitectura MNB-activo, se basa fundamentalmente en reubicar las entidades que componen el protocolo MAC.

## Arquitectura de Protocolos MNB-activo

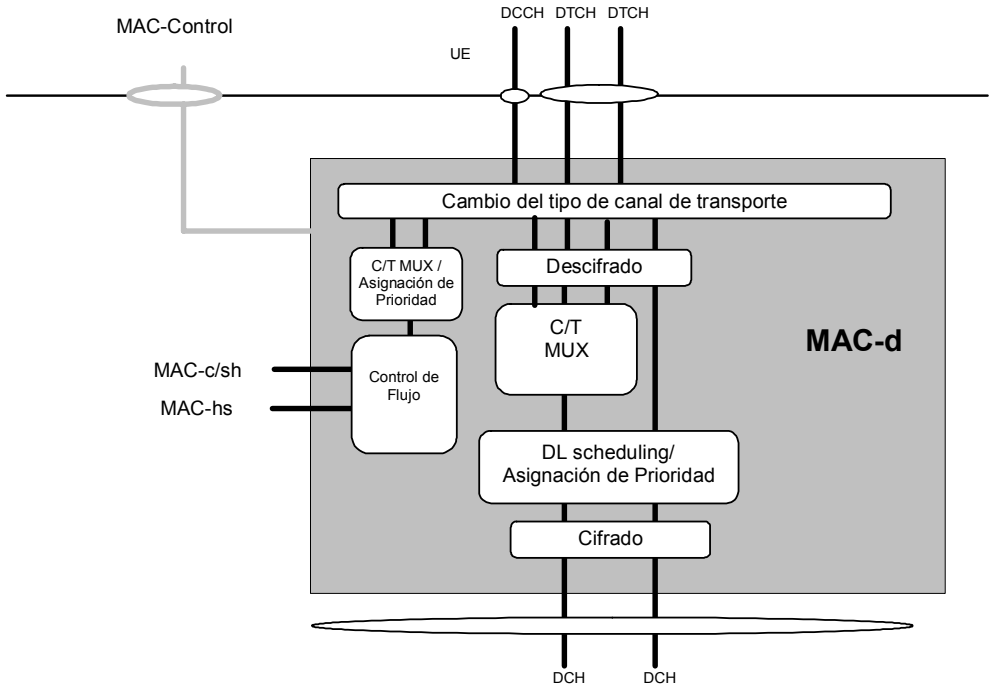
Para entender esta nueva arquitectura de la interfaz radio es importante analizar en profundidad el funcionamiento del protocolo MAC.

En primer lugar el protocolo MAC dentro de la transmisión DSCH se encuentra dividido en dos subcapas: la capa MAC-d y la capa MAC-c/sh.

En enlace descendente se asocia una capa MAC-d a cada usuario de manera que todos los canales lógicos destinados a un usuario pasan por su entidad MAC-d. La capa MAC-d realiza las siguientes acciones:

- Conmuta el tipo de canal de transporte para cada canal lógico en función de la información de control enviada por el protocolo RRC. De esta manera, un canal lógico de datos puede estar transmitiéndose sobre un canal dedicado mientras que un canal lógico de control asociado al mismo usuario puede ser transmitido por canal común.
- Multiplexa los distintos canales lógicos. Si varios canales lógicos se multiplexan sobre el mismo canal de transporte entonces se añade una cabecera de manera que se pueda distinguir qué canal lógico está siendo transmitido sobre el canal de transporte. Normalmente esta cabecera consta de 4 bits por lo que se puede discriminar hasta 15 canales lógicos.
- Define la prioridad entre los canales de transporte dedicados mediante la selección del TFC.
- Cifra la información en caso de que se transmita en modo transparente, realizando el descifrado en enlace ascendente.
- Realiza un control de flujo para los canales comunes o compartidos para evitar el desbordamiento en la siguiente subcapa del nivel MAC.
- En el caso de canales compartidos o comunes decide a qué entidad MAC va a reenviar el MAC-d-PDU.

La funcionalidad de la capa MAC-d se encuentra resumida en la Figura 2.9.



**Figura 2.9 Entidad MAC-d asociada a cada usuario [4]**

Según el estándar [4], todos los canales comunes o compartidos se pasan a la entidad MAC-c/sh salvo los canales de HSDPA que se envían directamente a la entidad MAC-hs. Como el canal DSCH es un canal compartido, la siguiente entidad a la que se envía la MAC-d-PDU será la entidad MAC-c/sh.

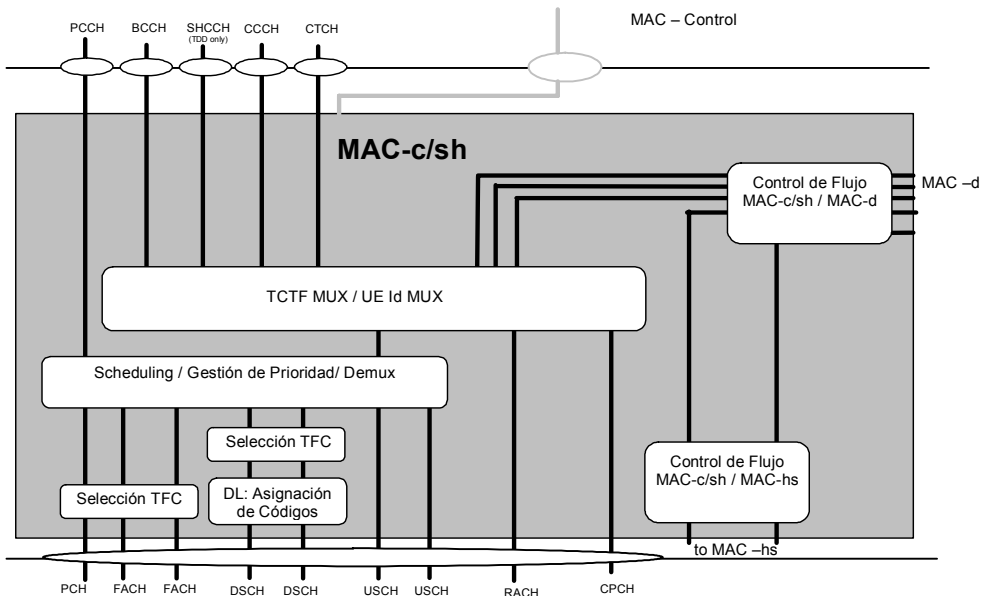
Existe una entidad MAC-c/sh asociada a cada celda de manera que en cada RNC habrá un elevado número de entidades MAC-c/sh. Además, es importante destacar que las entidades MAC-d y MAC-c/sh no tienen porqué estar localizadas en el mismo RNC. De esta manera los paquetes MAC-d-PDU se pueden transmitir vía Iur a través del protocolo *Frame Protocol* [10] y, utilizando un camino ATM virtual para cada usuario, se puede enviar toda la información necesaria a la entidad MAC-c/sh de cualquier RNC remoto.

De cualquier manera, la entidad MAC-c/sh realiza las siguientes funciones:

- Controla el flujo de datos desde/a la entidad MAC-d.

- Indica en la cabecera del paquete MAC-PDU a través de los bits TCTF si los datos contenidos provienen de un canal lógico común o dedicado.
- Si el canal lógico es dedicado, además incluye un identificador del usuario y por tanto de la entidad MAC-d origen de los datos.
- Se encarga de las labores de *scheduling* de los paquetes almacenados, es decir, selecciona la prioridad de los distintos flujos de datos.
- Selecciona el TFC adecuado para el canal de transporte seleccionado.
- Además, en caso de transmitir mediante DSCH, selecciona el código utilizado por cada canal de transporte.

La Figura 2.10 muestra todas las funciones llevadas a cabo por la entidad MAC-c/sh.



**Figura 2.10 Entidad MAC-c/sh convencional asociada a cada celda [4]**



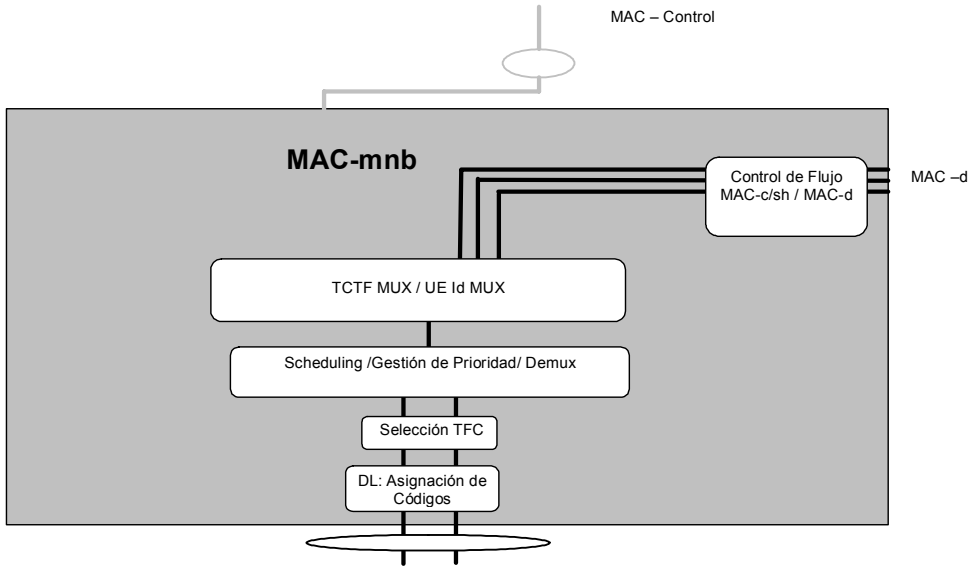
Antes de describir la nueva arquitectura de protocolos, hay que resaltar que la información de los canales comunes o de *broadcast* de la celda siempre se transmiten vía la entidad MAC-c/sh del SRNC, mientras que es solamente la información de datos del usuario la que puede ser redirigida hacia otra entidad remota MAC-c/sh, entidad remota que bien pudiera ser el MNB.

La nueva arquitectura de protocolos propuesta MNB-activo se basa en replicar la funcionalidad de la entidad MAC-c/sh en una nueva entidad situada en el Multinodo B y denominada MAC-mnb. Este nuevo subnivel del MAC realizará la gestión de todos los canales DSCH con la ventaja de que podrá configurar simultáneamente la asignación de recursos y la potencia transmitida por las antenas. La gestión del resto de canales comunes se dejará tal cual está, en la entidad MAC-c/sh convencional.

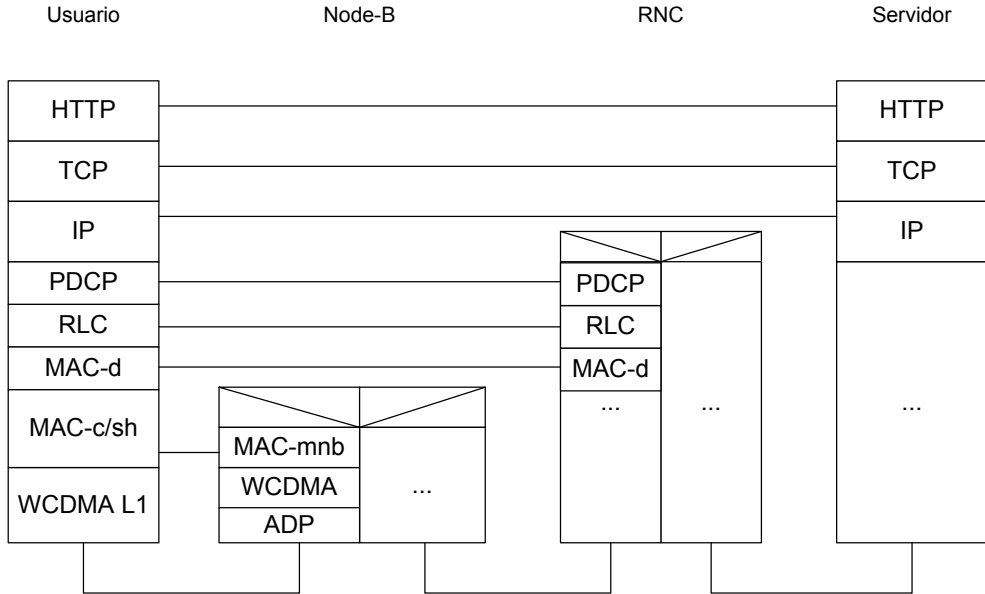
El cambio de arquitectura propuesto es muy sencillo de implementar en el RNC ya que el protocolo que se utiliza para comunicar el subnivel MAC-d con un MAC-c/sh remoto, el *Frame Protocol*, es el mismo que comunica al RNC y al Nodo-B a través de la interfaz Iub, por lo que en el RNC no habría que realizar ningún cambio salvo reconfigurar el establecimiento de la comunicación de los usuarios DSCH de un MNB para que su MAC-d esté conectado remotamente con el MNB que le presta servicio. Este cambio es tan simple como cambiar el camino virtual ATM vinculado al MAC-c/sh de manera que se seleccione el del MNB.

Aunque al tratarse de una reconfiguración *software* la nueva arquitectura de protocolos no impone un cambio importante en el RNC, sí que supone una profunda transformación en el Nodo-B ya que en éste se tendría que crear la nueva entidad MAC-mnb cuya funcionalidad está resumida en la Figura 2.11. La principal complejidad de esta nueva entidad radica precisamente en el *scheduling* y la asignación de códigos. Además, habrá que modificar ligeramente el nivel físico puesto que las MAC-PDU de los usuarios a los que se les asigne el uso de un canal compartido DSCH ya no le llegarán del RNC sino directamente del mismo Nodo-B.

La arquitectura de protocolos MNB-activo para la transmisión DSCH y la localización de cada uno de los protocolos se resume en la Figura 2.12.



**Figura 2.11 Entidad MAC-mnb**



**Figura 2.12 Estructura de protocolos del MNB-activo para DSCH**

## 2.2.5 Obtención de parámetros de la red

Para la toma de decisiones dentro de la gestión de recursos radio del MNB es necesario conocer cierta información del enlace descendente. Esta información sólo puede ser obtenida por el propio terminal, por lo que se debe establecer algún mecanismo para que el usuario envíe las medidas realizadas al sistema.

El protocolo RRC establece en la especificación [8] la posibilidad de configurar el terminal para que realice un determinado número de medidas y las envíe a la red de manera periódica cada 200 ms como mínimo. Las medidas son transmitidas por el UE en sentido ascendente utilizando el protocolo RRC. El problema es que, como se aprecia en la Figura 2.2, el protocolo RRC es de nivel 3 y está localizado físicamente en el RNC, por lo que, si se pretende que la toma de decisiones la realice el Multinodo B, hay que modificar de alguna manera la arquitectura de protocolos para permitir que el MNB conozca la información enviada por el terminal.

Para dar solución a este problema se plantean dos opciones. (1) Aprovechar que el protocolo de señalización entre RNC y Nodo-B permite el envío de información propietaria entre las dos entidades para transmitir toda la información reportada por los usuarios desde el RNC hasta el Nodo-B. Esta solución es muy sencilla y su único inconveniente es que habría que modificar el RNC para que, en primer lugar, configurara los terminales para realizar periódicamente las medidas necesarias y posteriormente reenviara esas medidas al Nodo-B. Esta opción no es tecnológicamente muy descabellada puesto que en la actualidad, con HSDPA, se está procediendo de manera similar, lo que garantiza que esta solución no es muy compleja.

(2) Replicar parcialmente la pila de protocolos RRC dentro del Nodo-B. Esta solución es relativamente sencilla y aporta muchas ventajas. A continuación, se describe en profundidad esta segunda alternativa.

## Modificación del Nodo-B

Siempre que se establece una comunicación de datos con un usuario y éste se encuentra en estado de conexión RRC se configura un canal de transporte DCH para soportar la transmisión de la información de control. Dentro de este canal de transporte se multiplexa 4 canales lógicos DCCH denominados SRB#1,..., SRB#4 respectivamente y cuyo formato a nivel MAC y RLC es constante para todos los servicios de datos convencionales [11]. De estos 4 canales lógicos el primero, el SRB#1, es el que transporta la información de control RRC y, por tanto, es el utilizado para configurar las medidas del usuario en enlace descendente y para enviar las medidas realizadas en enlace ascendente. La Tabla 2.2 muestra las características del transporte de señalización en UMTS.

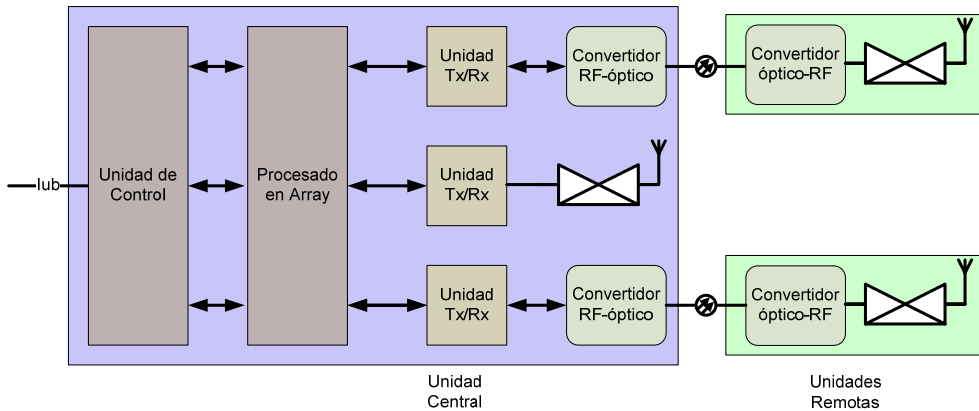
Por lo tanto, el MNB debe ser capaz de, una vez establecida la comunicación, enviar un comando de configuración de medida periódica al usuario identificando el procedimiento de medida con un identificador conocido. Cada paquete enviado por los usuarios en el canal de transporte asociado a la señalización tiene que ser analizado por el MNB. Si el canal lógico coincide con el del RRC y, además, es un informe de medidas correspondiente a las medidas solicitadas por el MNB, entonces el paquete será procesado directamente extrayendo los datos necesarios, en caso contrario será reenviado al RNC.

Este procedimiento de evaluar si la información transmitida en enlace ascendente es relativa al canal lógico #1 del canal de transporte dedicado de datos, no es en absoluto complejo. Simplemente, se deberá evaluar en primer lugar el TFCI, comprobando si existe información relativa a dicho canal de transporte. Esta labor siempre la realiza el Nodo-B con el fin de separar los canales de transporte, así que no requiere de un *software* adicional. Una vez localizado el DCH de control, se mirará en la cabecera de datos de la MAC-PDU si el canal lógico se corresponde con el primero. En caso afirmativo se decodificará el mensaje extrayendo la información que se considere importante.

Higher layer	RAB/signalling RB	<b>SRB#1</b>	<b>SRB#2</b>	<b>SRB#3</b>	<b>SRB#4</b>	
	User of Radio Bearer	RRC	RRC	NAS_DT High prio	NAS_DT Low prio	
RLC	Logical channel type	DCCH	DCCH	DCCH	DCCH	
	RLC mode	UM	AM	AM	AM	
	Payload sizes, bit	136	128	128	128	
	Max data rate, bps	13600	12800	12800	12800	
	AMD/UMD PDU header	8	16	16	16	
MAC	MAC header, bit	4	4	4	4	
	MAC multiplexing	4 logical channel multiplexing				
Layer 1	TrCH type	DCH				
	TB sizes, bit	148 (alt 0, 148) (note)				
	TFS	TF0, bits	0x148 (alt 1x0) (note)			
		TF1, bits	1x148			
		TF2, bits	2x148			
	TTI, ms	10				
	Coding type	CC 1/3				
	CRC, bit	16				
	Max number of bits/TTI before rate matching	1008				
	RM attribute	155 to 230				

**Tabla 2.2 Características del transporte de señalización en UMTS [9]**

## 2.3 Entidades Funcionales del Multinodo B



**Figura 2.13 Entidades Funcionales del Multinodo B**

La Figura 2.13 muestra todas las entidades funcionales que componen la estructura MNB. El MNB deberá ser transparente para el resto de la red por lo que, en primer lugar, se deberá contar con una unidad de control encargada de toda la conversión de protocolos; este es el bloque realmente inteligente. A continuación, el bloque de procesado en *array* distribuirá la información entre todas las antenas del MNB, determinando asimismo la potencia transmitida por cada una de ellas. El bloque de transmisión será el encargado de conformar la señal radio en sentido descendente, ensanchando los bits de información y modulando la señal convenientemente. Por su parte el bloque de recepción realizará el proceso inverso, filtrando convenientemente la señal radio recibida y extrayendo la información útil de cada uno de los usuarios.

Finalmente, la señal de RF se enviará a un amplificador bidireccional cercano a la antena, o bien pasando primero por enlace de fibra óptica o bien directamente, utilizando para ello un cable coaxial de bajas pérdidas. La antena será la encargada de radiar la señal al medio aéreo.

A continuación, se analizarán cada una de estas entidades funcionales empezando desde la unidad de control. De cada bloque se identificarán sus características y las principales tareas asignadas.

### 2.3.1 Unidad de control

Tanto esta entidad como la de procesado en *array* son las principales novedades de la nueva arquitectura radio MNB. La unidad de control tendrá asignadas unas funciones distintas en caso de que se opte por la configuración MNB-activo o MNB-pasivo.

En ambos casos, la unidad de control deberá actuar como *sniffer*<sup>2</sup> para procesar las medidas realizadas por el UE y extraer el estado actual de los canales de propagación entre las antenas implicadas en el MNB y el usuario.

En el MNB-pasivo la unidad de control envía directamente la información proveniente de los usuarios, una vez procesada, a la entidad de Procesado en *Array* que se encargará de decidir la potencia transmitida por cada antena. Además, en el MNB-pasivo la unidad de control es la encargada de recibir la información relativa a la asignación de recursos enviada desde el RNC, tanto de canales dedicados como de canales compartidos, y transmitirla tal cual a la entidad de Procesado, indicándole que es ella la encargada de decidir las antenas por las que se transmitirá la información.

En caso de MNB-activo, la información de control en enlace ascendente se deberá procesar igualmente. Sin embargo, en esta ocasión el MNB será el encargado de decidir sobre la asignación de recursos. Por una parte, a través del canal convencional NBAP, se recibirán los paquetes MAC-PDU y la información sobre los TFCI de los usuarios de voz y de la información de los canales comunes y de control. Por otro lado, se recibirá por el canal ATM virtual configurado con el RNC las RLC-PDU de los usuarios transmitiendo sobre el canal de transporte compartido DSCH. Por lo tanto, la información disponible en la unidad de control es la siguiente:

- Los niveles relativos de potencia de CPICH recibidos por todos los usuarios así como los niveles de interferencia experimentados en sentido descendente.
- El estado de las colas de los usuarios DSCH así como los retardos acumulados.

---

<sup>2</sup> En informática, un *packet sniffer* es un programa de captura de las tramas de red.

- Los datos transmitidos a los usuarios de voz.

A partir de esta información, la unidad de control deberá ajustar la transmisión realizada por todos los usuarios, configurando la prioridad en el uso de los canales compartidos, disminuyendo, si fuera necesario, la tasa de transmisión de los usuarios de voz y seleccionando sobre qué antena o antenas se va a transmitir y con qué relación de potencia entre ellas.

Una vez tomada esta decisión se deberá mandar las nuevas MAC-PDU a la unidad de Procesado en *Array*, indicándole los niveles de potencia asignados a cada antena. En este caso, el protocolo ADP está integrado en el propio MAC-mnb puesto que la asignación de recursos se hará teniendo ya en cuenta la distribución de antenas y de usuarios.

### **2.3.2 Procesado en *array***

En enlace descendente, la entidad de Procesado en *Array* se encarga de distribuir la información de las capas superiores a las distintas unidades remotas y ajustar los niveles de potencia en aquellos casos en los que la señal dirigida a un usuario se transmita simultáneamente por más de una antena. Si fuera necesaria introducir algún desfase entre las señales transmitidas, también se realizaría en este elemento.

En el caso del MNB-pasivo, será la propia entidad de Procesado en *Array* la que tome esta decisión, para lo que deberá disponer de cierta información de los usuarios y de la red, información que será suministrada por la Unidad de Control. En el caso del MNB-activo, la decisión habrá sido tomada con anterioridad por la Unidad de Control, de manera que únicamente se encargará de acatar las indicaciones recibidas y distribuir los datos a las unidades de Tx/Rx correspondientes.

En enlace ascendente, la entidad de Procesado en *Array* recibirá desde una o más antenas la señal transmitida por cada uno de los usuarios. Será por tanto en esta entidad en la que se realizará la combinación MRC, gracias a la cual se obtendrá una



importante ganancia por diversidad en frecuencia, o micro-diversidad. Tras la combinación se extraerán los bits de información y se comenzará toda la cadena de recepción de las MAC-PDU.

### 2.3.3 Unidad Tx/Rx

A partir de los bloques de datos recibidos de la capa MAC, el Nodo B deberá realizar una serie de labores de conformación de la señal de RF, desde la concatenación de bloques, pasando por la codificación, fragmentación y ensanchado, hasta la modulación.

La funcionalidad de las unidades de transmisión puede dividirse en 3 etapas o fases. En la primera, el Nodo B se encarga de crear los canales físicos. Una vez los bits de datos están ya ordenados en *slots*, en la segunda fase se ensancha la señal, utilizando los códigos de ensanchado, y se modula la señal resultante en QAM para, posteriormente, aplicar los códigos de aleatorización. Finalmente, en la tercera fase se modula la señal en frecuencia, ajustándola a la portadora correspondiente.

Sin duda la fase más compleja es la primera, ya que la unidad de Tx debe configurar todos los canales físicos a partir de la información de datos y de control recibida de la entidad de Procesado en *Array*. Las distintas etapas seguidas por el equipo de Tx en esta fase se recogen en la Figura 2.14.

Respecto a la recepción, la unidad Rx se encarga de filtrar la señal recibida y aplicar los códigos de aleatorización de los usuarios conectados. Una vez obtenida la señal de cada usuario, ésta se pasa a la entidad de Procesado en *Array* que será la encargada de combinar la señal procedente de las distintas antenas, como si de un *softer handover* se tratara.

Por cada unidad remota se deberá utilizar una unidad de Tx/Rx ya que la información transmitida por cada antena no será la misma. Por último, es importante destacar que las unidades de transmisión y recepción serán las mismas que las utilizadas en cualquier Nodo B convencional.

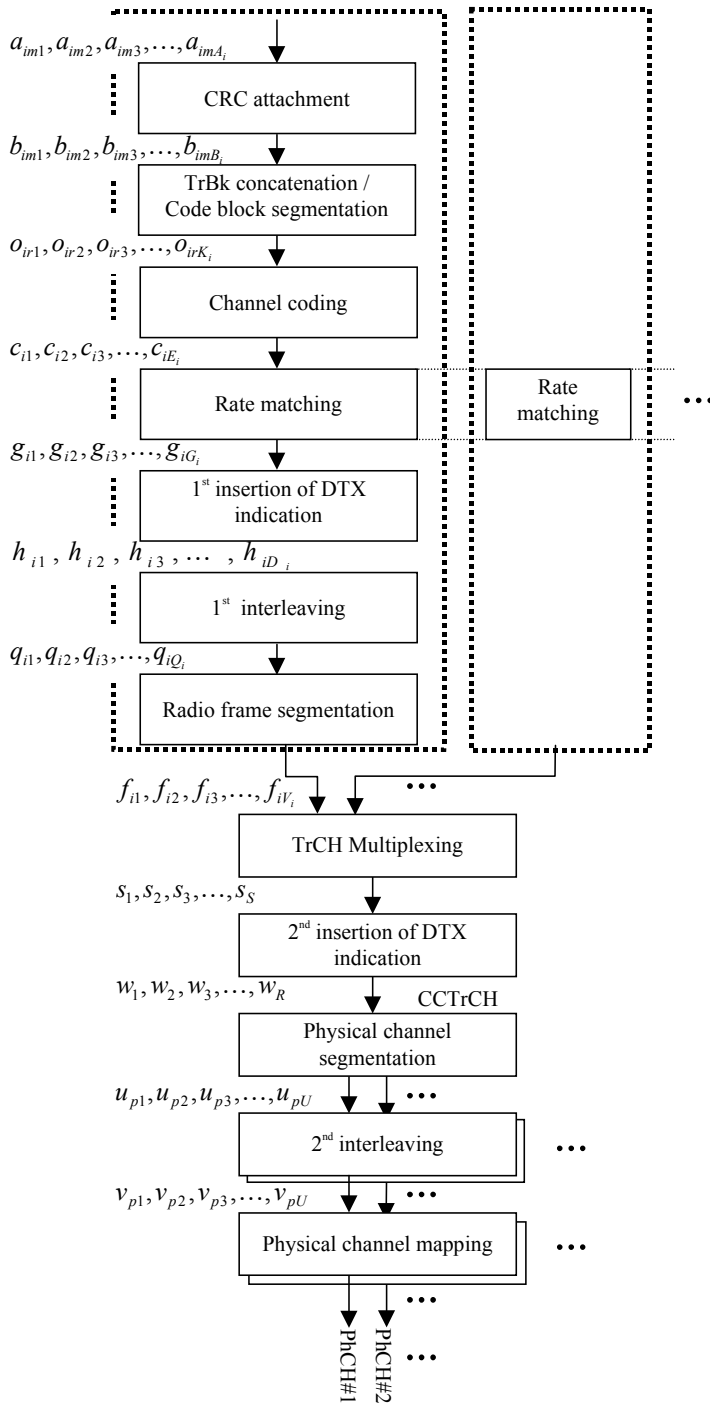


Figura 2.14 Proceso de conformación de los canales físicos [12]

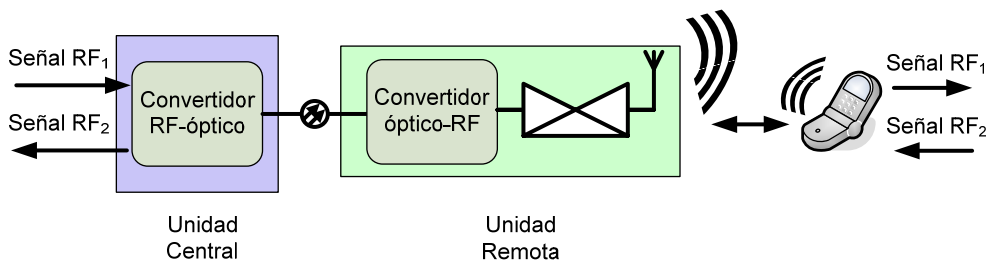
### 2.3.4 Convertidores de radio - fibra

Las características de baja dispersión y bajas pérdidas de la fibra óptica monomodo la han hecho el medio de transmisión elegido en muchas aplicaciones de sistemas de distribución de señal. La Figura 2.15 muestra el diagrama de bloques de un sistema radio-fibra básico con topología punto a punto y comunicación bidireccional.

En el caso del MNB, la antenna transmisora está situada en una posición remota respecto de la unidad central. La señal a transmitir por la estación base se modula sobre el transmisor óptico y se lleva mediante un enlace de fibra, hasta el receptor óptico, situado en la localización remota de la antenna emisora. Después de realizar la fotodetección, la señal eléctrica es amplificada y radiada a través de la antenna.

La función del sistema radio sobre fibra será pues doble. En primer lugar deberá transportar la señal de RF desde la unidad central hasta la unidad remota. En segundo lugar deberá amplificar la señal para que alcance los niveles de potencia necesarios para su transmisión.

Es importante destacar de nuevo que la generación de la señal de información, modulada ya adecuadamente a la frecuencia portadora UMTS (en torno a los 2 GHz), se realiza en la unidad central, mientras que la antenna transmisora se ubica en la posición idónea para la emisión de la señal vía radio. Es aquí donde los enlaces de fibra óptica juegan un papel importante, proporcionando la solución óptima para comunicar la unidad central con la antenna transmisora.



**Figura 2.15 Diagrama de bloques del sistema Radio-Fibra del MNB**

El uso de la fibra óptica como medio de transmisión de señales de radiofrecuencia lleva asociado una serie de importantes ventajas frente al cable coaxial. De entre ellas cabe destacar su baja figura de ruido, alto margen dinámico, inmunidad frente a interferencias, transporte de señales en su formato RF original contando con que el procesamiento complejo se encuentra en la localización remota, seguridad frente a posibles intervenciones de línea, tamaño y peso del cable menores, mayor capacidad de ancho de banda, menores pérdidas de señal y flexibilidad en cuanto a las configuraciones de sistema.

Como ejemplo de convertidor de radio sobre fibra se ha elegido el enlace óptico para aplicaciones de banda ancha de la casa ORTEL, división de EMCORE, el cual es una alternativa de bajo coste y altas prestaciones al cable coaxial en una banda que va desde los 5 MHz hasta los 2700 MHz, con lo que es válido para UMTS. Este equipo permite distancias de transmisión mucho mayores que las posibles mediante cables de cobre, a la vez que mantiene un alto nivel de calidad de la señal de radiofrecuencia. De entre sus características generales cabe mencionar su facilidad de instalación, inmunidad frente a interferencias electromagnéticas, distancias de transmisión superiores a 40 km y ganancia ajustable para la señal de radiofrecuencia.

Al diseñar los enlaces de fibra óptica se debería considerar la distancia entre los convertidores y las características de los equipos de recepción y transmisión del Nodo B para así establecer la potencia necesaria de la señal de RF a introducir en el enlace. Sin embargo, dado que el objetivo de este capítulo es únicamente mostrar la viabilidad tecnológica de la nueva arquitectura radio Multinodo B, este diseño queda fuera de los objetivos de esta Tesis Doctoral.



Figura 2.16 Equipo de transmisión radio sobre fibra, cortesía de ORTEL

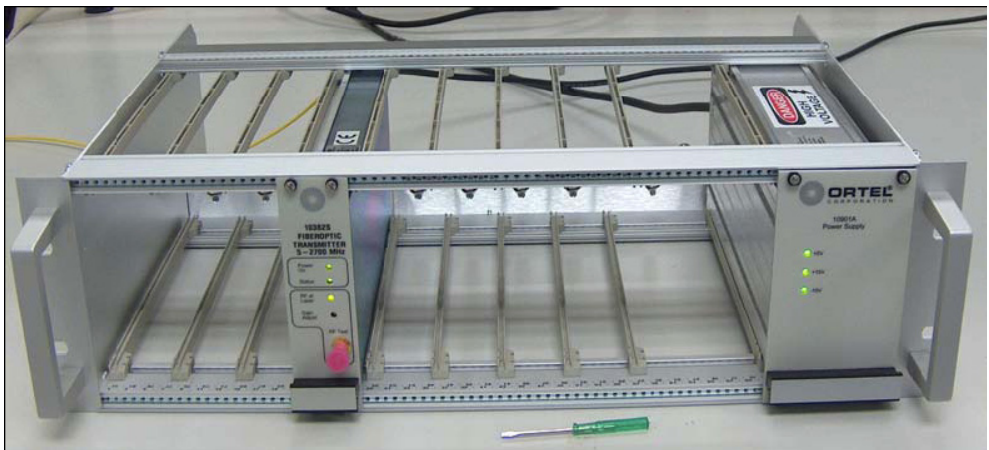


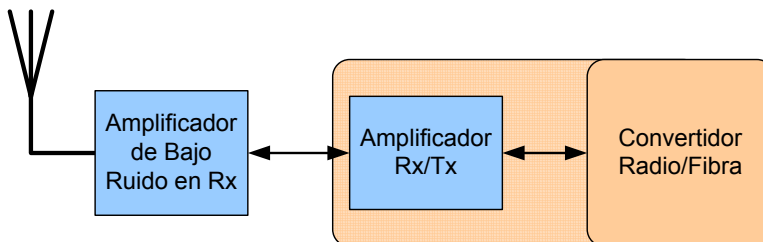
Figura 2.17 Chasis con fuente de alimentación y transmisor, cortesía del GCO-UPV

### 2.3.5 Amplificador bidireccional

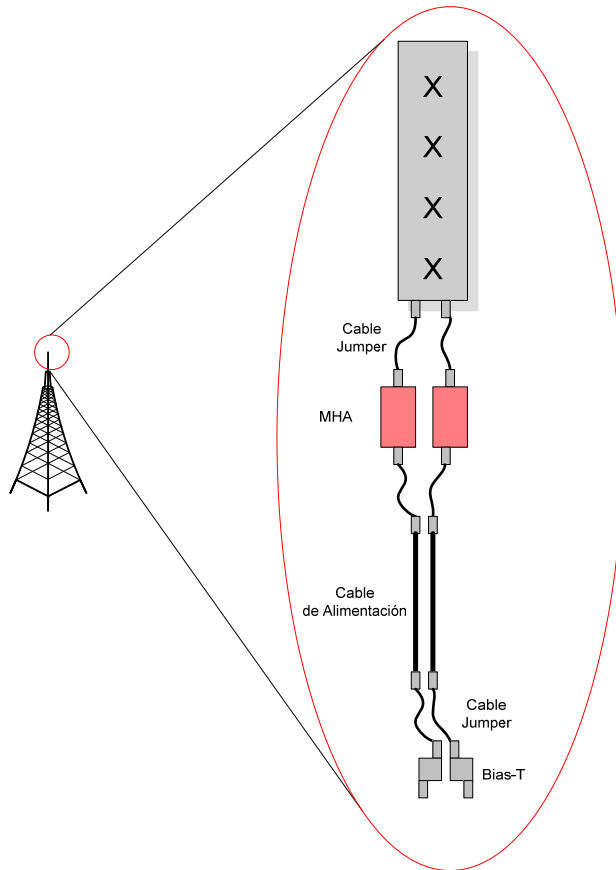
A pesar de que el convertidor radio sobre fibra es capaz de amplificar la señal en ambos sentidos de la comunicación, es conveniente utilizar en el enlace ascendente un amplificador previo de alta calidad que permita obtener un mejor comportamiento del sistema y sobre todo una menor figura de ruido. Si el conversor radio-fibra no tuviera integrado un sistema de amplificación habría que añadir un amplificador lineal

convencional a la entidad remota, similar al utilizado en cualquier Nodo B. La estructura final de amplificación se muestra en la Figura 2.18. En sentido descendente el amplificador deberá ajustar los niveles de potencia a los requeridos, tal y como se hace en un Nodo-B convencional. Por el contrario, en sentido ascendente, será necesario introducir un amplificador con muy baja figura de ruido, situado lo más cercano posible a la antena, conocido como *Mast Head Amplifier* (MHA). Gracias a la utilización de este tipo de amplificadores se reducirá la figura de ruido de la cascada de dispositivos que la señal atraviesa desde la antena en recepción hasta su procesamiento en banda base en la unidad central. De esta manera se mitigará ostensiblemente el incremento en la figura de ruido que se experimentaría al utilizar el sistema MNB a costa de una reducción mínima en la cobertura de la celda en enlace descendente, reducción debida a la inserción de un factor de pérdidas en este enlace.

El montaje del MHA se muestra en la Figura 2.19. Es importante destacar que realmente el primer elemento que se encuentra la señal en recepción tras la antena es el cable *jumper* que deberá ser de altísima calidad para así evitar cualquier pérdida. De hecho se suele suponer que las pérdidas de este cable son exactamente de 0dB. Por su parte el cable de alimentación lleva simultáneamente los datos a la antena y alimentación al MHA, siendo el Bias-T el encargado de introducir la alimentación en el cable.



**Figura 2.18 Estructura del amplificador**



**Figura 2.19 Montaje del MHA**

La reducción de la figura de ruido conseguida por la utilización del MHA se justifica examinando la ecuación de Friis característica de un sistema en cascada de dispositivos en recepción.

La Tabla 2.3 recoge valores típicos de ganancia y figura de ruido de los distintos elementos integrados en el sistema de recepción MNB. Con los datos recogidos en la tabla se pasa de una figura de ruido sin MHA de 4.6 dB (considerando sólo las pérdidas del cable y la figura de ruido del conversor) a un valor de la figura de ruido de 2.6 dB gracias a la utilización del MHA. Esta ganancia de 2 dB en la figura de ruido afecta directamente a las prestaciones del sistema en enlace ascendente suponiendo un aumento considerable de la cobertura y la capacidad.



**Figura 2.20 Ejemplo de MHA, cortesía de Unity Wireless Systems**

Finalmente, se muestra en la Figura 2.20 un dispositivo MHA. Sus reducidas dimensiones (320x170x76 mm) y su bajo peso (<8 kg) hacen que su montaje sea muy sencillo.

Elemento	Ganancia UL	NF	Pérdidas en DL
MHA	12.0 dB	1.3 dB	0.3 dB
Cable Alimentación	-2.0 dB	2.0 dB	2.0 dB
Bias-T	-0.3 dB	0.3 dB	0.3 dB
Radio-Fibra	26 dB	2.5 dB	-
BTS	-	3 dB	-

**Tabla 2.3 Parámetros típicos del sistema de recepción**

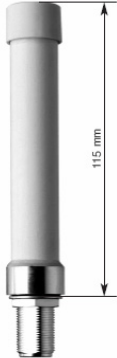
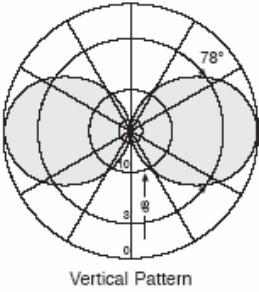
### 2.3.6 Antena


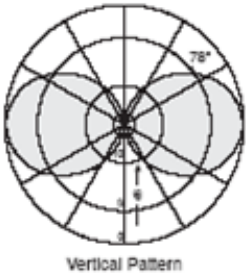

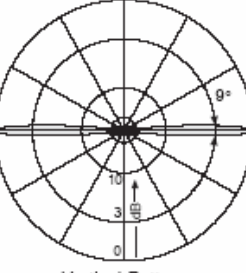
Cuando se trata de un emplazamiento macro-celular, es decir, una estación base situada encima de una azotea, se suele trabajar con antenas directivas, de ancho de haz típico de 60°-65°, configurando un mapa de cobertura sectorial. Sin embargo, cuando se



trata de un despliegue micro-celular las antenas están en las propias calles con lo que existe mucha reflexión en los propios edificios. Una configuración sectorial tendría una gran cantidad de interferencias intra-celulares que perjudicarían gravemente el funcionamiento del sistema. Por eso en entornos micro-celulares se suele utilizar antenas omnidireccionales.

Las operadoras utilizan micro-celdas para dar cobertura a radios celulares de no más de 200 o 300 metros. La antena de una micro-celda es de forma tubular y alargada y tiene una altura que oscila desde apenas diez centímetros hasta tres metros, dependiendo de la ganancia requerida (ver Tabla 2.4). Generalmente se colocan en las propias fachadas de los edificios, detrás de postes publicitarios o sobre postes de iluminación. La potencia que alimenta una antena micro-celular es relativamente pequeña, de no más de cinco vatios, de manera que, apenas dos metros por debajo del plano medio de la antena, los campos electromagnéticos radiados estén por debajo de los límites establecidos en las directrices europeas. Sin embargo, algunos estudios recomiendan no transmitir con más de 1W de potencia, valor que será utilizado como cota máxima en esta Tesis Doctoral.

Aspecto	Diagrama de Radiación	Ganancia	Longitud
		<p>2 dBi</p>	<p>11.5 cm</p>

	 <p style="text-align: center;">Vertical Pattern</p>	<p style="text-align: center;">2 dBi</p>	<p style="text-align: center;">21.6 cm</p>
	 <p style="text-align: center;">Vertical Pattern</p>	<p style="text-align: center;">10 dBi</p>	<p style="text-align: center;">303.3 cm</p>

**Tabla 2.4 Modelos de antena omnidireccional para UMTS cortesía de Kathrein**

Una de las ventajas del uso de este tipo de antenas es que su perfil alargado y tubular permite camuflarlas con gran facilidad, por lo que no suponen una contaminación estética de las ciudades y tienen una mayor aceptación popular. Además, se puede hacer uso de técnicas de camuflaje utilizando para ello papel cartón u otro tipo de materiales que no degraden la señal transmitida.

## Referencias

- [1] “Mejora de Arquitecturas de Acceso Radio UMTS Mediante Mutinodos B”, Comisión Interministerial de Ciencia y Tecnología, CYCIT TIC2002-02678.
- [2] 3GPP TS 25.401 V5.9.0, “UTRAN Overall Description”, 2004.
- [3] 3GPP TS 25.301 V7.0.0, “Radio Interface Protocol Architecture”, 2006.
- [4] 3GPP TS 25.321 V7.0.0, “Medium Access Control (MAC) Protocol Specification”, 2006.
- [5] 3GPP TS 25.322 V7.0.0, “Radio Link Control (RLC) Protocol Specification”, 2006.
- [6] 3GPP TS 25.323 V7.0.0, “Packet Data Convergence Protocol (PDCP) Specification”, 2006.
- [7] 3GPP TS 25.324 V7.0.0, “Broadcast/Multicast Control (BMC)”, 2006.
- [8] 3GPP TS 25.331 V6.0.1, “Radio Resource Control (RRC)”, 2004.
- [9] 3GPP TS 34.108 V5.3.0, “Common Test Environments for User Equipment (UE); Conformance Testing”, 2004.
- [10] 3GPP TS 25.427 V5.4.0, “UTRAN Iub/Iur interface user plane protocol for DCH data streams”, 2005.
- [11] 3GPP TR 25.993 V6.12.0, “Typical Examples of Radio Access Bearers (RABs) and Radio Bearers (RBs) Supported by Universal Terrestrial Radio Access (UTRA)”, 2005.
- [12] 3GPP TS 25.212 V5.6.0, “Multiplexing and Channel Coding (FDD)”, 2003.



# Capítulo 3

## *Evaluación de las Prestaciones de la Arquitectura MNB*

*No hay asunto, por complicado que sea, que estudiado con  
paciencia e inteligencia no se complique más.*

*New Speakers Handbook*

### **3.1 Introducción**

La arquitectura Multinodo B propuesta permite crear una red fácilmente configurable de antenas espacialmente distribuidas. Cada emplazamiento adicional que se introduzca en el MNB representará una nueva área de cobertura, área que podrá estar solapada con otros elementos del MNB o no. Si los nuevos emplazamientos se sitúan sobre la azotea de edificios de altura significativa se dispondrá de una cobertura macro-celular, mientras que si están localizados a nivel de calle se crearán celdas micro-celulares. En general, parece obvio que la utilización del MNB supondrá una mejora considerable en cuanto a cobertura. Además, puesto que la distancia entre transmisor y receptor se reduce, la potencia transmitida será considerablemente menor, lo que supone una gran mejora en UMTS, al tratarse ésta de una tecnología limitada por interferencias.

Así pues, parece claro que en términos de cobertura y potencia transmitida la nueva arquitectura MNB supone una importante ganancia, ganancia que se evaluará en el capítulo 5 cuando se estudie el comportamiento global del sistema. Sin embargo, UMTS también está limitado por su capacidad en términos de carga máxima o *pole*

*capacity* que puede soportar el sistema. Así pues, un usuario cualquiera introduce en el sistema un factor de carga en enlace descendente igual a:

$$\lambda_i = \frac{\alpha_i}{\alpha_i + \frac{G_{pi}}{(E_b/N_0)_i}} \quad (3.1)$$

donde  $\alpha_i$  es el factor de actividad del usuario,  $G_{pi}$  es su ganancia de procesado (también denominada SF) y  $(E_b/N_0)_i$  es la relación de energía de bit a densidad espectral de ruido más interferencias necesaria para tener una calidad de servicio determinada. Finalmente, la carga total del sistema en el enlace descendente será:

$$\eta_{DL} = [f + (1 - o)] \sum_{i=1}^N \lambda_i \quad (3.2)$$

siendo  $f$  la relación entre interferencias intra e inter-sistema,  $o$  el factor de ortogonalidad y  $N$  el número de usuarios en el sistema. Lo que se deberá garantizar en el sistema es que ni se supera el factor de carga máximo, normalmente establecido en el 75%, ni se sobrepasa la potencia disponible en la estación base.

Observando la expresión (3.2), se puede concluir que para reducir el factor de carga se debe: reducir el factor de interferencia intercelular, o aumentar la ortogonalidad de los códigos o bien reducir el factor de carga de un usuario. Un escenario micro-celular se caracteriza por tener normalmente visión directa (LOS, *Line of Sight*) entre transmisor y receptor, por lo que se suele considerar que el factor de ortogonalidad es mayor. El factor de interferencia intercelular también se verá reducido puesto que la señal queda confinada en las calles y no hay visión directa (NLOS, Non LOS) entre las estaciones micro-celulares. Por último, disponer de más de una estación transmitiendo simultáneamente a un usuario aumenta la diversidad en frecuencia y crea macro-diversidad por tener dos enlaces totalmente incorrelados en cuanto a las pérdidas de propagación. Esta diversidad adicional llevará consigo una reducción importante de los requisitos de  $E_b/N_0$ , como se verá a lo largo de este capítulo.

A priori, la nueva arquitectura MNB parece que aporta todos los elementos necesarios para no solamente aumentar la cobertura y reducir por tanto el consumo de

potencia del sistema en ambos sentidos de la comunicación, sino para también aumentar la capacidad del sistema.

Este capítulo de la Tesis Doctoral tiene como objetivo demostrar que, efectivamente, el uso de la estructura distribuida de antenas permite reducir la  $E_b/N_0$  objetivo. Para ello, se ha desarrollado un simulador de enlace completo, cuya descripción se verá de manera resumida a lo largo del capítulo y más exhaustivamente en el Anexo A.

Una vez descrito el simulador se validará su funcionamiento, comparando los resultados obtenidos con otros disponibles en la literatura. Ya validado, en el apartado 3.3 se estudiará el efecto de transmitir con más de una antena. Uno de los problemas observados en esta transmisión simultánea es precisamente que las contribuciones de ambas antenas pueden solaparse temporalmente, lo que supone una interferencia indeseable. En el apartado 3.3.4 se presentará una solución a este problema, solución que permitirá combinar la transmisión sincronizada de información de ambas antenas. Además, será necesario llevar a cabo una estimación de canal específica, analizada en el apartado 3.3.5. Una vez ajustada la transmisión entre las antenas del MNB, se estudiará la ganancia experimentada por la diversidad en frecuencia adicional que introduce el MNB. Se estudiarán por separado dos escenarios: en el primero ambas antenas formarán una celda macro-celular, mientras que en el segundo una de las antenas estará situada en un emplazamiento micro-celular. Finalmente, en el apartado 3.4, se recogen las *Look-Up Tables* (LUTs) para todos los servicios y escenarios considerados en esta Tesis Doctoral.

## **3.2 Simulación de Enlace de la Red MNB**

Las prestaciones de cualquier modificación realizada sobre el sistema UMTS deben ser evaluadas antes de poder implementar estos cambios en redes reales. Dada la enorme complejidad de los sistemas de comunicaciones móviles UMTS, resulta inviable llevar a cabo estos estudios por medio de métodos analíticos. El empleo de simuladores, capaces de reproducir el comportamiento de los sistemas de

comunicaciones a partir del modelado de sus distintos bloques funcionales, constituye una buena alternativa para la evaluación de estos sistemas, logrando un compromiso entre complejidad, coste, tiempo de desarrollo y precisión.

Además de permitir la evaluación de nuevas técnicas, el uso de simuladores tiene otra serie de ventajas adicionales:

- Permite, de manera sencilla, mostrar el comportamiento del sistema, favoreciendo el proceso de formación.
- Posibilita la realización de pruebas con coste reducido, permitiendo una planificación y despliegue más eficaz.
- Proporciona una potente herramienta de previsión del comportamiento de la red, haciendo posible la toma de acciones proactivas.

Un posible enfoque para la simulación de una red celular sería considerar un modelo global de la red, que reprodujese la interacción entre un número elevado de estaciones base y de terminales móviles, teniendo en cuenta aspectos de tráfico y movilidad de los usuarios, y considerando, además, de manera detallada los aspectos de codificación, modulación y propagación asociados a cada uno de los enlaces base-móvil. Sin embargo, una simulación directa de este tipo supondría una carga computacional prohibitiva. Para reducir la complejidad de la simulación se suele descomponer ésta en dos etapas o niveles de abstracción, conocidos como nivel de enlace y nivel de sistema o de red.

En las simulaciones a nivel de enlace se obtienen las prestaciones de un solo enlace radio, estudiando el efecto del canal móvil sobre la transmisión de datos. Se modela un único enlace base-móvil (o móvil-base), incluyendo en la simulación los procesos de modulación, codificación, demodulación, caracterización del canal, algoritmo de control de potencia, etc. En ocasiones se incluyen también aspectos de más bajo nivel como la sincronización de código o recuperación de portadora en el receptor. Los principales resultados de estas simulaciones son las funciones que relacionan la tasa de error de bit, denominada BER (*Bit Error Rate*), o de trama, denominada FER (*Frame*



*Error Rate*), con la relación potencia de señal a potencia de ruido mas interferencias, SIR (*Signal to Interference Ratio*), a la salida del receptor, y ésta con la relación potencia de portadora a potencia de ruido e interferencia, CIR (*Carrier to Interference Ratio*), a la entrada del receptor. Estas tablas que relacionan la calidad media de la señal recibida (en forma de CIR o  $E_b/N_0$  medio) con la tasa de error de bit se conocen en la literatura como *Look-Up Tables*.

Por otro lado, las simulaciones a nivel de sistema permiten evaluar las prestaciones de la red global. El modelo empleado en este nivel considera un conjunto de estaciones base y todos los usuarios asociados a las mismas. Se modela la señal recibida por cada usuario, así como las interferencias de otros usuarios y el ruido, teniendo en cuenta las pérdidas de propagación, el efecto del *shadowing* y la movilidad y generación del tráfico de los usuarios. Se obtiene así la relación entre la potencia de portadora y la potencia interferente y de ruido, CIR, para cada usuario, teniendo en cuenta la configuración de la red y los algoritmos de gestión de recursos radio empleados en el sistema. Estos valores de CIR pueden traducirse a valores de BER o FER empleando los resultados obtenidos en las simulaciones de enlace. De esta manera, se puede obtener una medida de la calidad de servicio proporcionada al usuario y de la capacidad del sistema.

Como se ha visto, aunque ambos niveles de simulación se implementan y evalúan independientemente, los resultados obtenidos en las simulaciones de enlace actúan como interfaz entre ambos niveles.

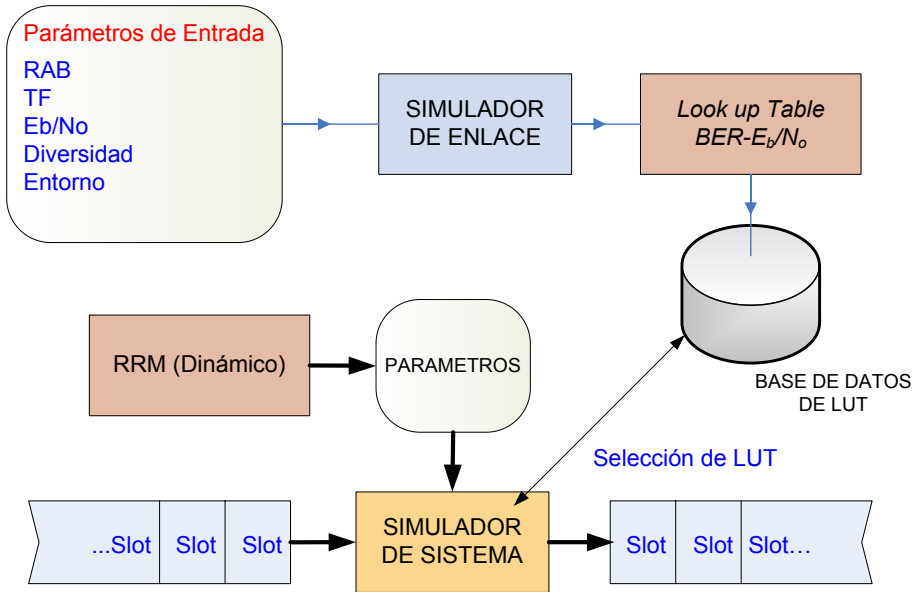
Uno de los principales motivos para descomponer la simulación en dos niveles es la diferencia entre las resoluciones temporales requeridas en cada uno de ellos. En las simulaciones a nivel de sistema es necesario extender las simulaciones durante al menos 10-20 minutos [1][2], para obtener unos estadísticos aceptables de los modelos aleatorios de generación de tráfico y de movilidad empleados. Por otra parte, en el nivel de enlace, para obtener una caracterización adecuada del enlace radio, se requiere una resolución a nivel de *chip* o de símbolo. Si se considerase esta resolución temporal en las simulaciones a nivel de sistema el coste computacional del simulador resultante sería inviable.

Como ya se ha comentado, uno de los primeros objetivos de esta Tesis Doctoral es desarrollar un simulador a nivel de enlace que permita modelar la red de acceso radio de UMTS, UTRAN, en sentido descendente y en el modo FDD.

Como consecuencia de su flexibilidad, la red UTRA-FDD puede configurarse de múltiples maneras. Para organizar los servicios ofrecidos por la UTRAN los organismos de estandarización especificaron un conjunto de portadoras radio o *Radio Access Bearers* (RAB), [3]. Una RAB queda inequívocamente especificada por una tabla que contiene el conjunto de parámetros que configuran los protocolos de capa física y capa de nivel de enlace de datos (MAC y RLC) para un servicio concreto con una determinada QoS. Todos los parámetros más importantes, como el formato de transporte o *Transport Format* (TF), el intervalo de tiempo de transmisión o *Transmission Time Interval* (TTI), el tipo de canal lógico asociado, el tipo de codificación de canal, el tamaño del bloque de transporte o *Transport Block* (TB), etc. se especifican en dichas tablas RAB, tal y como se puede observar en el ejemplo mostrado en la Tabla 2.1. Hay un elevado número de RABs especificadas en [3] pero solamente un conjunto de ellas se han considerado relevantes para esta Tesis Doctoral.

El conjunto de parámetros que definen unívocamente las prestaciones de nivel de enlace son:

- La tabla RAB específica para el enlace descendente.
- El formato de transporte dentro de dicha RAB.
- El valor de  $E_b/N_0$  medio (donde  $N_0$  incluye todas las fuentes de interferencia más el ruido térmico).
- Los procedimientos de transmisión/recepción, especificando si hay mecanismos de diversidad.
- Entorno de propagación: *outdoor/indoor, pedestrian/vehicular, etc.*



**Figura 3.1 Interoperabilidad entre simulaciones de Enlace y de Sistema [4]**

La Figura 3.1 muestra claramente los dos pasos necesarios en la simulación de un sistema: primero se obtienen los estadísticos de primer orden de la BER para un valor medio de  $E_b/N_0$ , definiendo el conjunto de LUTs necesarias, para posteriormente utilizar dichas tablas en un simulador de sistema.

En este momento podría surgir la duda acerca de la suficiencia de los estadísticos de primer orden para caracterizar correctamente el nivel radio. Hay varias razones que justifican esta idea [4]:

- El entrelazado reduce la correlación del error (memoria del canal) dentro de un TTI.
- Para velocidades superiores a 13 km/h el tiempo de coherencia del canal es menor de 10 ms, así que no hay correlación entre TTI consecutivas.

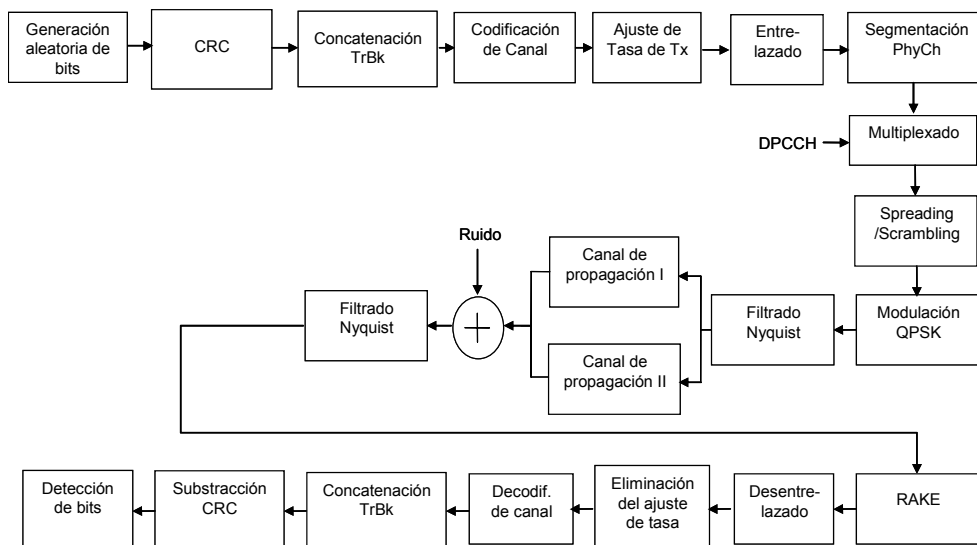
Sin embargo, la duda surge ya que, por debajo de 13 km/h, el control de potencia aplicado 1500 veces por segundo es capaz de cancelar el efecto de los

desvanecimientos rápidos o *fast fading* y del *shadowing*, por lo que teóricamente el nivel de  $E_b/N_0$  permanecería prácticamente constante de *slot* a *slot*. Esta idea, junto con el hecho de que el canal a estas velocidades bajas no está incorrelado, justificaría la necesidad de correlar los valores obtenidos de BER para *slots* consecutivos, es decir justificaría un estudio de los estadísticos de segundo orden de la BER para un  $E_b/N_0$  medio dado. Sin embargo, la utilización del primer estadístico sigue siendo válida puesto que en entornos con elevada carga las variaciones en las interferencias y en los TFs empleados son altas, por lo que se producen continuos cambios de la  $E_b/N_0$  independientemente del comportamiento del canal.

En este apartado no se abordará la descripción detallada del simulador implementado, reservando esta descripción al Anexo A. Lo que se incluye, además de una descripción breve del simulador y sus principales bloques funcionales, es la validación del mismo, comparando los resultados obtenidos con ciertos modelos teóricos y con otros simuladores encontrados en la literatura. Además, se analizarán distintas técnicas utilizadas para mejorar la estimación de canal en UMTS.

### 3.2.1 Estructura del simulador de enlace

Básicamente, el simulador de enlace está compuesto por dos tipos de elementos: los necesarios para llevar a cabo la configuración de la capa física y los relacionados con los fenómenos de degradación. La configuración de la capa física incluye aspectos tales como la codificación necesaria para el FEC (*Forward Error Correction*) y entrelazado, los códigos de ensanchado (OVSF y Gold), el filtrado *Nyquist*, el receptor *Rake*, la estimación de canal, los mecanismos de control de potencia y las técnicas de diversidad. Entre los fenómenos de degradación se encuentran las interferencias inter e intra-celulares, el ruido térmico, las interferencias entre símbolos o ISI (*Inter Symbol Interferente*), el *fast fading* debido al multicamino y el efecto *doppler* asociado a la velocidad del usuario.



**Figura 3.2 Diagrama de Bloques del simulador de Enlace**

La Figura 3.2 muestra el diagrama de bloques implementado del simulador de enlace. La descripción detallada de cada uno de estos bloques se encuentra en el Anexo A. Se trata de un diagrama de bloques conceptual, en el sentido de que algunos bloques están implementados con sumo detalle en el simulador mientras que otros se han considerado solamente de manera indirecta. Por ejemplo, los bits de CRC no se han calculado y añadido en el simulador al bloque de transporte puesto que su función de detección de errores no es necesaria ya que se conoce exactamente la información enviada a través del canal, por lo que la detección de errores se realiza por comparación directa de la secuencia transmitida y la recibida. Sin embargo, el efecto del CRC se ha considerado al calcular el tamaño en bits del bloque TTI y también a efectos de la eficiencia energética global necesaria para definir la potencia de ruido interferente introducida.

En el diagrama de bloques se han representado dos bloques de canal de propagación que representa el caso en el que se dispone de dos antenas del MNB transmitiendo simultáneamente. De esta forma, será posible modelar los distintos caminos de propagación experimentados en ambos enlaces. En caso de tener solamente una antena bastaría con eliminar uno de los bloques.

### 3.2.2 Suposiciones y simplificaciones

Aunque algunas suposiciones se exponen a lo largo de la descripción de los distintos bloques funcionales en el Anexo A, es importante resumir todas las simplificaciones asumidas, de manera que sean fácilmente detectables las limitaciones del simulador.

#### Implementación de múltiples usuarios

Por el Teorema Central del Límite, que postula que la suma de múltiples procesos aleatorios sigue una distribución gaussiana, el conjunto de ruido más interferencias provenientes de la propia celda y de las celdas vecinas sigue una distribución gaussiana. Por esto, en las simulaciones de enlace solamente se ha considerado un único usuario afectado por ruido blanco aditivo gaussiano (AWGN – *Additive White Gaussian Noise*).

#### Implementación y efecto del control de potencia

El control de potencia es un mecanismo fundamental en cuanto a las prestaciones del sistema, sobre todo a velocidades bajas. Como se comenta en el Anexo A, las interferencias y el ruido térmico se añaden como un ruido gausiado de varianza  $\sigma_N^2$ . El control de potencia se ha implementado de la siguiente forma: si la potencia de la señal respecto a la potencia del ruido no satisface la  $E_b/N_0$  objetivo en un *slot*, el sistema reacciona incrementado la potencia de la señal transmitida en el siguiente *slot* y viceversa. Es importante destacar que en muchos trabajos la implementación de este control de potencia se realiza de manera perfecta, ajustando exactamente el nivel de la señal [4][5]. En el simulador desarrollado en esta Tesis Doctoral se ha implementado también este control de potencia perfecto, pero además se ha incluido otra variante en la que se respetan las especificaciones del 3GPP [6], realizando incrementos de potencia de 1dB. Gracias a este segundo mecanismo de control de potencia, se evitan consideraciones optimistas del sistema, tal y como se demuestra en [7].

Otro aspecto importante a destacar es que la aplicación del control de potencia supone un incremento en la potencia media transmitida, incremento conocido en la

literatura como margen de *fast fading* [1]. En muchas ocasiones dicho margen no se considera en las simulaciones de enlace, sino que se añade a posteriori a partir de estudios prácticos. Sin embargo, en esta Tesis Doctoral se ha preferido medir empíricamente la  $E_b/N_0$  media, de manera que se consigan unos resultados más fidedignos que incluyan directamente dicho margen. Así pues, los valores de  $E_b/N_0$  calculados siempre incluyen el margen de *fast fading* exacto.

### **Seguimiento perfecto del canal**

En las simulaciones se ha supuesto que las ramas del *Rake* se actualizan dinámicamente para ajustarse temporalmente a los puntos de máxima energía de la respuesta al impulso del canal, asumiendo pues un conocimiento perfecto de los retardos asignados a cada componente multicamino. Además, el canal se considera invariante en cuanto a los retardos temporales. Al tener el receptor *Rake* una resolución temporal de un periodo de *chip*, si dos componentes coinciden en dicho periodo el receptor seleccionará aquella de la que disponga una mayor potencia.

### **No existencia de *Shadowing* ni de pérdidas de propagación**

Las pérdidas de propagación asociadas al “sombreado” o *shadowing* de edificios, montañas u otros grandes obstáculos no se tienen en cuenta en el simulador, ya que el control de potencia es capaz de compensar las pérdidas lentas asociadas a dicho desvanecimiento. Por otra parte, tampoco hay que considerar las pérdidas asociadas a la propagación, puesto que la potencia de ruido insertada en el canal se ajusta externamente para garantizar la  $E_b/N_0$  objetivo. De esta manera se independiza la  $E_b/N_0$  de las pérdidas de propagación.

### **No se producen fallos en el canal de control**

De los tres campos que se transmiten en el canal de control, únicamente los bits piloto son utilizados realmente en recepción para realizar la estimación de canal. El

resto de campos, es decir, el TFCI y los bits de control de potencia, se rellenan aleatoriamente pues es el simulador quien asume la responsabilidad de llevar a cabo estas funciones. Por lo tanto, se considera que la recepción de los comandos del control de potencia está libre de fallos.

### **Simplificaciones del modelo de canal**

Por una parte se ha supuesto *scatterers* incorrelados en las distintas componentes multicamino del canal, de modo que cada una de ellas se puede generar de manera independiente. Además, se ha considerado *scattering* isotrópico y antenas receptoras omnidireccionales, lo que ha permitido la utilización del espectro *doppler* clásico o de *Clarke*. Por otro lado, se ha supuesto que el canal es invariante en cuanto a los retardos temporales y que se comporta como un filtro FIR (*Finite Impulse Response*). Finalmente, al suponer que no hay visión directa entre el transmisor y el receptor, se ha seleccionado un modelo de canal *Rayleigh*, aplicando el Teorema Central del Límite.

Otro aspecto importante a destacar es que se ha discretizado la evolución del canal generando 100 muestras del mismo por tiempo de coherencia ( $T_c$ ), con

$$T_c = \sqrt{\frac{9}{16\pi f_D^2}} \quad (3.3)$$

siendo  $f_D$  la frecuencia *doppler* máxima.

### **Traslación de frecuencia**

En la simulación se ha omitido la etapa de conversión de frecuencia ya que al trabajar con la envolvente compleja se elimina la portadora. Además, no se ha incluido el efecto del ruido de fase o de las no linealidades de los conversores y amplificadores.

### **Estimación de canal**

Considerando únicamente los bits piloto transmitidos por el canal físico de control, en el simulador se han implementado tres diferentes técnicas de estimación de canal:



estimación óptima, estimación subóptima y estimación por correlación. En la estimación óptima el receptor *Rake* conoce exactamente las atenuaciones y desfases producidos por el canal en cada uno de sus *fingers* o ramas, de manera que, cada vez que cambia el canal (100 veces por  $T_c$ ), el receptor *Rake* se adapta al cambio.

En cuanto a la estimación subóptima, el principio de funcionamiento es el mismo salvo que lo único que conoce el *Rake* es el estado exacto del canal en el instante en que recibe los bits piloto. De esta manera lo que se hace es interpolar entre las dos muestras de canal del *slot* actual y anterior.

Finalmente, en la estimación por correlación se hace uso de los símbolos piloto introducidos en el canal físico de control DPCCCH. Estos símbolos piloto se utilizan para estimar el canal analizando la variación de fase y amplitud producida por el canal en esos símbolos conocidos. De igual manera que en la técnica anterior, una vez determinada la atenuación y el desfase producidos por el canal sobre los símbolos piloto, el resto de símbolos se corrigen con el valor resultante de la interpolación entre la muestra obtenida en ese *slot* y la del *slot* precursor.

Además de la estimación de canal por correlación sobre los bits transmitidos por el canal físico de control dedicado, en el simulador también se ha incluido la transmisión del canal piloto común, o CPICH, de manera que la estimación del canal se puede realizar a partir de esta secuencia de bits que también es conocida.

### **Duración de la simulación**

Para obtener unos resultados válidos se han prolongado las simulaciones hasta un tiempo siempre superior a 1000 veces el tiempo de coherencia del canal. Además, las simulaciones no finalizan hasta detectar al menos 100 errores en la transmisión de los bits de información.

Cada simulación se repite siempre al menos con tres semillas de aleatorización distintas, de manera que se garantice unos intervalos de confianza lo suficientemente pequeños como para disponer de una medida rigurosa de los resultados obtenidos.

### **3.2.3 Validación del simulador**

Cuando se desarrolla una herramienta de simulación es necesario validar su funcionamiento, con el objetivo de determinar si los resultados que esta herramienta proporciona describen con fidelidad el comportamiento del sistema simulado. En este apartado se presentan las pruebas realizadas para verificar el funcionamiento del simulador desarrollado en la presente Tesis Doctoral.

La validación de la herramienta se ha llevado a cabo en tres fases. En primer lugar, se encuentra la validación de la implementación en el simulador de los bloques funcionales. Esta validación se puede leer directamente junto a la descripción de dichos bloques en el Anexo A. A continuación, se ha estudiado el comportamiento teórico del sistema y el efecto producido por diversos factores, como la velocidad del usuario o la estimación de canal. Por último, se ha verificado el funcionamiento conjunto del simulador llevando a cabo una serie de simulaciones para comprobar que los resultados que pueden obtenerse con la herramienta son coherentes con los que otros autores han obtenido con anterioridad.

#### **3.2.3.1 Validación respecto al comportamiento teórico del sistema**

Para validar el programa de simulación es importante comprobar los resultados respecto a una predicción teórica. Aplicando ciertas simplificaciones es posible encontrar una fórmula teórica para el valor de la BER en la red UTRAN. Es importante destacar que en todo este apartado se evalúa el comportamiento del sistema a nivel físico, es decir, calculando la tasa de error de bit físico transmitido. Se ha resaltado este aspecto puesto que en el apartado siguiente se evaluarán los resultados a nivel de red, es decir, considerando la tasa de error de bit de información transmitida. En este apartado se evalúa el sistema a nivel físico puesto que así se puede obviar el efecto producido por la codificación, efecto que no se puede incluir de manera analítica puesto que no existen modelos al respecto.

Si se asume una estimación perfecta del canal, el receptor *Rake* se comporta como un combinador MRC de  $L$  componentes, siendo  $L$  el número de ramas o *fingers* del *Rake*. Debido al ancho de banda limitado del receptor, los caminos de propagación, o *taps*, con diferencias de retardo menores que un intervalo de *chip* son interpretados como una única contribución y no son resolubles individualmente. De esta manera, solamente los *taps* que están separados al menos un intervalo de *chip* son discriminables. Si el número de *taps* útiles en el PDP que modela el entorno de propagación es mayor que  $L$ , entonces el *Rake* se ve obligado a seleccionar las componentes de mayor energía, combinando después los *taps* seleccionados mediante la técnica MRC. Este caso no se puede considerar como un procesamiento MRC puro. Sin embargo, si el número de caminos resolubles es menor o igual que  $L$ , entonces existe una expresión analítica para la probabilidad de error de bit físico [8]:

$$P_b = \frac{1}{2} \sum_{k=1}^L \pi_k \left[ 1 - \sqrt{\frac{\gamma_k}{1 + \gamma_k}} \right] \quad (3.4)$$

donde  $P_b$  es la BER media,  $\gamma_k$  es la  $E_b/N_0$  media del  $k$ -ésimo camino de propagación y

$$\pi_k = \prod_{\substack{i=1 \\ i \neq k}}^L \frac{\gamma_k}{\gamma_k - \gamma_i} \quad (3.5)$$

En la Figura 3.3 se muestran los resultados en forma de BER- $E_b/N_0$  para el enlace descendente con diferentes valores de SF. En este caso se ha simulado con las siguientes condiciones: la velocidad del terminal es de 120 km/h, no se ejecuta el control de potencia y se realiza estimación óptima del canal. De esta manera se asegura la mínima ISI. El PDP se corresponde con un entorno *outdoor* de 4 caminos de propagación separados un periodo de *chip* (260 ns) y con una potencia relativa de 0, -3, -6 y -9 dB (véase Figura 3.4). Este escenario está propuesto como caso de estudio número 3 del Anexo B de las especificaciones de la UTRAN [9] para la evaluación de las prestaciones de la red en entorno *vehicular*.

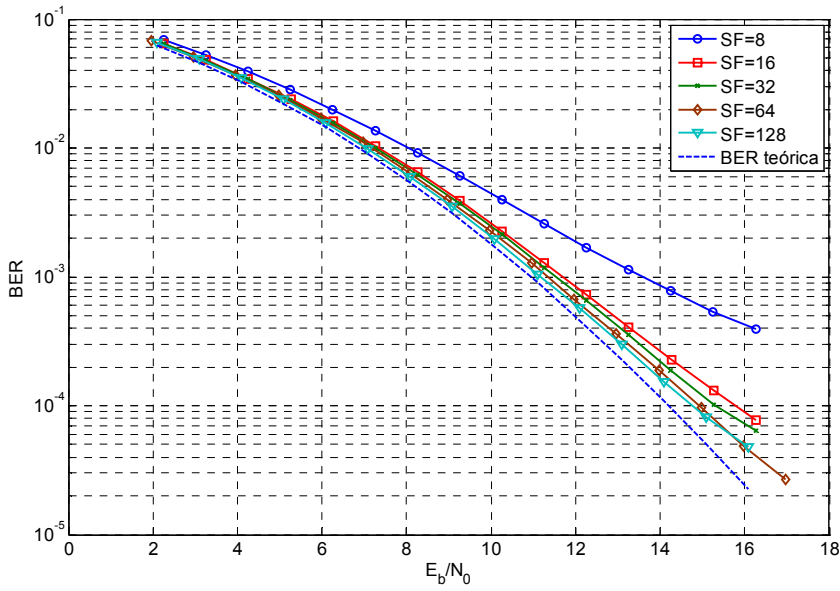


Figura 3.3 BER para una red UTRA-FDD en sentido descendente a 120 km/h con estimación óptima de canal

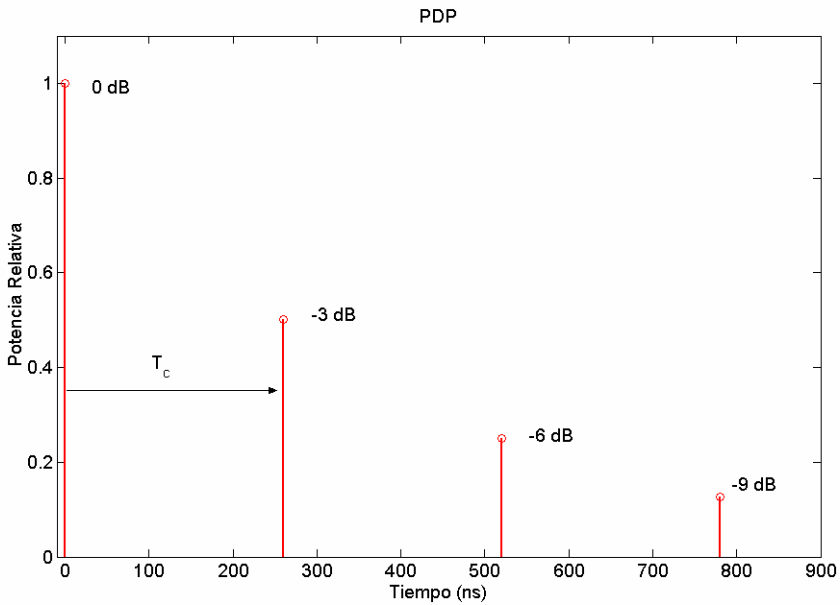
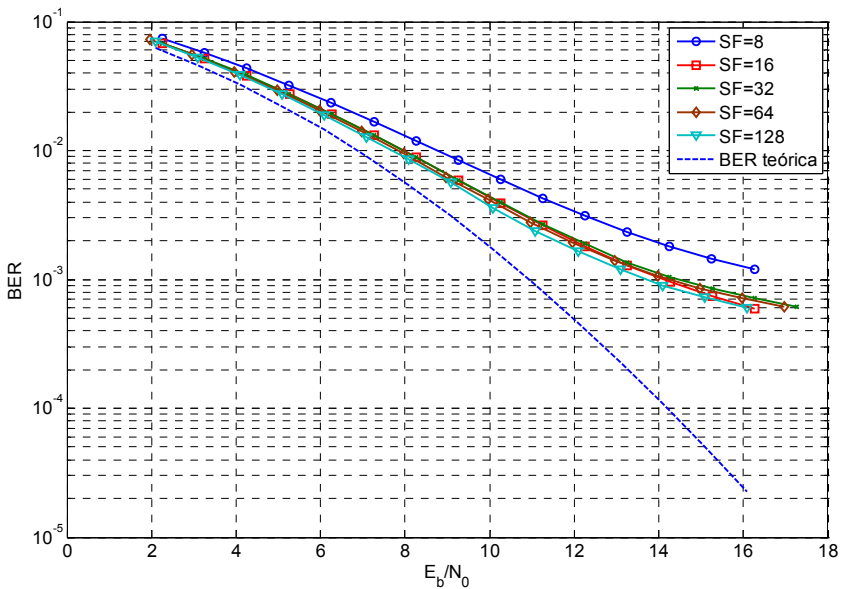


Figura 3.4 Perfil de retardo del escenario simulado

La línea discontinua de la Figura 3.3 es la BER teórica obtenida sin interferencia entre símbolos con cuatro *fingers* en el combinador MRC. Se puede ver como las curvas de BER con un alto SF tienden a coincidir con la fórmula teórica. Esto es debido a que con baja tasa de transmisión, es decir con alto SF, la duración del bit es mucho mayor que el *delay spread* del canal y no hay apenas interferencia entre símbolo. Por ejemplo, con un SF de 128 la duración de un símbolo es de 128 *chips* mientras que solamente la potencia contenida en tres *chips* actuará como potencia interferente, es decir, la proporción de interferencia respecto a la señal útil es muy pequeña por lo que los resultados se parecen a los teóricos, obtenidos sin ISI. Esta interferencia se hace más patente al disminuir en SF, efecto mostrado también en la Figura 3.3. Por ejemplo, para un SF de 8 la ISI es muy significativa y esto hace que la tendencia de la curva se separe de su comportamiento teórico sin ISI. Es importante destacar que los resultados mostrados en la figura, además de estar en consonancia con los resultados teóricos, coinciden con un estudio similar presentado en [5].



**Figura 3.5 BER física para una red UTRA-FDD en sentido descendente a 120 km/h con estimación subóptima de canal**

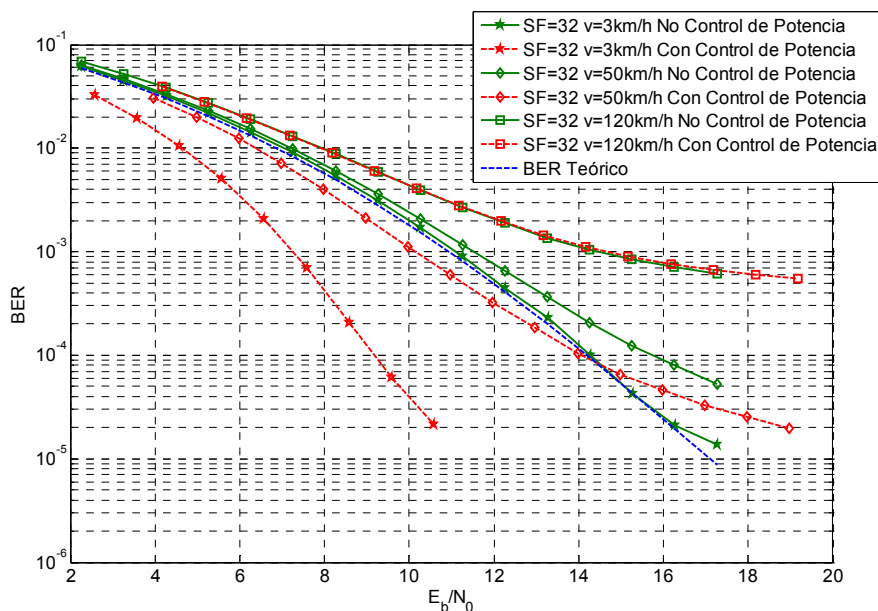
Otro aspecto importante que queda reflejado en la Figura 3.3 es cómo un SF superior no sólo afecta directamente a la sensibilidad del terminal, puesto que reduce el nivel de las interferencias intra e inter-sistema, sino que también protege más a la señal de las propias interferencias entre símbolos provocadas por el multicamino. Este hecho habrá que tenerlo en cuenta, puesto que la  $E_b/N_0$  objetivo para SF distintos será diferente.

Además del SF hay otros factores, como la estimación de canal, que afectan claramente a la correcta recepción de la señal. La Figura 3.5 muestra el efecto producido por una estimación subóptima del canal, es decir, se dispone de información exacta del canal en el preciso momento en que se reciben los bits piloto, mientras que para el resto de muestras se interpolan las estimaciones. Como se puede apreciar en la figura, en este caso la ISI está mucho más acentuada en comparación con la estimación óptima, ya que sólo se actualizan los coeficientes del receptor *Rake* una vez por *slot*. Al comparar estos resultados con los mostrados en [5] se observa que las prestaciones mostradas por el simulador de enlace implementado son ligeramente peores. Esto es debido a una suposición realizada en [5], donde se considera una actualización adicional de los coeficientes del *Rake* a mitad de *slot*. Esta suposición no está fundamentada en ningún aspecto tecnológico por lo que se ha optado por su no implementación en el simulador. De esta manera los resultados obtenidos son mucho más realistas.

Otro de los efectos que necesita una validación rigurosa es el producido por el mecanismo de control de potencia en bucle cerrado. Tal y como se explica en las especificaciones, el control de potencia en bucle cerrado es un mecanismo de realimentación que tiene como objetivo mantener la relación señal a interferencia sobre un determinado nivel, de manera que permita asegurar la calidad de la comunicación. Este mecanismo se ejecuta *slot a slot*, es decir, con una frecuencia de 1500 Hz. Gracias a su frecuencia de funcionamiento, el control de potencia rápido puede compensar el *fast fading*, además del *shadowing*, para velocidades del terminal móvil moderadas, de hasta 70 u 80 km/h. Para velocidades superiores el control de potencia rápido deja de ser efectivo y para mejorar la calidad se ha de recurrir a estrategias combinadas de codificación y entrelazado. Al incluir el control de potencia se mejorará claramente el

funcionamiento del sistema lo que se reflejará en los resultados del simulador de enlace. Siempre que se puedan corregir las fluctuaciones producidas por el *fast fading* se evitará en gran medida los problemas relativos al efecto multicamino, lo que mejorará el comportamiento del receptor.

La Figura 3.6 muestra los resultados obtenidos respecto al control de potencia para un SF igual a 32 y una estimación de canal subóptima. En primer lugar se puede observar como, al incrementar la velocidad del terminal, el canal presenta un mayor efecto *doppler*, lo que se traduce en una mayor variabilidad del *fast fading* y, por tanto, una estimación de canal con más errores y un peor rendimiento del sistema. Aunque se conozca el estado del canal de manera perfecta en el momento de la recepción de los bits piloto, si la velocidad es alta este valor no será válido para las muestras separadas una determinada distancia, lo que hará que el error cometido sea mayor con velocidades crecientes. La interpolación permite seguir el valor medio pero no las variaciones rápidas del canal, variaciones características del *fast fading* a alta velocidad.



**Figura 3.6 Efecto del control de potencia a distintas velocidades. BER física para una red UTRA-FDD con estimación de canal subóptima en sentido descendente**

Observando las líneas continuas, que reflejan el comportamiento del sistema sin control de potencia, se puede apreciar el efecto que la velocidad tiene sobre las prestaciones del sistema. A 3 km/h el sistema se aproxima mucho al comportamiento teórico esperado. Al aumentar la velocidad del terminal y cometer más fallos en la estimación de canal por ser ésta subóptima, las curvas se levantan alejándose del comportamiento teórico esperado. Así, se puede ver la presencia de una ver residual para valores de  $E_b/N_0$  superiores a 20 dB.

En la Figura 3.6 también se observa cómo el control de potencia mejora el rendimiento del sistema. Las líneas discontinuas muestran los resultados obtenidos al utilizar el control de potencia real implementado. Dicha mejora que es muy importante para 3km/h y va decreciendo con la velocidad, de manera que a 120km/h el control de potencia no supone ningún efecto en las prestaciones del sistema. Todas estas conclusiones están en total consonancia con los conceptos teóricos recogidos en las especificaciones del 3GPP y con los resultados recogidos en [5].

Como contrapartida, el uso del control de potencia supone un incremento en la  $E_b/N_0$  media, un incremento conocido como margen de *fast fading* o margen del control de potencia. Es importante destacar que los resultados mostrados en la Figura 3.6 se han representado considerando la  $E_b/N_0$  experimentada en el receptor tras aplicar el ajuste discreto de la potencia transmitida.

La Tabla 3.1 resume el incremento producido en la  $E_b/N_0$  media en función de la velocidad. Esta tabla está también en consonancia con los valores típicos de dicho margen del control de potencia [1].

v= 120 km/h	v= 50 km/h	v= 3 km/h
1.90 dB	1.71 dB	0.31 dB

**Tabla 3.1 Margen del control de potencia para distintas velocidades**



### 3.2.3.2 Validación de los resultados obtenidos en RABs reales

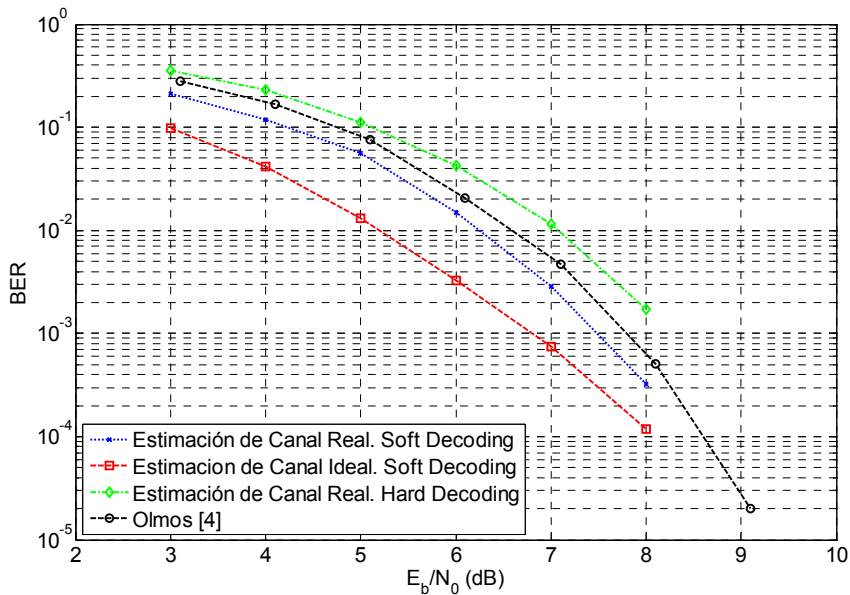
Una vez validado el simulador en la capa física es necesario alcanzar el nivel RLC para determinar el verdadero comportamiento del sistema UMTS y validar el funcionamiento del simulador incluyendo codificación de canal y adición de cabeceras. Para realizar esta validación a nivel de sistema se han comparado los resultados obtenidos con los mostrados en [4],[5] y [10]. Todos estos trabajos han sido realizados por el mismo grupo de investigación pero cada una de las publicaciones evalúa distintos aspectos de las simulaciones realizadas. Con el fin de que la comparación sea lo más fiable posible, se ha seleccionado el mismo entorno de simulación empleado en estos estudios, en concreto un entorno *vehicular* a 50 km/h. Este escenario de propagación se ha caracterizado como hasta ahora, mediante un modelo de 4 contribuciones retardadas cada una de ellas un periodo de *chip* y con la distribución de potencias mostrada en la Figura 3.4. Además, se han seleccionado dos RABs características: el servicio de voz a 12.2 kbps, que utiliza codificación convolucional, y el servicio de vídeo conferencia por conmutación de circuitos a 64 kbps, que usa turbo-codificación. Utilizando estas dos RABs ejemplo será posible evaluar y validar las dos técnicas de codificación implementadas en UMTS y en el simulador de enlace.

Por otro lado, se ha seleccionado la estimación de canal más realista, la basada en correlación, puesto que es la que se utiliza en los sistemas reales y, además, fue la seleccionada en la herramienta de simulación con la que se pretende contrastar los resultados obtenidos. Para que los resultados sean exactamente los mismos, se ha realizado la estimación de canal a partir de los bits piloto transmitidos en el canal de control dedicado, interpolando las muestras de canal entre dos estimaciones consecutivas. Respecto al control de potencia, como en [5] se implementó el control perfecto de potencia, en primer lugar se utilizará dicho modelo para después evaluar el efecto de considerar un modelo no perfecto basado en saltos discretos de 1dB.

En primer lugar se han estudiado las prestaciones del servicio de voz a 12.2 kbps descrito en [3] y resumido en la Tabla 3.2. La Figura 3.7 muestra los resultados obtenidos para este servicio considerando un control de potencia perfecto. Para este

servicio el estándar especifica la utilización de una codificación convolucional. Tal y como se explica en el Anexo A, la decodificación utilizada influye de manera muy significativa en los resultados obtenidos. Así pues, la diferencia observada al utilizar una decodificación firme o indecisa es bastante notable. Además el decodificador implementado en el simulador desarrollado en esta Tesis obtiene mejores resultados que el utilizado como referencia (ver Anexo A), lo que justifica que, con las mismas condiciones, los resultados sean sensiblemente mejores.

Otro aspecto mostrado en la Figura 3.7 es el efecto que tiene el uso de una estimación de canal real. Incluso para 50 km/h el incremento en la  $E_b/N_0$  necesaria para garantizar una BER de  $10^{-3}$  es de cerca de 0.5 dB, en comparación con lo necesario en caso de conocer perfectamente el estado del canal. Dicho efecto será incluso más importante para velocidades más altas.

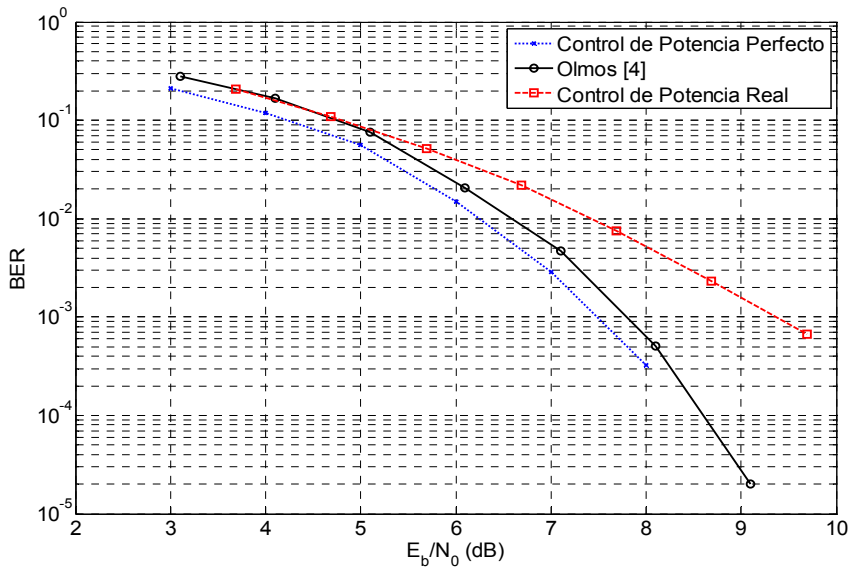


**Figura 3.7 BER para el servicio de voz.  
Efecto de la estimación de canal y la decodificación**

Higher layer	RAB/Signalling RB	RAB subflow #1	RAB subflow #2	RAB subflow #3	
RLC	Logical channel type	DTCH			
	RLC mode	TM	TM	TM	
	Payload sizes, bit	0	103	60	
		39			
	81				
Max data rate, bps	12 200				
TrD PDU header, bit	0				
MAC	MAC header, bit	0			
	MAC multiplexing	N/A			
Layer 1	TrCH type	DCH	DCH	DCH	
	TB sizes, bit	0	103	60	
		39			
		81			
	TFS (note 1)	TF0, bits	1x0 (note 2)	0x103	0x60
		TF1, bits	1x39	1x103	1x60
		TF2, bits	1x81	N/A	N/A
	TTI, ms	20	20	20	
	Coding type	CC 1/3	CC 1/3	CC 1/2	
	CRC, bit	12	N/A	N/A	
	Max number of bits/TTI after channel coding	303	333	136	
RM attribute	180-220	170-210	215-256		
SF	128				

**Tabla 3.2 RAB para el servicio de voz a 12.2 kbps [3]**

La Figura 3.8 muestra la importancia que tiene el modelado del control de potencia. Al realizar un control de potencia discreto basado en incrementos y decrementos de 1dB en la potencia transmitida, se aumenta el margen del control de potencia y se cometen más errores en la transmisión, lo que finalmente redundará en un aumento más que significativo de la potencia requerida en transmisión. En lugar de trabajar con un valor de  $E_b/N_0$  objetivo al que se le suma el margen de control de potencia, se ha considerado más apropiado introducir directamente en las simulaciones de enlace el efecto introducido por el uso del control de potencia real. También resulta destacable resaltar cómo el margen del control de potencia observado a nivel físico (ver Tabla 3.1) se corresponde perfectamente con el incremento de la  $E_b/N_0$  del servicio de voz, considerando una BER objetivo de  $10^{-3}$  y una velocidad del terminal de 50 km/h. En concreto, en la Figura 3.8 se ve que este incremento es de aproximadamente 1.7 dB, coincidiendo con lo mostrado en la Tabla 3.1 para dicha velocidad.



**Figura 3.8 BER para el servicio de voz. Efecto del modelo de control de potencia**

Los resultados obtenidos con control de potencia perfecto, mostrados en la Figura 3.8, están en consonancia con otros resultados previos como los mostrados en [5],[11] y [12]. A modo de ejemplo, para una BER objetivo igual a  $10^{-3}$  la  $E_b/N_0$  media necesaria se sitúa en 7.5 dB, valor que coincide aproximadamente con los mostrados en los trabajos previos citados.

Es importante destacar que todos los parámetros relevantes del servicio como son el formato de transporte (TFCI), el intervalo de tiempo de transmisión (TTI), el canal lógico asociado, la codificación de canal, el tamaño del bloque de transporte (TB), etc. han sido implementados tal y como se especifica en la RAB correspondiente al servicio (Tabla 3.2). Tanto la capa física como la MAC y RLC se han incluido apropiadamente, constituyéndose los diferentes subflujos de información de tráfico de red y procesando los datos tal y como viene especificado por el 3GPP [3].

Como ejemplo de utilización de la turbo-codificación se ha seleccionado el servicio de vídeo conferencia por conmutación de circuitos a 64 kbps. Este servicio, cuyas características se resumen en la Tabla 3.3, es uno de los servicios básicos de UMTS y

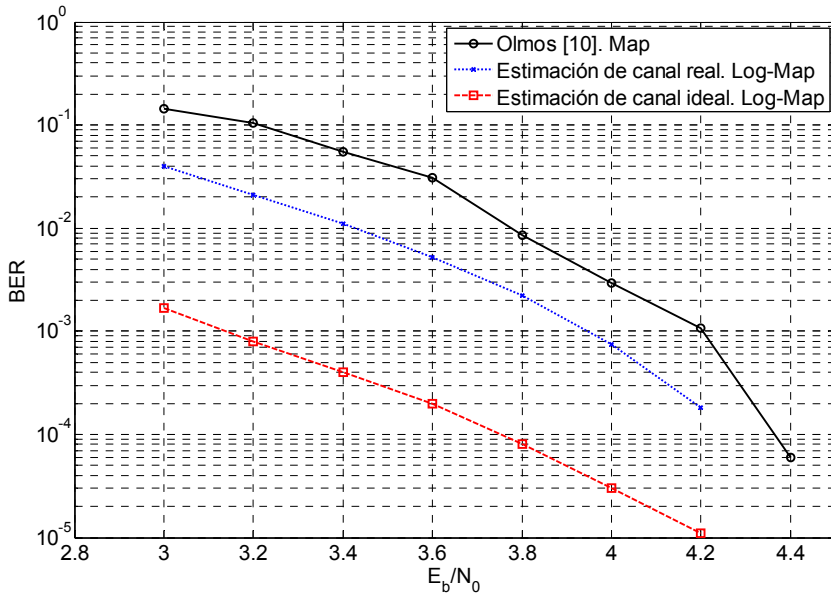
se ha caracterizado en muchos trabajos. Es por esto por lo que se le ha seleccionado como RAB ejemplo puesto que así se dispone de distintas fuentes para su validación. En la Figura 3.9 se comparan los resultados obtenidos con el simulador desarrollado en esta Tesis Doctoral y los de referencia, considerando el control de potencia perfecto.

Tal y como se comenta en el Anexo A, la turbo-decodificación se ha implementado utilizando el algoritmo de decodificación más potente, el Log-Map. Sin embargo, la decodificación implementada en el simulador utilizado como referencia se basa en el algoritmo Map, más antiguo y con peores prestaciones, tal y como se muestra en el Anexo A, coincidiendo esta conclusión con [13]. Se observa por tanto una cierta mejora en los resultados obtenidos con el simulador respecto a los esperados. Esta mejora es pequeña, en torno a los 0.2 dB, valor muy similar al encontrado a nivel físico (ver Anexo A). También se observa en la Figura 3.9 como, de nuevo, es fundamental realizar una estimación del canal real para poder obtener una valoración fidedigna de las prestaciones del sistema. La estimación perfecta ofrece siempre unos resultados mucho más optimistas.

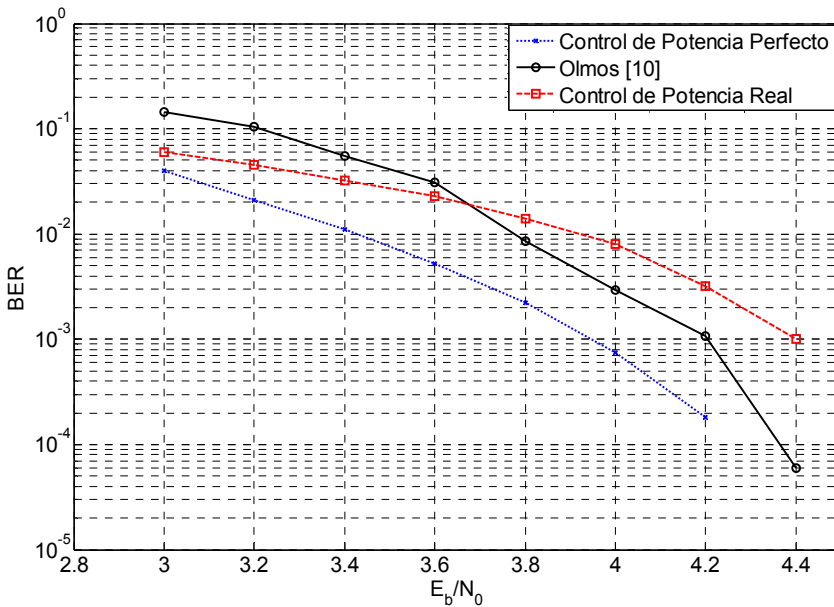
Finalmente, en la Figura 3.10 se muestra el efecto del modelo de control de potencia.

Higher layer	RAB/Signalling RB		RAB
RLC	Logical channel type		DTCH
	RLC mode		TM
	Payload sizes, bit		640
	Max data rate, bps		64000
	TrD PDU header, bit		0
MAC	MAC header, bit		0
	MAC multiplexing		N/A
Layer 1	TrCH type		DCH
	TB sizes, bit		640
	TFS	TF0, bits	0x640
		TF1, bits	2x640(alt. 4x640)
	TTI, ms		20(alt. 40)
	Coding type		TC
	CRC, bit		16
	Max number of bits/TTI after channel coding		3948(alt. 7884)
	RM attribute		150-195

**Tabla 3.3 RAB para el servicio de vídeo conferencia a 64 kbps**



**Figura 3.9** BER para el servicio de vídeo conferencia a 64 kbps.  
Efecto de la estimación de canal y la decodificación



**Figura 3.10** BER para el servicio de vídeo conferencia a 64 kbps.  
Efecto del modelo de control de potencia

### 3.2.4 Mejoras en la estimación de canal

La estimación de canal es una técnica fundamental para el buen funcionamiento del sistema UMTS. Los fundamentos de la estimación de canal por correlación se explicaron en el Capítulo 1. Los trabajos previos que se han utilizado como referencia para validar el comportamiento del simulador realizan la estimación del canal a partir de los bits piloto transmitidos cada *slot* sobre el canal físico de control dedicado. A posteriori, utilizando la estimación de canal realizada sobre *slots* consecutivos, se realiza una interpolación para calcular las muestras de canal necesarias para corregir el desvanecimiento sufrido por la señal. Esta interpolación reduce la BER de los bits más alejados de la secuencia piloto.

Sin embargo, en esta Tesis Doctoral se han utilizado dos técnicas adicionales para mejorar la estimación de canal. La primera se basa en realizar un filtrado paso bajo de las estimaciones de canal realizadas en una trama, de manera que se elimine gran parte del error cometido en las mismas. Esta técnica de filtrado, propuesta inicialmente en [14], es más efectiva a velocidades bajas ya que el canal varía lentamente y, por tanto, el promediado elimina gran parte del ruido experimentado en la estimación. El filtro utilizado en el simulador es no causal y toma dos muestras anteriores y posteriores junto con la muestra actual para realizar el promediado ponderado. En concreto los coeficientes del filtro son:

$$(\alpha(-2), \alpha(-1), \alpha(0), \alpha(1), \alpha(2)) = (0.0315, 0.2389, 0.4592, 0.2389, 0.0315)$$

De esta manera se pondera más la muestra actual de estimación pero también considerando las dos muestras anteriores y posteriores. Se ha demostrado que este filtro es adecuado para velocidades moderadas [14], siendo necesario aumentar la longitud del mismo conforme se disminuye la velocidad del terminal.

La Figura 3.11 muestra el efecto de incluir el filtrado de las muestras de canal. En la parte superior de la imagen se muestran los resultados obtenidos sin filtrar las estimas del canal, mientras que abajo se recogen los mismos resultados cuando se aplica un

filtrado previo a la interpolación. Las gráficas de la izquierda representan, en línea discontinua, la evolución de la fase experimentada por el canal debido al efecto multicamino durante 10 tramas (100ms). La línea continua representa la estimación de canal obtenida por el estimador basado en correlación. A mano derecha se recoge la función densidad de probabilidad (pdf) del error en fase cometido en la estimación. El escenario considerado ha sido el mismo que en los apartados anteriores, siendo el servicio utilizado el de voz y la  $E_b/N_0 = 8\text{dB}$ .

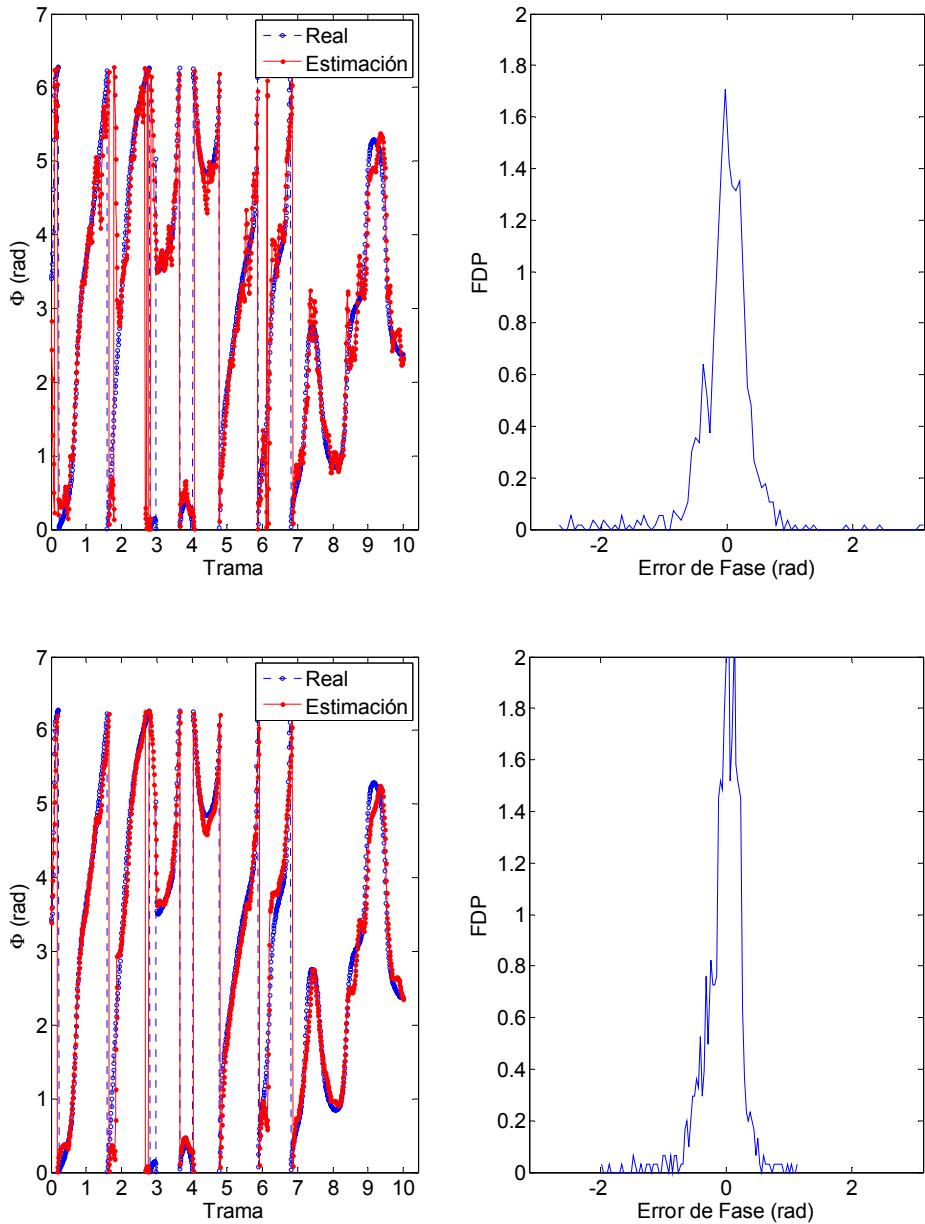
Se puede observar claramente como, gracias al filtrado, se suaviza el ruido experimentado en la estimación, siguiéndose de manera mucho más acertada las variaciones del canal. En la función densidad de probabilidad también se distingue cómo el error cometido disminuye, comprimiéndose la forma de la pdf. En ambos casos el error cometido tiene media cero, lo que era de esperar, siendo la varianza de dicho error de 0.22 para el caso sin filtrado y de 0.0926 cuando se aplica el filtrado correspondiente.

Una vez conocido el efecto del filtrado sobre la estimación de canal, se ha analizado qué consecuencias tiene esta mejora en la estimación sobre los resultados obtenidos con el simulador de enlace.

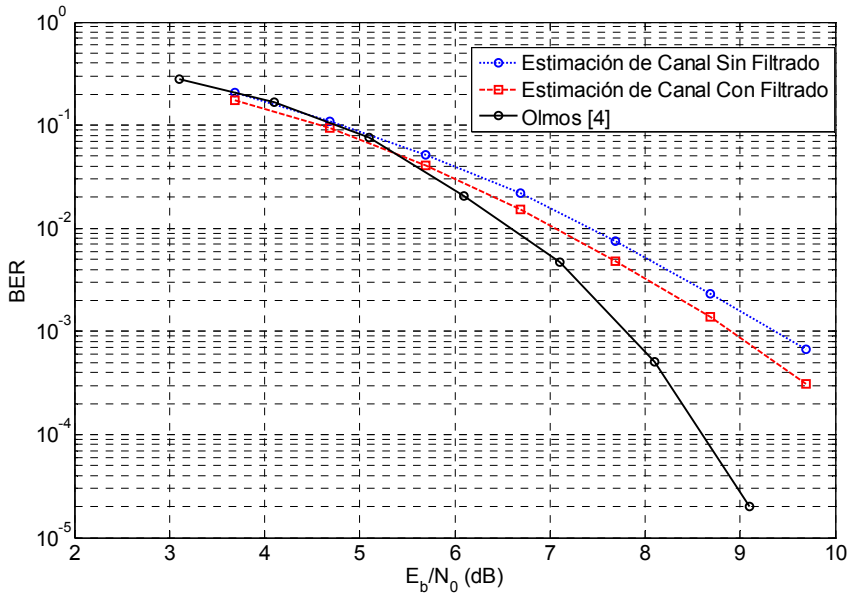
En la Figura 3.12 se observa que el empleo del filtrado de las muestras de canal implica una ganancia de 0.5 dB en la  $E_b/N_0$  objetivo, valor muy aproximado al obtenido en [14].

El seguimiento del canal también se puede mejorar utilizando los bits pilotos transmitidos en el canal CPICH para realizar la estimación. Las ventajas de utilizar los bits del canal piloto para estimar el canal son varias. En primer lugar, el canal CPICH se ensancha con un SF=256, lo que le confiere el mejor comportamiento frente a interferencia entre símbolos, tal y como se mostró en la Figura 3.3. Por otro lado, en cada *slot* se envía una secuencia de 10 símbolos piloto, secuencia mucho más larga que la enviada en el canal de control dedicado, por lo que la estimación siempre será de mayor calidad. Por último, el nivel de potencia recibido del canal piloto suele ser bastante alto, por lo que se dispone de una fuente fiable de conocimiento del canal.





**Figura 3.11** Estimación de canal sin filtrado de muestras (arriba) y con filtrado (abajo). Se representa una secuencia de canal de 10 tramas (izquierda) y la pdf del error de fase cometido en la estimación (derecha)



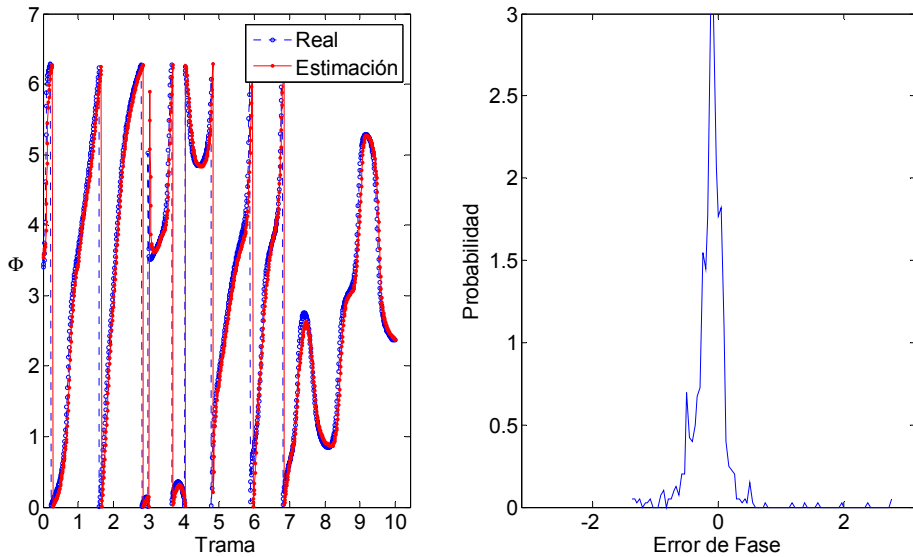
**Figura 3.12 BER para el servicio de voz. Efecto del filtrado sobre las estimas de canal**

En la Figura 3.13 se muestra el comportamiento de la estimación de canal al considerar el peor estado de recepción posible del canal de CPICH, es decir, considerando una  $E_c/I_0 = -12\text{dB}$ . Incluso en este estado, la estimación de canal es mejor que la mostrada en la Figura 3.11, donde se utilizaban los bits del canal de control dedicado y un nivel alto de  $E_b/N_0$ . Además, la estimación de canal por CPICH es muy robusta frente al nivel de ruido. La

Tabla 3.4 muestra la evolución de la varianza del error respecto al nivel de  $E_c/I_0$  recibido. La diferencia entre un nivel bueno de señal piloto (-4dB) y el peor nivel (-12dB), es apenas apreciable.

	$E_c/I_0 = -4\text{dB}$	$E_c/I_0 = -8\text{dB}$	$E_c/I_0 = -10\text{dB}$	$E_c/I_0 = -12\text{dB}$
Varianza del Error	0.0748	0.077	0.079	0.082

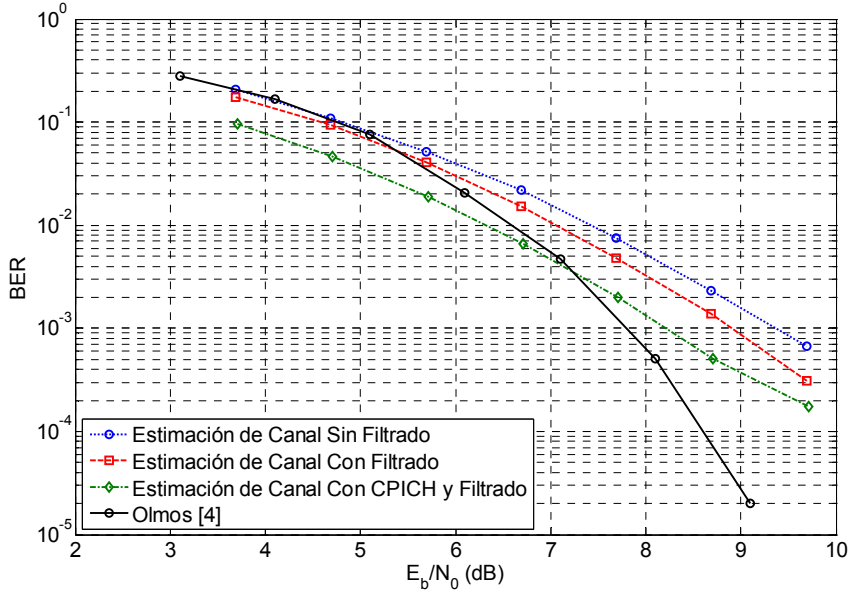
**Tabla 3.4 Varianza del error cometido en la estimación de canal en función del nivel de calidad de CPICH experimentado**



**Figura 3.13** Estimación de canal utilizando el CPICH. Se representa una secuencia de canal de 10 tramas (izquierda) y la pdf del error cometido en la estimación (derecha)

Gracias a este resultado se puede concluir que no hace falta considerar el efecto que tiene la distancia al transmisor en la estimación de canal mediante CPICH. Así pues, siempre que se evalúe las prestaciones del MNB considerando estimación de canal mediante CPICH, se calculará para un valor fijo de  $E_c/I_0$ .

Finalmente, se ha evaluado de nuevo el efecto que tiene sobre la relación  $E_b/N_0$  – BER utilizar la estimación mediante CPICH. En la Figura 3.14 se comparan las tres técnicas de estimación utilizadas. Como era de esperar, la mejora en la estimación de canal conseguida gracias a la utilización del CPICH supone una importante ganancia en la  $E_b/N_0$  objetivo. Es importante resaltar cómo, cuando se incrementa mucho la  $E_b/N_0$ , la respuesta con CPICH tiende a acercarse a las otras. Esto es debido a que se está considerando una calidad fija en la señal recibida del CPICH, mientras que las otras gráficas, al mejorar la relación señal a ruido, seguirán mejorando su estimación de canal hasta equipararse e incluso mejorar la obtenida mediante el uso del CPICH.



**Figura 3.14 BER para el servicio de voz.**  
**Efecto de la estimación de canal a partir de la señal de CPICH**

### 3.3 Prestaciones de la arquitectura MNB

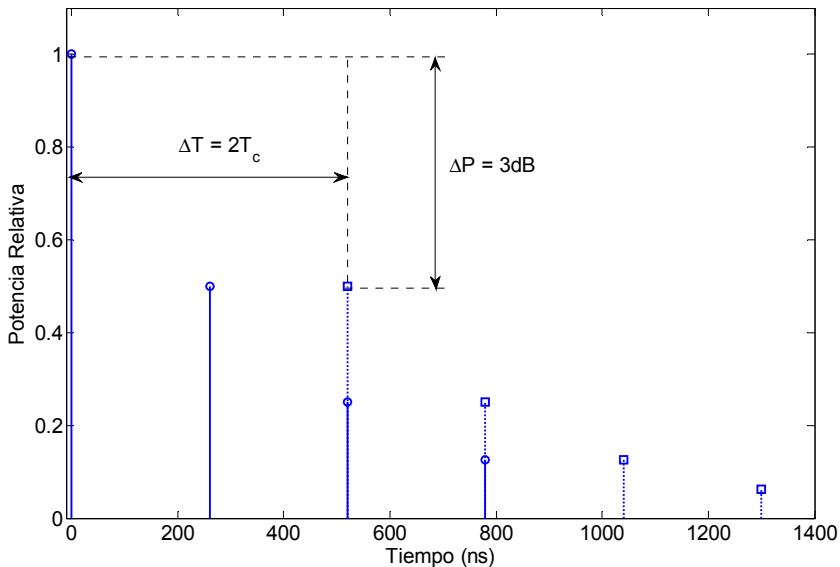
Tal y como se explicó en el Capítulo 2, la nueva arquitectura Multinodo B propuesta en esta Tesis Doctoral consiste en desplegar un conjunto de antenas configurando una única celda lógica. Por el sólo hecho de considerar un escenario mixto macro-micro celular se reduce la interferencia inter-celular, al igual que la distancia entre los usuarios y las antenas del sistema, por lo que parece obvio esperar una ganancia producida por la dispersión de los elementos transmisores. Sin embargo, transmitir simultáneamente desde dos emplazamientos introduce en el sistema una micro-diversidad adicional que implica una ganancia, sobre todo teniendo en cuenta la absoluta decorrelación entre los enlaces implicados. El receptor *Rake* podrá disponer de más caminos resolubles, lo que le garantiza un mayor aprovechamiento de la potencia transmitida y una importante reducción de la interferencia producida por el efecto multicamino.

Este apartado profundiza en este aspecto específico, analizando los pros y los contras de la transmisión simultánea desde dos antenas y demostrando de qué manera se puede aprovechar al máximo esta transmisión. A lo largo de todo el apartado se ha utilizado un único escenario, intentando discriminar una a una las principales componentes que afectan al comportamiento del MNB. Se ha considerado como caso de estudio un entorno de propagación macro-celular y el servicio de voz a 12.2 kbps, siendo la velocidad del móvil de 50 km/h. La estimación de canal se realizará mediante los bits del canal de control dedicado, realizando asimismo el posterior filtrado e interpolación de las muestras para mejorar las prestaciones.

Para estudiar la posible ganancia por micro-diversidad obtenida gracias al MNB es necesario evaluar el comportamiento del sistema a nivel de enlace, viendo el efecto producido al recibir múltiples replicas de la señal de interés provenientes de dos antenas del MNB situadas en distintos emplazamientos. Las dos señales se recibirán con retardos temporales y diferencias de potencia aleatorias.

Siendo el escenario de simulación el mismo que el utilizado en el anterior apartado, en esta ocasión el perfil de retardo se ve alterado al recibir las componentes multicamino de la segunda antena. La segunda agrupación quedará perfectamente definida simplemente especificando el retardo y la diferencia de potencia ( $\Delta T$  e  $\Delta P$  respectivamente) con respecto a la primera señal, tal y como se ve en la Figura 3.15.

En primer lugar se considerará un escenario óptimo, en el que las señales se reciben con igual potencia y sin solapamiento entre ellas. Lo que se analizará será el mínimo número de ramas que deberá disponer el receptor para sacar el mayor provecho de la diversidad introducida. En segundo lugar, se verá cuál es el efecto de que las dos señales no lleguen con el mismo nivel de potencia, para después analizar qué ocurre cuando las señales recibidas se solapan en el tiempo. A partir de este último análisis, se concluirá que la transmisión debe sincronizarse de alguna manera, siendo por tanto necesario mejorar la estimación de canal para poder discriminar convenientemente las contribuciones recibidas. Por ello, el último apartado de esta sección profundiza en la estimación de canal y en las técnicas propuestas en esta Tesis Doctoral para optimizar el funcionamiento del MNB.



**Figura 3.15 PDP con dos antenas en el Multinodo B**

### 3.3.1 Efecto del número de *fingers* considerado

El receptor *Rake* fue diseñado para extraer la mayor cantidad posible de energía de las componentes multicamino recibidas en recepción. En su funcionamiento el *Rake* es capaz de sincronizar hasta un número máximo de caminos, número que viene limitado por los recursos *hardware* disponibles. Cada una de los caminos se recibe en lo que se conoce como rama o *finger* del receptor y, por tanto, el número de caminos resolubles es directamente igual al número de *fingers* disponibles en el *Rake*. Normalmente, un terminal suele disponer de entre 4 y 8 *fingers*, mientras que en un entorno macrocelular como el considerado se suelen recibir de 2 a 4 caminos [12].

Como primera aproximación, se ha analizado el comportamiento del sistema al introducir una segunda antena cuando ambas contribuciones no se solapan temporalmente en el receptor y son recibidas con el mismo nivel de potencia. En este caso, en el escenario de simulación considerado habrá 8 contribuciones recibidas que el receptor *Rake* será capaz de discriminar. Al comparar con la solución sin MNB, es importante destacar que la potencia transmitida por las dos antenas será la misma que la transmitida anteriormente por una sola antena. En este caso, como se recibe justo la misma potencia en ambas secuencias, eso quiere decir que cada una de ellas transmite con la mitad de potencia. Como se puede apreciar tanto en la Figura 3.16 como en la Figura 3.17, el número de ramas empleadas en el receptor *Rake* es un parámetro fundamental para determinar el comportamiento final del receptor. El menor número de *fingers* considerado es cuatro, ya que los terminales actuales siempre tienen al menos ese número, mientras que el máximo número de *fingers* se ha establecido en 8, ya que, obviamente, no parece razonable considerar más de 8 *fingers* dado que la señal no presenta más de ese número de contribuciones.

La Figura 3.16 muestra la relación  $E_b/N_0$  – BER comparando la solución sin MNB con el sistema MNB con un número creciente de *fingers* en el *Rake* del receptor. Lo primero que hay que destacar es cómo, efectivamente, el MNB supone una importante ganancia para el sistema ya que, para una BER objetivo dada, se reduce

considerablemente la  $E_b/N_0$  requerida, lo que, por consiguiente, supondrá una menor potencia transmitida para garantizar la misma QoS. Queda por tanto de manifiesto cómo se puede aumentar la capacidad del sistema gracias a la arquitectura MNB propuesta. Además, la mayor ganancia se consigue por el simple hecho de transmitir con diversidad incluso considerando el mismo número de 4 *fingers* en el receptor. Aumentar el número de *fingers* siempre supone un aumento en las prestaciones, aunque esta mejora se reducirá hasta llegar a ser casi inapreciable para 8 *fingers*.

La Figura 3.17 muestra con más detalle la evolución de la  $E_b/N_0$  media necesaria para satisfacer una BER objetivo de  $10^{-3}$ , BER característica para el servicio de voz [15]. Al aumentar el número de ramas del receptor, el sistema es capaz de extraer del canal más energía de la señal transmitida. Además, es importante destacar que las últimas dos contribuciones no tienen prácticamente energía en comparación con la componente principal, por lo que es razonable hacer uso de un receptor *Rake* de 6 ramas. Esta configuración será la empleada para el resto de análisis llevados a cabo.

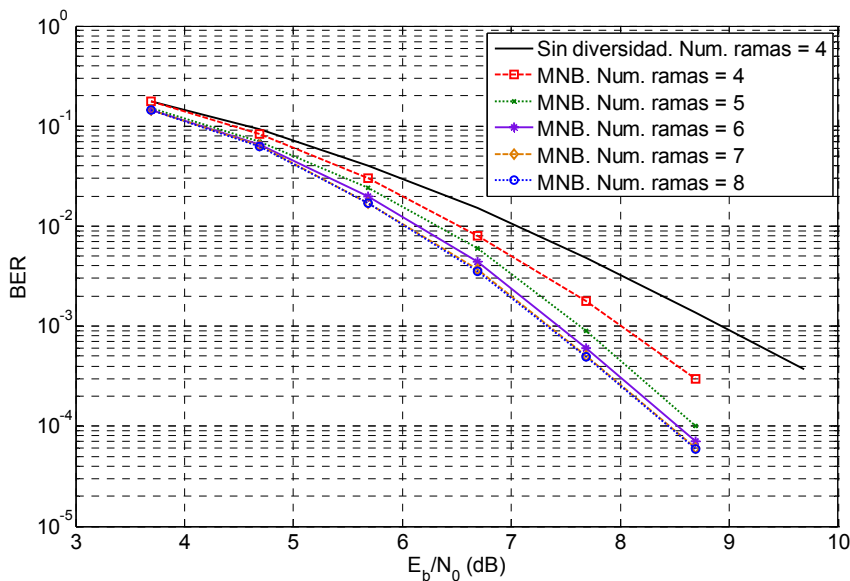
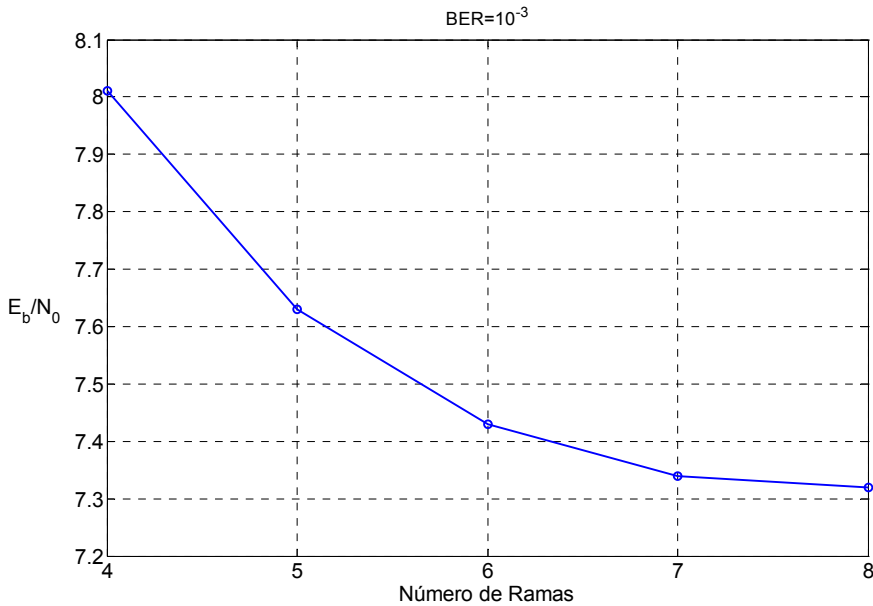


Figura 3.16 BER vs  $E_b/N_0$  con y sin diversidad espacial. Efecto del número de ramas del receptor *Rake*





**Figura 3.17 Efecto del número de ramas del receptor *Rake* en el  $E_b/N_0$  objetivo para un valor de BER fija**

De nuevo es importante recordar que al transmitir desde dos emplazamientos no se incrementa el consumo de potencia respecto a un solo transmisor. La ganancia observada es solamente debida al aumento de micro-diversidad en el receptor *Rake*. El sistema es capaz de combinar coherentemente más muestras de la señal recibida, mientras que el ruido e interferencias están incorreladas y se eliminan. De esta manera, el disponer de más caminos resolubles aumenta la eficacia de la recepción, idea que está en consonancia con la ecuación (3.4).

La Tabla 3.5 resume la ganancia obtenida en la  $E_b/N_0$  gracias a la diversidad adicional proporcionada por el MNB. Como se puede observar, la ganancia puede alcanzar valores de hasta 1.61 dB, ganancia muy importante para un entorno *vehicular* a 50 km/h en comparación con otras técnicas empleadas, como la analizada en [10].

También se puede destacar que, para un valor fijo de  $E_b/N_0 = 8$  dB, la mejora en las prestaciones del sistema en términos de BER es prácticamente de un orden de magnitud.

	No Diversidad	4 Ramas	6 Ramas	8 Ramas
$E_b/N_0$ Objetivo	8.93 dB	8.01 dB	7.43 dB	7.32 dB
Ganancia	0	0.92 dB	1.5 dB	1.61 dB

Tabla 3.5  $E_b/N_0$  medio para una BER objetivo igual a  $10^{-3}$

### 3.3.2 Efecto de la diferencia relativa de potencia

El comportamiento de la técnica de combinación MRC implementada en el *Rake* depende en gran medida de la diferencia relativa de potencia entre los caminos de señal o *taps* recibidos. Utilizando la expresión (3.4) se ha analizado cuáles son las prestaciones del sistema al ecualizar progresivamente los cuatro *taps* considerados en el escenario utilizado como referencia, manteniendo siempre la misma potencia recibida total.

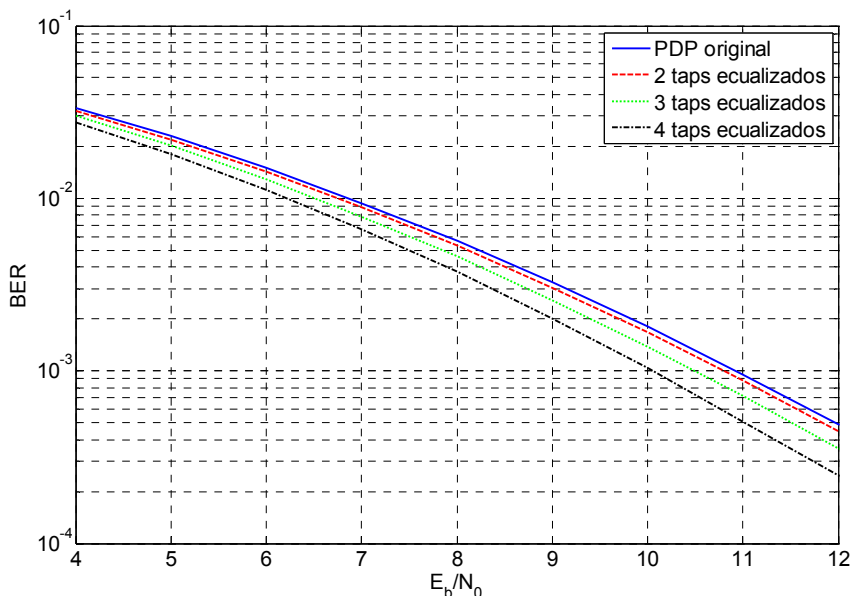


Figura 3.18 Efecto a nivel físico de la ecualización de los caminos recibidos

Efectivamente, la Figura 3.18 demuestra cómo recibir los cuatro caminos con igual potencia relativa es mucho mejor que recibirlos con potencias dispares.

En el apartado anterior se ha considerado que la potencia recibida desde ambas antenas era idéntica. En este escenario hay dos contribuciones, las de mayor nivel de potencia, con idéntico valor, lo que constituye por tanto el mejor escenario posible. Este apartado analiza cuál es el efecto producido por tener una diferencia relativa de potencia entre las contribuciones y hasta qué punto puede afectar esta diferencia a las prestaciones del sistema.

Bajo el mismo escenario de simulación y considerando que las dos contribuciones recibidas no se solapan en el tiempo, la Figura 3.19 muestra el efecto producido por la diferencia de potencia entre las señales recibidas desde ambas antenas. Como se puede observar en la figura, cuanto mayor es la diferencia de potencia más pequeña es la ganancia por diversidad experimentada. Sin embargo, aún con diferencias de potencia del orden de 6 dB la ganancia sigue siendo considerable.

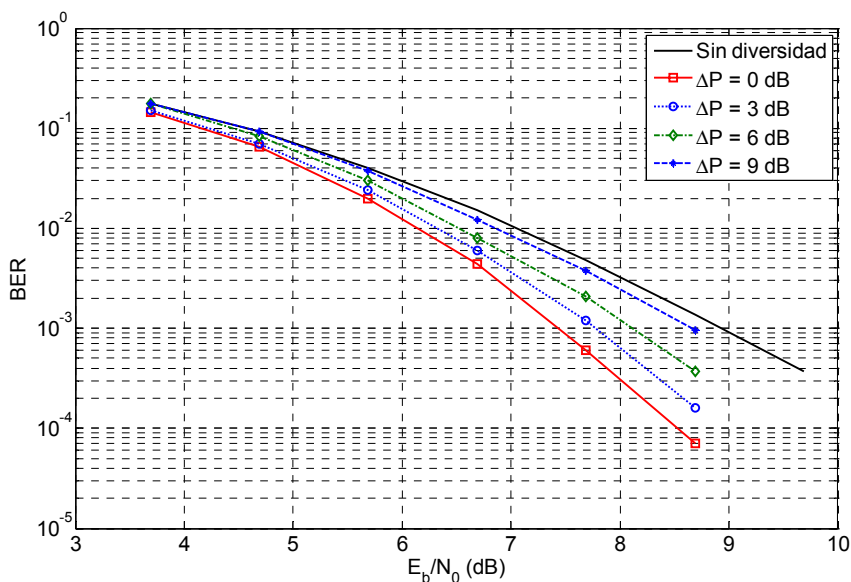


Figura 3.19 BER vs  $E_b/N_0$  con y sin MNB. Efecto de la diferencia relativa de potencia

De nuevo hay que reincidir en el hecho de que la potencia total recibida en el receptor *Rake* es la misma. De esta manera, si se continuara aumentando la diferencia de potencia relativa entre las estaciones base implicadas, llegaría un momento en que la señal de una de ellas sería despreciable y el sistema presentaría las mismas prestaciones que en el caso de no considerar diversidad.

Finalmente, se puede concluir que las mejores prestaciones del sistema MNB se obtienen cuando se recibe idéntico nivel de potencia desde dos antenas. Además, la diferencia de potencia existente en ningún caso provoca una degradación en el sistema en comparación con la configuración tradicional.

### 3.3.3 Efecto del retardo relativo

En este apartado se va a considerar que la potencia recibida desde ambas antenas transmisoras es la misma y lo que se evalúa es el efecto producido por el posible solape de las contribuciones recibidas de ambas antenas. A priori, el sistema radio sobre fibra introduce un retardo adicional que no sufre la señal en su propagación aire. Si este retardo no fuera corregido las contribuciones de ambas señales podrían llegar al usuario muy desfasadas, incurriendo en interferencia entre símbolos. Sin embargo, este desfase puede corregirse de manera sencilla conociendo simplemente la composición de la fibra óptica, la distancia del emplazamiento y el retardo introducido por los equipos de radio sobre fibra [16]. Con estos datos es relativamente sencillo calcular qué retardo adicional es necesario introducir en la transmisión desde la unidad central para sincronizar la recepción de la señal. Una vez hecho el ajuste grueso, aún se puede experimentar una diferencia temporal entre las señales recibidas debido a los diferentes caminos de propagación seguidos en la transmisión radio. Sin embargo, este desfase también estará acotado por la distancia entre los emplazamientos. Teniendo en cuenta que un desfase temporal de un *chip* se corresponde con 80m de distancia a la velocidad de la luz, el retraso máximo que se puede experimentar estará limitado por no más de 4 o 5 *chips*.

Considerando el caso de perfecta sincronización, en el que las dos contribuciones están totalmente solapadas ( $\Delta T = 0$ ), el receptor *Rake* no será capaz de discriminar una de las dos señales transmitidas por lo que, siendo la potencia recibida igual, entonces los resultados serán exactamente los mismos que se obtendrían sin ninguna diversidad en transmisión. Conforme las señales estén menos solapadas, aparecerán nuevos caminos sobre los que extraer señal, aumentará la diversidad y se mejorará la calidad de la señal recibida. La Figura 3.20 muestra claramente la degradación en las prestaciones del sistema debido a la coincidencia temporal entre ambas señales, obteniéndose efectivamente los mismos resultados que sin utilizar el MNB cuando ambas señales están totalmente solapadas. Por otro lado, este efecto se reduce al aumentar el retardo entre las contribuciones, de manera que para señales no solapadas (a partir de  $\Delta T = 4$ ) se vuelven a obtener unas buenas prestaciones, las mismas que las mostradas en el apartado 3.3.1.

Al aumentar aun más la diferencia temporal entre las dos señales (en la gráfica  $\Delta T = 10$ ) el sistema no parece acusar en gran medida el efecto de tal retardo, aunque sí que se aprecia cómo para valores altos de  $E_b/N_0$  comienza a haber una mayor separación con respecto a  $\Delta T = 4$  fruto del mayor nivel de interferencia entre símbolo que se experimenta con *delays spread* más altos. Para retardos muy acusados, o para SF más bajos, cabe esperar que este efecto sea aún más evidente.

Tras estudiar el efecto de la diferencia temporal y de amplitud de las contribuciones recibidas desde dos antenas del MNB, se puede concluir que, incluso en el peor de los casos, el MNB nunca empeora la calidad experimentada por un usuario. Por otro lado, y dado que tras la sincronización temporal entre las antenas del MNB existe una gran probabilidad de que las dos contribuciones se reciban solapadamente en el tiempo, es necesario desarrollar algún mecanismo que permita que el receptor pueda discriminar las distintas señales recibidas, de manera que, aún solapadas, se pueda conseguir una ganancia por diversidad adicional en el sistema MNB. La siguiente sección está dedicada precisamente a este cometido.

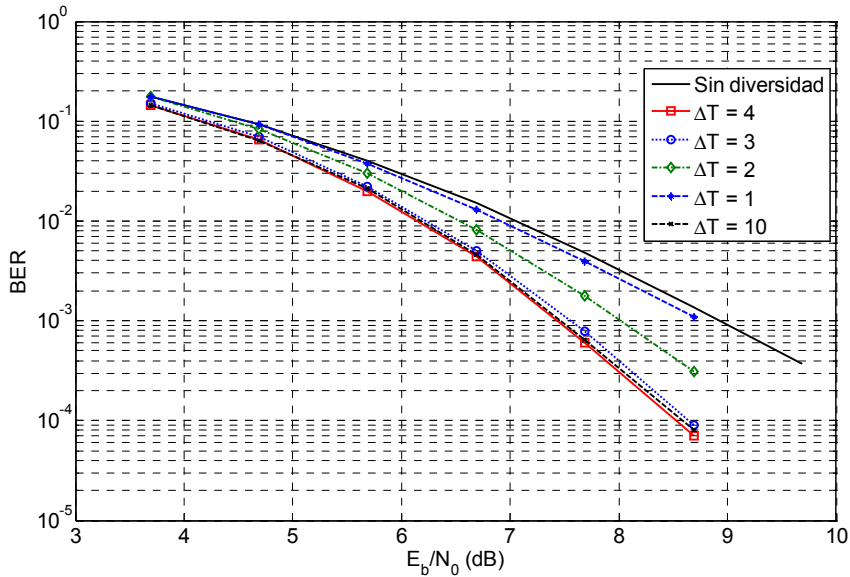


Figura 3.20 BER vs  $E_b/N_0$  con y sin MNB. Efecto de la diferencia relativa de tiempo

### 3.3.4 Procesado en *array* y combinación coherente de la señal en recepción

Los resultados anteriores han mostrado claramente que se puede obtener una importante ganancia en recepción si se consigue una combinación coherente de la señal transmitida.

La ganancia máxima que se puede conseguir mediante la combinación coherente de dos señales de igual potencia es de 3dB. La demostración de dicha cota máxima se puede realizar fácilmente de la siguiente manera. Inicialmente se supone la transmisión de una única señal compleja,  $\mathbf{v}$ , de  $n$  muestras, recibéndose la señal  $\mathbf{v}$  con 1W de potencia. A continuación, se envía la misma información desde dos antenas suficientemente separadas, de manera que por cada antena se transmita la mitad de potencia, es decir, se envía una señal de tensión  $\mathbf{v}/\sqrt{2}$ . Si la combinación es coherente,

es decir, las dos señales se ajustan exactamente en fase y se suman, la señal recibida será  $2\mathbf{v}/\sqrt{2}$ , cuya potencia resulta ser de  $2W$ . La potencia útil recibida es por tanto el doble que la recibida desde una única antena, lo que implica una ganancia de 3dB.

Además, por transmitir desde dos emplazamientos se tiene una ganancia adicional relativa a la diversidad de *fast fading* introducida en el sistema. Esto quiere decir que al transmitir desde dos emplazamientos la probabilidad de que ambos sufran simultáneamente un fuerte desvanecimiento es mucho menor que la probabilidad de que lo sufra un solo camino de propagación. El efecto de esta ganancia por diversidad dependerá de las características del canal y del funcionamiento de las técnicas de entrelazado y codificación pero, de cualquier manera, también incrementará la ganancia experimentada. A pesar de que la ganancia teórica que se puede obtener mediante la transmisión desde dos antenas es de más de 3dB, muchos estudios, como los mostrados en [1] y [10], aseguran que dicha ganancia en escenarios reales no es tal alta, debido a la no perfecta estimación del canal y al efecto de la interferencia entre símbolos.

El diseño del MNB permite explotar esta ganancia de combinación coherente y el incremento de la diversidad de *fast fading*. En el caso de que las dos señales no estén solapadas, el receptor *Rake* está diseñado para sumar coherentemente las señales recibidas, ya que en recepción se pueden discriminar las señales, ajustándolas en fase por separado. Sin embargo, cuando las señales llegan simultáneamente, como se utiliza el mismo código de aleatorización, el receptor es incapaz de separar ambas contribuciones, por lo que no se puede alcanzar ninguna ganancia, tal y como se ha demostrado en el apartado anterior.

Como las señales transmitidas por las antenas del MNB están a priori sincronizadas para reducir la ISI, es necesario desarrollar algún tipo de técnica de procesado en *array* que permita aprovechar al máximo las capacidades de combinación coherente que ofrece el MNB, independientemente de que las señales estén o no solapadas en el tiempo. Con este objetivo en mente, en esta Tesis Doctoral se propone hacer uso de un campo disponible en el canal ascendente de control que el terminal debe rellenar en caso de que así se lo solicite la red, reportando cierta información referente a la

relación de desfase experimentada por el usuario entre dos canales piloto transmitidos por el mismo Nodo-B. La información contenida en estos bits, denominada FBI (*Feedback Information*) [17], puede ser utilizada por el Nodo-B para transmitir en diversidad. Este mecanismo, denominado por el estándar del 3GPP CLMTD (*Closed Loop Mode Transmit Diversity*), está totalmente especificado respecto a la información reportada por el usuario, dejando en manos del fabricante del terminal cómo calcular el desfase que es necesario introducir en la transmisión desde ambas antenas para optimizar la calidad de servicio.

La técnica de CLMTD fue descrita ya en el Capítulo 1, aunque para una mayor información siempre es preferible acudir directamente a las especificaciones [6]. El usuario es capaz de conocer el canal experimentado por cada antena puesto que éstas transmiten un canal piloto distinto, con distinto factor de ensanchado. A partir de esta información y utilizando los bits del FBI, se informa al Nodo-B del desfase que hay que introducir entre ambos canales. Respecto al envío de los bits de información es importante recalcar que cada *slot* el usuario sólo envía un bit de información. Además la granularidad de la diferencia de fase reportada es de  $\pi/4$ , ya que el FBI está compuesto por 3 bits con los que se puede identificar cualquiera de las ocho posibles diferencias de fase introducidas entre las antenas. Para no disponer solamente de una medida de desfase cada 3 *slots*, se ha ideado un mecanismo correctivo a partir de la secuencia de bits recibida, de manera que el último bit enviado corrija la última información reportada. Este mecanismo, que permite hacer un mejor seguimiento del canal, ha sido implementado en el simulador tal y como está especificado en [6]. De cualquier manera, desde que el usuario mide el estado de canal hasta que el sistema es capaz de transmitir conforme esa información, transcurren más o menos 2 *slots*, es decir, 1.3ms. Este tiempo para velocidades bajas será suficiente pero para velocidades altas, donde el tiempo de coherencia del canal es más pequeño, puede llegar a ser dramático. Sin embargo, como el MNB está diseñado para ser desplegado en entornos urbanos con una velocidad máxima de usuario de 50km/h, este tiempo será suficiente, ya que el tiempo de coherencia del canal es de 2ms a esta velocidad.



A partir de la información más o menos precisa enviada por el usuario, el Nodo-B multiplica la misma secuencia de datos por un peso  $w_1$  y  $w_2$ , transmitiendo cada señal resultante por una antena distinta. En general, el factor de ponderación es un número complejo.

El problema de la selección de los pesos óptimos para la maximización de la calidad percibida por el usuario es un problema clásico de *beamforming* [18]. Muchos algoritmos intentan maximizar la relación señal a ruido experimentada por el usuario, pero para ello se debe tener un conocimiento preciso de la matriz de covarianza del ruido e interferencias experimentado, información de la que no se puede disponer en un sistema como el WCDMA donde la variabilidad de las interferencias es muy alta. Ya que no se puede maximizar la relación señal a ruido, los pesos utilizados en la transmisión deben seleccionarse de manera que se maximice la potencia de señal útil recibida, independientemente de las interferencias experimentadas.

Se define  $\mathbf{h}_1$  como el vector complejo de la respuesta estimada del canal al impulso entre la primera antena del MNB y el usuario y  $w_1$  el peso asignado a la información transmitida desde la primera antena. De igual manera se definen  $\mathbf{h}_2$  y  $w_2$  para la segunda antena. En general, si  $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2]$  y  $\mathbf{w} = [w_1, w_2]$ , se deberá maximizar la potencia recibida  $P$ , siendo

$$P = E[\mathbf{w}^H \mathbf{H}^H \mathbf{H} \mathbf{w}] \quad (3.6)$$

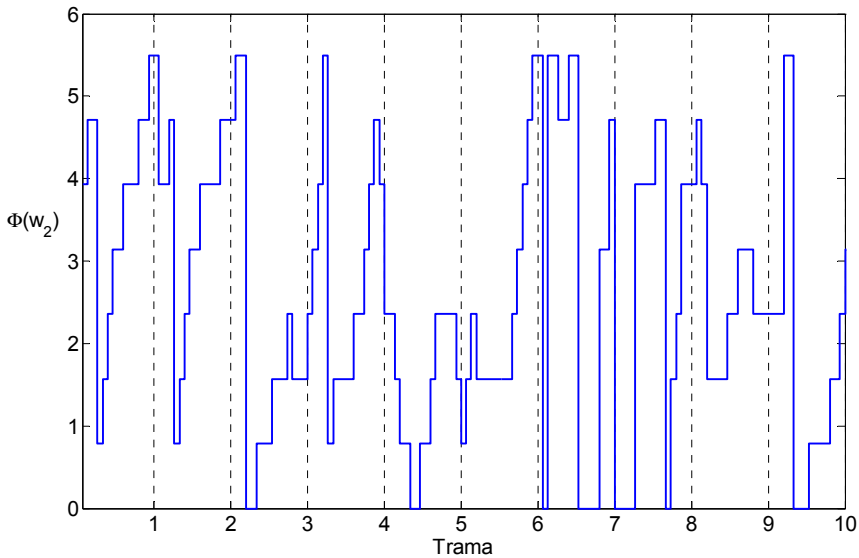
donde  $H$  es el operador hermético, es decir, la matriz transpuesta y conjugada. Para resolver este problema de maximización de manera matemática se puede usar el teorema minimax [19] de manera que el vector de pesos óptimo,  $\mathbf{w}^{opt}$ , es el vector propio asociado al valor propio más alto de la matriz  $\mathbf{H}^H \mathbf{H}$ . A nivel práctico, es más eficaz asignar  $w_1 = 1$  y probar directamente con los ocho valores posibles  $w_2 = e^{j\Phi}$  hasta encontrar el que maximiza (3.6). El valor óptimo de la fase,  $\Phi$ , es el que el usuario mandará codificado en tres bits, transmitiendo el Nodo-B con los correspondientes pesos,  $\hat{\mathbf{w}}^{opt}$ .

La transmisión mediante CLMTD ha sido ya estudiada en la literatura, pero siempre bajo ciertas simplificaciones, o bien considerando un canal plano en frecuencia, o bien suponiendo una estimación perfecta del canal. En esta Tesis Doctoral se ha tenido en cuenta la selectividad en frecuencia del canal y sobre todo las imperfecciones de la estimación del canal, problema que se aborda en profundidad en el siguiente apartado.

### **3.3.5 Estimación de canal en el MNB**

Tal y como se ha descrito en el apartado anterior, el método propuesto para conseguir la máxima ganancia por transmitir desde las dos antenas del MNB se basa en que el terminal ajuste la fase entre las dos señales transmitidas, de manera que la suma de todas las contribuciones recibidas tenga potencia máxima. Este ajuste de fase se envía al Nodo-B quien de manera discreta modificará los pesos utilizados en la transmisión. Si la estimación del canal fuera perfecta y no hubiera ningún retraso entre el canal reportado y el ajuste en la transmisión, se conseguiría una suma coherente de la señal y por tanto una importante ganancia en el receptor.

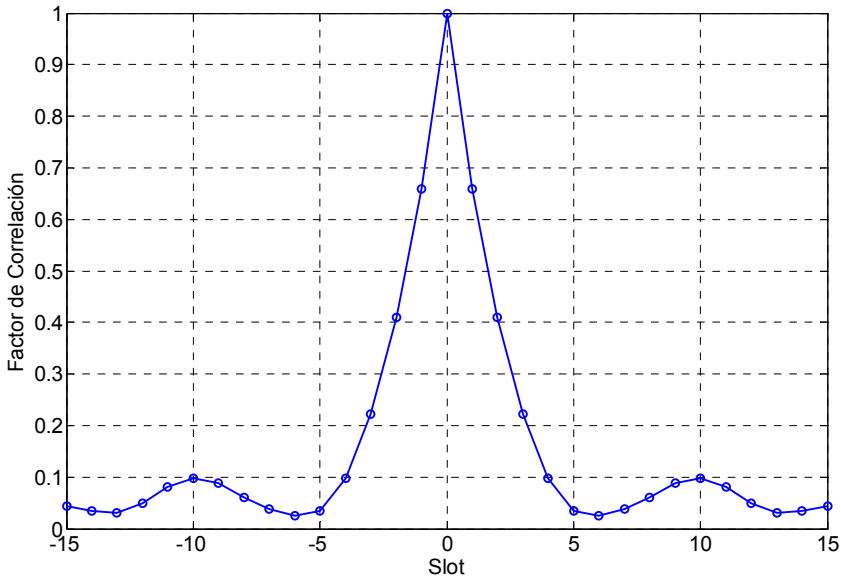
De manera intuitiva los pesos empleados tienden a alinear en fase las dos contribuciones recibidas con mayor potencia. Así pues, no se trata únicamente de un seguimiento del desfase introducido por el canal, sino que ahora la atenuación de cada camino influirá notablemente en el cálculo de los pesos. Esto hace que la variabilidad de los pesos sea mayor que la propia variabilidad de un camino de propagación. Lo primero que se ha de comprobar es cuál es el comportamiento de los pesos,  $w$ , a la máxima velocidad de un terminal en el escenario MNB, es decir, 50 km/h. El modelo de canal es el mismo modelo macro-celular utilizado en apartados anteriores, recibándose dos contribuciones provenientes de las dos antenas implicadas. Además, inicialmente se va a considerar que ambas contribuciones llegan simultáneamente al receptor, de manera que los caminos están solapados.



**Figura 3.21 Evolución del desfase aplicado a la segunda antena durante 10 tramas**

La Figura 3.21 muestra la evolución de la fase del peso aplicado a la segunda antena,  $w_2$ , durante 10 tramas, es decir, 100ms. El otro peso,  $w_1$ , no se representa puesto que, como se explicó en el apartado anterior, se ha optado por transmitir siempre con  $w_1 = 1$ .

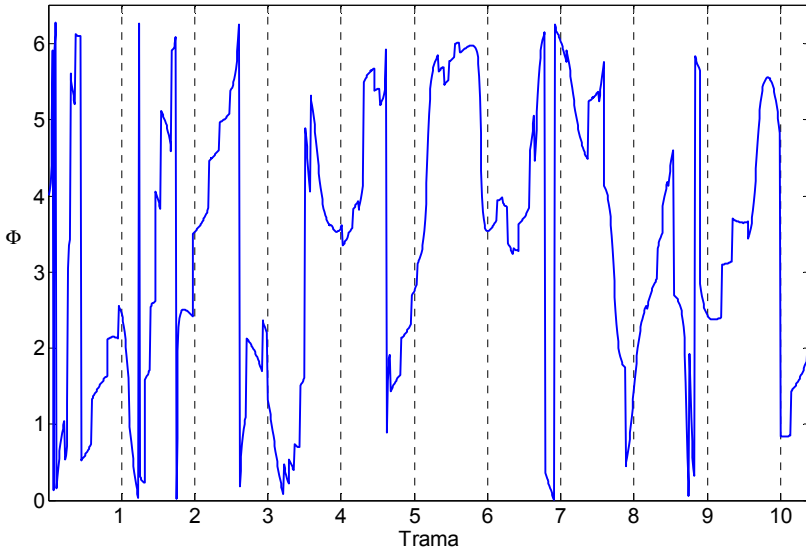
En la imagen se puede observar claramente cómo sólo existen ocho diferentes niveles de fase, puesto que se dispone de 3 bits para el ajuste de fase. Además se distinguen los 15 valores utilizados por trama, correspondientes a los 15 *slots*. Existe una gran variabilidad de los mismos aunque bien es cierto que para ciertos *slots* se repite exactamente el mismo valor del peso reportando en el *slot* anterior. La función de autocorrelación de la fase de  $w_2$  se muestra en la Figura 3.22. Como se puede comprobar, para 50km/h considerar dos *slots* de retraso supone que las muestras tienen una correlación de apenas 0.41, lo que implica que el error cometido en el ajuste no será despreciable. Además, esta alta variabilidad de los pesos aplicados también tendrá un efecto considerable en el canal finalmente experimentado.



**Figura 3.22** Función de autocorrelación del desfase utilizado en transmisión

Debido al efecto multicamino, un canal convencional experimenta una importante variabilidad pero, aún así, a velocidades moderadas, como es a 50km/h, el canal tiene un comportamiento relativamente suave, de manera que la estimación de canal puede seguir sus variaciones, tal y como se mostró en el apartado 3.2.4. Sin embargo, al añadir otro elemento aleatorio de mayor variabilidad, la estimación de canal será incapaz de hacer un seguimiento preciso del canal, al menos utilizando los bits piloto del canal de control dedicado y las técnicas de procesamiento convencionales.

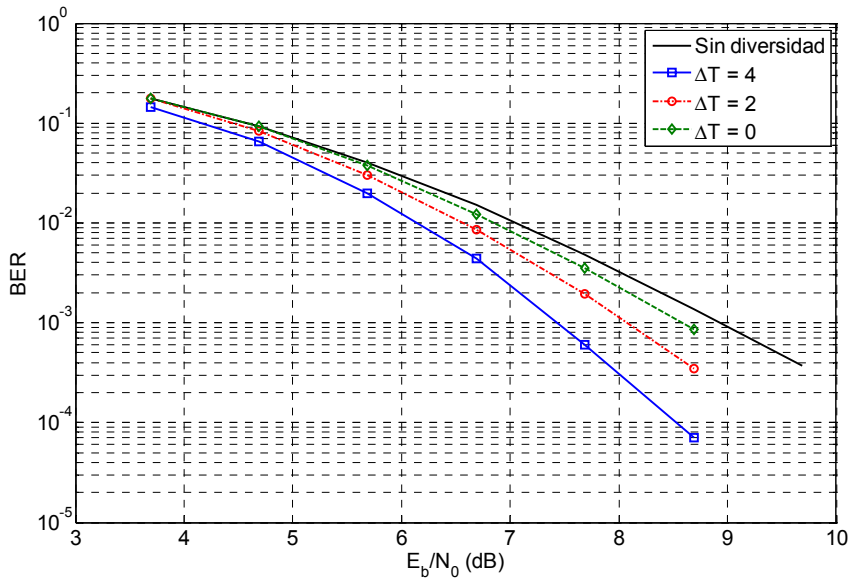
Es importante recordar que, para mejorar las prestaciones de la estimación de canal en UMTS, se suelen aplicar conjuntamente dos mecanismos: el pre-filtrado de las estimas de canal, para reducir el error cometido en la estimación, y una posterior interpolación de las muestras de canal entre dos *slots*. La interpolación se realiza siempre, mientras que en anteriores apartados ya se demostró que el pre-filtrado mejoraba sensiblemente las prestaciones del receptor, sobre todo tratándose de velocidades no muy altas.



**Figura 3.23 Evolución del desfase introducido por el canal utilizando la técnica descrita de procesado en *array* en la transmisión**

Sin embargo, ahora el canal cuenta con un elemento adicional multiplicativo que introduce una importante aleatorización en su evolución. Tal y como se muestra en la Figura 3.23, por aplicar las técnicas de procesado en *array* existen muchos cambios bruscos en el canal, cambios provocados por los ajustes discretos de fase realizados en transmisión. Estos cambios se percibirán como una fuente de ruido por el estimador de canal, de manera que las técnicas antes mencionadas, es decir, interpolación y filtrado, al eliminar dicho ruido no harán más que perjudicar al proceso de seguimiento del canal, teniendo un efecto negativo sobre las prestaciones finales del sistema.

La Figura 3.24 representa la relación BER –  $E_b/N_0$  del servicio de voz cuando se transmite simultáneamente desde ambas antenas aplicando la técnica de procesado en *array*, explicada en el apartado anterior, para garantizar la combinación coherente. Se ha supuesto, además, que al receptor le llega el mismo nivel de potencia desde ambas antenas, es decir que  $\Delta P = 0$ . En la figura se muestra el efecto que tiene la diferencia temporal entre las señales recibidas.



**Figura 3.24 BER vs  $E_b/N_0$  con MNB y CLMTD y una estimación de canal basada en los bits piloto del canal de control**

En la obtención de estos resultados, la estimación de canal se ha realizado a partir de los bits pilotos del canal de control, filtrando las muestras e interpolando entre *slots* consecutivos. Lo primero que se puede apreciar es que el procesado en *array* no afecta a la ganancia experimentada gracias a la utilización del MNB en el caso de que las señales no estén solapadas. Si no existe solape, cualquier desfase introducido en transmisión implica la misma potencia recibida, por lo que la selección inicial no se modifica y  $w_1 = w_2 = 1$ . De ahí que los resultados observados con  $\Delta T = 4$  sean los mismos que los mostrados en la Figura 3.20.

Cuando la señal está totalmente solapada, es decir  $\Delta T = 0$ , se observa una cierta ganancia, siendo dicha ganancia muy inferior a lo esperado. Este efecto negativo es debido a los errores cometidos en la estimación de canal.

Como se ha podido comprobar, cuando las señales están solapadas realizar una estimación del canal directamente de los bits piloto del canal de control dedicado supone cometer muchos errores en la estimación, debido principalmente a las técnicas

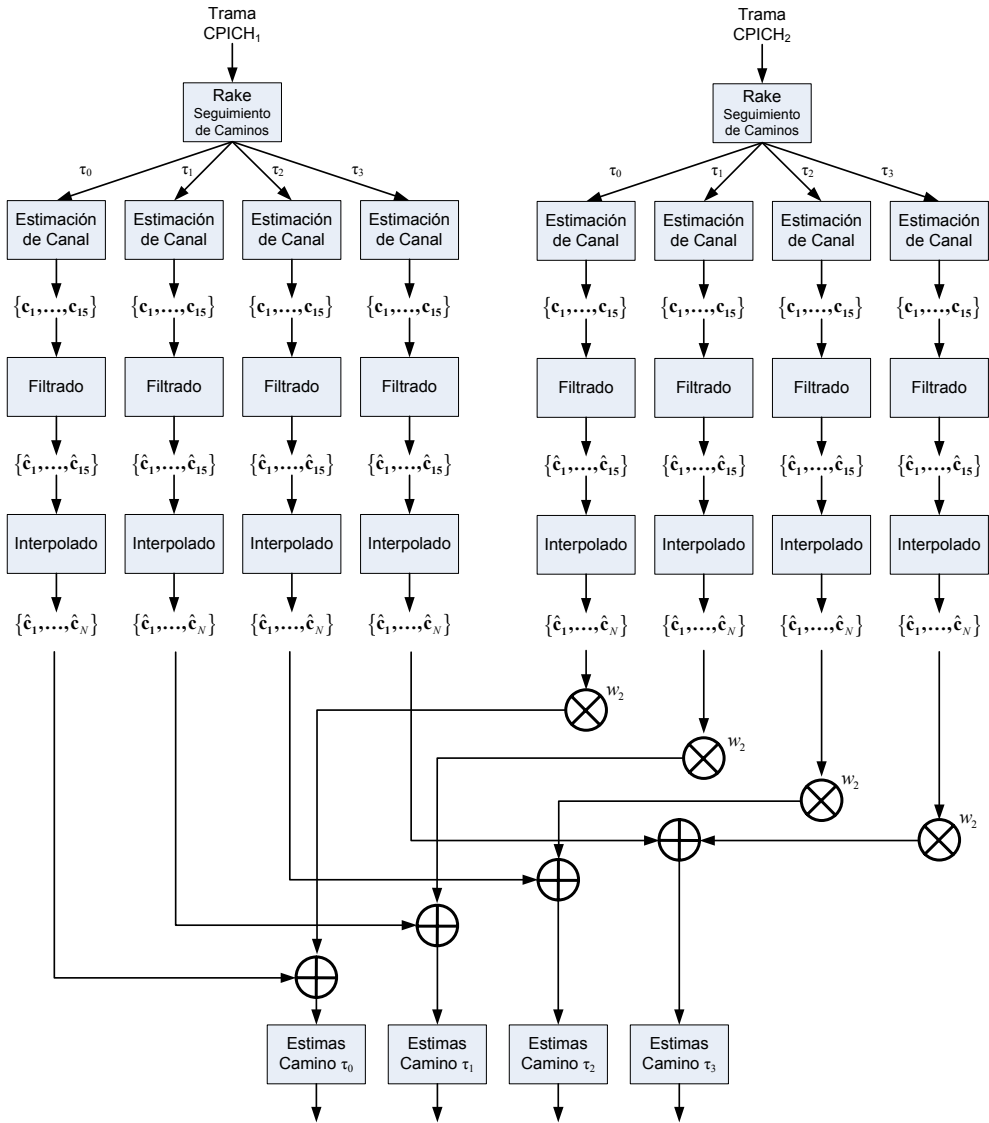
de filtrado e interpolación utilizadas. Una posible solución sería, precisamente, deshabilitar dichas funciones, lo que a priori funcionaría mejor pero, por el contrario, se perdería toda la ganancia que dichas técnicas aportan a la estimación de canal, por lo que no parece muy recomendable.

Existe otra solución que es hacer uso del canal piloto común para realizar la estimación. Por una parte, y tal y como se demostró en el apartado 3.2.4, la utilización del canal piloto común incrementa ostensiblemente la calidad de la estimación de canal. Por otra parte, ambos canales piloto se transmiten utilizando un código de ensanchamiento distinto, lo que implica que el usuario puede decodificarlos por separado. El mecanismo propuesto en esta Tesis Doctoral se basa en realizar la estimación de canal por separado para ambos canales de CPICH, de manera que sobre las muestras de canal recibidas en la trama se apliquen todas las técnicas de filtrado e interpolado comentadas. Como sobre el canal de CPICH no se realiza el ajuste de fase, las señales presentarán una gran continuidad y no se experimentarán los problemas observados anteriormente. El canal final experimentado por la señal recibida en un determinado *finger* será la suma de los canales de la primera y segunda antena que coincidan temporalmente en ese *chip*, con la salvedad de que las contribuciones de la segunda antena estarán previamente multiplicadas por  $w_2$ . Como los pesos utilizados son precisamente los reportados por el usuario, se conoce perfectamente su valor. Además, como el usuario conoce la relación temporal entre los canales de CPICH, será también sencillo calcular qué caminos están solapados y solamente restará multiplicar y sumar convenientemente.

En la Figura 3.25 se muestra esquemáticamente cuál sería el funcionamiento del estimador de canal cuando las dos señales transmitidas por el MNB llegan sincronizadas. En este caso el primer *tap* de la primera antena coincide temporalmente con el primer *tap* de la segunda antena y así sucesivamente, por lo que ambas contribuciones se suman. Además se ha supuesto que el TTI es de 10ms, es decir, que está compuesto de 15 *slots*. Por último decir que  $N$  es el número de símbolos por TTI.

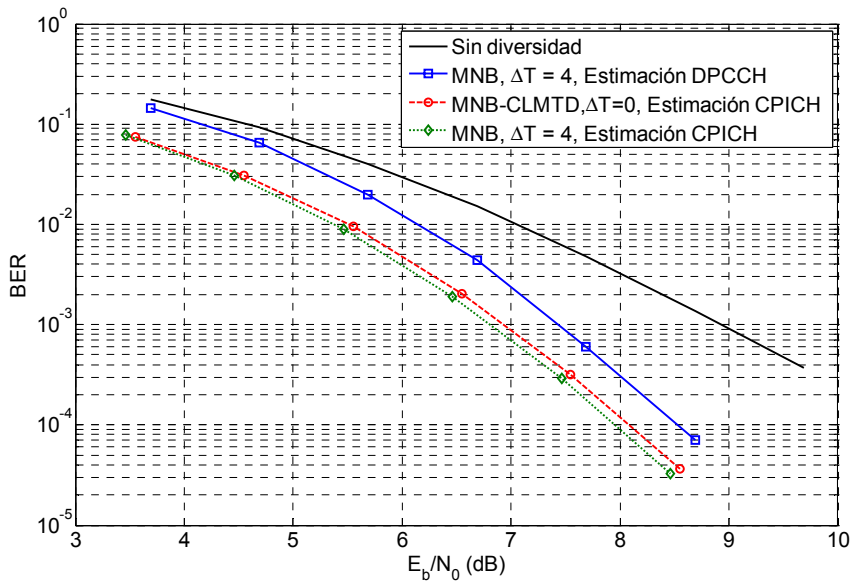
La Figura 3.26 muestra la relación BER -  $E_b/N_0$  obtenida con el nuevo mecanismo de estimación de canal propuesto en esta Tesis Doctoral. Como se puede observar, las

prestaciones obtenidas mejoran ostensiblemente al incluir este nuevo mecanismo, en comparación con los resultados obtenidos con el método de estimación de canal clásico.



**Figura 3.25** Procedimiento de estimación de canal cuando las señales recibidas están perfectamente sincronizadas





**Figura 3.26 BER vs  $E_b/N_0$  con MNB y CLMTD. Efecto del nuevo mecanismo de estimación de canal basado en el uso del canal CPICH**

La mejora respecto a cuando las señales no llegan solapadas se explica por el hecho en sí de que la estimación de canal por CPICH es mucho más potente que la estimación de canal realizada a partir de los bits pilotos del canal de control dedicado, DPCCH. Este hecho ya se explicó anteriormente en el apartado 3.2.4, por lo que los nuevos resultados corroboran tal afirmación. Así pues, aun sin estar la señal solapada, estimar el canal mediante CPICH mejora hasta 0.5 dB los resultados obtenidos gracias a la utilización del MNB. La enorme ventaja es que ahora, gracias al procesado en *array* propuesto en esta Tesis, cuando la señal está totalmente solapada en el tiempo las prestaciones apenas se degradan, pudiéndose conseguir un comportamiento prácticamente constante e independiente del valor de  $\Delta T$ .

Finalmente, es importante destacar que para el escenario considerado, la ganancia obtenida es de hasta 2.1dB, valor sin duda muy significativo y que por sí mismo ya justificaría la utilización de la nueva arquitectura MNB.

### 3.4 Look-up Tables Consideradas

En la concepción del MNB se identifican dos tipos de escenarios de utilización. En el primero, las antenas que conforman el MNB están emplazadas sobre azoteas de modo que se constituyen dos celdas macro-celulares. El modelo de *Power Delay Profile* (PDP) para este tipo de escenarios es el *vehicular*, el mismo que se ha utilizado en todos los análisis realizados hasta ahora. En el segundo escenario, las antenas adicionales introducidas por el MNB se sitúan a nivel de calle, configurando una micro-celda. El modelo de propagación para micro-celdas es el *pedestrian* [20], caracterizado por tener menos caminos resolubles de propagación. La Tabla 3.6 resume el PDP para ambos escenarios de simulación considerados.

<i>Pedestrian</i> (3km/h)		<i>Vehicular</i> (50km/h)	
Tiempo Relativo [ns]	Potencia Media [dB]	Tiempo Relativo [ns]	Potencia Media [dB]
0	0	0	0
976	-10	260	-3
		521	-6
		781	-9

**Tabla 3.6 Power Delay Profile de los escenarios considerados [20]**

Aparte de estos escenarios de uso, también es necesario caracterizar las prestaciones a nivel radio de UMTS y del MNB considerando distintos tipos de servicio. Cada servicio, descrito totalmente a nivel radio por la tabla de su *Radio Bearer*, se caracteriza por un factor de ensanchado determinado, al igual que por un tamaño concreto de los bloques de datos, que al final determinan la longitud de la trama codificada y por tanto sus prestaciones. En el sistema UMTS la voz siempre se transmite por conmutación de circuitos, mientras que el resto de servicios, tales como vídeo conferencia, web o descarga FTP, se pueden transmitir o bien por conmutación de circuitos o bien por conmutación de paquetes. En esta Tesis Doctoral se ha decidido transmitir únicamente la voz por conmutación de circuitos mientras que el resto de servicios se transmitirán haciendo uso del canal compartido DSCH (véase el Capítulo 4 para una descripción más minuciosa del uso del canal compartido).

Además, tal y como se ha mostrado en el apartado anterior, gracias a la técnica de procesado en *array* propuesta en esta Tesis, es posible independizar los resultados obtenidos de la diferencia temporal entre las dos contribuciones recibidas. De los otros dos factores que determinaban las prestaciones del MNB, el número de *fingers* del receptor *Rake* se va a considerar fijo e igual a seis, mientras que el efecto de la diferencia de potencia relativa se deberá evaluar para todos los escenarios.

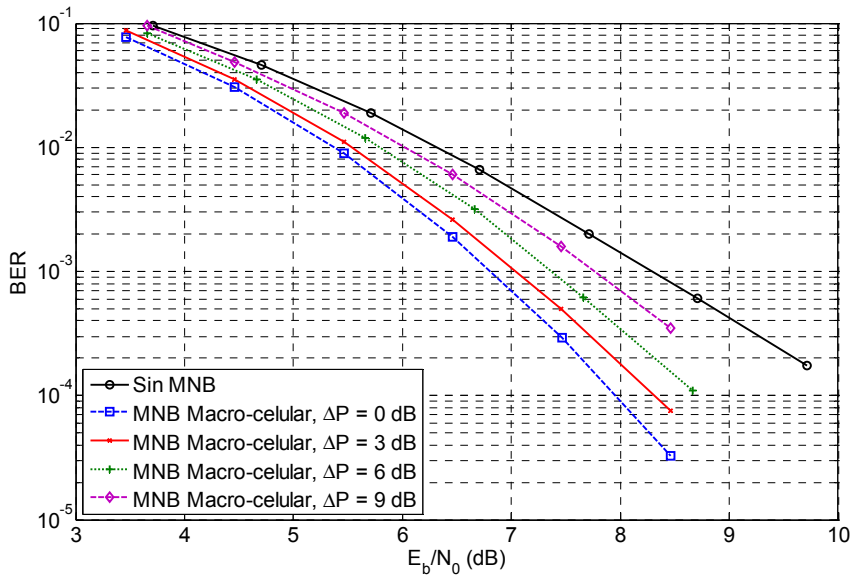
Respecto a la estimación de canal, es importante destacar que a lo largo de todo este apartado se va a considerar su cálculo a partir de la señal piloto del canal común CPICH, ya que esta estimación permite obtener mejores resultados en el sistema, tal y como se ha demostrado en apartados anteriores.

A continuación, se describirán todos los servicios, sus características, los requerimientos de  $E_b/N_0$  sin considerar MNB y la ganancia obtenida en los dos escenarios MNB planteados, teniendo en cuenta distintas diferencias de potencia.

### **3.4.1 Servicio de voz a 12.2 kbps**

Las características del servicio de voz por conmutación de circuitos a 12.2 kbps ya se mostraron en la Tabla 3.2. En primer lugar, se ha analizado el comportamiento radio del sistema MNB cuando se dispone de dos antenas transmitiendo desde emplazamientos macro-celulares. En este caso los caminos resolubles son 8, aunque el *Rake* se sincronizará con los seis de mayor potencia, y se experimentará el mejor comportamiento del MNB. Así pues, tal y como se muestra en la Figura 3.27, para una BER objetivo de  $10^{-3}$  se necesita garantizar una relación  $E_b/N_0$  de apenas 6.8 dB. Conforme la diferencia de potencia entre las contribuciones aumenta, se pierde parte de la diversidad introducida en el sistema y el comportamiento del MNB se acerca al experimentado desde una sola antena en transmisión.

En el segundo experimento realizado se ha considerado la transmisión simultánea desde un emplazamiento macro-celular y uno micro-celular, configurándose lo que se ha denominado un escenario MNB mixto.



**Figura 3.27 Servicio de voz. Relación BER- $E_b/N_0$  con MNB macro-celular. Efecto de la diferencia de potencia relativa entre las señales recibidas**

En este caso se dispone de 5 caminos de propagación útiles, ya que el segundo *tap* del entorno *pedestrian* se recibe con muy poca potencia por lo que apenas aporta diversidad al sistema. En el primer análisis realizado en este escenario se ha considerado que el transmisor dominante es el macro-celular, es decir, se recibe siempre más potencia de la antena principal que de la secundaria.

Los resultados mostrados en la Figura 3.28 demuestran cómo, en este caso, la máxima ganancia conseguida es mucho menor, incrementándose la  $E_b/N_0$  objetivo de 6.8 a 7.67, es decir, casi 1dB menos de ganancia.

Además, la diferencia de potencia relativa entre las señales recibidas no se acusa de igual manera. Así, para una diferencia de 3dB apenas se distingue un empeoramiento en las prestaciones del sistema MNB, igual que para 6 dB.

Por último, se ha considerado el caso restante, es decir, un escenario de transmisión mixto donde el emplazamiento dominante es el micro-celular. La Figura 3.29 muestra los resultados obtenidos en este caso.

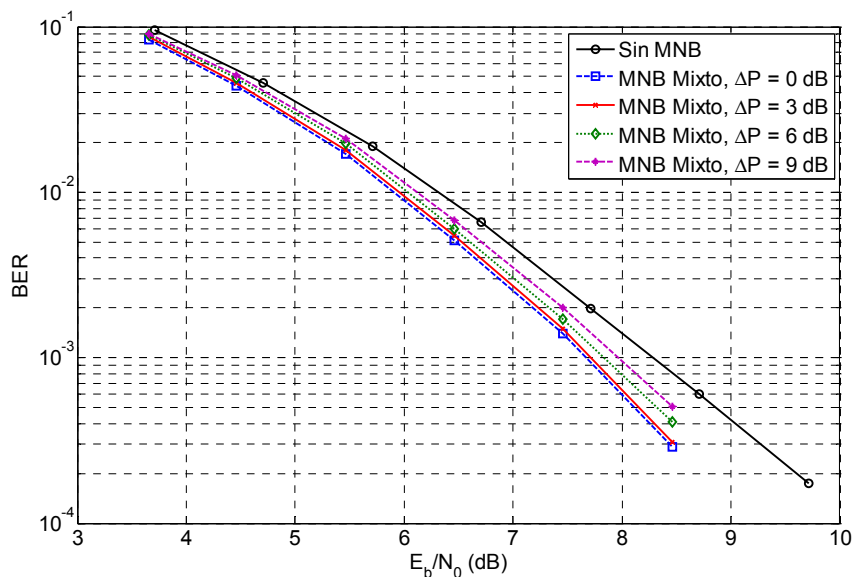


Figura 3.28 Servicio de Voz. Relación BER- $E_b/N_0$  con MNB mixto macro/micro celular. Efecto de la diferencia de potencia relativa entre las señales recibidas siendo mayor el nivel de potencia recibido de la antena macro-celular

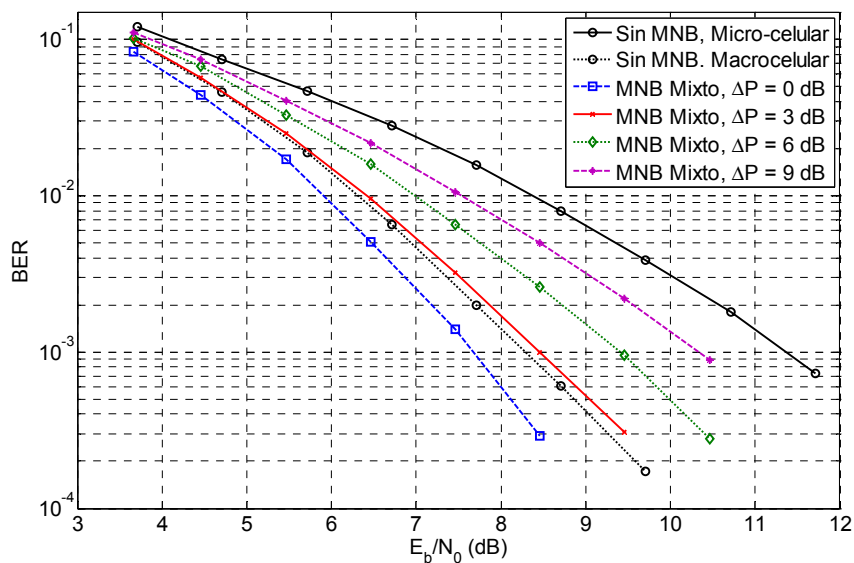


Figura 3.29 Servicio de voz. Relación BER- $E_b/N_0$  con MNB mixto micro/macro celular. Efecto de la diferencia de potencia relativa entre las señales recibidas siendo mayor el nivel de potencia recibido de la antena micro-celular

En primer lugar, es importante comparar los resultados obtenidos sin MNB para ambos modelos de propagación. Obviamente, cuando se considera un canal de propagación como el *pedestrian*, sin apenas diversidad multicamino, el comportamiento es mucho peor, ya que obviamente no se dispone de diversidad con la que trabajar en el receptor *Rake*. El efecto queda más patente al comparar las  $E_b/N_0$  objetivo para ambos escenarios. En el caso del escenario macro-celular la  $E_b/N_0$  objetivo es de 8.28dB, mientras que para el escenario micro-celular este valor objetivo sube hasta 11.36dB, es decir, el disponer de menor diversidad multicamino supone un empeoramiento de las prestaciones de más de 3 dB, lo que implica tener que transmitir con el doble de potencia.

Así pues, cuando una transmisión micro-celular se complementa con la transmisión macro-celular la mejora experimentada en el sistema es espectacular. En concreto, para el servicio de voz considerado, se pasa de la  $E_b/N_0$  objetivo de 11.36dB a 7.67dB, lo que supone una ganancia de 3.69dB.

Los resultados también muestran que, en este caso, la diferencia de potencia entre las señales recibidas tiene un efecto muy importante sobre esta ganancia, por lo que deberá calcularse para cada  $\Delta P$  cuál es la ganancia específica.

Tras este análisis de todos los escenarios considerados, se puede concluir que es necesario obtener distintas *Look-up tables* en función del tipo de transmisor y la relación de potencia entre las contribuciones recibidas.

### 3.4.2 Servicios de datos

Para los servicios de datos se ha considerado la transmisión a través del canal de transporte DSCH, mapeado sobre el canal físico PDSCH (*Physical Downlink Shared Channel*). Hay que recalcar que el canal de transporte DSCH no soporta *soft handover*, por lo que el MNB aportará una ganancia inédita en este tipo de transmisión. Además, hay que recordar que la codificación utilizada en este canal es turbo-codificación, mucho más potente sobre todo para secuencias de datos largas.

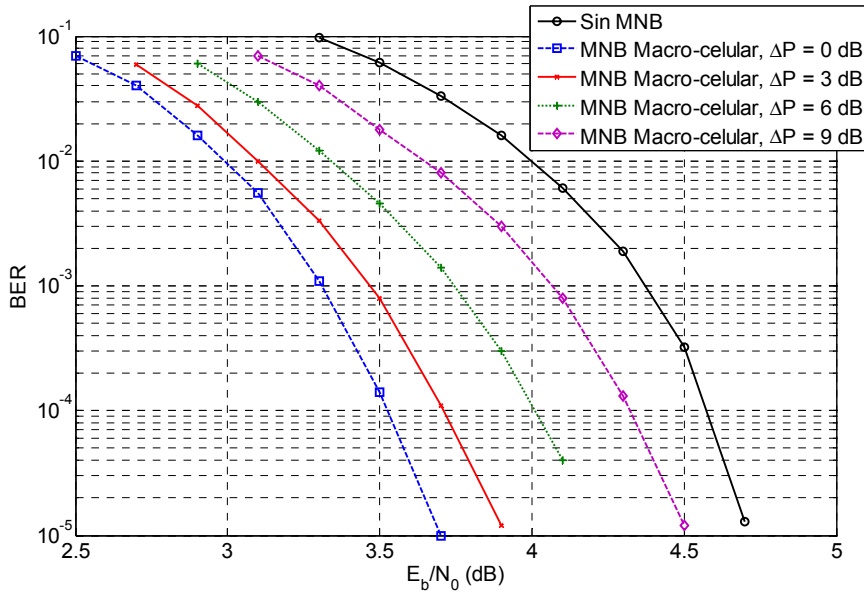
Higher layer	RAB/Signalling RB	RAB	
RLC	Logical channel type	DTCH	
	RLC mode	AM	
	Payload sizes, bit	320	
	Max data rate, bps	16000	
	AMD PDU header, bit	16	
MAC	MAC header, bit	0	
	MAC multiplexing	N/A	
Layer 1	TrCH type	DSCH	
	SF	128	
	TB sizes, bit	336	
	TFS	TF0, bits	0x336
		TF1, bits	1x336
		TF2, bits	2x336
	TTI, ms	40	
	Coding type	TC	
	CRC, bit	16	
	Max number of bits/TTI after channel coding	2124	
RM attribute	135-175		

**Tabla 3.7 RAB para el servicio de datos a 16 kbps [3]**

El primer servicio de datos analizado transmite a una tasa máxima de 16kbps. Las características de este servicio se resumen en la Tabla 3.7. Todas las características de la RAB se han implementado detalladamente en el simulador de enlace, de manera que se garantice la validez de los resultados.

En primer lugar, se han analizado los requerimientos del servicio en el escenario MNB macro-celular. Tal y como se observa en la Figura 3.30, los resultados para servicios de datos son mucho mejores que los obtenidos para el servicio de voz. Esto es debido principalmente al mejor funcionamiento de la turbo-codificación, tal y como queda demostrado en el Anexo A.

Para el servicio de datos, se ha considerado una BER objetivo de  $10^{-4}$ , valor que está en consonancia con las especificaciones [15]. Tal y como se puede observar, la ganancia en valor absoluto obtenida por utilizar MNB en el escenario macro-celular se ha reducido en el servicio de datos a 16kbps. Este valor resulta ser de 1.05dB para datos mientras que para voz era de 1.48dB. Aunque a nivel absoluto la ganancia conseguida por el MNB es menor para el servicio de datos, a nivel relativo resulta más alta, por lo que no puede considerarse como un mal resultado.



**Figura 3.30 Servicio de datos a 16kbps. Relación BER- $E_b/N_0$  con MNB macro-celular. Efecto de la diferencia de potencia relativa entre las señales recibidas**

Por último, se puede comprobar cómo, al igual que en el servicio de voz, el aumento de la diferencia relativa de potencia afecta considerablemente a las prestaciones ofrecidas a nivel radio por el MNB. Aún con 9dB de diferencia, la ganancia sigue siendo significativa. Conforme aumenta la diferencia de potencia el comportamiento del MNB se aproxima al mostrado por el sistema convencional.

Respecto al MNB mixto con mayor potencia recibida del emplazamiento macro-celular, los resultados obtenidos se resumen gráficamente en la Figura 3.31. Como se puede observar de nuevo, la ganancia introducida por el nuevo *tap* del enlace micro-celular no es muy significativa, y menos en el caso de transmisión de datos, donde esta ganancia se reduce hasta los 0.37dB. La reducción en esta ganancia no es muy significativa conforme aumenta la diferencia de potencia relativa, pudiéndose considerar una ganancia de 0 dB por encima de los 6dB de diferencia de potencia.



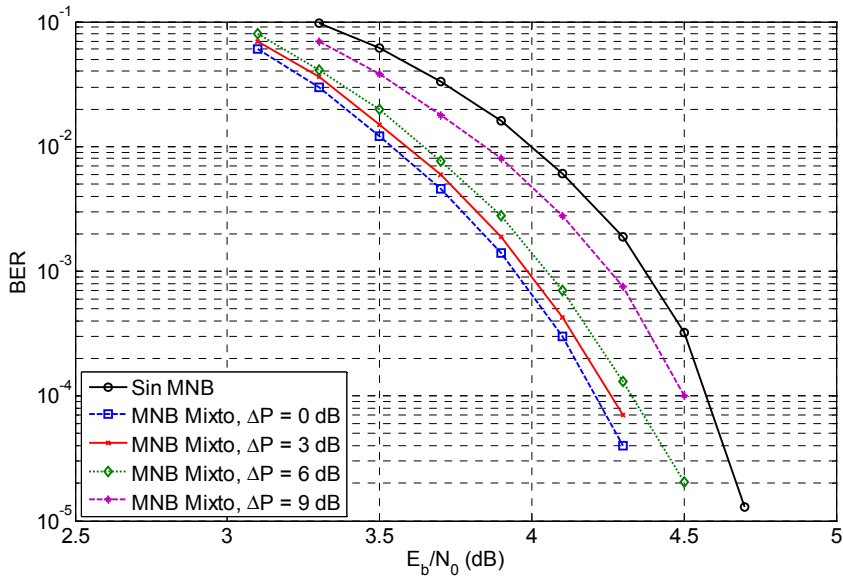


Figura 3.31 Servicio de datos a 16kbps. Relación BER- $E_b/N_0$  con MNB Mixto macro/micro celular. Efecto de la diferencia de potencia relativa entre las señales recibidas siendo mayor el nivel de potencia recibido de la antena macro-celular

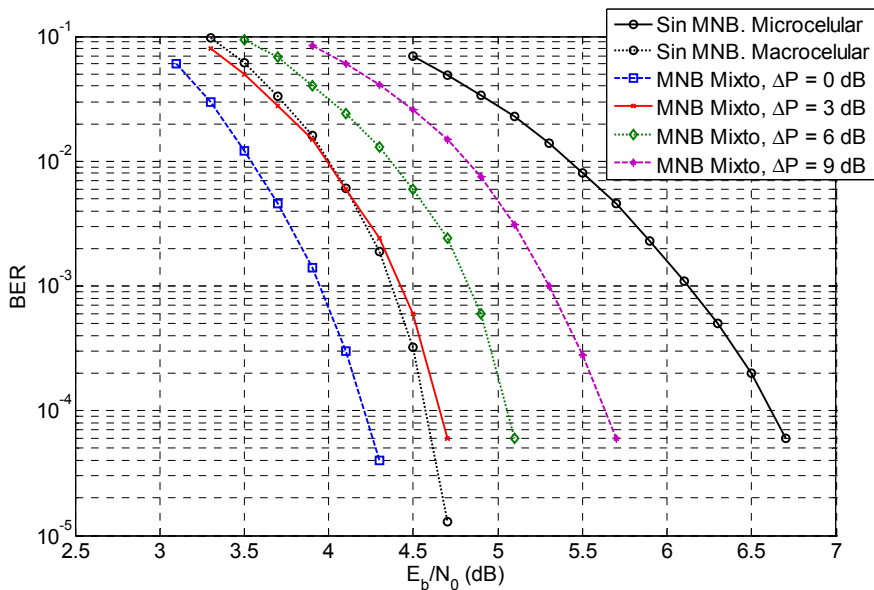


Figura 3.32 Servicio de datos a 16kbps. Relación BER- $E_b/N_0$  con MNB Mixto micro/macro celular. Efecto de la diferencia de potencia relativa entre las señales recibidas siendo mayor el nivel de potencia recibido de la antena micro-celular

Finalmente, en la Figura 3.32 se ha dibujado el comportamiento radio del sistema en el escenario mixto, siendo en este caso la señal de mayor potencia la proveniente del emplazamiento micro-celular. En este caso, la ganancia del MNB es mucho mayor ya que la carencia de diversidad por multicamino que caracteriza al entorno de propagación *pedestrian* se compensa gracias al uso del MNB. Además, es importante destacar que, incluso con diferencias de potencia de 9dB, la ganancia conseguida por el MNB es bastante significativa.

Para el resto de servicios de datos no se va a representar gráficamente las prestaciones del MNB ya que el comportamiento es muy similar al ya observado.

En la Tabla 3.8 se muestran las características del servicio de transmisión de datos a 128 kbps como resumen de las características del resto de servicios considerados. La única diferencia entre el servicio de transmisión de datos a 128 kbps, a 64 y a 32 es el número máximo de TB transmitidos por TTI y el *Spreading Factor*. En el caso de transmisión de datos a 64 kbps el máximo número de TB por TTI es de 4 y el SF es 32, mientras que para datos a 32kbps el número máximo de TB es 2 y el SF 64. Respecto al servicio de datos a 256 kbps la tabla de descripción de la RAB es exactamente igual a la de 128 salvo que el TTI es justo la mitad, es decir 10ms, de ahí que se transmita exactamente el doble de bits por segundo.

Finalmente, la Tabla 3.9 resume el comportamiento del MNB para los distintos tipos de servicios considerados y escenarios. Lo que se muestra es la  $E_b/N_0$  objetivo para los valores de BER considerados para cada tipo de servicio, es decir,  $10^{-3}$  para voz y  $10^{-4}$  para datos. El valor de G representa la ganancia respecto a no considerar MNB.

Observando la tabla se puede apreciar que la ganancia obtenida por transmitir en un escenario macro-celular con la diversidad introducida por el MNB está acotada entre 1 y 1.5dB, dependiendo del servicio. Esta ganancia es, sin duda, muy importante, sobre todo tratándose de un escenario *vehicular* a 50 km/h. Para velocidades más bajas se puede esperar un incremento adicional en la ganancia.

Higher layer	RAB/Signalling RB	RAB	
RLC	Logical channel type	DTCH	
	RLC mode	AM	
	Payload sizes, bit	320	
	Max data rate, bps	128000	
	AMD PDU header, bit	16	
MAC	MAC header, bit	0	
	MAC multiplexing	N/A	
Layer 1	TrCH type	DSCH	
	SF	16	
	TB sizes, bit	336	
		TF0, bits	0x336
		TF1, bits	1x336
		TF2, bits	2x336
		TF3, bits	4 x336
		TF4, bits	8 x336
	TTI, ms	20	
	Coding type	TC	
	CRC, bit	16	
Max number of bits/TTI after channel coding	8460		
RM attribute	120-160		

**Tabla 3.8 RAB para el Servicio de Datos a 128 kbps [3]**

Por otra parte, la ganancia conseguida en un entorno mixto dependerá de con respecto a qué se esté comparando el comportamiento del MNB. Si se compara respecto a transmitir desde un emplazamiento macro-celular la ganancia es poca, próxima a los 0.5dB. Sin embargo, para un escenario micro-celular la ganancia conseguida gracias al Multinodo B es muy importante y está siempre por encima de los 2dB, llegando incluso a alcanzar una ganancia de 3.69 dB para el servicio de voz.

Respecto a los servicios de datos, en general las prestaciones de los servicios de alta tasa binaria no son mucho peores que los de baja tasa. Esto es debido a que, tal y como se muestra en el Anexo A, las prestaciones de los turbo-códigos son mucho mejores con tamaños de bloque crecientes y precisamente los servicios de mayor tasa binaria transmiten bloques más grandes. Obviamente, a partir de este hecho tampoco se puede generalizar y concluir que los servicios de alta tasa requieren menor  $E_b/N_0$  que los de menor tasa, ya que el factor de ensanchado también influye en las prestaciones del receptor.

ESCENARIO MACRO-CELULAR												
	Voz		16kbps		32kbps		64kbps		128kbps		256kbps	
	$E_b/N_0$	G	$E_b/N_0$	G	$E_b/N_0$	G	$E_b/N_0$	G	$E_b/N_0$	G	$E_b/N_0$	G
SIN MNB	8.28	0	4.59	0	4.33	0	4.42	0	4.61	0	4.79	0
$\Delta P=0$	6.8	1.48	3.54	1.05	3.42	0.91	3.47	0.95	3.6	1.01	3.69	1.1
$\Delta P=1$	6.86	1.42	3.57	1.02	3.44	0.89	3.52	0.9	3.65	0.96	3.82	0.97
$\Delta P=2$	6.93	1.35	3.65	0.94	3.52	0.81	3.59	0.83	3.72	0.89	3.89	0.90
$\Delta P=3$	7.04	1.24	3.73	0.86	3.61	0.72	3.66	0.76	3.79	0.82	3.96	0.83
$\Delta P=4$	7.14	1.14	3.82	0.78	3.71	0.62	3.76	0.66	3.87	0.74	4.06	0.73
$\Delta P=5$	7.25	1.03	3.92	0.67	3.78	0.55	3.84	0.58	3.98	0.63	4.15	0.64
$\Delta P=6$	7.36	0.92	4.03	0.56	3.86	0.47	3.92	0.50	4.09	0.52	4.24	0.55
$\Delta P=7$	7.51	0.77	4.14	0.45	3.94	0.39	4	0.42	4.18	0.43	4.33	0.46
$\Delta P=8$	7.63	0.65	4.25	0.34	4.02	0.31	4.09	0.33	4.27	0.34	4.43	0.36
$\Delta P=9$	7.77	0.51	4.34	0.25	4.12	0.21	4.19	0.23	4.37	0.24	4.52	0.27
ESCENARIO MIXTO MACRO-CELULAR-MICRO-CELULAR												
	Voz		16kbps		32kbps		64kbps		128kbps		256kbps	
	$E_b/N_0$	G	$E_b/N_0$	G	$E_b/N_0$	G	$E_b/N_0$	G	$E_b/N_0$	G	$E_b/N_0$	G
SIN MNB	8.28	0	4.59	0	4.33	0	4.42	0	4.61	0	4.79	0
$\Delta P=0$	7.67	0.61	4.15	0.44	3.91	0.42	4	0.42	4.17	0.44	4.33	0.46
$\Delta P=1$	7.68	0.6	4.17	0.42	3.92	0.41	4.02	0.40	4.19	0.42	4.35	0.44
$\Delta P=2$	7.7	0.58	4.19	0.40	3.94	0.39	4.04	0.38	4.22	0.39	4.37	0.42
$\Delta P=3$	7.72	0.56	4.22	0.37	3.97	0.36	4.07	0.35	4.24	0.37	4.4	0.39
$\Delta P=4$	7.76	0.52	4.25	0.34	4	0.33	4.09	0.33	4.26	0.35	4.44	0.35
$\Delta P=5$	7.79	0.49	4.28	0.31	4.02	0.31	4.11	0.31	4.29	0.32	4.47	0.32
$\Delta P=6$	7.83	0.45	4.3	0.29	4.05	0.28	4.14	0.28	4.31	0.3	4.49	0.3
$\Delta P=7$	7.88	0.4	4.36	0.23	4.1	0.23	4.2	0.22	4.37	0.24	4.54	0.25
$\Delta P=8$	7.92	0.36	4.42	0.17	4.15	0.18	4.27	0.15	4.44	0.17	4.61	0.18
$\Delta P=9$	7.97	0.31	4.5	0.09	4.24	0.09	4.34	0.08	4.51	0.1	4.68	0.11
ESCENARIO MIXTO MICRO-CELULAR-MACRO-CELULAR												
	Voz		16kbps		32kbps		64kbps		128kbps		256kbps	
	$E_b/N_0$	G	$E_b/N_0$	G	$E_b/N_0$	G	$E_b/N_0$	G	$E_b/N_0$	G	$E_b/N_0$	G
SIN MNB	11.36	0	6.62	0	6.12	0	6.37	0	6.70	0	7.09	0
$\Delta P=0$	7.67	3.69	4.21	2.41	3.96	2.16	4.05	2.32	4.20	2.50	4.2	2.89
$\Delta P=1$	7.83	3.53	4.32	2.3	4.05	2.07	4.14	2.23	4.31	2.39	4.33	2.76
$\Delta P=2$	8.1	3.26	4.45	2.17	4.14	1.98	4.22	2.15	4.47	2.23	4.49	2.6
$\Delta P=3$	8.46	2.9	4.66	1.96	4.26	1.86	4.41	1.96	4.65	2.05	4.65	2.44
$\Delta P=4$	8.73	2.63	4.78	1.84	4.38	1.74	4.56	1.81	4.79	1.91	4.78	2.31
$\Delta P=5$	9.07	2.29	4.93	1.69	4.51	1.61	4.71	1.66	4.96	1.74	4.96	2.13
$\Delta P=6$	9.41	1.95	5.06	1.56	4.66	1.46	4.87	1.5	5.09	1.61	5.15	1.94
$\Delta P=7$	9.67	1.69	5.23	1.39	4.84	1.28	5.03	1.34	5.26	1.44	5.36	1.73
$\Delta P=8$	10.02	1.34	5.41	1.21	5.07	1.05	5.19	1.18	5.41	1.29	5.55	1.54
$\Delta P=9$	10.33	1.03	5.63	0.99	5.24	0.88	5.33	1.04	5.59	1.11	5.75	1.34

**Tabla 3.9 Resumen de las prestaciones del Multinodo B**

Así pues, al transmitir con mayor tasa el factor de ensanchado será menor, lo que supondrá una menor protección frente a interferencias y una mayor ISI. Como se puede

observar, el comportamiento del servicio de transmisión de datos a 32 kbps y a 64 kbps es mejor que el de 16 kbps, debido a la mejora en la codificación. Sin embargo, para 128 y 256 kbps esta mejora no compensa el empeoramiento sufrido en las prestaciones debido a reducir el factor de ensanchado, por lo que la  $E_b/N_0$  objetivo es mayor.

## 3.5 Conclusiones

Este capítulo se ha centrado en analizar la ganancia conseguida gracias a la nueva arquitectura radio Multinodo B en relación al incremento de diversidad en transmisión introducido. Con este objetivo, se ha desarrollado un simulador de enlace completo del sistema UMTS. La descripción del mismo no se aborda en este capítulo pero sí su validación a partir de otros trabajos previos encontrados en la literatura.

Tras un primer análisis de las prestaciones del MNB sin realizar ningún procesado previo adicional de la señal transmitida, se demuestra que el MNB es capaz de ofrecer una mejora importante en el comportamiento del sistema, siempre que las señales recibidas no se solapen en el tiempo. Sin embargo, si las señales llegan solapadas no se puede garantizar la combinación coherente, por lo que se pierde la ganancia en diversidad multicamino.

Para solucionar esta limitación, en este capítulo se ha propuesto hacer uso de una técnica de procesado en *array* que permite maximizar la potencia recibida de señal útil. Para que esta técnica funcione es necesario recibir información del enlace descendente, información que solamente el terminal puede reportar. Para enviar dicha información de *feedback* se ha propuesto utilizar unos bits específicos de realimentación que todo terminal UMTS tiene que ser capaz de procesar y enviar al Nodo-B. Sin embargo, el uso de esta técnica de procesado por sí sola no es capaz de solucionar plenamente el problema, si no se la acompaña de un mecanismo de estimación de canal apropiado. Por ello, se ha propuesto un mecanismo sencillo de estimación de canal a partir de los bits piloto enviados por el canal CPICH.

Utilizando conjuntamente la técnica de procesado en *array* y el mecanismo adaptado de estimación de canal, se ha evaluado el comportamiento del MNB cuando las señales se reciben temporalmente solapadas. Los resultados obtenidos muestran que, gracias a los cambios introducidos, en cualquier caso el MNB es capaz de combinar coherentemente las señales recibidas desde ambas antenas.

Una vez comprobada la capacidad de aumentar la diversidad que ofrece la nueva arquitectura MNB, se ha evaluado el valor exacto de esta ganancia para un conjunto de servicios y escenarios específicos. Además, se ha caracterizado el comportamiento del receptor dependiendo de la diferencia relativa de potencia entre las señales recibidas, puesto que esta diferencia de potencia tiene un efecto muy notable en las prestaciones del sistema.

Observando los resultados se puede afirmar que el MNB es capaz de mejorar notablemente el comportamiento del sistema, sobre todo en entornos micro-celulares a los que se complementa con una transmisión macro-celular. En este caso la ganancia conseguida puede llegar a alcanzar hasta los 3.67 dB.

La nueva arquitectura radio Multinodo B se presenta, por tanto, como una gran evolución de la arquitectura radio convencional, capaz de dotar de una importante diversidad adicional a los entornos más desfavorecidos. Además, ya que en la transmisión de datos mediante el canal de transporte DSCH no se puede utilizar *soft handover*, el MNB permitirá cubrir esta notable carencia del estándar.

## Referencias

- [1] H. Holma, A. Toskala, "WCDMA for UMTS: Radio Access For Third Generation Mobile Communications", John Wiley & Sons, 2001.
- [2] S. Hämäläinen, H. Holma, K. Sipilä, "Advanced WCDMA Radio Network Simulator", in Proceedings of the IEEE International Symposium on Personal, Indoor and Mobile Radio Conference, pp. 951-955, 1999.
- [3] 3GPP TS 34.108 V5.3.0, "Common Test Environments for User Equipment; Conformance Testing", 2004.
- [4] J.J.Olmos, S.Ruiz, COST273: MORANS, "Description of UTRA-FDD link level simulator", May 2003.
- [5] IST-2000-25133, "Advanced Radio Resource Management for Wireless Services (ARROWS)".
- [6] 3GPP TS 25.214 V5.9.0, "Universal Mobile Telecommunications System (UMTS); Physical Layer Procedures (FDD)", 2004.
- [7] J. Monserrat, R. Patelli and N. Cardona, "Performance Comparison of Two Dynamic System Simulators Over COST273-MORANS Reference Scenarios", in Proceedings of the Wireless Personal and Mobile Communications, 2005.
- [8] J.G.Proakis, "Digital Communications", Ed. Mc.Graw Hill, 1983, pp. 486.
- [9] 3GPP TS 25.101 V7.7.0, "UE Radio Transmission and Reception (FDD)", 2007.
- [10] J.J.Olmos, S.Ruiz, "Transport block error rates for UTRA-FDD downlink with transmission diversity and turbo coding", in Proceedings of the IEEE International Symposium on Personal, Indoor and Mobile Radio Conference, 2002.
- [11] B. Melis, G. Romano, "UMTS W-CDMA: Evaluation of radio performance by means of link level simulations", in Proceedings of the IEEE Personal Communications Conference, pp. 42-49, 2000.
- [12] E.T.S.I., "UMTS Terrestrial Radio Access; Concept Evaluation," UMTS 30.06, 1997.
- [13] L. Hanzo, "Turbo Coding, Turbo Equalisation and Space-Time Coding for Transmission over Fading Channels", John Wiley & Sons, 2002.
- [14] H. Andoh, M. Sawahashi, F. Adachi, "Channel Estimation Using Time Multiplexed Pilot Symbols for Coherent *Rake* Combining for DS-CDMA Mobile Radio", in Proc.



- of the IEEE Personal Indoor and Mobile Radio Communications Conference, vol. 3, pp. 954 - 958, 1997.
- [15] 3GPP TS 23.107 V3.9.0, "Quality of Service (QoS) Concept and Architecture", 2002.
- [16] H. Al-Raweshidy, "Radio Over Fiber Technologies for Mobile Communications Networks", Ed. Artech House, 2002.
- [17] 3GPP TS 25.211, "Physical channels and mapping of transport channels onto physical channels (FDD)".
- [18] B.D. Van Veen, K.M. Buckley, "Beamforming: a Versatile Approach to Spatial Filtering", IEEE Signal Processing Magazine, vol. 5, no. 2, pp. 4 - 24, 1988.
- [19] S. Haykin, "Adaptive Filter Theory", Prentice-Hall, 1996.
- [20] 3GPP TS 25.101 V7.7.0, "User Equipment (UE) radio transmission and reception (FDD)", 2007.
- [21] J. F. Monserrat, D. Gómez-Barquero, D. Calabuig, L. Rubio, and N. Cardona, "Evaluation of soft *handover* Micro Diversity Gain on the UMTS System Capacity and QoS", in Proc. of the IEEE Personal Indoor and Mobile Radio Communications Conference, Helsinki, September 2006.



## **Parte II**

*Gestión de Recursos en la red MNB*

---



# Capítulo 4

## *Gestión de Recursos en Redes UMTS*

*Si buscas resultados distintos, no hagas siempre lo mismo.*

*Albert Einstein*

En el Capítulo 3 se ha analizado en profundidad cuál es la ganancia por diversidad que la nueva arquitectura propuesta en esta Tesis Doctoral introduce en el sistema UMTS. Aun queda por resolver otro de los retos que plantea esta nueva arquitectura: el reparto óptimo de los recursos disponibles en la red, de manera que se pueda aprovechar al máximo esta ganancia. Si se diseña un algoritmo de reparto que sepa desde qué antena o qué combinación de antenas deben transmitir a cada usuario, se podrá configurar una red mucho más flexible que minimice el nivel de interferencia introducido en el sistema y que, por tanto, permita maximizar la capacidad de la red.

El diseño óptimo de la gestión de recursos en redes MNB va ser el objetivo de los siguientes dos capítulos. En éste se aborda, de forma genérica, la optimización de la gestión de recursos en UMTS, mientras que en el siguiente se extenderán los resultados obtenidos a la nueva arquitectura radio MNB.

### **4.1 Introducción**

La historia ha demostrado que las redes móviles están en constante evolución y que la demanda de los usuarios de servicios multi-media, incluyendo distintos tipos de servicios, siempre está en alza. La calidad de servicio o *Quality of Service* (QoS) de

tales servicios varía considerablemente ya que los requerimientos de cada tipo son muy distintos. En general, los servicios se pueden clasificar en servicios de tiempo real o *Real Time* (RT) y de no tiempo real o *Non Real Time* (NRT) [1]. Los servicios RT, como por ejemplo voz y vídeo conferencia, se caracterizan por sus estrictas restricciones temporales, ya que los paquetes se han de transmitir antes de un determinado tiempo conocido como *deadline*. Por otra parte, los servicios NRT, como correo electrónico, navegación web o descarga FTP, se caracterizan por un patrón de tráfico a ráfagas y una restricción temporal no estricta. Con respecto al retardo en la recepción de los servicios NRT, normalmente no se suele definir un retardo máximo pero aún así el tiempo de respuesta del servicio, definido como el tiempo transcurrido desde el instante de la petición hasta que se produce la recepción completa y correcta de la información, es una buena medida de la satisfacción percibida por el usuario, ya que, tal y como indican las especificaciones [1], “*el usuario espera la recepción del mensaje en un tiempo acotado*”.

Sin embargo, la toma de esta decisión no es un proceso trivial por el elevado número de condiciones a tener en cuenta simultáneamente. Cada uno de los usuarios de la celda estará caracterizado por el servicio que se le esté prestando, las condiciones particulares de satisfacción de ese servicio y la posición relativa del usuario. Además, hay que tener en cuenta la particularidad que plantea UMTS, ya que las interferencias introducidas afectan al resto de usuarios y, además, estas interferencias dependen de la posición relativa de los mismos.

Los principales objetivos de un algoritmo de gestión de recursos son: maximizar la utilización de los recursos disponibles, distribuir los mismos de la manera más equitativa posible y satisfacer simultáneamente los requerimientos de QoS de cada usuario particular [2][3]. Se trata, por tanto, de un problema de optimización multi-objetivo, lo que implica un espacio de búsqueda N-dimensional enorme. Además, el gestor de recursos, o *scheduler*, tiene que operar en tiempo real, lo que complica más si cabe la solución del problema.

Existen dos enfoques distintos para realizar la asignación de recursos: en el primero, conocido como asignación fija de recursos, o *Fixed Resource Allocation* (FRA), se

reserva durante toda la comunicación una cantidad fija de recursos para cada sesión activa en el sistema. Esta política es la más utilizada para servicios RT aunque resulta tremendamente inadecuada para servicios NRT, ya que en este caso los recursos reservados estarán infrautilizados. Por otro lado, los algoritmos basados en una asignación dinámica de recursos, o *Dynamic Resource Allocation* (DRA), establecen varias conexiones sobre los mismos recursos, de manera que es necesario disponer de alguna política de cronoejecución o *scheduling* para determinar qué usuario tiene que hacer uso del recurso en cada instante. El uso de esquemas basados en DRA para mejorar la capacidad de los sistemas de comunicaciones móviles ha sido ampliamente discutido en la literatura [1]-[10]. Además, recientemente se han propuesto distintas técnicas inteligentes de DRA que utilizan métodos computacionales basados en redes neuronales [11]-[15], introduciendo así una nueva clase de esquemas de DRA neuronales.

Los algoritmos de DRA varían en función de la regla de decisión utilizada para aceptar o rechazar una asignación de recursos específica, siempre que ésta satisfaga las restricciones del sistema, es decir, ancho de banda disponible y potencia máxima transmitida. En [4] El-Kadi propuso un esquema que, en caso de que un nuevo usuario necesite un ancho de banda adicional, pide ‘prestado’ ancho de banda a otras conexiones ya existentes, de manera que se minimice el rechazo de conexiones. Epstein y Schwartz [2] añadieron a esta solución el uso de técnicas de predicción para realizar la reserva de recursos. En general, las técnicas de reserva de ancho de banda son útiles para fuentes de tráfico constantes pero no para servicios a ráfagas como los NRT. Otras técnicas, como las propuestas por Mihailescu [5] y Forkel [6], asignan recursos en función de una mejor calidad en el enlace radio. Este tipo de técnicas, basadas en asignar recursos primero a aquellos usuarios con máximo valor de SIR, maximizan la tasa binaria media o *throughput* medio del sistema pero a costa de una distribución no equitativa de los recursos, lo que causa altos retardos en transmisión para aquellos usuarios que experimentan malas condiciones del canal radio. Hernández y Valdovinos [7] establecieron una política de prioridades que considera diferentes clases de usuarios y una asignación dinámica de dichas prioridades en función del

tiempo de espera de los paquetes en cola. Este tipo de algoritmos puede mejorar el tiempo medio de retardo en la transmisión de los paquetes pero, por el contrario, no son capaces de maximizar la utilización de los recursos disponibles en el sistema. Otros autores, como por ejemplo Almajano y Pérez-Romero [8], definieron un algoritmo para garantizar la asignación de una determinada tasa mínima, pero sin preocuparse ni por la maximización del uso de los recursos ni por el retardo en la transmisión de los paquetes.

Por otro lado, el problema de la asignación óptima de recursos se puede formular como un problema de optimización de una determinada función de coste que caracteriza al problema y que está definida en un espacio N-dimensional. Los problemas de optimización N-dimensionales no pueden ser resueltos mediante heurísticos ni mediante técnicas convencionales de optimización. Sherif [9] y Papavassiliou [10] hicieron uso de algoritmos genéticos para encontrar una solución sub-óptima al problema, mientras que Xiao [3] propuso el uso de técnicas de programación lineal para su resolución. El problema es que tampoco se pueden utilizar estas técnicas matemáticas que, aún potentes, no son computacionalmente óptimas, ya que el tiempo de toma de decisión del sistema tiene que ser extremadamente rápido para garantizar el buen funcionamiento del sistema celular.

Por lo tanto, la mayor parte de los algoritmos propuestos en la literatura o son incapaces de satisfacer simultáneamente los requerimientos de QoS de los usuarios, es decir, máximo retardo y mínima tasa binaria, a la vez que se maximiza el *throughput* global del sistema, o no pueden operar en tiempo real.

Recientemente se ha propuesto el uso de técnicas en tiempo real basadas en redes neuronales de Hopfield, o *Hopfield Neural Networks* (HNN), para distribuir equitativamente los recursos y maximizar su uso [12]-[15]. El proceso de optimización de las HNN se basa en la minimización de una función de energía [16], de manera que, con una descripción adecuada, se puede conseguir que dicha función represente el problema de asignación de recursos en redes UMTS. Una de las principales ventajas de las técnicas neuronales es su alta velocidad computacional obtenida gracias a su implementación *hardware*, lo que resulta especialmente importante para aplicaciones



industriales. Por ejemplo, un sistema HNN puede obtener una solución en pocos microsegundos [12], lo que permite tomar una decisión trama a trama sobre la distribución de los recursos entre los usuarios, al menos en las tecnologías móviles actuales.

Una de las desventajas de las redes Hopfield es su posible mal funcionamiento cuando la función de energía no es convexa en el hiper-cubo bajo estudio [17]. Esta condición provoca ciertos problemas adicionales relacionados con la convergencia a mínimos locales no deseados. Estos problemas se pueden resolver fácilmente mediante un ajuste minucioso de los parámetros que controlan la función de energía, penalizando adecuadamente los estados no deseados [15]. Además, en la literatura se pueden encontrar otras soluciones para escapar de mínimos locales, permitiendo la convergencia hasta el mínimo global [18]. Un trabajo reciente de Ahn y Ramakrishna [12] demuestra que, con un diseño adecuado, un esquema de asignación de recursos DRA basado en HNN puede ser una solución práctica al problema. En [13] García aplicó este modelo a la transmisión en enlace ascendente en una celda DS-CDMA (*Direct Sequence Code Division Multiple Access*) incluyendo un modelo realista de las pérdidas de propagación y la movilidad de los usuarios. De igual manera, en [14] se aplicó el mismo esquema pero en este caso en el enlace descendente.

Hasta la fecha, los trabajos basados en el uso de las redes Hopfield para la asignación dinámica de recursos han tratado de resolver el problema desde un punto de vista centrado en el *throughput*, es decir, maximizando únicamente el ancho de banda asignado en el sistema. Sin embargo, con este enfoque el problema del DRA se resuelve parcialmente, no garantizando la calidad de servicio global del usuario ya que, en el proceso de *scheduling*, no se ha tenido en cuenta el tiempo de respuesta del servicio.

En esta Tesis Doctoral se propone un esquema de asignación de recursos basado en el uso de las redes neuronales de Hopfield e inspirado en el trabajo de Ahn y Ramakrishna [12]. En la red neuronal se ha introducido una nueva técnica de control del retardo así como un mecanismo para reducir el nivel de interferencias del sistema. El algoritmo resultante, denominado *Delay-Centric HNN DRA* (DC-HNN-DRA), es

capaz de maximizar el uso de los recursos a la vez que minimiza el retardo medio de los paquetes transmitidos. La transmisión de los datos se hará considerando el canal más flexible de UMTS, el canal compartido DSCH. Finalmente se han comparado los resultados obtenidos con otras técnicas de gestión de recursos basadas también en la minimización del retardo. Además, en este capítulo se evalúa el efecto de considerar simultáneamente en el sistema distintos tipos de servicios. Para ello se han modelado distintos servicios tanto de tiempo real como de no tiempo real.

Es importante destacar que en este capítulo no se tendrá en cuenta el uso de la nueva arquitectura radio MNB, sino que se transmitirá considerando la arquitectura UMTS convencional. Será en el Capítulo 5 cuando se extenderán la formulación de la red Hopfield para incluir, además, la distribución de recursos entre las antenas del Multinodo B.

El capítulo se ha estructurado de la siguiente manera: en el apartado 4.2 se describe sucintamente las características del sistema considerado. El apartado 4.3 analiza las distintas técnicas de gestión de recursos incluidas en la tecnología UMTS, pormenorizando su implementación en el simulador de sistema utilizado. Tras una descripción matemática del funcionamiento de las redes neuronales, realizada en el apartado 4.4, el apartado 4.5 se centra en la formulación de la red Hopfield propuesta en esta Tesis Doctoral, evaluando sus prestaciones respecto a otros algoritmos de *scheduling* de referencia.

## 4.2 Modelo del Sistema

Un modelo de simulación de sistema, o también conocido como simulador de nivel de red, considera un conjunto de estaciones base y todos los usuarios asociados a las mismas. Se modela asimismo la señal recibida por cada usuario, así como las interferencias de otros usuarios y el ruido, teniendo en cuenta las pérdidas de propagación, el efecto del desvanecimiento lento del canal, o *shadowing*, y la movilidad y generación de tráfico de los usuarios. Se obtiene, combinando toda esta

información, el nivel de calidad de señal experimentado por cada usuario, medido como la relación entre la señal útil y la señal interferente, SIR. Para realizar esta medida se deberá tener en cuenta la configuración de la red y los algoritmos de gestión de recursos radio empleados en el sistema.

Finalmente, se traduce la calidad de señal, expresada como un cierto nivel de *CIR* (*Carrier to Interference Ratio*) o  $E_b/N_0$  (energía media de bit respecto a densidad espectral de ruido más interferencia), a valores de BER (*Bit Error Rate*) o FER (*Frame Error Rate*). Para ello, se han de emplear los resultados obtenidos en la caracterización del enlace radio realizada mediante simulaciones de enlace, como las mostradas en el Capítulo 3. De esta manera, se puede obtener una medida de la calidad de servicio proporcionada al usuario y de la capacidad del sistema.

Dentro de las simulaciones a nivel de sistema pueden distinguirse dos tipos de simulaciones, en función de que se modele o no la evolución temporal del sistema. Según esta clasificación, las simulaciones pueden ser dinámicas o estáticas.

Las simulaciones dinámicas modelan la evolución del sistema a lo largo del tiempo, empleando una determinada resolución temporal. Para estudiar el funcionamiento de los algoritmos de gestión de recursos es fundamental evaluar el efecto que una determinada decisión tiene sobre la calidad de servicio experimentada por el usuario. Por eso, para analizar la gestión de recursos de las redes MNB, se ha considerado un modelo de simulación dinámico cuyo paso temporal es fijo y se corresponde con un *slot* UMTS.

Con el objetivo de evaluar no solamente la capacidad y calidad de servicio del sistema MNB, sino también el comportamiento del algoritmo de *scheduling* propuesto, en esta Tesis Doctoral se ha desarrollado completamente un simulador dinámico programado en Matlab y C. La descripción de los detalles y modelos utilizados en el simulador se han recogido en el Anexo B, para así facilitar la lectura de este capítulo en el que se va a describir en profundidad la implementación de los distintos algoritmos de gestión de recursos, centrandó la atención sobre todo en el mecanismo de *scheduling*, tan importante para las prestaciones finales del sistema UMTS.

## 4.2.1 Métricas de calidad de servicio

En la sección de resultados de este capítulo y los dos capítulos siguientes se van a presentar una serie de métricas de calidad, a partir de las cuales se podrán comparar las prestaciones de los distintos algoritmos de gestión de recursos comparados. Este apartado tiene por objetivo describir cada una de las métricas utilizadas.

- **Potencia total consumida:** es la suma de la potencia transmitida por todas las antenas que conforman el Nodo B. Esta potencia incluye también la potencia transmitida en los canales comunes de control, como por ejemplo el canal piloto.
- **Número de usuarios activos o en cola del sistema:** es el número de usuarios que el sistema tiene que considerar en el reparto de recursos, o bien porque tienen paquetes en la cola de entrada o bien porque presentan un estado de actividad de voz.
- **Número de usuarios servidos por el sistema:** representa el número de usuarios a los que efectivamente se ha asignado recursos.
- **Tasa binaria total del sistema:** es la suma de las tasas binarias asignadas a todos los usuarios activos en el sistema. Esta tasa binaria se considera no como tasa binaria física sino como tasa binaria a nivel de aplicación.
- **Tasa binaria efectiva del sistema:** por varios motivos la tasa binaria asignada a un usuario puede no ser utilizada efectivamente. Si, por ejemplo, el usuario experimenta un mal estado del canal y no se reciben correctamente los datos transmitidos esa información tendrá que retransmitirse, por lo que se desperdiciará la capacidad asignada. Otro motivo por el que se reduce la tasa efectiva es porque un usuario vacíe completamente su cola de transmisión. El resto de recursos asignados a ese usuario permanecerán inutilizados hasta que se proceda de nuevo al reparto de los mismos.
- **Retraso por objeto:** siempre que se genera una página web o se inicia una descarga FTP o se genera una nueva trama de vídeo para un usuario de vídeo

conferencia se inicializa un temporizador. Una vez el objeto ha sido enviado íntegramente al usuario y éste lo ha recibido correctamente, entonces se mide el tiempo que el objeto ha permanecido en el Nodo-B. Éste tiempo es el retraso del objeto.

- Probabilidad de error de bloque de los usuarios de voz: en principio, siempre que un usuario de voz está activo debe poder transmitir, ya que se trata de un usuario en tiempo real. Sin embargo, puede que las condiciones de propagación sean tan desfavorables o el nivel de interferencia tan elevado que ni siquiera con la tasa binaria más baja se pueda establecer una conexión satisfactoria con el usuario. En esos casos se supone que los bloques transmitidos por el usuario no llegan correctamente a su destino y que, por tanto, se han perdido. La probabilidad de error de bloque recoge el porcentaje de bloques no recibidos correctamente respecto al número de bloques transmitidos.
- Probabilidad de eliminación de tramas de vídeo: al ser la vídeo conferencia un servicio en tiempo real, el Nodo-B mantiene la trama actual de vídeo destinada a un usuario hasta que ésta se transmite o hasta que llega una nueva trama. En este último caso la trama no transmitida se elimina puesto que no tiene sentido almacenarla. La probabilidad de eliminación de tramas de vídeo recoge la proporción de tramas eliminadas respecto al total de tramas transmitidas correctamente.
- Probabilidad de *Outage*: una vez asignados los recursos, los usuarios continúan su movimiento a lo largo de la celda. Las condiciones de propagación pueden variar bruscamente, sobre todo si el usuario se aleja de la estación base a alta velocidad, por lo que puede que el control de potencia en bucle rápido no sea capaz de seguir las variaciones de las pérdidas de propagación. Un usuario se dice que está en *outage* cuando la calidad de señal recibida no es la requerida. La probabilidad de *outage* recoge la proporción de *slots* UMTS en que un usuario activo presenta mala calidad.

## 4.3 Algoritmos de Gestión de Recursos

Los sistemas de tercera generación, como UMTS, han sido diseñados para ofrecer a los usuarios una gran variedad de servicios, cada uno de ellos con requerimientos distintos. La técnica de acceso WCDMA empleada en la UTRAN ofrece la versatilidad necesaria para conseguir este objetivo. No obstante, para aprovechar las múltiples posibilidades que brinda el acceso radio de UMTS se requiere el uso de un conjunto de técnicas que permitan gestionar adecuadamente los recursos del sistema. Esto es lo que se conoce como gestión de recursos radio. El principal objetivo de las técnicas de gestión de recursos es asegurar una utilización óptima de los recursos limitados con los que cuenta la red de acceso, con la finalidad de satisfacer los requerimientos de los distintos usuarios, así como de obtener las máximas prestaciones del sistema.

Gran parte de las técnicas de gestión de recursos no han sido estandarizadas en las especificaciones de UMTS. Con ello se persigue que los distintos fabricantes y operadores puedan desarrollar y optimizar sus propios algoritmos de gestión y así obtener una importante ventaja competitiva respecto al resto de agentes del mercado.

En una red tan compleja como UMTS pueden identificarse gran cantidad de mecanismos relacionados con la gestión de recursos radio. En los siguientes apartados se describen cada uno de estos mecanismos de gestión, incidiendo en la implementación realizada de los mismos en el simulador de sistema utilizado en esta Tesis Doctoral.

### 4.3.1 Control de potencia

Tal y como se explica en las especificaciones, el control de potencia es una técnica fundamental para el buen funcionamiento del sistema UMTS. Existen dos procedimientos de control de potencia en UMTS: el control de potencia en bucle abierto, responsable de asignar una potencia inicial a la comunicación, y el control de potencia en bucle cerrado, que trata de asegurar la satisfacción de una determinada

relación señal a ruido objetivo. Ambos mecanismos de control de potencia se ha considerado en el simulador. A continuación se procede a la descripción de la implementación de los mismos.

#### **4.3.1.1 Control de potencia en bucle abierto**

La implementación del control de potencia en lazo abierto se realizará de la siguiente manera: cada vez que se reasignan recursos se debe iniciar la comunicación con unas condiciones diferentes, por lo que el algoritmo de gestión de recursos debe estimar la potencia inicial transmitida a cada usuario. Esta potencia se puede conocer utilizando la ecuación  $E_b/N_0$  correspondiente al servicio y a las condiciones particulares de propagación. Todos los términos incluidos en dicha expresión son conocidos por el sistema gracias a la información enviada por el usuario en enlace ascendente dentro de los *Measurement Reports*. De la expresión de la  $E_b/N_0$ , y asumiendo un valor objetivo de ésta, se podrá despejar la potencia asignada al usuario. Éste valor de potencia será con el que se iniciará la comunicación con el usuario.

#### **4.3.1.2 Control de potencia en bucle cerrado**

Cada cierto tiempo el algoritmo de *scheduling*, haciendo uso de la información obtenida del control de potencia en bucle abierto, establecerá la asignación de recursos para el siguiente periodo. Esta asignación de recursos permanecerá invariable durante todo este periodo, mientras que las condiciones del canal y de las interferencias estarán en constante cambio, dependiendo de la movilidad de los usuarios y de las fluctuaciones en la transmisión. El control de potencia en lazo cerrado tendrá que hacer frente a estos cambios.

Cada *slot* se actualizará, considerando el nivel de interferencia experimentado en el *slot* anterior, la potencia asignada a cada usuario, subiendo o bajando dicha potencia 1 dB, tal y como establecen las especificaciones. Si el nivel de potencia es suficiente

como para garantizar el valor de  $E_b/N_0$  objetivo, entonces se considerará que la trama se ha recibido correctamente. Por el contrario, si no se satisface el objetivo de calidad el paquete quedará descartado, debiendo proceder a su retransmisión.

Por su parte, el control de potencia en bucle externo se implementará en el simulador al considerar un valor de  $E_b/N_0$  objetivo diferente para cada uno de los servicios y de las tasas binarias asignadas. La movilidad no afectará a este valor de calidad objetivo que será constante durante toda la comunicación.

Para más detalles teóricos acerca del control de potencia en UMTS se remite al lector a la extensa bibliografía disponible al respecto.

### **4.3.2 Selección inicial de celda**

Sin considerar la nueva arquitectura radio Multinodo B, y al no existir más que un transmisor activo en el área bajo estudio, la selección de celda es trivial, es decir, la celda central del sistema será la encargada de transmitir a todos los usuarios.

### **4.3.3 Algoritmo de traspaso**

El mecanismo de traspaso también está especificado por el 3GPP, por lo que no se van a dar más detalles sobre su funcionamiento.

Aunque en el escenario de simulación se considera un anillo de celdas interferentes, éstas no colaboran con la celda central en procesos de *handover*. Además, el modelo de movilidad asegura que los usuarios nunca abandonan la celda central, por lo que definitivamente no se va a considerar ningún mecanismo de traspaso entre celdas.



Por último, es importante señalar que además de no modelarse el traspaso de usuarios entre celdas, tampoco se han considerado procedimientos de transmisión en diversidad mediante *soft handover*.

#### 4.3.4 Control de admisión

El control de admisión de llamadas, o CAC (*Call Admission Control*) es la función de gestión de recursos encargada de controlar el acceso de los usuarios al sistema. Un algoritmo de control de admisión debe, ante la llegada de una nueva llamada, estimar el impacto que tendría la admisión de dicha llamada sobre la carga del sistema y tomar la decisión de admitirla o bloquearla basándose en un criterio determinado.

En el simulador implementado se introduce un número fijo de usuarios de cada uno de los servicios considerados, estando cada usuario activo o inactivo en función de su patrón de tráfico (para más detalles sobre los patrones de tráfico considerados léase el Anexo B). Dado que los usuarios ya se encuentran dentro del sistema, no es necesario implementar ningún algoritmo de control de admisión de llamadas.

Por otro lado, la decisión de qué usuarios van a transmitir en cada momento será realizada por el algoritmo de cronoejecución, o *packet scheduler*, descrito más adelante en el apartado 4.5.

#### 4.3.5 Control de congestión

Una tarea importante en la gestión de recursos radio es asegurar que el sistema no se sobrecargue y se mantenga estable. Si el sistema ha sido planificado adecuadamente y la gestión de recursos funciona correctamente, las situaciones de sobrecarga se producirán con poca frecuencia. No obstante, en los casos en los que no se pueda evitar la sobrecarga de la red es necesario contar con mecanismos que permitan devolver la estabilidad al sistema. El control de carga, o control de congestión, engloba los

algoritmos y mecanismos encargados de resolver las situaciones de sobrecarga que pueden producirse, devolviendo al sistema al nivel de carga objetivo definido en la fase de planificación.

Existen multitud de acciones que pueden realizarse para reducir la carga de la red ante una situación de congestión. A continuación se listan algunas de ellas:

- Denegar las peticiones de incremento de potencia procedentes del móvil en el control de potencia en bucle cerrado en el enlace ascendente o reducir la  $E_b/N_0$  objetivo utilizada por el control de potencia en bucle cerrado en el enlace descendente. Estas dos acciones se ejecutan en la estación base con una periodicidad de *slot*, es decir con una frecuencia de 1.5 kHz, por lo que presentan una velocidad de actuación muy rápida. Además, este mecanismo de control de carga permite limitar los recursos asignados a algunas comunicaciones, a costa de sacrificar su calidad, con lo que se consigue una priorización dinámica entre distintos servicios e incluso entre diferentes comunicaciones pertenecientes al mismo tipo de servicio.
- Reducir el *throughput* de los servicios de datos. Con esta acción se consigue liberar parte de los recursos asignados a un usuario a costa de un mayor retardo en la transmisión.
- Reducir la tasa binaria de los servicios en tiempo real, por ejemplo utilizando el codificador de voz AMR.
- Traspasar comunicaciones a otra portadora WCDMA o a otro sistema, por ejemplo GSM. Con estas acciones se consigue combatir las situaciones de congestión, balanceando la carga entre las distintas portadoras asignadas a un operador o derivando parte de la carga a otro sistema menos congestionado.
- Interrumpir llamadas en curso de una manera controlada. Esta acción supone un gran perjuicio para los usuarios, por lo que sólo debe utilizarse como último recurso cuando la carga del sistema se mantenga elevada incluso después de haber llevado a cabo otras acciones de control menos drásticas.

#### 4.3.5.1 Implementación del control de congestión

Muchas de las acciones llevadas a cabo por el algoritmo de control de congestión se realizan automáticamente en el algoritmo de gestión de recursos propuesto en esta Tesis Doctoral (ver apartado 4.5). Por ejemplo, el algoritmo de gestión de recursos se preocupa de no exceder nunca la potencia máxima disponible en el sistema y también puede reducir la tasa binaria asignada a los usuarios interactivos para poder dar servicio a usuarios de mayor prioridad. De igual manera, la codificación de los usuarios de voz también puede adaptarse para reducir el nivel de interferencia del sistema en caso de que se detecte congestión.

Sin embargo, una vez asignados los recursos el sistema evoluciona en función de las condiciones de propagación de los usuarios. A pesar de que el control de potencia se ajuste a dichos cambios, es posible que el sistema, que inicialmente estaba controlado, entre en una fase de congestión.

Dicha congestión se detecta por dos motivos principalmente: o porque la potencia requerida por un usuario supera la potencia máxima transmitida por usuario o bien porque la estación base se queda sin recursos. En el primer caso, directamente al usuario se le niega un incremento de la potencia por encima de la potencia máxima permitida por usuario, por lo que probablemente la calidad requerida por el usuario no sea la necesaria y, por tanto, se declare al usuario en situación de *outage*. Si lo que pasa es que la situación actual del sistema requeriría superar la potencia máxima disponible en la estación base, entonces lo primero que se hace es seleccionar uno a uno los usuarios interactivos que mayor consumo de potencia estén realizando y dejar de transmitirles temporalmente en ese *slot*, hasta que el sistema esté controlado. Si incluso cuando los usuarios interactivos no transmiten nada el sistema sobrepasa la potencia máxima, entonces se iniciaría el mismo proceso para los usuarios en tiempo real, es decir, para los usuarios de vídeo conferencia y de voz, por este orden.

### **4.3.6 Cronoejecución o *scheduling***

El sistema UMTS soporta una mayor capacidad y velocidad de transmisión en comparación con los sistemas de comunicaciones móviles precedentes. Este aumento en la velocidad de transmisión ha permitido el desarrollo de nuevos servicios móviles, como la vídeo telefonía, el acceso a Internet o la transmisión de vídeo bajo demanda. Fruto del imparable crecimiento que ha experimentado la transmisión de paquetes sobre los sistemas de comunicaciones móviles, ha surgido un creciente interés, tanto por parte de las grandes compañías de telefonía móvil como de los fabricantes, por desarrollar algoritmos eficientes de gestión adaptativa de paquetes, también denominada cronoejecución o *scheduling*.

Este apartado no pretende describir el algoritmo de gestión de paquetes propuesto en esta Tesis Doctoral (ver sección 4.5) sino más bien profundizar, en primer lugar, en el concepto general de cronoejecución o gestión adaptativa de paquetes, para pasar posteriormente a analizar los distintos métodos de transmisión de paquetes estandarizados en el sistema UMTS. Los algoritmos de *scheduling* desarrollados en esta Tesis Doctoral se basan principalmente en el uso de un canal de datos flexible, por lo que se ha seleccionado como canal por defecto el DSCH, estandarizado en UMTS. Por ello, a lo largo de esta explicación se prestará especial atención a este canal de transporte y sus peculiaridades.

#### **4.3.6.1 Concepto de cronoejecución de paquetes**

Tal y como se explicó anteriormente, los servicios RT se caracterizan porque los datos se generan de forma continua y deben ser entregados al receptor por orden de generación y con un retardo máximo garantizado. Para ofrecer una determinada calidad de servicio (QoS) a los servicios RT es necesario reservar una serie de recursos fijos durante toda la comunicación. Se trata, por tanto, de servicios que habitualmente se ofrecen por conmutación de circuitos, es decir, reservando un canal exclusivo para cada

comunicación. Por otro lado, los servicios NRT se caracterizan porque el tráfico se genera a ráfagas y porque los paquetes de información han de llegar sin errores pero sin un orden establecido al receptor. Por todo esto, los servicios NRT se suelen ofrecer por conmutación de paquetes, es decir, varias comunicaciones comparten el mismo canal y una entidad externa decide quién hace uso del canal en cada momento. Si, por el contrario, los servicios NRT se ofrecieran por conmutación de circuitos habría una clara ineficacia en la utilización de los recursos puesto que, al tratarse de tráfico a ráfagas, durante grandes periodos de tiempo no se estaría haciendo uso de los recursos reservados.

El algoritmo de cronoejecución de paquetes, o *packet scheduler*, es el encargado de gestionar el acceso al canal de todos los servicios de paquetes. La cronoejecución es, sin duda, uno de los algoritmos básicos en la gestión de recursos radio, o RRM (*Radio Resource Management*), dentro de los sistemas de comunicaciones móviles. El algoritmo de *scheduling* decide cuándo se inicia una transmisión de paquetes así como su velocidad. La cronoejecución tiene, por tanto, un fuerte impacto en las prestaciones de la red y es uno de los algoritmos fundamentales de gestión de recursos radio en los sistemas móviles que admiten conmutación de paquetes.

Las tareas más importantes del algoritmo de *scheduling* son:

- Determinar los recursos disponibles para servicios de paquetes.
- Repartir los recursos disponibles entre los usuarios.
- Monitorizar la asignación de dichos recursos.
- Monitorizar la carga del sistema, es decir, la relación entre recursos consumidos y recursos disponibles en la celda.
- Ejecutar, cuando corresponda, acciones de control de la congestión para los servicios de paquetes.

Una parte de la cronoejecución de paquetes se realiza en el terminal móvil y otra en la red. El terminal móvil es el encargado de la gestión de los paquetes en el enlace

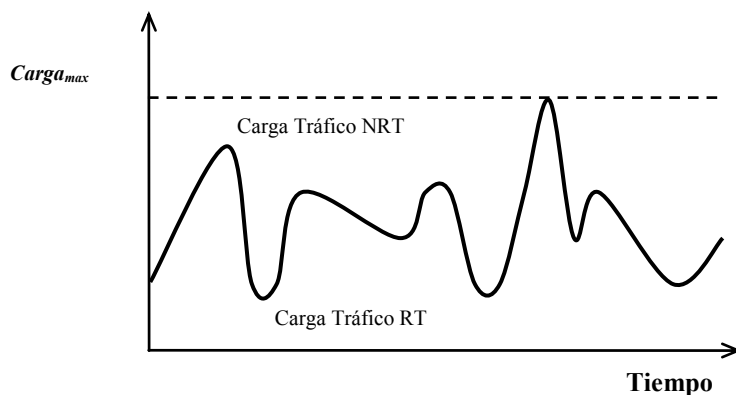
ascendente mediante un proceso FIFO (*First In First Out*). Por otro lado, el móvil provee a la red de todas las medidas necesarias para realizar la cronoejecución en enlace descendente. La parte realizada por la red en el *packet scheduler* se encarga de estimar la capacidad disponible para servicios de paquetes y distribuirla entre los usuarios activos.

Desde el punto de vista de la gestión de recursos radio, el sistema se encontrará en un punto de funcionamiento óptimo cuando la carga del sistema sea igual al nivel máximo tolerable, ni más, lo cual provocaría una degradación de la calidad de servicio por bloqueo de llamadas o pérdidas de paquetes, ni menos, lo que implicaría malgastar recursos. El objetivo del algoritmo de cronoejecución es utilizar eficientemente toda la capacidad disponible para servicios de paquetes, manteniendo la calidad necesaria en las conexiones en tiempo real, como por ejemplo, comunicaciones de voz.

La Figura 4.1 muestra la relación directa que existe entre los recursos disponibles para la transmisión de paquetes o servicios NRT y los recursos consumidos por los servicios en tiempo real, prioritarios para el sistema.

Es importante destacar que el tráfico de paquetes que gestiona el algoritmo de *scheduling* es normalmente un tráfico no garantizado, es decir, sin restricciones de retardo y sin una tasa mínima garantizada, en la medida en que la capacidad disponible tampoco está garantizada. Debido a esta restricción, en general no se podrá transmitir toda la información de datos disponible en un determinado momento, por lo que habrá que establecer criterios de priorización en las transmisiones. Las informaciones referentes a las diferentes clases de servicio (*interactive, background*) o grados de servicio (*premium*), serán utilizadas en la definición de dichas prioridades.

El objetivo principal de la gestión de paquetes en los sistemas de comunicaciones móviles es maximizar la capacidad del sistema (en términos de número de usuarios activos soportados por la red) proporcionando los niveles de cobertura y calidad de servicio deseados, para una cantidad de recursos radio e infraestructura desplegada dadas.



**Figura 4.1 Ilustración del papel de la cronoejecución en la carga**

Los mecanismos de cronoejecución en la red no están estandarizados, por lo que son un elemento diferenciador importante entre equipos de distintos fabricantes. Con ello se persigue que los fabricantes puedan desarrollar y optimizar sus propios algoritmos con el objetivo de obtener una ventaja competitiva respecto a los usuarios finales. Es por tanto de suma importancia en la optimización de los sistemas de comunicaciones móviles actuales el desarrollo de potentes y eficaces algoritmos de *scheduling*.

#### 4.3.6.2 Cronoejecución en UMTS

El *packet scheduler* en UMTS se encuentra, según el estándar, en el RNC y recibe medidas de la carga de la red de las estaciones base y medidas del tráfico en enlace ascendente de los móviles. El *packet scheduler* trabaja periódicamente en intervalos temporales con valores típicos entre 10ms y 1s [19]. El *packet scheduler* utiliza la siguiente información:

- Potencia total utilizada por las estaciones base o Nodos B. Esta información representa una medida de la carga utilizada en enlace descendente.

- Capacidad utilizada por las conexiones en tiempo real. Esta información se puede obtener a partir de la estimación de la carga en enlace ascendente basada en promediado estadístico.
- Umbral máximo para el nivel de carga establecido durante la fase de planificación. Este parámetro define el nivel de interferencia máximo que puede existir en la celda para que las conexiones en tiempo real no vean degradada su QoS.
- Peticiones de incremento de tasa binaria de los móviles.

Los Nodos B proporcionan periódicamente al RNC información sobre la potencia total utilizada así como las potencias dedicadas a cada conexión. Con esta información el *packet scheduler* puede estimar la potencia total utilizada para tráfico no controlable, que consiste en tráfico en tiempo real e interferencias intercelulares. Esta parte de la interferencia no puede ser afectada por el *packet scheduler*. La capacidad restante es compartida por los usuarios de paquetes activos.

#### **4.3.6.3 Métodos de transmisión de paquetes en UMTS**

Existen tres tipos de canales de transporte en UMTS que pueden utilizarse para transmitir información de paquetes: canales comunes, canales dedicados y canales compartidos. A continuación se describen las propiedades y ventajas de cada tipo de canal para transmitir información de paquetes.

##### ***Canales Comunes (RACH/FACH/CPCH)***

Los canales comunes son el RACH (*Random Access Channel*) y el CPCH (*Common Packet Channel*) en el enlace ascendente o *uplink* y el FACH (*Forward Access Channel*) en enlace descendente o *downlink*. El RACH y el FACH suelen ser utilizados



para transmitir información de señalización, pero también datos. Normalmente sólo suele haber unos pocos canales RACH y FACH por sector y a veces incluso sólo uno.

La principal ventaja de los canales comunes es que no se realiza un proceso de establecimiento de la conexión, por lo que apenas hay un tiempo de espera antes de iniciar la transmisión.

Como los canales comunes son unidireccionales no tienen un canal de retorno, por lo que en ellos no se puede utilizar el control de potencia rápido característico de UMTS y se transmite simplemente con una potencia fija determinada por el bucle abierto del control de potencia. Además, en estos canales tampoco se pueden realizar traspasos. Por todo ello, el rendimiento a nivel de enlace de los canales comunes es peor que el de los canales dedicados, por lo que introducen más interferencias.

Los canales comunes FACH y RACH son adecuados para transmitir paquetes IP pequeños, como por ejemplo durante la fase de establecimiento de la conexión de TCP, ya que el establecimiento de la conexión mediante TCP implica la transmisión de varios paquetes IP pequeños, de tamaño alrededor de los 40 bytes.

El CPCH es una versión mejorada del canal RACH en el que se puede utilizar control de potencia rápido pero no *soft handover*. Además, se pueden asignar hasta 64 tramas, es decir 640 ms, a un usuario, con lo que se puede transmitir más información que en el RACH. El inconveniente del CPCH es que el tiempo de establecimiento de la conexión es mayor que el del RACH.

### ***Canales Dedicados (DCH)***

A pesar de que la transmisión de paquetes siempre es más eficiente mediante conmutación de paquetes, UMTS también permite la posibilidad de transmitir servicios NRT mediante conmutación de circuitos haciendo uso de canales dedicados DCH (*Dedicated Channel*). Los canales dedicados son canales bidireccionales tanto en *uplink* como en *downlink*, y permiten realizar el control de potencia rápido y *soft handovers*, por lo que generan menos interferencia que los canales comunes. El principal inconveniente de los canales dedicados es que es necesario realizar un

proceso de establecimiento de la conexión, lo cual implica un mayor retardo en el acceso comparado con el de los canales comunes.

Los canales dedicados permiten tasas binarias desde unos pocos kbps hasta 384 kbps. La tasa binaria de un canal dedicado no es un parámetro fijo, sino que puede variar a lo largo de la conexión. Una vez se termina de transmitir toda la información en cola, el usuario mantiene el canal dedicado unos pocos segundos antes de liberar el canal para que pueda ser asignado a otro usuario. Esto implica que los canales dedicados no sean eficientes para tráfico a ráfagas, sino que su uso está recomendado para transmitir cantidades de información medias o grandes.

Dada la alta complejidad del *packet scheduler*, en las primeras versiones comerciales de los RNC-UMTS no se incluía su implementación, por lo que los servicios NRT siempre se soportaban mediante canales DCH. Desde mediados de 2004 las nuevas RNCs incorporan ya los mecanismos de cronoejecución para su aplicación a los canales compartidos.

### ***Canales Compartidos (DSCH)***

En el enlace descendente se puede agotar rápidamente el número de códigos de canalización libres dentro de un mismo código de aleatorización, puesto que el árbol de códigos se comparte entre todos los usuarios activos, los canales comunes y los canales de control.

Este problema se podría solucionar de dos maneras: la primera solución sería utilizar más de un código de aleatorización por celda, pero esto ocasionaría unas pérdidas adicionales debido al incremento de la interferencia. La segunda solución es emplear un canal compartido, lo que reduce el consumo de recursos de códigos en el enlace descendente. Aunque al compartir recursos es imposible garantizar al 100% la disponibilidad del canal, esta técnica posee un gran rendimiento para servicios de paquetes. Por ello, en UMTS se estandarizó el canal DSCH (*Downlink Shared Channel*) y su posterior evolución en HSDPA. En esta Tesis Doctoral sólo se ha

considerado la transmisión mediante DSCH, dejando la evolución a HSDPA como una línea futura de trabajo.

El DSCH se utiliza para transmitir tráfico de paquetes a ráfagas en el enlace descendente. La idea es compartir un único canal físico, es decir, un código, entre muchos usuarios, multiplexándolos temporalmente. Si se utilizaran canales dedicados para este tipo de tráfico, todos los usuarios requerirían un código que vendría determinado por sus tasas binarias máximas. Cuando un usuario termina su información a transmitir, los recursos del DSCH se asignan a otro usuario inmediatamente, incrementando la eficiencia respecto a los canales dedicados.

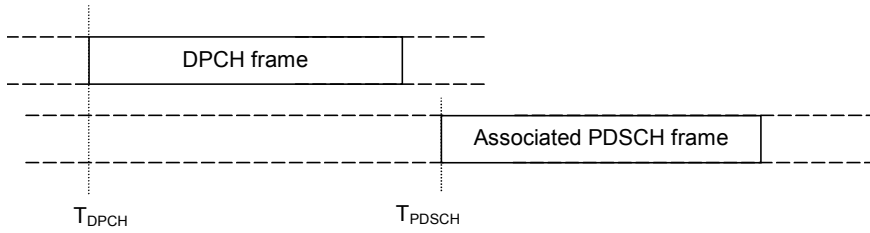
El canal de transporte DSCH se mapea siempre sobre el canal físico PDSCH (*Physical Downlink Shared Channel*).

Cuando se decide transmitir a un usuario mediante DSCH se establece un canal dedicado DCH de baja tasa binaria, para controlar el uso del canal compartido. Por el canal DCH se transmite el canal de control físico, la señalización y los comandos de control de potencia. Además, el canal DCH lleva un indicador para informar al usuario de que la siguiente trama DSCH debe ser decodificada, así como la información necesaria para tal decodificación y el código de ensanchamiento del DSCH. El envío de esta información se puede realizar de dos maneras:

- Usando el campo del TFCI trama a trama
- Mediante señalización de los niveles superiores

Como es necesario que el usuario primero decodifique la información transmitida en el canal físico dedicado para después ser capaz de decodificar el PDSCH, el canal DPCH se transmite siempre antes. En concreto, el estándar [20] define que la transmisión del canal PDSCH se realizará al menos tres *slots* después de finalizar la transmisión del correspondiente canal dedicado. La Figura 4.2 muestra esta relación temporal.

Uno de los grandes inconvenientes del DSCH es que no admite *soft handovers*. Por eso, la utilización de la nueva arquitectura MNB permitirá solventar este problema.



**Figura 4.2 Relación temporal entre el canal dedicado y el compartido [20]**

La Tabla 4.1 resume los distintos tipos de canales disponibles para la transmisión de paquetes, al igual que las características de cada uno de ellos.

Canal de Transporte	DCH	RACH	FACH	CPCH	DSCH
Tipo de canal	Dedicado	Común	Común	Común	Compartido
Dirección	UL/DL	UL	DL	UL	DL
Código empleado	Función de la tasa max.	Asignación fija	Asignación fija	Asignación fija	Asignación fija
PC	FPC	OPC	OPC	FPC	FPC
SHO	Si	No	No	No	No
Volumen de tráfico	Medio o alto	Bajo	Bajo	Bajo o Medio	Medio o alto
Idoneidad para tráfico a ráfagas	Pobre	Buena	Buena	Buena	Buena
Tiempo de establecimiento	Alto	Bajo	Bajo	Bajo	Bajo
Prestación	Alta	Baja	Baja	Media	Media o Alta
Soportado en las primeras versiones	Si	Si	Si	No	No

**Tabla 4.1 Comparativa de los distintos canales de transporte utilizados para la transmisión de paquetes**

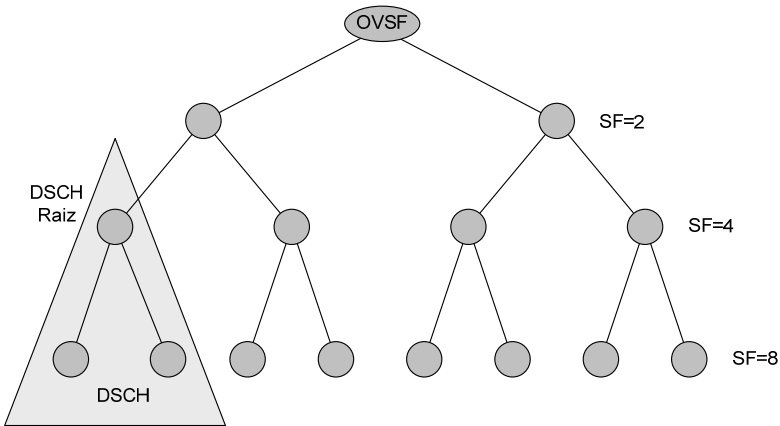
#### 4.3.6.4 Gestión de paquetes mediante DSCH

En sentido ascendente, el acceso al medio físico para la transmisión de servicios NRT mediante conmutación de paquetes se realiza por competición, por lo que no se aplica ningún algoritmo de cronoexecución. En el enlace descendente, sobre todo si se utiliza el canal DSCH, sí que se hace uso de estos algoritmos.

Todos los usuarios NRT están vinculados a uno o varios canales DSCH. Cada canal DSCH habilitado estará por tanto compartido por múltiples usuarios. El *packet scheduler* se encarga de seleccionar el uso que hacen los usuarios de cada canal DSCH. Teniendo en cuenta las anteriores restricciones de potencia y asignación de códigos, el algoritmo de cronoexecución asigna los canales DSCH a los distintos usuarios siguiendo dos principios básicos: la prioridad de los usuarios y la disponibilidad de recursos. La priorización es un proceso fundamental en el que los usuarios se ordenan según un determinado criterio que podrá tener en cuenta la tasa media asignada a cada usuario por contrato, el tamaño de los datos pendientes de transmitir o el retardo acumulado por estos datos. Una vez ordenados los usuarios, el siguiente paso consiste en decidir qué tasa de transmisión va a reservarse a cada comunicación en el siguiente periodo de *scheduling* (desde 10 ms hasta 1s). Recorriendo la lista de usuarios priorizados se van asignando recursos, comprobando que las limitaciones en códigos y en potencia se satisfacen. Una vez se supere alguna de las dos limitaciones se dará por finalizada la asignación de recursos.

#### 4.3.7 Gestión de códigos

Como en cualquier sistema CDMA, en UMTS los distintos usuarios se multiplexan utilizando un conjunto finito de códigos ortogonales. En UMTS estos códigos se denominan códigos OVSF (*Orthogonal Variable Spreading Factor*) y se generan mediante una estructura en forma de árbol como la que se muestra en la Figura 4.3.



**Figura 4.3 Rama de código para el canal DSCH**

Los árboles OVSF tienen la propiedad de que dos o más códigos pertenecientes a ramas distintas son ortogonales, mientras que códigos que pertenecen a la misma rama no guardan dicha ortogonalidad. De esta manera, se puede disponer de 4 códigos con factor de ensanchado (SF) 4, 8 con SF=8, etc.

Normalmente, para la transmisión de paquetes en UMTS se asigna una rama de código OVSF para el canal DSCH, como se aprecia en la Figura 4.3. La elección de qué ramas de código OVSF se asignan para DSCH se realiza de forma dinámica en función de la carga de los distintos servicios. Normalmente los sistemas reales asignan un número determinado de ramas de SF 8. Por ejemplo, si no hubiera ninguna comunicación en tiempo real se podría asignar para la transmisión de paquetes hasta 7 de las 8 ramas SF 8, ya que la octava está reservada para la transmisión de los canales comunes necesarios para el funcionamiento de la red (señal de *broadcast*, canal piloto etc.) y para los canales dedicados asociados a cada usuario DSCH.

Dado un número  $N$  de conexiones simultáneas, en [21] se establece una condición necesaria para que esa distribución de recursos sea realizable, es decir, para que se pueda afirmar que hay suficientes códigos para permitir esa asignación.

$$\sum_{i=1}^N \frac{1}{SF_i} \leq \frac{1}{SF_{Raiz}} \tag{4.1}$$

Según esta expresión, si se tiene en cuenta todo el árbol de códigos el  $SF_{Raiz} = 1$  por lo que

$$\sum_{i=1}^N \frac{1}{SF_i} < 1 \quad (4.2)$$

Si, por otro lado, se consideran dos partes del árbol de códigos diferenciadas (los códigos reservados a  $M$  transmisiones de conmutación de circuitos DCH y los códigos reservados para  $L$  conexiones por conmutación de paquetes con DSCH) se deberá cumplir:

$$\begin{aligned} \sum_{i=1}^L \frac{1}{SF_i} + \sum_{i=1}^M \frac{1}{SF_i} &< 1 \\ \sum_{i=1}^L \frac{1}{SF_i} &< 1 - \sum_{i=1}^M \frac{1}{SF_i} \end{aligned} \quad (4.3)$$

O lo que es lo mismo, teóricamente todos los códigos no utilizados por los servicios RT podrán ser utilizados por los servicios NRT. En este sentido, en la asignación de códigos, el comportamiento de UMTS es el reflejado en la Figura 4.1.

Es también importante destacar cómo se realiza la asignación de códigos a un usuario concreto. Como se explicó anteriormente, cada canal DSCH asignado a un usuario lleva asociado un canal dedicado DCH. En el establecimiento del canal se especifica cuál es el código OVSF raíz y que subcódigos se pueden asignar. Por ejemplo, si se especifica que la raíz es un determinado código con SF 8, entonces al usuario se le puede asignar o bien todo el código o uno de los dos códigos SF 16 que de él derivan o uno de los cuatro SF 32, etc. Cada una de las combinaciones posibles se especifica en una tabla, de manera que, mediante el TFI enviado en el DCH, se marca qué código va a utilizar el usuario en el siguiente TTI [22].

## 4.4 Funcionamiento de las Redes Neuronales

### 4.4.1 Introducción

Los algoritmos recursivos o iterativos se caracterizan porque una nueva solución o iteración del algoritmo se calcula a partir de la actualización de los valores antiguos. Este comportamiento se emplea asimismo en los algoritmos recursivos de minimización, en los que, cada iteración, la solución se aproxima más a la solución óptima final. Algoritmos como el de Newton o sobre todo el *steepest-descent* [23] han mostrado su buen comportamiento en la localización del mínimo de una función en  $\mathbb{R}^N$ . Muchos problemas prácticos de optimización se pueden formular utilizando una expresión de segundo grado, o en su defecto se puede aproximar por el polinomio de Taylor de orden 2, y variables de decisión binarias, lo que confina el espacio de búsqueda a las esquinas de un hiper-cubo  $N$ -dimensional, problema de complejidad *NP-hard*. Este hecho motivó el uso de las redes neuronales como alternativa a los métodos clásicos de optimización en  $\mathbb{R}^N$ , ya que precisamente la neurona presenta la característica de tener acotada su salida.

Hopfield propuso el uso de redes neuronales realimentadas en problemas de optimización [24] donde, por primera vez, las salidas de las neuronas podrían tener valores continuos acotados entre un valor máximo y mínimo. A nivel *hardware*, la red neuronal recursiva propuesta por Hopfield es una red analógica compuesta por múltiples amplificadores operacionales interconectados mediante resistencias [25]. Cada neurona de la red, que representa cada una de las variables que definen el problema, se considera activa o inactiva si la salida analógica sobrepasa un determinado umbral o no. La estabilidad de esta red se garantiza siempre y cuando la ecuación de energía que describa el problema sea de Lyapunov [16], ya que de este modo se asegura la continuidad del gradiente de la función. A pesar de su estabilidad, las redes Hopfield presentan algunos problemas adicionales relacionados con la pobre calidad de la solución encontrada por el algoritmo. Estos problemas se solucionan



ajustando los parámetros de la función de energía para penalizar convenientemente los estados no deseados. Además, se han propuesto soluciones muy útiles para escapar de posibles mínimos locales en la evolución de las neuronas [18].

Las redes neuronales Hopfield se han utilizado para solucionar múltiples problemas prácticos (véase como ejemplo [11]-[15] aunque en la literatura se pueden encontrar muchos más), dada la rápida convergencia de estas redes a una solución válida. De este modo, las redes Hopfield pueden reemplazar el uso de heurísticos, uso que únicamente venía motivado por la necesidad de una rápida respuesta.

Por otro lado, la cronoejecución es en la actualidad un proceso fundamental dentro de la gestión de recursos radio de las redes móviles celulares para aprovechar al máximo la capacidad de cada tecnología, garantizando una determinada calidad de servicio QoS a todos los usuarios. En la literatura se han propuesto diversas técnicas de DRA pero, o no son lo suficientemente potentes, o no son capaces de obtener una solución óptima en tiempo real ya que, matemáticamente, el DRA es un problema de optimización NP-complejo. En [12] se postuló que un problema de tales dimensiones solamente podría ser resuelto, con un tiempo de resolución muy bajo, del orden de microsegundos, mediante la utilización de redes neuronales Hopfield (HNN). Además, en este relevante trabajo se demostró que era posible maximizar el uso de los recursos y garantizar un reparto equitativo de los mismos, minimizando así la probabilidad de bloqueo y terminación forzosa. A partir de este trabajo, varios grupos de investigación han iniciado un estudio más en profundidad de la aplicación de las redes Hopfield a DRA en UMTS [13]-[15]. Una continuación de este trabajo se expone también en esta Tesis Doctoral.

Dada la enorme complejidad que plantea la cronoejecución en las redes MNB, se ha optado por la utilización de las redes Hopfield dada su sencilla implementación *hardware* y su enorme rapidez en la toma de decisiones, del orden de microsegundos.

En esta sección se inicia el estudio de las redes neuronales, describiendo los distintos tipos de arquitectura y centrandose sobre todo la atención en las redes neuronales Hopfield. Se analizará la estructura tanto *hardware* como matemática de estas redes y

se describirá su dinámica y la discretización necesaria para simular su comportamiento en el tiempo.

## **4.4.2 Redes neuronales**

Las redes neuronales se caracterizan por su capacidad de procesamiento en paralelo que les permite resolver problemas complejos en un corto periodo de tiempo. Esta característica es especialmente interesante en el contexto de las comunicaciones móviles avanzadas, ya que la gestión de recursos es una tarea extraordinariamente compleja que necesita ser realizada en un periodo de tiempo muy breve.

Las redes neuronales Hopfield (HNN, *Hopfield Neural Networks*) no son más que una variante de red neuronal realimentada y se han utilizado en la gestión óptima de recursos radioeléctricos en numerosos trabajos encontrados en la literatura [26]-[29]. En todos estos estudios las redes HNN se han mostrado mucho más potentes que otros esquemas tradicionales.

El resto de esta sección provee una revisión rápida de las características más relevantes de las redes neuronales, lo que permitirá entender con más claridad el funcionamiento del algoritmo propuesto de gestión de recursos basado en redes Hopfield.

### **4.4.2.1 Características de las redes neuronales**

Una red neuronal intenta modelar el funcionamiento del cerebro humano de manera que la red actúe como un modelo computacional para cierto tipo de problemas, como por ejemplo problemas de optimización combinatoria. El modo en que el cerebro actúa es totalmente diferente al sistema digital. El altísimo número de interconexiones entre las neuronas permite que el cerebro se comporte como un computador paralelo, no lineal y de alta complejidad capaz de realizar cálculos mucho más rápido que el más rápido de los ordenadores digitales.

Las redes neuronales se diseñan para modelar el funcionamiento del cerebro y según [16] se pueden definir como:

*“Un procesador distribuido paralelo que tiene una capacidad natural para almacenar conocimiento experimental y hacerlo útil. Una red neuronal se parece al cerebro en dos aspectos:*

- *El conocimiento se adquiere a través de un proceso de aprendizaje.*
- *El grado de acoplamiento de las conexiones inter-neuronales, conocido como peso sináptico, se utiliza para almacenar dicho conocimiento.”*

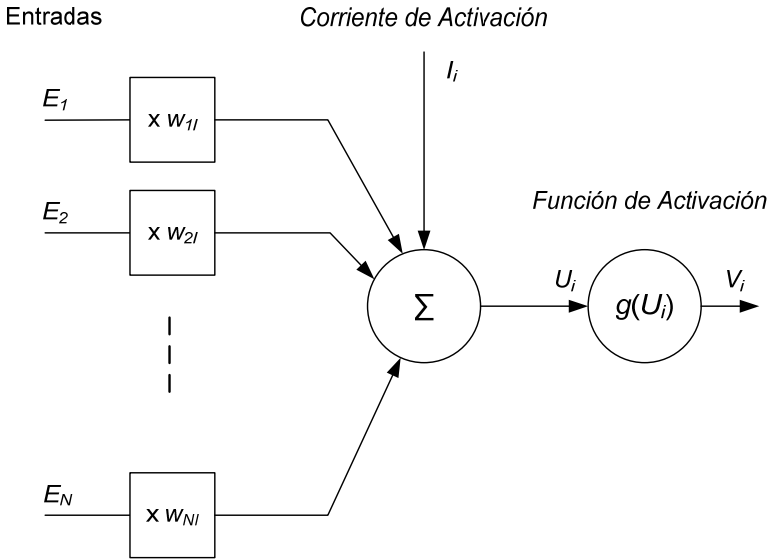
Algunas redes neuronales siguen un proceso de modificación de los pesos sinápticos para adaptar la red a los objetivos deseados. Este procedimiento se conoce como proceso de aprendizaje.

El enorme potencial de las redes neuronales se basa en su estructura distribuida paralela y en su capacidad de generalización, es decir, de producir salidas razonables ante entradas no recibidas en el proceso de aprendizaje.

#### **4.4.2.2 Modelo de neurona**

La neurona es la unidad de procesado de información dentro de las redes neuronales. La Figura 4.4 muestra el modelo general de una neurona en que se pueden distinguir tres elementos básicos:

1. Un conjunto de enlaces de conexión o uniones sinápticas caracterizados por su peso de interconexión.
2. Un sumador que, combinado con las uniones sinápticas, constituye un combinador lineal.
3. Una función de activación de la neurona que, a partir de la entrada, obtiene una determinada salida.



**Figura 4.4 Modelo de neurona**

Matemáticamente, la entrada a la función de activación,  $U_i$ , se puede representar como:

$$U_i = \sum_{j=0}^N w_{ji} E_j + I_i \quad (4.4)$$

siendo  $w_{ji}$  el peso de la interconexión de la entrada  $j$ -ésima en la neurona  $i$ -ésima. El resto de términos aparecen reflejados en la Figura 4.4. Análogamente, la ecuación (4.4) se puede generalizar de manera matricial para una red neuronal. Suponiendo un conjunto de  $K$  neuronas y  $N$  entradas:

$$\bar{U} = \bar{\bar{W}} \cdot \bar{E} + \bar{I} \quad (4.5)$$

Donde  $\bar{U}$  es un vector columna de  $K$  elementos correspondientes a las entradas de la función de activación de todas las neuronas del sistema,  $\bar{\bar{W}}$  es una matriz  $K \times N$  de la siguiente forma:

$$\bar{W} = \begin{pmatrix} w_{11} & w_{21} & w_{31} & \cdots & w_{N1} \\ w_{12} & w_{22} & w_{32} & \cdots & w_{N2} \\ w_{13} & w_{23} & w_{33} & \cdots & w_{N3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ w_{1K} & w_{2K} & w_{3K} & \cdots & w_{NK} \end{pmatrix} \quad (4.6)$$

y, finalmente,  $\bar{E}$  es un vector columna de  $N$  elementos que representa las entradas a todas las neuronas e  $\bar{I}$  es un vector columna con las  $K$  corrientes de entrada de las  $K$  neuronas.

Las salidas de las  $K$  neuronas se calcularían sin más que aplicar la función de activación a todos los elementos del vector  $\bar{U}$ .

$$\bar{V} = g(\bar{U}) \quad (4.7)$$

Asimismo, se pueden distinguir tres tipos de función de activación:

1. Función umbral. La salida de la neurona es máxima si la entrada es positiva y mínima en caso contrario.

$$V_i = g(U_i) = \begin{cases} V_{\max} & \text{si } U_i \geq 0 \\ V_{\min} & \text{si } U_i < 0 \end{cases} \quad (4.8)$$

2. Función lineal. Este tipo de función de activación se describe mediante la siguiente ecuación:

$$V_i = g(U_i) = \begin{cases} V_{\max} & \text{si } U_i \geq a \\ \frac{U_i \cdot (V_{\max} - V_{\min})}{2a} + \frac{V_{\max} + V_{\min}}{2} & \text{si } -a \leq U_i \leq a \\ V_{\min} & \text{si } U_i \leq -a \end{cases} \quad (4.9)$$

La función de activación es lineal entre  $-a$  y  $a$ , tal y como se aprecia en la Figura 4.5 (a).  $V_{\max}$  y  $V_{\min}$  representan respectivamente la máxima y mínima salida de la neurona, que estará limitada por la tensión máxima de alimentación del circuito *hardware* que implementa la red neuronal.

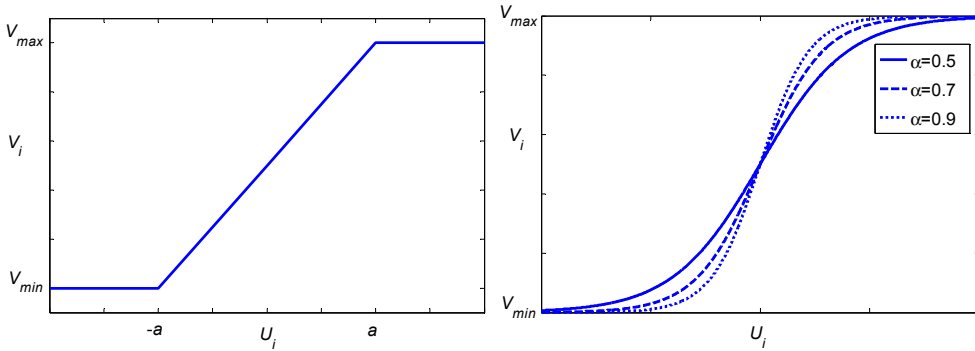


Figura 4.5 Tipos de función de activación de la neurona: (a) Lineal (b) Sigmoidal

3. Función sigmoidal. Esta función de activación presenta la ventaja de ser continua y diferenciable en todo  $\mathbb{R}$ . La función sigmoidal viene descrita por la siguiente ecuación:

$$V_i = g(U_i) = V_{\min} + \frac{V_{\max} - V_{\min}}{1 + e^{-\alpha U_i}} \quad (4.10)$$

El factor  $\alpha$ , o factor de forma de la sigmoide, controla la pendiente de la función, de forma que un valor  $\alpha = \infty$  supone que la función sigmoidal se transforme en la función umbral.

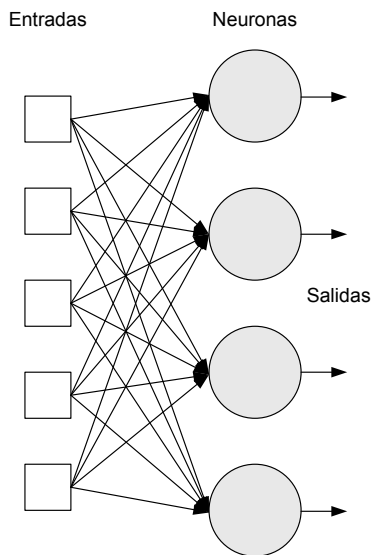
#### 4.4.2.3 Arquitecturas de redes neuronales

Según [16] se distinguen tres clases diferentes de arquitectura de red.

1. Redes de una sola capa sin realimentación. Una red neuronal se puede estructurar en capas de neuronas, en un conjunto de neuronas que comparten la misma fuente de datos. Las redes neuronales de una sola capa presentan un conjunto de entradas que se procesan por un número determinado de neuronas, la salida de estas neuronas es directamente la salida de la red neuronal. Sin

duda, este es el modelo de red neuronal más sencillo y se corresponde con el mostrado en la Figura 4.6.

2. Redes multi-capa sin realimentación. La idea es que en este caso las salidas de una capa de neuronas se convierten a su vez en entradas de la siguiente capa de neuronas, conformando una estructura multi-capa. La Figura 4.7 muestra esta arquitectura multi-capa en la que el número de neuronas de cada capa no tiene porqué ser el mismo.
3. Redes realimentadas. Como su propio nombre indica, las redes realimentadas se basan en que la propia salida de una red mono-capa se realimenta, conformando la entrada de las neuronas en la siguiente iteración. Surge, por tanto, la necesidad de introducir algún elemento que permita independizar las salidas de las entradas de las neuronas, de manera que se pueda producir una evolución iterativa. Este elemento suele ser un simple registro de desplazamiento, como el mostrado en la Figura 4.8.



**Figura 4.6 Red neuronal de una sola capa sin realimentación**

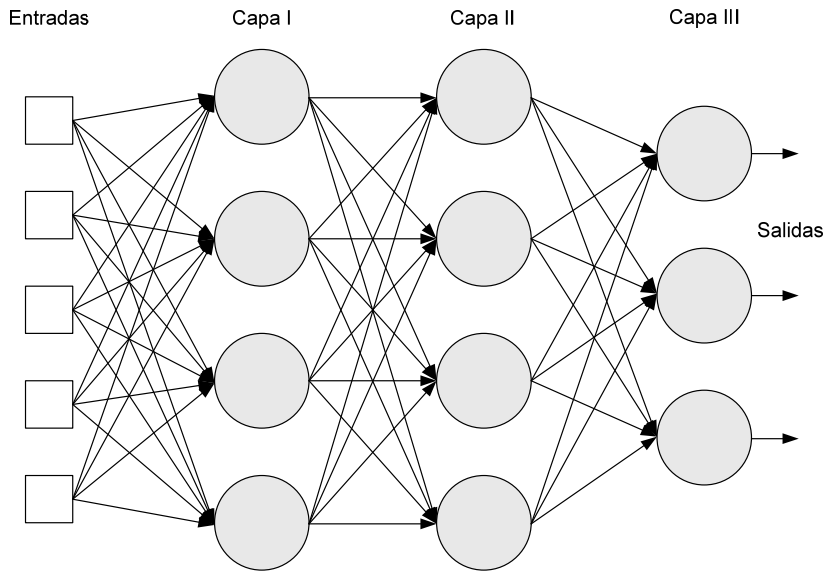


Figura 4.7 Red neuronal multi-capa

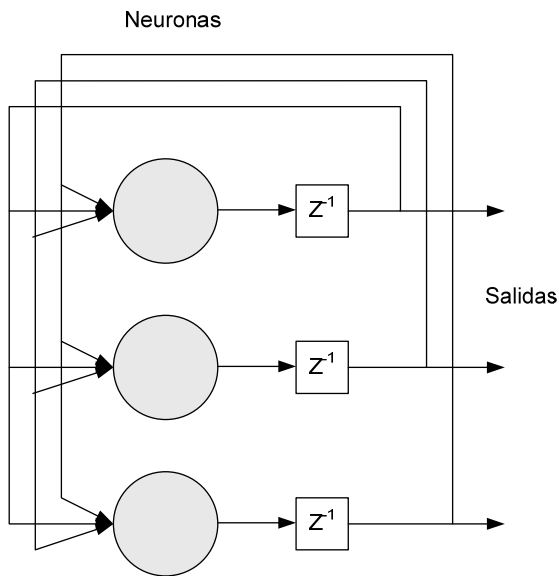


Figura 4.8 Red neuronal realimentada



### 4.4.3 Redes neuronales Hopfield

Las redes neuronales Hopfield no son más que un tipo específico de red neuronal mono-capa realimentada. Por la propia definición de la función de energía que caracteriza la red neuronal, son especialmente indicadas para la optimización de problemas  $N$ -dimensionales complejos. A continuación, se procede a la descripción de este tipo de redes neuronales, las utilizadas en esta Tesis Doctoral.

#### 4.4.3.1 Estructura *hardware*

Hopfield describió en [25] una estructura *hardware* analógica capaz de implementar una red neuronal realimentada. Los registros de desplazamiento fueron sustituidos por el retardo introducido por la carga de un condensador y la función de activación se implementó utilizando la respuesta de un amplificador operacional sin realimentación.

El esquema *hardware* propuesto por Hopfield es el siguiente:

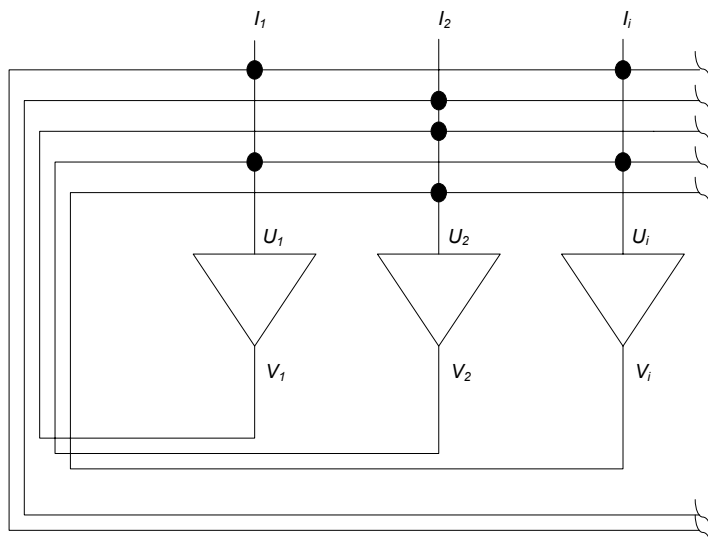
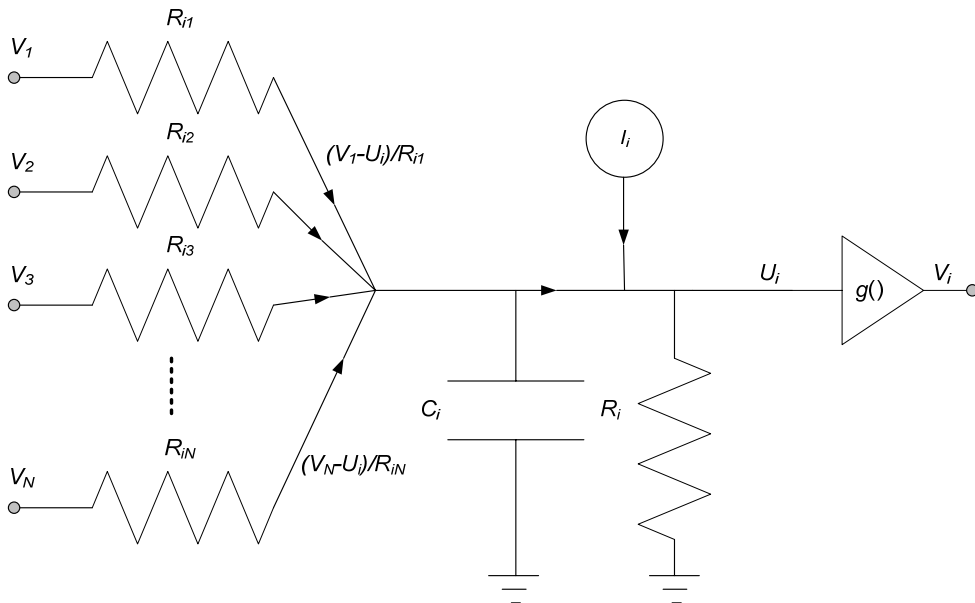


Figura 4.9 Circuito eléctrico de la red neuronal propuesta por Hopfield [25]

Los triángulos se corresponden con amplificadores operacionales mientras que los puntos representan las interconexiones entre las líneas, que finalmente se traducirán en una mayor o menor resistencia.

Cada uno de los operacionales de cada neurona recibe una serie de corrientes: su corriente de activación más las corrientes provenientes de las salidas del resto de neuronas a las que esté interconectada. El circuito correspondiente a una sola de las neuronas, considerando la capacidad e impedancia parásita del operacional, se podría modelar tal y como se muestra en la Figura 4.10.

En este modelo las resistencias  $R_{ij}$  representan las uniones sinápticas entre las neuronas, es decir, la unión entre la salida de la neurona  $j$ -ésima y la entrada de la neurona  $i$ -ésima. Es lo que en la Figura 4.9 se representaba como un punto. Estas resistencias también se pueden escribir en términos de su conductancia  $T_{ij}=1/R_{ij}$ . A la entrada de cada neurona se suman las corrientes provenientes del conjunto de neuronas realimentadas hacia esa neurona más la corriente de activación,  $I_i$ .



**Figura 4.10** Modelo eléctrico de una neurona

La capacidad  $C_i$  y la resistencia  $R_i$  son la capacidad y resistencia parásitas del amplificador operacional y modelan la capacidad y resistencia reales de la membrana celular. Debido al proceso de carga del condensador  $C_i$  la salida de las neuronas no se actualiza instantáneamente ante cambios en las entradas, sino que hay un determinado retardo. Aplicando las leyes de *Kirchoff* al nodo de entrada de la neurona, se puede obtener fácilmente la siguiente expresión:

$$C_i \frac{dU_i(t)}{dt} + \frac{U_i(t)}{R_i} = \sum_{j=1}^N \frac{(V_j(t) - U_i(t))}{R_{ij}} + I_i \quad (4.11)$$

Escribiendo  $R_i$  en términos de conductancia,  $T_i$ , resulta:

$$C_i \frac{dU_i(t)}{dt} + U_i(t) \left( T_i + \sum_{j=1}^N T_{ij} \right) = \sum_{j=1}^N T_{ij} V_j(t) + I_i \quad (4.12)$$

Se define  $T_{Ei}$  como la conductancia equivalente del paralelo entre  $R_i$  y todas las  $R_{ij}$ , con lo que la expresión anterior se puede reescribir como:

$$C_i \frac{dU_i(t)}{dt} + U_i(t) \cdot T_{Ei} = \sum_{j=1}^N T_{ij} V_j(t) + I_i \quad (4.13)$$

o lo que es lo mismo:

$$C_i \frac{dU_i(t)}{dt} + \frac{U_i(t)}{R_{Ei}} = \sum_{j=1}^N T_{ij} V_j(t) + I_i \quad (4.14)$$

#### 4.4.3.2 Ecuación de energía

Hasta ahora, tras un análisis sencillo de su estructura *hardware*, se ha caracterizado el comportamiento de una red genérica de Hopfield. El objetivo ahora es relacionar la ecuación que describe la red con una determinada función de energía. Si fuera posible

encontrar una relación entre ambas, entonces la propia red Hopfield serviría para caracterizar el problema de minimización de la función de energía.

El primer paso es definir una ecuación de energía de manera arbitraria, cuyos términos estén directamente relacionados con la estructura *hardware* mostrada en el anterior apartado. Hopfield propuso la siguiente [25]:

$$E = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N T_{ij} V_i V_j + \sum_{i=1}^N \frac{1}{R_{Ei}} \int_0^{V_i} g_i^{-1}(V) dV - \sum_{i=1}^N I_i V_i \quad (4.15)$$

Es importante destacar que se trata de una función  $N$ -dimensional, donde las  $N$  variables implicadas se corresponden directamente con las  $N$  salidas de la red neuronal.

Analizando (4.15), la derivada parcial de  $E$  respecto  $V_i$  resulta ser:

$$\frac{\partial E}{\partial V_i} = -\sum_{j=1}^N T_{ij} V_j + \frac{U_i}{R_{Ei}} - I_i \quad (4.16)$$

siempre y cuando la matriz  $T$  sea simétrica, es decir que  $T_{ij} = T_{ji}$ . A partir de (4.14) se puede reescribir la derivada parcial de  $E$  respecto a  $V_i$  como:

$$\frac{\partial E}{\partial V_i} = -C_i \frac{dU_i(t)}{dt} \quad (4.17)$$

Resultando el gradiente de  $E$ :

$$\nabla E = \begin{pmatrix} \frac{\partial E}{\partial V_1} \\ \frac{\partial E}{\partial V_2} \\ \vdots \\ \frac{\partial E}{\partial V_N} \end{pmatrix} = - \begin{pmatrix} C_1 \frac{dU_1(t)}{dt} \\ C_2 \frac{dU_2(t)}{dt} \\ \vdots \\ C_N \frac{dU_N(t)}{dt} \end{pmatrix} \quad (4.18)$$

Partiendo de estas expresiones se puede establecer un procedimiento iterativo de aproximación al punto de gradiente cero. Cuando las entradas de las neuronas no varíen, y por tanto las salidas tampoco, el gradiente de  $E$  será igual a cero por lo que el

la red se encontrará en un posible máximo, mínimo o punto de inflexión de la función de energía. Es importante destacar que aunque la función de energía,  $E$ , sea  $N$ -dimensional, se trata de una ecuación de segundo grado, por lo que sólo contará con un punto de gradiente cero. Para saber la naturaleza de este punto se debe analizar el Hessiano de la ecuación de energía. En concreto, los valores propios de la matriz de Hessel,  $H$ , son los que determinan si el punto de gradiente cero es un mínimo o no. En caso de que todos los autovalores de la matriz  $H$  sean positivos entonces la ecuación de energía presenta un mínimo en ese punto. Si, por el contrario, todos los valores son negativos entonces  $E$  presenta un máximo. Si los valores propios son de signo tanto negativo como positivo la función presentará un punto de silla.

A continuación, se muestra la formulación del Hessiano de la ecuación de energía propuesta en (4.15). En la expresión obtenida se puede ver cómo se relaciona este Hessiano con los parámetros característicos de la red de Hopfield. Como conclusión a este análisis, se puede decir que antes de minimizar cualquier ecuación de energía sería conveniente analizar los valores propios de su Hessiano, de manera que se conozca si el gradiente tenderá efectivamente a acercarse a un mínimo.

$$\begin{aligned}
 H &= \begin{pmatrix} \frac{\partial^2 E}{\partial V_1^2} & \frac{\partial^2 E}{\partial V_1 \partial V_2} & \cdots & \frac{\partial^2 E}{\partial V_1 \partial V_N} \\ \frac{\partial^2 E}{\partial V_2 \partial V_1} & \frac{\partial^2 E}{\partial V_2^2} & \cdots & \frac{\partial^2 E}{\partial V_2 \partial V_N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 E}{\partial V_N \partial V_1} & \frac{\partial^2 E}{\partial V_N \partial V_2} & \cdots & \frac{\partial^2 E}{\partial V_N^2} \end{pmatrix} = \\
 &= \begin{pmatrix} T_{11} - \frac{dU_1}{R_{E1} \cdot dV_1} & T_{12} & \cdots & T_{1N} \\ T_{21} & T_{22} - \frac{dU_2}{R_{E2} \cdot dV_2} & \cdots & T_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ T_{N1} & T_{N2} & \cdots & T_{NN} - \frac{dU_N}{R_{EN} \cdot dV_N} \end{pmatrix} \tag{4.19}
 \end{aligned}$$

En su dinámica, la red neuronal *hardware* tenderá a estabilizarse en un punto, punto en el que no variarán las entradas ni las salidas de las neuronas. A partir de (4.18), si las neuronas no varían su estado es porque este punto exacto se corresponde con el punto en el que el gradiente es cero. En caso de que el sistema esté bien diseñado y el punto de gradiente cero se corresponda con un mínimo y éste se encuentre dentro del hiper-cubo definido, la red neuronal convergerá al mínimo de la ecuación de energía. Esto demuestra que la red es capaz de optimizar automáticamente el problema planteado en la función de energía, obteniendo en sus salidas el valor de las  $N$  variables,  $V_i$ , que minimizan (4.15).

#### 4.4.3.3 Dinámica del sistema

A partir de (4.17) y aplicando el método de Euler para la resolución de sistemas de ecuaciones diferenciales, se puede obtener la siguiente expresión:

$$U_i(t + \Delta t) = U_i(t) - \Delta t \cdot \left( \frac{1}{C_i} \frac{\partial E}{\partial V_i} \right) \quad (4.20)$$

De esta manera, gradientes negativos, es decir, siempre que la red se acerque a un mínimo, harán que la salida de la neurona aumente, mientras que gradientes positivos reducirán el valor de la neurona. Por lo tanto, el sistema evolucionará siempre hacia puntos de mínima energía, optimizando por tanto el problema planteado.

A partir de los valores de  $U_i$  se calculan las salidas de las neuronas,  $V_i$ . Las redes Hopfield están diseñadas para que obtengan una salida binaria, es decir, o bien la  $V_i$  es igual a  $V_{max}$ , en cuyo caso se dice que la neurona está activa o en estado ON, o bien la salida es igual a  $V_{min}$ , es decir, la neurona está inactiva o en estado OFF. Estos valores de actividad o inactividad se interpretan a la luz del problema como afirmación o negación de una pregunta. Por ejemplo, en un problema de asignación de recursos, un estado ON podría representar que a un determinado usuario se le da recursos mientras que el estado OFF de su neurona representa la negación de esta asignación.

Independientemente de que el punto de  $\mathbb{R}^N$  que presente gradiente cero sea máximo, mínimo o punto de silla, en el cubo  $N$ -dimensional formado por las  $N$  variables – desde  $V_{min}$  hasta  $V_{max}$  – que constituyen las  $N$  neuronas, el gradiente de la función de energía hará tender las salidas de las neuronas hacia uno de los extremos, siempre bajo la dirección que marquen los gradientes parciales. Es conveniente analizar el comportamiento de la red en los siguientes casos planteados:

- El gradiente cero está dentro del hipercubo y es un máximo. Salvo que el punto aleatorio inicial coincida exactamente con el máximo (caso que ocurre con probabilidad cero pues se trata de un único punto dentro del espacio infinito  $\mathbb{R}^N$ ), en cuyo caso el gradiente sería cero y las neuronas no modificarían su valor, el valor de las neuronas divergirá de este máximo hasta alcanzar la arista del hipercubo que diste más del máximo, en términos de su nivel de energía. Ante el peligro de quedarse estancado en alguna de las aristas, la ecuación de energía deberá disponer de algún término que fuerce al sistema a alcanzar las esquinas del hipercubo.
- El gradiente cero está dentro del hipercubo y es un mínimo. En este caso el sistema evolucionará irremediamente hacia el mínimo hasta alcanzarlo. Siempre lo alcanzará puesto que, al ser una ecuación de segundo grado, no existen mínimos locales. El sistema tiene que devolver una solución discreta, es decir neurona activa o inactiva, por lo que, a partir de la solución proporcionada por la red neuronal, el sistema devolverá la esquina del hipercubo más cercana, en términos de distancia euclídea, al mínimo global de la ecuación de energía. En principio no hay manera de asegurar que la esquina más cercana sea la que presente menor energía, pero parece razonable que así sea. Esta es sin duda una de las limitaciones de la redes Hopfield, la necesidad de obtener una solución localizada en una esquina del hipercubo.
- El gradiente cero está dentro del hipercubo y es un punto de silla. En este caso en el punto de gradiente cero hay algunas neuronas que presentan un mínimo mientras que otras presentan un máximo. El sistema en principio nunca se quedará bloqueado en este punto de silla, puesto que para que esto ocurriera la

aproximación a ese punto se debería realizar desde un hiper-plano en el que todas las variables que en el punto de silla presentasen un máximo estuvieran exactamente en ese máximo y como ese hiper-plano respecto a  $\mathbb{R}^N$  es un subespacio de medida cero la probabilidad de que esto ocurra es cero.

- El gradiente cero está fuera del hipercubo. Realmente este es el caso más común. Si la ecuación de energía se ha diseñado convenientemente para que el mínimo se encuentre en las esquinas, el sistema siempre tenderá a una de ellas alcanzando una solución que garantice todas las restricciones impuestas. Sin embargo, habrá un nivel de incertidumbre importante respecto a la solución final en función del punto aleatorio de partida del sistema. Esto es debido a que, por lo general, la matriz de Hessel de la ecuación de energía es de grandes dimensiones por lo que rara vez es definida positiva. Por ello, lo más probable en un problema real es que el punto de gradiente cero sea un punto de silla. Debido a esto, es muy probable que dentro del hipercubo se definan como ‘cordilleras’ de energía máxima que partan el hipercubo, de manera que, según el lado en que se tome el punto inicial, se evolucione hacia un lado u otro del hipercubo y, por tanto, hacia soluciones distintas. De cualquier modo, definiendo convenientemente la ecuación de energía se puede garantizar que la zona a la que acabe convergiendo la red neuronal sea la correcta. A partir de la solución entregada por la red neuronal y mediante un algoritmo de prueba y error sencillo, como el algoritmo de greedy, será posible encontrar en un subespacio de soluciones mucho más reducido la solución óptima.

#### **4.4.3.4 Simplificación de la ecuación de energía**

A partir de la definición de la ecuación de energía del problema que se intente optimizar y mediante comparación directa con la expresión mostrada en (4.15), se deberían poder extraer los parámetros físicos que caracterizan la red *hardware*, es decir, las corrientes de activación y el valor de las resistencias de interconexión de las



neuronas. Una vez extraídos estos parámetros se podría implementar el circuito analógico correspondiente para optimizar el problema dado.

El principal problema de la ecuación (4.15) es precisamente el segundo término, ya que es prácticamente imposible definir una ecuación de energía que incluya un término integral similar. Por eso mismo, Hopfield consideró una pendiente infinita de la función  $g()$ , de manera que la función de activación fuera directamente la función umbral.

Analizando el caso de una función de activación sigmoideal (4.10):

$$\int_0^{V_i} g_i^{-1}(V) dV = \frac{1}{\alpha_i} \left[ (V_i - V_{\min}) \ln(V_i - V_{\min}) + (V_{\max} - V_i) \ln(V_{\max} - V_i) \right] \quad (4.21)$$

Por lo que si  $\alpha_i \rightarrow \infty$  la integral mostrada en (4.15) directamente desaparece quedando:

$$E = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N T_{ij} V_i V_j - \sum_{i=1}^N I_i V_i \quad (4.22)$$

Lo mismo ocurre para la función de activación lineal. Considerando esta nueva expresión de la energía, la derivada parcial de  $E$  resulta ser:

$$\frac{\partial E}{\partial V_i} = -\sum_{j=1}^N T_{ij} V_j - I_i \quad (4.23)$$

Por lo que, por comparación directa con (4.14), se puede reescribir la derivada parcial de  $E$  respecto a  $V_i$  como:

$$\frac{\partial E}{\partial V_i} = -C_i \frac{dU_i(t)}{dt} - \frac{U_i(t)}{R_{Ei}} \quad (4.24)$$

Podría parecer, analizando esta última expresión, que con esta aproximación ya no es válida la afirmación de que cuando las neuronas no cambien de estado (la derivada de  $U_i$  respecto del tiempo sea cero) la energía presente un gradiente cero. Sin embargo, considerando  $\alpha_i \rightarrow \infty$ , en el momento en que el sistema está evolucionando y las  $U_i$

están cambiando de valor, el valor absoluto de  $U_i$  es prácticamente cero (aumentar la pendiente hace que la evolución de las  $U_i$  se produzca cerca del cero) por lo que el último término de (4.24) se puede despreciar quedando una ecuación idéntica a (4.17).

#### 4.4.3.5 Modelo discreto de las redes neuronales Hopfield

La discretización de la dinámica de las redes Hopfield es fundamental para poder simular la evolución de las mismas. A partir de (4.17) y conociendo el valor del gradiente de la energía (4.16), se puede obtener la expresión siguiente:

$$\frac{dU_i(t)}{dt} = \frac{1}{C_i} \left( \sum_{j=1}^N T_{ij} V_j(t) + I_i - \frac{U_i(t)}{R_{Ei}} \right) \quad (4.25)$$

y aplicando euler

$$U_i(t + \Delta t) = U_i(t) + \Delta t \cdot \frac{1}{C_i} \left( \sum_{j=1}^N T_{ij} V_j(t) + I_i - \frac{U_i(t)}{R_{Ei}} \right) \quad (4.26)$$

Agrupando términos convenientemente resulta:

$$U_i(t + \Delta t) = A_i \cdot U_i(t) + \left( \sum_{j=1}^N W_{ij} V_j(t) + B_i \right) \quad (4.27)$$

siendo

$$A_i = 1 - \frac{\Delta t}{C_i \cdot R_{Ei}}, \quad W_{ij} = \frac{T_{ij} \cdot \Delta t}{C_i}, \quad B_i = \frac{I_i \cdot \Delta t}{C_i} \quad (4.28)$$

Si en lugar de la evolución temporal de las variables se considera una sucesión, donde el evento  $t + \Delta t$  corresponde con el elemento  $n+1$  de la secuencia, resulta que:

$$U_i^{[n+1]} = A_i \cdot U_i^{[n]} + \left( \sum_{j=1}^N W_{ij} V_j^{[n]} + B_i \right) \quad (4.29)$$

Si, además, se considera una función de activación lineal de forma que  $V_i = \alpha \cdot U_i$  y se ajusta convenientemente el valor de los pesos, se puede demostrar fácilmente que:

$$\bar{U}^{[n+1]} = \bar{U}^{[n]} - \mu \cdot \nabla f(\bar{U}^{[n]}) = \bar{U}^{[n]} + \mu (\bar{R} \bar{U}^{[n]} + \bar{C}) \quad (4.30)$$

con

$$\begin{aligned} f(\bar{U}) &= -\frac{1}{2} \cdot \bar{U}^T \cdot \bar{R} \cdot \bar{U} - \bar{C}^T \cdot \bar{U} \\ \nabla f(\bar{U}) &= -\bar{R} \cdot \bar{U} - \bar{C} \end{aligned} \quad (4.31)$$

$$\bar{R} = \begin{pmatrix} \frac{W_{11} \cdot \alpha - 1 + A_1}{\mu} & W_{12} \cdot \frac{\alpha}{\mu} & \dots & W_{1N} \cdot \frac{\alpha}{\mu} \\ W_{21} \cdot \frac{\alpha}{\mu} & \frac{W_{22} \cdot \alpha - 1 + A_2}{\mu} & \dots & W_{2N} \cdot \frac{\alpha}{\mu} \\ \vdots & \vdots & \ddots & \vdots \\ W_{N1} \cdot \frac{\alpha}{\mu} & W_{N2} \cdot \frac{\alpha}{\mu} & \dots & \frac{W_{NN} \cdot \alpha - 1 + A_N}{\mu} \end{pmatrix} \quad (4.32)$$

y

$$\bar{C} = \left( \frac{B_1}{\mu} \quad \frac{B_2}{\mu} \quad \dots \quad \frac{B_N}{\mu} \right)^t \quad (4.33)$$

En [23] se demuestra que una aproximación sucesiva de la forma que tiene (4.30) converge hacia el mínimo de la función  $f(\bar{U})$ . Por comparación directa se puede observar como  $f(\bar{V})$  es prácticamente igual a la expresión mostrada en (4.15) salvo por el término de la integral, término que el propio Hopfield eliminó suponiendo el uso de una función de activación umbral.

Por lo tanto, la ecuación de energía propuesta por Hopfield se puede minimizar utilizando la siguiente aproximación sucesiva:

$$V_i^{[n+1]} = V_i^{[n]} + \mu \left( \sum_{j=1}^N T_{ij} V_j^{[n]} + I_i \right) = V_i^{[n]} - \mu \frac{\partial E}{\partial V_i} \quad (4.34)$$

y acotar  $V_i$  entre  $V_{min}$  y  $V_{max}$ .

Como se ha comentado anteriormente, la función a minimizar, aún siendo  $N$ -dimensional, es de segundo grado por lo que, si el rango de la matriz  $\overline{\overline{T}}$  es  $N$ , sólo existe un único punto crítico. A partir de [23] se puede demostrar que si  $\overline{\overline{T}}$  es simétrica – restricción ya impuesta por Hopfield – y, además, es semidefinida negativa, entonces la sucesión  $\overline{V}^{[n]}$  converge al mínimo de la función en  $\mathbb{R}^N$ ,  $\overline{V}^*$ , siempre que el factor de convergencia  $\mu$  sea positivo y menor que la inversa del mayor valor absoluto de los valores propios de  $\overline{\overline{T}}$ :

$$0 < \mu < \frac{2}{|\lambda|_{\max}} \quad (4.35)$$

Esta afirmación coincide con lo expuesto en [30].

Sin embargo, cuando el espacio de búsqueda se reduce a un hipercubo  $N$ -dimensional si el punto crítico se encuentra fuera del hipercubo, independientemente de que este punto sea un máximo, un mínimo o un punto de silla, el gradiente siempre forzará al sistema a evolucionar hacia el borde del hipercubo. En caso de que el punto crítico esté dentro del hipercubo, éste deberá ser necesariamente un máximo o un punto de silla y nunca será un mínimo ya que la ecuación de energía se diseña para presentar mínima energía en las esquinas del hipercubo. Como tras la discretización el espacio de búsqueda es también discreto, la presencia de un máximo o un punto de silla dentro del hipercubo puede acarrear problemas, ya que el punto inicial de búsqueda del algoritmo iterativo puede estar a una distancia menor que  $\rho$ , siendo  $\rho$  la distancia máxima al punto crítico a la que se desea estabilizarse. De este modo, el diseño de la ecuación de energía debe asegurar que el punto crítico está o bien fuera del hipercubo o, como caso peor, justo en el borde del mismo. Como ejemplo, decir que la función de energía diseñada en el apartado 4.5 se caracteriza por presentar un punto de silla y por encontrarse éste fuera del hipercubo de búsqueda. Lo único que se puede realizar para

tratar de asegurar la convergencia es que, al menos, el factor de convergencia cumpla lo establecido en (4.35).

#### 4.4.3.6 Implementación en dispositivos *hardware*

La implementación del proceso iterativo mostrado en la ecuación (4.30) o (4.34) resulta muy sencilla ya que se reduce a la multiplicación de una matriz por un vector y a la posterior suma de otros dos vectores.

Para comprobar la fácil implementación de este procedimiento de optimización, se ha seleccionado un kit de desarrollo de Texas Instruments para implementar una red neuronal, en concreto el DSKC6711 mostrado en la Figura 4.11.

Tras implementar un ejemplo sencillo de red neuronal utilizando el Code Composer Studio proporcionado por el propio kit de desarrollo, se ha determinado que el número de operaciones (multiplicación-acumulación o suma o comparación) realizadas en cada iteración es aproximadamente de  $5N+N^2$ , siendo  $N$  el número de neuronas.

El DSP de TI integrado en la placa de desarrollo DSKC6711 permite realizar algunas operaciones en paralelo siempre que la información esté correctamente localizada en la memoria caché de la placa.

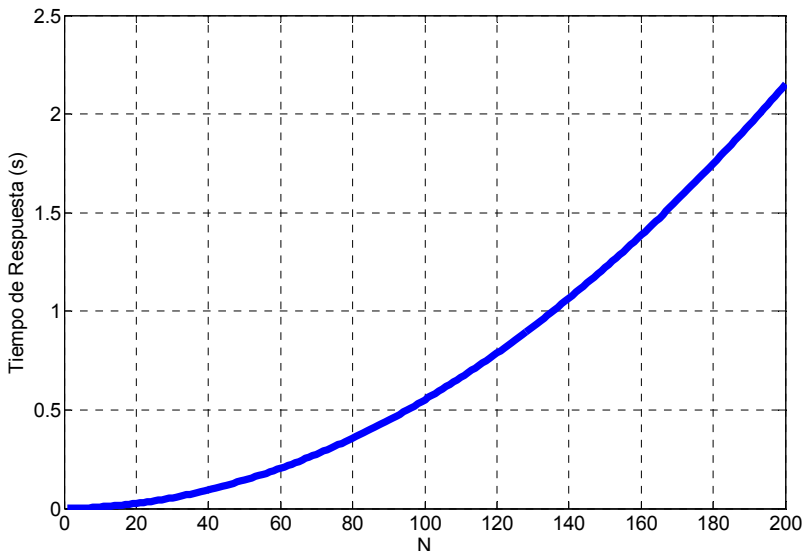


**Figura 4.11** Kit de desarrollo de Texas Instruments

De cualquier manera, al trabajar en coma flotante, el procesador necesita de dos ciclos de reloj solamente para realizar una multiplicación, además de los necesarios para cargar los datos en la ALU desde la caché. Optimizando la implementación, y gracias a la programación paralela llevada a cabo, el número de ciclos de reloj medio por operación resultó ser de 2.1.

Con estos datos se puede obtener una estimación del tiempo de respuesta de la red neuronal para un número determinado de iteraciones y de neuronas. La Figura 4.12 muestra estos resultados para un número creciente de neuronas y un valor fijo de 5000 iteraciones del algoritmo. La frecuencia de trabajo del kit de desarrollo DSKC6711 es de 200 MHz.

Aunque los resultados mostrados son suficientemente buenos, podrían ser mejorados considerablemente implementando la red neuronal en dispositivos FPGA (*Field Programmable Gate Array*). Estos dispositivos están especialmente diseñados para realizar de manera sencilla una implementación en paralelo de operaciones independientes, lo que se ajusta perfectamente a las redes neuronales debido al comportamiento independiente de las neuronas.



**Figura 4.12** Tiempo de respuesta del sistema en función del número de neuronas

Dado que el objetivo de la presente Tesis Doctoral no se centra en la implementación *hardware*, los aspectos relacionados con la implementación en paralelo se enmarcan también en líneas futuras de trabajo.

## 4.5 Utilización de Redes Hopfield Para Asignar Recursos en UMTS

En la transmisión de datos en sistemas celulares, el enlace descendente soporta una mayor carga de tráfico y, por tanto, será en enlace limitante en la capacidad de la red. Debido a esto, el objetivo de este apartado es formular y evaluar una red neuronal Hopfield capaz de gestionar los recursos radio de una red UMTS convencional en sentido descendente. Mientras que este apartado se centrará únicamente en el caso de una red UMTS convencional, en el Capítulo 5 se extenderá esta formulación, adaptándola a la nueva arquitectura radio Multinodo B.

La principal ventaja del uso de una red neuronal es que la asignación de recursos a los usuarios no se hace secuencialmente, uno detrás de otro, sino que los usuarios competirán entre ellos simultáneamente por los recursos. Esto se ajusta perfectamente a los sistemas limitados por interferencias, como UMTS, en los que la asignación a un usuario de un determinado recurso afecta a todos los demás. En UMTS, una asignación de recursos secuencial es muy engañosa puesto que no se tiene en cuenta el efecto provocado por cada asignación al conjunto del sistema y se puede desaprovechar mucha capacidad de la red al asignar inconvenientemente los recursos disponibles. En la asignación llevada a cabo mediante redes neuronales, los usuarios compiten simultáneamente por los recursos, de manera que cada usuario aumenta su demanda hasta que incumple alguna de sus limitaciones o hasta que el sistema se queda sin recursos. Al incrementar todos los usuarios simultáneamente sus peticiones, aquellos que consuman más recursos serán los primeros en exceder la limitación de recursos máxima y, por tanto, serán los primeros en ser penalizados, por lo que el sistema se asegura de que, en igualdad de condiciones de prioridad, aquellos usuarios que consumen más recursos se vean perjudicados frente a otros con mejores condiciones de propagación.

Como se ha explicado en el apartado anterior, otra ventaja de las redes Hopfield es su rápida velocidad de respuesta, que permite realizar una asignación de recursos bastante sofisticada en tiempo real. Además, el comportamiento de la red neuronal se controla



mediante su función de energía característica. Añadiendo los términos necesarios y ponderándolos convenientemente se pueden garantizar un comportamiento adecuado y una correcta priorización de los distintos objetivos marcados, muchas veces contradictorios entre sí.

Este apartado está estructurado de la siguiente manera. En primer lugar es necesario plantear cuáles son los objetivos o requisitos que se desea satisfacer con el algoritmo desarrollado. En el apartado 4.5.2 se describe cuál es la nueva formulación de la red Hopfield que permitirá resolver el problema de gestión de recursos radio planteado. La evaluación de los resultados y discusión correspondiente se realizará en el apartado 4.5.3.

### **4.5.1 Requisitos del algoritmo de gestión de recursos**

El problema de la asignación dinámica de recursos en UMTS consiste en encontrar la asignación óptima de tasas binarias para todos los usuarios activos dadas unas determinadas restricciones del sistema. Las implicaciones que tienen estas restricciones deben reflejarse cuidadosamente en el diseño del algoritmo utilizado. Algunas de estas restricciones son realmente estrictas, y deben ser siempre satisfechas, como por ejemplo la máxima cantidad de recursos que se pueden asignar. Obviamente, no tiene sentido asignar más recursos de los que realmente existe en el sistema. En contraposición con las restricciones estrictas, existe otro tipo de restricciones, denominadas flexibles, que, aunque pueden no ser satisfechas, son recomendables. El retardo máximo de los paquetes web es un claro ejemplo de este último tipo de restricción flexible.

A continuación, se listan todas las restricciones que el algoritmo desarrollado tendrá que cumplir.

### Restricción de carga

Todos los sistemas de comunicaciones móviles tienen una cantidad limitada de recursos a distribuir entre los usuarios. Cada tasa binaria tiene asociada una cierta cantidad de recursos necesaria, que depende no sólo de la correspondiente tasa sino también de otros factores, tales como la posición de los usuarios, las condiciones de interferencia, etc. Obviamente, asignar tasas binarias más altas siempre implica un consumo mayor de recursos. El tipo de recurso que se comparte entre los usuarios depende en gran medida de la tecnología de acceso radio considerada, por ejemplo, *slots* temporales en GSM, tasa binaria en WLAN o potencia transmitida y número de códigos de ensanchado en UMTS. Es importante destacar que, en ocasiones, el sistema no puede hacer uso de todos los recursos disponibles, puesto que parte de los mismos están reservados para otros usos, como por ejemplo para permitir el acceso de usuarios traspasados.

UMTS es una tecnología especialmente compleja en la gestión de la carga puesto que, en primer lugar, tiene más de una restricción estricta de carga: la potencia y el número de códigos. Además, tiene la particularidad de que su capacidad depende en gran medida de las condiciones de propagación de los usuarios, por lo que para determinar el consumo de recursos de una determinada asignación de tasa binaria se tiene que conocer el estado del resto de los usuarios.

En general, la carga del sistema se puede definir como la proporción de recursos consumidos por el sistema:

$$\eta = \frac{\text{Recursos Asignados}}{\text{Recursos Disponibles}} \quad (4.36)$$

Resumiendo, en cuanto a las restricciones de carga impuestas por el sistema UMTS, se deberá garantizar

- No exceder la potencia máxima disponible en el Nodo-B.
- No exceder el número de códigos disponibles en el árbol de códigos OVSF.

### Restricción de tasa binaria

En esta Tesis Doctoral se ha asumido que dentro de UMTS se puede asignar un número finito de tasas binarias  $\mathfrak{R} \subset \mathbb{R}$ . Del mismo modo, cada usuario se puede caracterizar por un subconjunto de posibles tasas binarias  $\mathfrak{R}_i \subset \mathfrak{R}$ , específicas del tipo de servicio al que el usuario esté suscrito. El algoritmo de DRA desarrollado deberá evitar que un usuario transmita con tasas binarias que no le están permitidas, es decir, si  $R_j$  es la tasa binaria asignada al usuario  $i$ , entonces el algoritmo de DRA deberá verificar que  $R_j \in \mathfrak{R}_i$ .

En resumidas cuentas, respecto a la limitación de tasa binaria impuesta por cada usuario, el gestor de recursos deberá

- Garantizar que a un usuario no se le asignan tasas binarias no permitidas por el sistema, bien por ser muy elevadas (usuarios de baja calidad) o bien por ser muy bajas (usuarios de alta calidad).

Con una definición adecuada de  $\mathfrak{R}_i$  será muy sencillo garantizar una determinada tasa mínima a cualquier usuario.

### Restricción de retraso en la transmisión de paquetes

Respecto a la calidad de servicio de los usuarios, el algoritmo de gestión de recursos será responsable de

- Asegurar que el retraso experimentado por cada usuario en la transmisión de un objeto no excede el máximo retraso permitido por el servicio.

Este último apartado merece especial atención puesto que el retraso en la transmisión es un parámetro fundamental en la medida de la satisfacción de los usuarios. Además, dependiendo del tipo de servicio, el retraso de paquetes, a partir de ahora denominado indistintamente *packet delay*, se considerará como una restricción estricta o flexible. Para los servicios de tiempo real se trata de una restricción estricta ya que el *delay* máximo no puede excederse. Por otro lado, para los servicios de no tiempo real el *packet delay* debe ser minimizado a pesar de que no exista un *delay* máximo a partir

del cual se deba descartar la información. Un tiempo de espera menor implica una mayor satisfacción del usuario.

Para introducir el *delay* dentro del proceso de asignación de recursos, se ha definido la tasa mínima que se debe asignar al usuario  $i$ -ésimo para garantizar la transmisión del paquete dentro de los márgenes temporales establecidos,  $R_{\min,i}$ . La siguiente expresión muestra cómo calcular la tasa mínima para todos los servicios, incluida la navegación web y la descarga FTP, donde puede darse el caso de tener más de una descarga simultánea:

$$R_{\min,i} \leq P \cdot \max_b \left( \frac{B_b}{t_{\max} - t_b} \right) \quad (4.37)$$

$B_b$  y  $t_b$  representan el número de bits pendientes de transmitir del  $b$ -ésimo objeto en espera y el tiempo de permanencia en la cola del mismo,  $t_{\max}$  es el *delay* máximo permitido (y para servicios NRT el *delay* máximo deseable por definición de calidad de servicio) y  $P$  es el número de objetos que se transmite simultáneamente, que será igual a uno para los servicios RT. Respecto a esto, hay que remarcar que en vídeo conferencia nunca habrá más de un objeto esperando a ser transmitido, por lo que tanto  $b$  como  $P$  serán siempre igual a uno.

Para el servicio web, un objeto podrá ser el objeto HTML principal de la página o cualquiera de sus objetos incrustados. Para FTP un objeto es un archivo descargado mientras que para vídeo conferencia un objeto es una trama de vídeo.

En la expresión anterior se supone que para tráfico *web* y FTP, si se dispone de varias descargas simultáneas, la tasa binaria asignada al usuario se divide equitativamente entre todas las descargas. Es por esto que la expresión utilizada da una cota superior a la tasa mínima, ya que una vez acabada la transmisión de uno de los objetos la tasa binaria liberada se distribuiría equitativamente entre los restantes. La desigualdad se transforma en igualdad si el paquete con mayor restricción temporal es precisamente el más pequeño. En cualquier caso, la definición utilizada para la tasa

mínima se utiliza únicamente para asignar prioridades a los usuarios, por lo que (4.37) es una buena expresión cerrada para este fin.

Finalmente, si cualquier objeto excede su máximo *delay*, es decir,  $t_b > t_{\max}$ , entonces  $R_{\min,i} = \infty$ .

## 4.5.2 Optimización del problema DRA mediante redes Hopfield

El problema de la asignación dinámica de recursos es un problema NP (*Non-deterministic Polynomial time*) lo que implica que es prácticamente imposible encontrar una solución analítica cuando aumenta el número de usuarios o el número de posibles tasas binarias consideradas. Tal y como se ha comentado anteriormente, una de las principales ventajas de las HNN es que es capaz de encontrar una solución semi-óptima en un tiempo muy reducido, gracias sobre todo al inherente paralelismo de la red. Este apartado describe el algoritmo de DRA propuesto, basado en HNN y orientado a satisfacer el máximo *delay*. A dicho algoritmo se le ha asignado el acrónimo DC-HNN-DRA, es decir, *Delay Centric Hopfield Neural Network-based Dynamic Resource Allocation Algorithm*. A continuación, se describirá el modelo de red considerado, la formulación empleada, su dinámica y el cálculo de sus constantes.

### 4.5.2.1 Modelo de red HNN

Tal y como se explicó en el apartado 4.4.3, las redes neuronales Hopfield están constituidas por un conjunto de neuronas interconectadas. Las neuronas irán cambiando sus salidas dinámicamente hasta alcanzar un punto de equilibrio. Hopfield mostró que la función de energía,  $E$ , representa la dinámica de la red Hopfield y, por tanto, el problema de encontrar el equilibrio de la red se puede resolver encontrando el mínimo local de la función de energía y viceversa [24][25].

Tal y como se describió antes, la dinámica de la red Hopfield se puede expresar como:

$$\frac{dU_i(t)}{dt} = -\frac{1}{C_i} \frac{\partial E}{\partial V_i} \quad (4.38)$$

En la implementación de la red Hopfield llevada a cabo en esta Tesis Doctoral, la relación entre las entradas de las neuronas,  $U_i$ , y las salidas,  $V_i$ , viene dada por la función sigmoideal mostrada en (4.10). Haciendo  $V_{\min}=0$  y  $V_{\max}=1$ , la relación es:

$$V_i = \frac{1}{1 + e^{-\alpha U_i}} \quad (4.39)$$

definiendo  $U_i \in [-10,10]$ , para así ser más fiel a las limitaciones *hardware* de la posible implementación del problema.

A partir de (4.38), suponiendo  $C_i = 1$  y aplicando la técnica de Euler para resolución de ecuaciones diferenciales, se puede encontrar la siguiente formulación para la dinámica de la red HNN:

$$U_i^{[n+1]} = U_i^{[n]} - \Delta t \left( \frac{\partial E}{\partial V_i} \right) = U_i^{[n]} + \Delta t \left( \sum_{j=1}^N T_{ij} V_j^{[n]} + I_i \right) \quad (4.40)$$

donde  $\Delta t$  es el periodo de muestreo de la evolución de las tensiones de las neuronas.

El mínimo de la función de energía se encuentra en una de las  $2^L$  esquinas del hipercubo  $L$ -dimensional definido en  $V_i \in [0,1]$ , siendo  $L$  el número total de neuronas del problema [24]. Por lo tanto, el problema de optimización se transforma en definir la función de energía que mejor se ajuste al problema, ya que la propia dinámica de la red neuronal llevará las neuronas al estado de mínima energía, o estado de equilibrio del problema. Una vez alcanzado el equilibrio, todas las neuronas estarán en estado ON si su salida es mayor o igual que 0.5, o en estado OFF en caso contrario.

### 4.5.2.2 Formulación del problema

Antes de pasar a definir la función de energía es interesante describir la formulación empleada. Se ha considerado una agrupación lógica de las neuronas en forma de matriz bidimensional. De cara a su implementación física, esta agrupación carece de sentido pero para la simulación simplifica bastante la obtención de resultados. Se define  $\overline{V}$  como una matriz de dimensiones  $N \times M$  donde  $N$  es el número de usuarios del sistema (ya no el número de neuronas) y  $M$  es el máximo número de tasas binarias disponibles, es decir, el tamaño máximo de los conjunto de tasas disponibles de todos los usuarios  $\mathfrak{R}_i$ . Por ejemplo, para usuarios de canales compartidos DSCH se ha considerado el siguiente conjunto de tasas binarias disponibles,  $\mathfrak{R}_i = \{0 \text{ kbps}, 16 \text{ kbps}, 32 \text{ kbps}, 64 \text{ kbps}, 128 \text{ kbps}, 256 \text{ kbps}\}$ , mientras que para usuarios de voz se ha considerado el siguiente conjunto  $\mathfrak{R}_i = \{4.75 \text{ kbps}, 5.9 \text{ kbps}, 7.95 \text{ kbps}, 12.2 \text{ kbps}\}$ , por lo que finalmente  $M=6$ . El estado final de las neuronas representa la asignación de recursos llevada a cabo. De este modo, si la neurona  $V_{ij}$  está en estado ON esto quiere decir que al usuario  $i$ -ésimo se le ha asignado la tasa binaria  $j$ -ésima, mientras que el resto de neuronas del mismo usuario, que coincide con el resto de elementos de la misma fila de la matriz  $\overline{V}$ , deben estar en estado OFF. No debe confundirse la nomenclatura  $(i,j)$  utilizada ahora con la empleada, por ejemplo, en la ecuación (4.40) donde  $T_{ij}$  representaba la interconexión entre la neurona  $i$  y la neurona  $j$ . La ecuación (4.40) debe modificarse para adecuarse a la nueva formulación de la siguiente manera:

$$U_{ij}^{[n+1]} = U_{ij}^{[n]} - \Delta t \left( \frac{\partial E}{\partial V_{ij}} \right) = U_{ij}^{[n]} + \Delta t \left( \sum_{k=1}^N \sum_{l=1}^M T_{ij,kl} V_{kl}^{[n]} + I_{ij} \right) \quad (4.41)$$

La matriz  $\overline{T}$  sigue siendo una matriz de dos dimensiones que representa la interconexión entre neuronas. La  $M$  primeras entradas se corresponden con el primer usuario, las  $M$  siguientes con el segundo y así sucesivamente. De esta manera, la fila  $ij$

se corresponden con la fila  $(i-1)M + j$  de la matriz  $\overline{T}$  y de igual forma se procede para calcular la columna correspondiente.

Se define  $R_j$  como la tasa binaria  $j$ -ésima del conjunto  $\mathfrak{R}_i$  y  $R_{\max,i}$  como la tasa binaria máxima del usuario  $i$ , es decir,  $R_{\max,i} = \max\{R_j, R_j \in \mathfrak{R}_i\}$ .

La función de energía del algoritmo DC-HNN-DRA propuesto en esta Tesis Doctoral para redes UMTS convencionales es la siguiente:

$$\begin{aligned}
 E = & -\mu_1 \sum_{i=1}^N O_i \sum_{j=1}^M C_{ij} V_{ij} - \mu_2 \sum_{i=1}^N O_i \sum_{j=1}^M \frac{R_j}{R_{\max,i}} V_{ij} + \frac{\mu_3}{2} \sum_{i=1}^N \sum_{j=1}^M \frac{R_j}{R_{\max,i}} \xi_{ij} V_{ij} + \\
 & + \frac{\mu_4}{2} \sum_{i=1}^N \sum_{j=1}^M \psi_{ij} V_{ij} + \frac{\mu_5}{2} \sum_{i=1}^N \sum_{j=1}^M V_{ij} (1 - V_{ij}) + \frac{\mu_6}{2} \sum_{i=1}^N \left( 1 - \sum_{j=1}^M V_{ij} \right)^2
 \end{aligned} \tag{4.42}$$

Las constantes  $\mu_1$  a  $\mu_6$  se han de ajustar convenientemente para garantizar la rápida convergencia a la solución deseada. A continuación, se va a proceder a justificar uno a uno todos los términos que conforman la ecuación, analizando el efecto introducido por cada uno de ellos.

### **Primer Término de la Función de Energía**

Este término presentará un valor máximo en el estado óptimo de asignación, por eso tiene un signo negativo en la función de energía, de manera que, en el óptimo, la energía se reduzca lo máximo posible. En general, este término se encarga de favorecer las tasas binarias que garanticen la transmisión de los paquetes antes de expirar su *delay* máximo. Para ello, se introduce la función de coste  $C_{ij}$  que mide la ganancia de asignar al usuario  $i$  la tasa binaria  $R_j$ . Cada coste individual depende de la tasa binaria del usuario,  $R_j$ , y de su correspondiente tasa mínima,  $R_{\min,i}$ , definida en la ecuación (4.37).

Para obtener un buen comportamiento de este término, sería deseable que la función  $C_{ij}$  fuera monótonamente creciente y que estuviera acotada superior e inferiormente



por 1 y 0, de manera que  $C_{ij}$  sea igual a 1 si  $R_j$  es la tasa binaria máxima del usuario. Además, es necesario que haya un importante incremento para  $R_j \geq R_{\min,i}$ . Con estas condiciones se asegura que la función de coste no crezca incontrolablemente y que, además, la función de energía se minimice considerablemente al transmitir con tasas binarias al menos iguales a la mínima necesaria. Una función monótona, acotada y que presenta un gran crecimiento a partir de un determinado valor es la función sigmoïdal:

$$S(x, s, r) = \frac{1}{1 + e^{-s(x+r)}} \tag{4.43}$$

A partir de esta función sigmoïdal, la función de coste se ha definido como:

$$C_{ij} = \frac{S(R_j, s_i, r_i) - S(0, s_i, r_i)}{S(R_{\max,i}, s_i, r_i) - S(0, s_i, r_i)} \tag{4.44}$$

Con esta definición, la función de coste toma valores en el intervalo [0,1] para todas las tasas binarias del sistema. Así, si  $R_j = 0$  entonces  $C_{ij} = 0$  mientras que si  $R_j = R_{\max,i}$  entonces  $C_{ij} = 1$ . Los valores de  $s_i$  y de  $r_i$  se han definido como:

$$s_i = \begin{cases} \frac{2 \ln(9)}{R_{\min,i}} & R_{\min,i} \leq R_{\max,i} \\ \frac{2 \ln(9) R_{\min,i}}{(R_{\max,i})^2} & R_{\min,i} > R_{\max,i} \end{cases} \tag{4.45}$$

$$r_i = \begin{cases} -\frac{R_{\min,i}}{2} & R_{\min,i} \leq R_{\max,i} \\ -R_{\max,i} + \frac{(R_{\max,i})^2}{2R_{\min,i}} & R_{\min,i} > R_{\max,i} \end{cases}$$

De forma que, si  $R_{\min,i} \leq R_{\max,i}$ , cuando  $R_j = R_{\min,i}$  se cumple que:

$$S(R_{\min,i}, s_i, r_i) = \frac{1}{1 + e^{-\frac{2 \ln(9)}{R_{\min,i}} \left( R_{\min,i} - \frac{R_{\min,i}}{2} \right)}} = \frac{1}{1 + e^{-\ln(9)}} \approx 0.9$$

$$S(0, s_i, r_i) = \frac{1}{1 + e^{-\frac{2 \ln(9)}{R_{\min,i}} \left( 0 - \frac{R_{\min,i}}{2} \right)}} = \frac{1}{1 + e^{\ln(9)}} = 0.1 \tag{4.46}$$

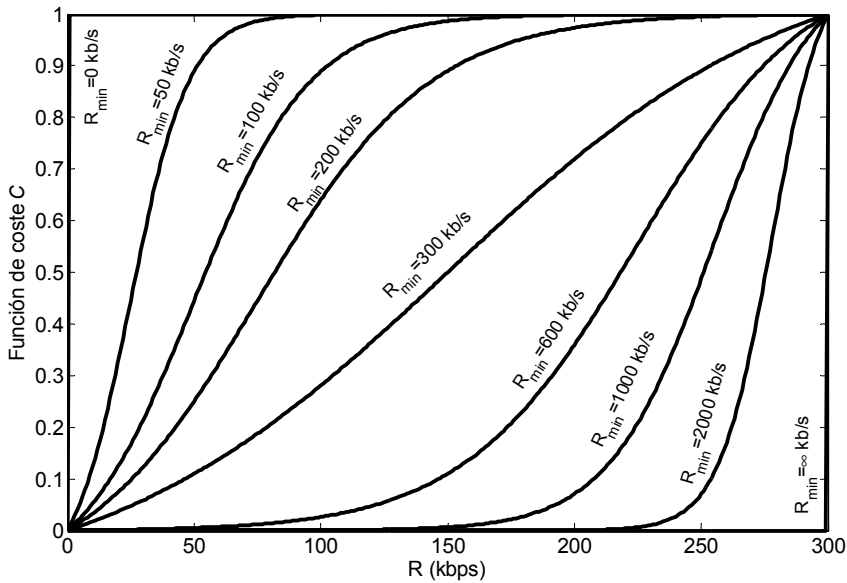


Figura 4.13 Efecto de  $R_{\min,i}$  en la función de coste para  $R_{\max,i} = 300$  kbps

mientras que  $S(R_{\max,i}, s_i, r_i)$  está siempre acotado entre 0.9 y 1, por lo que cuando se asigne exactamente la tasa mínima el valor de  $C_{ij}$  estará comprendido entre 0.889, cuando  $R_{\max,i}$  sea mucho mayor que  $R_{\min,i}$ , y 1 cuando  $R_{\min,i} = R_{\max,i}$ .

En la Figura 4.13 se ve claramente el efecto de la  $R_{\min,i}$  sobre la función de coste y cómo la asignación de una tasa igual a la tasa mínima tiene un coste que se encuentra entre los márgenes mencionados.

Además, en la Figura 4.13 se muestra cómo la función sigmoideal se escala sobre el eje de la tasa binaria asignada, presentando una forma de función escalón centrada en 0 kbps cuando la tasa mínima es de 0 kbps y de función escalón centrada en 300 kbps (máxima tasa binaria) para una tasa mínima infinita. Así, cuando se excede el *delay* máximo y  $R_{\min,i} = \infty$ , la función de coste toma un valor de cero para  $R_j < R_{\max,i}$  y de uno siempre que  $R_j \geq R_{\max,i}$ , de manera que así se garantiza que la única asignación de recursos válida sea la tasa binaria máxima del usuario.

Además, es importante destacar que, a pesar de que este término tiene como principal función garantizar el *delay*, al ser monótonamente creciente reduce la función de energía para valores crecientes de tasa binaria asignada, favoreciendo así la maximización del uso de los recursos.

Tanto el primer como el segundo término de la función de energía están multiplicados por el factor  $O_i$ . Este elemento fue concebido para ordenar a los usuarios en función de su consumo de potencia. Así, al usuario que experimenta las mejores condiciones de propagación se le asigna un factor  $O_i = 1$ , mientras que el usuario que consume más potencia tiene  $O_i = 0.5$ . En general, si se ordena  $N$  usuarios por orden de consumo creciente de recursos, al usuario  $i$ -ésimo se le asignará un factor  $O_i = 1 - (i-1)/2(N-1)$ . Ya que este factor de ordenación se ha introducido en los dos términos encargados de aumentar la salida de las neuronas, la red Hopfield favorecerá a aquellos usuarios que experimenten mejores condiciones de propagación en comparación con otros con funciones de coste similares.

### ***Segundo Término de la Función de Energía***

Este término se encarga de aumentar los recursos asignados. Al igual que el anterior, está acotado entre 0 y 1, ya que la tasa binaria asignada está dividida por la máxima tasa disponible del usuario  $i$ . No importa si el sistema tiene una cantidad de recursos limitada o no, y tampoco importa si tiene sentido el que un único usuario tenga varias neuronas activas (varias tasas binarias asignadas), este término siempre aumentará la salida de todas las neuronas, pero eso sí, cuanto mayor sea la tasa binaria que representa una neurona, más la aumentará.

Estos dos primeros términos serán los encargados de garantizar la calidad de servicio del usuario final, ya que el primer término se encarga de no exceder el *delay* máximo mientras que el segundo término maximizará el uso de los recursos.

### ***Tercer Término de la Función de Energía***

Este término penaliza las tasas binarias que provocan que el sistema asigne más recursos de los que realmente dispone. Gracias a él la red neuronal sólo alcanzará una situación estable en estados que satisfagan las restricciones de potencia disponible en la estación base y número de códigos de canalización disponibles en el árbol OVSF. Para ello se hace uso de la función  $\xi_{ij}$  que toma únicamente los valores 0 ó 1. El valor 1 indica que se ha de penalizar la asignación de una determinada tasa binaria mientras que el valor 0 indica lo contrario. La tasa binaria  $j$  del usuario  $i$  será penalizada, es decir  $\xi_{ij} = 1$ , si asignando dicha tasa binaria y sólo dicha tasa binaria al usuario  $i$ -ésimo ( $V_{ij} = 1$  y  $V_{il} = 0, \forall l \neq j$ ), y manteniendo la asignación de recursos del resto de usuarios según indique el estado actual de sus neuronas, el sistema sobrepasa el límite máximo de sus recursos.

La relación de energía de bit respecto a densidad espectral de potencia de ruido más interferencias para el usuario  $i$ -ésimo en UMTS viene dada por la siguiente expresión:

$$\left(\frac{E_b}{N_0}\right)_i = \frac{\left(\frac{P_i}{L_{p,i}}\right) \frac{W}{R_i}}{P_N + \chi_i + \rho \left(\frac{P_T - P_i}{L_{p,i}}\right)} \quad (4.47)$$

donde  $P_i$  es la potencia destinada al usuario  $i$  por la estación base,  $L_{p,i}$  representa las pérdidas de propagación del enlace entre estación base y terminal,  $P_N$  es el ruido térmico,  $\chi_i$  es la potencia interferente a la salida de la antena de recepción del usuario  $i$ ,  $\rho$  es el factor de ortogonalidad ( $\rho = 0$  para códigos perfectamente ortogonales) y  $P_T$  es la potencia total transmitida por el Nodo B. A partir de (4.47) y suponiendo una determinada relación  $E_b/N_0$  objetivo, se puede despejar la potencia asignada al usuario:

$$P_i = \frac{(P_N + \chi_i) \cdot L_{p,i} + \rho P_T}{\left( \rho + \frac{W}{\left( \frac{E_b}{N_0} \right)_i R_i} \right)} \quad (4.48)$$

La potencia total transmitida por el Nodo-B será exactamente la suma de todas las potencias transmitidas a los  $N$  usuarios más la potencia de los canales de control  $P_C$ :

$$P_T = \sum_{i=1}^N P_i + P_C = \sum_{i=1}^N \frac{(P_N + \chi_i) \cdot L_{p,i} + \rho P_T}{\left( \rho + \frac{W}{\left( \frac{E_b}{N_0} \right)_i R_i} \right)} + P_C \quad (4.49)$$

$$P_T = \sum_{i=1}^N \frac{(P_N + \chi_i) \cdot L_{p,i}}{\left( \rho + \frac{W}{\left( \frac{E_b}{N_0} \right)_i R_i} \right)} + \sum_{i=1}^N \frac{\rho P_T}{\left( \rho + \frac{W}{\left( \frac{E_b}{N_0} \right)_i R_i} \right)} + P_C$$

agrupando convenientemente

$$P_T = \frac{\sum_{i=1}^N \frac{(P_N + \chi_i) \cdot L_{p,i}}{\left( \rho + \frac{W}{\left( \frac{E_b}{N_0} \right)_i R_i} \right)} + P_C}{1 - \sum_{i=1}^N \frac{\rho}{\left( \rho + \frac{W}{\left( \frac{E_b}{N_0} \right)_i R_i} \right)}} \quad (4.50)$$

Finalmente, el factor de carga en enlace descendente se puede definir como:

$$\eta_{DL} = \frac{P_T}{P_{T \max}} \quad (4.51)$$

siendo  $P_{T_{\max}}$  la potencia máxima disponible en la estación base, que normalmente es de 43 dBm. El factor de carga máximo del sistema será  $\eta_{\max}$ , siendo por tanto la potencia transmitida máxima  $P_{T_{\max}} \cdot \eta_{\max}$ .

A partir de la expresión (4.50) y considerando el estado real de las neuronas, se puede calcular la potencia total consumida en la estación base ante una determinada asignación de recursos como:

$$P_T = \frac{\sum_{i=1}^N \sum_{j=1}^M \frac{(P_N + \chi_i) \cdot L_{P,i} V_{ij} + P_C}{D_{ij}}}{1 - \sum_{i=1}^N \sum_{j=1}^M \frac{\rho}{D_{ij}} V_{ij}} \quad (4.52)$$

donde

$$D_{ij} = \left( \rho + \frac{W}{\left( \frac{E_b}{N_0} \right)_{ij} R_{ij}} \right) \quad (4.53)$$

Con todo, el tercer término de la energía se podría definir inicialmente como:

$$\xi_{ij} = u \left( \frac{H_{ij}}{\eta_{\max} P_{T_{\max}}} - 1 \right) \quad (4.54)$$

$$H_{ij} = \frac{\sum_{l=1}^N \sum_{m=1}^M \frac{(P_N + \chi_l) \cdot L_{P,l} V_{lm} + (P_N + \chi_i) \cdot L_{P,i} + P_C}{D_{lm}}}{1 - \sum_{l=1}^N \sum_{m=1}^M \frac{\rho}{D_{lm}} V_{lm} - \frac{\rho}{D_{ij}}} \quad (4.55)$$

siendo  $u(\cdot)$  la función escalón:

$$u(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (4.56)$$

El análisis de la ecuación (4.54) muestra que si al asignar a un usuario una determinada tasa binaria, y por tanto una determinada  $D_{ij}$ , se sobrepasa la potencia máxima del sistema, entonces  $H_{ij} > \eta_{\max} P_{T \max}$ , por lo que  $\xi_{ij} = 1$ . Al ser el tercer término positivo, que se sobrepase la máxima potencia disponible en la estación base incrementará la función de energía, por lo que será penalizado. Por otro lado, si una determinada asignación de recursos no implica que se supere la potencia máxima entonces  $\xi_{ij} = 0$ , por lo que el tercer término se anulará.

Hasta ahora, según la definición que se ha realizado, el tercer término únicamente se encarga de controlar que no se exceda la potencia máxima disponible en el sistema. Sin embargo, las restricciones señaladas en el apartado 4.5.1 también incluían un concepto adicional: la limitación en el número de códigos disponibles.

La disponibilidad de códigos se puede modelar de manera sencilla con la correspondencia entre códigos y tasas binarias que establece la inecuación de Kraft [21]. Gracias a esta transformación, el problema de la disponibilidad de códigos se traduce en no superar la tasa máxima disponible en todo el sistema,  $R_{System, \max}$ . En caso de que una determinada asignación de recursos ( $i, j$ ) sobrepase este límite, la asignación será penalizada.

Finalmente quedará que:

$$\xi_{ij} = u \left( \frac{H_{ij}}{\eta_{\max} P_{T \max}} - 1 \right) \parallel u \left( \frac{\sum_{\substack{l=1 \\ l \neq i}}^N \sum_{m=1}^M R_m V_{lm} + R_j}{R_{System, \max}} - 1 \right) \quad (4.57)$$

siendo  $\parallel$  el operador or.

Adicionalmente, este tercer término se pondera por el factor  $R_j/R_{\max,i}$ , de manera que la penalización sea mayor para tasas binarias altas, ya que éstas consumen más recursos y son por tanto más críticas.

#### ***Cuarto Término de la Función de Energía***

Este término se encarga de evitar el uso de tasas binarias no permitidas para un determinado tipo de servicio. La matriz  $\psi_{ij}$  será igual a 1 para aquellas tasas no permitidas para el usuario  $i$  e igual a 0 para aquellas tasas sí permitidas. Matemáticamente:

$$\psi_{ij} = \begin{cases} 0 & R_j \in \mathfrak{R}_i \\ 1 & R_j \notin \mathfrak{R}_i \end{cases} \quad (4.58)$$

De esta forma, si se activan servicios no permitidos, su neurona correspondiente incrementará la función de energía a minimizar. Este término se podrá utilizar para definir distintos tipos de usuarios, con mayores o menores privilegios o para garantizar una determinada tasa mínima al usuario.

#### ***Quinto Término de la Función de Energía***

El quinto término fuerza a las neuronas a irse a los extremos, o bien  $V_{ij} = 0$  o bien  $V_{ij} = 1$ , ya que, para estos casos, este término es igual a 0, mientras que para el resto este término perjudica, o sea, aumenta la función de energía.

#### ***Sexto Término de la Función de Energía***

Si el quinto término empuja a las neuronas a los extremos del hipercubo, el sexto garantiza, para cada usuario, que solamente se asigne una tasa binaria. De esta manera, ambos, el quinto y sexto término, se combinan para obtener una única neurona activa al final de las iteraciones.



### 4.5.2.3 Dinámica del sistema y parámetros *hardware*

El algoritmo de Hopfield se inicia con todas las neuronas inactivas,  $V_{ij} = 0 + \varepsilon$ , siendo  $\varepsilon$  una variable aleatoria uniformemente distribuida en el intervalo  $[0, \varepsilon_m]$ .

La evolución de la tensión de las neuronas se realizará considerando la expresión (4.41), teniendo en cuenta que el gradiente de  $E$  resulta ser:

$$\frac{\partial E}{\partial V_{ij}} = -\mu_1 O_i C_{ij} - \mu_2 \frac{O_i R_j}{R_{\max,i}} + \frac{\mu_3}{2} \cdot \frac{R_j}{R_{\max,i}} \xi_{ij} + \frac{\mu_4}{2} \psi_{ij} + \frac{\mu_5}{2} (1 - 2V_{ij}) - \mu_6 \left( 1 - \sum_{l=1}^M V_{il} \right) \quad (4.59)$$

Todas las salidas,  $V_{ij}$ , se calculan cada iteración utilizando (4.39) y la solución de (4.41).

Ahora solo resta identificar los parámetros *hardware* de la red Hopfield propuesta. Considerando la estructura matricial de las neuronas, la nueva formulación del gradiente de la función de energía resulta ser:

$$\frac{\partial E}{\partial V_{ij}} = -\sum_{k=1}^N \sum_{l=1}^M T_{ij,kl} V_{kl} - I_{ij} \quad (4.60)$$

Ahora, por comparación directa entre (4.59) y (4.60) se puede obtener la expresión de la matriz  $\overline{\overline{T}}$  y del vector de corrientes:

$$\begin{aligned} T_{ij,kl} &= \mu_5 \cdot \delta_{ik} \cdot \delta_{jl} - \mu_6 \cdot \delta_{ik} \\ I_{ij} &= \mu_1 O_i C_{ij} + \mu_2 \frac{O_i R_j}{R_{\max,i}} - \frac{\mu_3}{2} \cdot \frac{R_j}{R_{\max,i}} \xi_{ij} - \frac{\mu_4}{2} \psi_{ij} - \frac{\mu_5}{2} + \mu_6 \end{aligned} \quad (4.61)$$

Como se puede apreciar, la matriz  $\overline{\overline{T}}$  es simétrica y el valor de las resistencias constante, mientras que las corrientes variarán dinámicamente en función del estado de la red y de las características de transmisión de cada usuario.

#### 4.5.2.4 Cálculo de las constantes

Para determinar el cálculo de las constantes se ha de analizar el comportamiento de la red en casos extremos, de manera que aún en estos se garantice el funcionamiento deseado de la red neuronal. Para empezar, el valor de  $\mu_1$  y  $\mu_2$  se puede elegir con bastante libertad mientras que el resto de coeficientes se derivarán de ellos. Sin embargo, es importante darse cuenta que será precisamente la relación entre estos dos parámetros la que controlará qué importancia se le da a los dos objetivos que se pretende conseguir: minimizar el *delay* y maximizar el uso de los recursos. Si minimizar el *delay* es más importante que la maximización del *throughput*, entonces  $\mu_1 > \mu_2$  y cuanta más diferencia haya entre estos dos factores, mayor será la importancia relativa dada al objetivo de minimizar el *delay*.

A continuación, se analizarán las condiciones que deberán cumplir cada una de las constantes. El orden del análisis es el necesario para poder ir despejando estas condiciones una tras otra.

##### ***Quinto Término***

Este término solo sirve para aumentar la velocidad de convergencia de la red neuronal. Lo que se deberá asegurar es que nunca este término evite que una neurona pase del estado 0 a 1, en caso de que el resto de términos así lo indiquen. Cuanto mayor sea este término más rápida y mejor será la convergencia de la red pero peores los resultados obtenidos. Para no perjudicar el funcionamiento del sistema se deberá garantizar que, incluso en el peor caso, este término no impida salir de un mínimo local.

El caso peor se presenta cuando una solución no óptima con una determinada tasa binaria para un usuario alcanza  $V_{ij}=1$  mientras que otra solución que podría maximizar más los recursos está en  $V_{ik}=0$ . Suponiendo que ambas tasas binarias están permitidas, que garantizan sobradamente la tasa mínima necesaria y que no sobrepasan la carga máxima del sistema, entonces el gradiente será para el primer caso:

$$\frac{\partial E}{\partial V_{ij}} = -\mu_1 O_i - \mu_2 \frac{O_i R_j}{R_{\max,i}} - \frac{\mu_5}{2} \quad (4.62)$$

siendo para el segundo caso:

$$\frac{\partial E}{\partial V_{ik}} = -\mu_1 O_i - \mu_2 \frac{O_i R_k}{R_{\max,i}} + \frac{\mu_5}{2} \quad (4.63)$$

Como la segunda opción maximiza los recursos deberá tener un gradiente más negativo que aumente el valor de la neurona por lo que:

$$\begin{aligned} -\mu_1 O_i - \mu_2 \frac{O_i R_k}{R_{\max,i}} + \frac{\mu_5}{2} < -\mu_1 O_i - \mu_2 \frac{O_i R_j}{R_{\max,i}} - \frac{\mu_5}{2} \\ \mu_5 < \mu_2 \cdot \frac{O_i (R_k - R_j)}{R_{\max,i}} \end{aligned} \quad (4.64)$$

En el escenario de simulación planteado el peor caso se dará para  $O_i = 0.5$ ,  $R_k = 32$  kbps y  $R_j = 16$  kbps quedando finalmente que:

$$\mu_5 < \mu_2 \cdot \frac{16}{512} \quad (4.65)$$

### ***Tercer Término***

El objetivo de la red neuronal es asignar la máxima tasa binaria disponible por el usuario sin exceder los recursos máximos. Para ello, al menos una de las neuronas del usuario debe estar más favorecida (incrementando su valor más rápidamente o por lo menos decrecentándose más lentamente) que las neuronas que sí que impliquen exceder los máximos recursos disponibles. Suponiendo que el usuario puede hacer uso de todas las tasas binarias disponibles, en el caso de la neurona favorecida ( $i, fav$ ) el gradiente de la energía es:

$$\frac{\partial E}{\partial V_{i,fav}} = -\mu_1 O_i C_{i,fav} - \mu_2 \frac{O_i R_{fav}}{R_{\max,i}} + \frac{\mu_5}{2} (1 - 2V_{i,fav}) - \mu_6 \left( 1 - \sum_l^M V_{il} \right) \quad (4.66)$$

Mientras que el gradiente de la energía de las neuronas que sí que exceden los máximos recursos del sistema ( $i, exc$ ) es:

$$\frac{\partial E}{\partial V_{i,exc}} = -\mu_1 O_i C_{i,exc} - \mu_2 \frac{O_i R_{exc}}{R_{max,i}} + \frac{\mu_3}{2} \frac{R_{exc}}{R_{max,i}} + \frac{\mu_5}{2} (1 - 2V_{i,exc}) - \mu_6 \left( 1 - \sum_{l=1}^M V_{il} \right) \quad (4.67)$$

La condición necesaria para garantizar que se realice una asignación correcta de recursos es que:

$$\frac{\partial E}{\partial V_{i,fav}} < \frac{\partial E}{\partial V_{i,exc}} \quad (4.68)$$

con lo que

$$\mu_3 > 2\mu_1 O_i \frac{R_{max,i}}{R_{exc}} (C_{i,exc} - C_{i,fav}) + 2\mu_2 \frac{O_i (R_{exc} - R_{fav})}{R_{exc}} + 2\mu_5 \frac{R_{max,i}}{R_{exc}} (V_{i,exc} - V_{i,fav}) \quad (4.69)$$

El caso peor se encuentra cuando  $C_{i,exc} = 1$ ,  $C_{i,fav} = 0$ ,  $R_{fav} = 0$ ,  $V_{i,exc} = 1$ ,  $V_{i,fav} = 0$  y  $O_i = 1$ . En este caso:

$$\mu_3 > 2\mu_1 \frac{R_{max,i}}{R_{exc}} + 2\mu_2 + 2\mu_5 \frac{R_{max,i}}{R_{exc}} \quad (4.70)$$

Dado el caso particular de las tasas binarias definidas en esta Tesis Doctoral,  $\mu_3$  deberá cumplir:

$$\mu_3 > 32\mu_1 + 2\mu_2 + 32\mu_5 \quad (4.71)$$

### Sexto Término

Aunque se disponga de suficientes recursos, nunca se debe asignar más de una tasa binaria a un único usuario. Para ello, el sexto término ha de ser más importante que los dos primeros. Sin embargo, este término depende en gran medida del estado actual y, además, si el sistema se encuentra en un punto de equilibrio se anulará, ya que

$$\sum_{l=1}^M V_{il} = 1 \quad (4.72)$$

Si el sistema se encuentra en un punto de equilibrio, los dos primeros términos lo sacarán de él ya que el sexto no tiene fuerza alguna (es nulo). Lo que se debe forzar es que el sexto término sea más importante que los dos primeros cuando el sistema se haya separado una distancia  $\delta$  del punto de equilibrio. En esta situación el peor caso se obtiene si en el punto de equilibrio hay alguna neurona correspondiente a la máxima tasa binaria a cero.

En este caso los dos primeros términos aumentarán esta neurona al ritmo:

$$-\mu_1 O_i - \mu_2 O_i \quad (4.73)$$

Mientras que al sumar la salida de todas las neuronas de un mismo usuario más que la unidad, el sexto término reducirá esta neurona con un factor:

$$\delta\mu_6 \quad (4.74)$$

Por lo tanto:

$$|-\mu_1 O_i - \mu_2 O_i| < \delta\mu_6 \quad (4.75)$$

con lo que, en el peor caso,  $\mu_6$  deberá garantizar:

$$\mu_6 > \frac{\mu_1 + \mu_2}{\delta} \quad (4.76)$$

#### **Cuarto Término**

Por último si  $\psi_{ij} = 1$ , el cuarto término ha de disminuir la salida de esa neurona sin importar lo que valgan el resto de términos.

El caso en el que el incremento de la neurona es mayor se produce cuando  $C_{ij} = 1$ ,  $R_{ij} = R_{\max,i}$  y  $\xi_{ij} = 0$ , siendo el gradiente de la energía en este caso:

$$\begin{aligned} \frac{\partial E}{\partial V_{ij}} &= -\mu_1 O_i - \mu_2 O_i + \frac{\mu_4}{2} + \frac{\mu_5}{2} (1 - 2V_{ij}) - \mu_6 \left( 1 - \sum_{l=1}^M V_{il} \right) > 0 \\ \mu_4 &> 2\mu_1 O_i + 2\mu_2 O_i - \mu_5 (1 - 2V_{ij}) + 2\mu_6 \left( 1 - \sum_{l=1}^M V_{il} \right) \end{aligned} \quad (4.77)$$

Como  $\mu_5 < \mu_6$ , el peor caso se da cuando  $V_{il} = 0, \forall l$ . Finalmente se ha de cumplir que:

$$\mu_4 > 2\mu_1 + 2\mu_2 - \mu_5 + 2\mu_6 \quad (4.78)$$

### ***Proceso de cálculo de todas las constantes***

Resumiendo, el proceso de selección de las constantes sería el siguiente

- Dar un valor a  $\mu_1$  y  $\mu_2$
- A partir de (4.65) elegir el valor de  $\mu_5$
- Utilizando la expresión (4.71) calcular el valor de  $\mu_3$
- Con (4.76) seleccionar  $\mu_6$
- Finalmente, calcular  $\mu_4$  a partir de (4.78)

En concreto en las simulaciones realizadas se ha considerado un valor de  $\mu_1 = \mu_2 = 500$  y un valor de  $\delta = 0.1$ , por lo que el resto de constantes tienen los siguientes valores:  $\mu_3 = 15$ ,  $\mu_3 = 17500$ ,  $\mu_6 = 10000$  y  $\mu_4 = 22000$ .

### **4.5.3 Evaluación de resultados y discusión**

Para evaluar las prestaciones del algoritmo DC-HNN-DRA propuesto en esta Tesis Doctoral para la gestión de recursos en el enlace descendente de UMTS, se ha seleccionado un escenario de simulación con cuatro tipos de servicios: voz, vídeo conferencia, navegación web y descarga FTP. Cada uno de estos servicios es suficientemente representativo como para estudiar el funcionamiento del algoritmo para distintos tipos de tráfico y para la combinación de ellos.

En este apartado se describirá, en primer lugar, algunos detalles importantes de los modelos de tráfico. A continuación, se resumirá el funcionamiento del resto de algoritmos de gestión de recursos encontrados en la literatura y utilizados para comparar las prestaciones del DC-HNN-DRA. Finalmente, se compararán los resultados obtenidos para cada una de las combinaciones de servicios consideradas.

#### **4.5.3.1 Tipos de servicios**

Una descripción detallada de los modelos de tráfico utilizados en esta Tesis se puede encontrar en el Anexo B. La navegación web se ha seleccionado como ejemplo clásico de servicio interactivo NRT, mientras que la descarga FTP se puede considerar como un servicio interactivo o como un servicio de *background*. Realmente la descarga FTP es interactiva porque normalmente el usuario espera a que finalice la descarga. Además, si el usuario aprecia que ésta se está realizando muy lentamente sentirá la denominada ‘ansiedad de descarga’. Obviamente, el tipo de interactividad no es el mismo que para la descarga de una página web, porque el tiempo de espera es mucho mayor dependiendo del tamaño del fichero. Como el tiempo de espera puede llegar a ser muy alto, algunos autores lo consideran como un servicio de *background*. En función de cómo se considere este servicio, los algoritmos de gestión de recursos le darán mayor o menor prioridad. Es por esto por lo que se ha decidido evaluar ambas variantes, comparando el comportamiento de los distintos algoritmos considerados.

Respecto al servicio de voz, claramente será considerado como el servicio más prioritario puesto que se trata de un servicio de tiempo real muy sensible a las variaciones del tiempo de transmisión y que requiere, por tanto, un tratamiento similar al ofrecido por conmutación de circuitos.

El servicio de vídeo conferencia se ha implementado mediante H.263. La vídeo conferencia también se considera un servicio de tiempo real, pero en este caso, al tratarse de la transmisión discontinua de tramas de vídeo, se puede servir utilizando técnicas de conmutación de paquetes. Es importante destacar que no se ha considerado

ninguna técnica de *buffering* en el receptor por lo que, en caso de que una trama de vídeo no se haya transmitido antes de la llegada de la siguiente trama, se descartará la trama antigua considerándola como perdida. Como la generación de tramas no se realiza a tasa constante, sino que se trata de una generación variable, habrá que ser bastante conservador en la especificación de los tiempos máximos de transmisión de la trama, para así evitar su pérdida.

Respecto a  $R_{\min,i}$ , se ha calculado según la expresión (4.37). En caso de que se considere el servicio FTP como servicio de *background*, su  $R_{\min,i} = 0$ , puesto que su máximo *delay* es infinito. Por otro lado,  $R_{\min,i} = \infty$  para usuarios de voz, ya que en este caso los datos se han de transmitir inmediatamente.

#### 4.5.3.2 Algoritmos de gestión de recursos de referencia

A continuación, se describen otros algoritmos de gestión de recursos para redes UMTS convencionales. Estos algoritmos servirán para comparar las prestaciones del nuevo algoritmo de cronoejecución, DC-HNN-DRA, desarrollado mediante redes neuronales de Hopfield.

##### ***Round Robin (RR)***

La técnica de *Round Robin* [31] es sin duda una de las más clásicas dentro de la gestión de recursos radio y, a pesar de su sencillez, es ampliamente utilizada en las comunicaciones móviles. Se basa en que todos los usuarios tienen el mismo nivel de prioridad y, por tanto, los recursos se van distribuyendo equánime y cíclicamente. De esta manera, se establece un intervalo de tiempo predeterminado que constituye el intervalo de asignación de recursos. Por otra parte, se construye una lista con todos los usuarios activos. En el primer intervalo de tiempo comienzan transmitiendo los  $N$  primeros usuarios de la lista a la máxima tasa binaria disponible en el sistema, estando  $N$  determinado por las restricciones de potencia y disponibilidad de códigos. Por



ejemplo, si el hecho de que el usuario  $N+1$  también transmita hace que se sobrepase el límite de recursos establecido, entonces transmitirán los  $N$  primeros usuarios. Tras este primer turno transmitirán el usuario  $N+1$  y los  $M$  siguientes, calculando  $M$  de la misma manera que  $N$  en el primer turno. Es importante destacar que, aunque a los usuarios se les debería asignar la máxima tasa binaria disponible en el sistema, esto no siempre es posible por la limitación establecida por la red en la potencia máxima. En caso de que se supere esta potencia máxima, el usuario conserva su turno pero reduciendo su tasa de transmisión para no sobrepasar la potencia máxima. El procedimiento se repetirá hasta que se finalice con todos los usuarios de la lista, momento en el que se empieza de nuevo desde el principio.

Cuando hay multi-servicio se gestionan dos listas, una para servicios RT, es decir, voz y vídeo conferencia, y otra para servicios NRT, web y FTP. Siempre se empieza asignando recursos a los usuarios RT y entre ellos a los de voz, puesto que son los más prioritarios. Un usuario de voz solamente se quedará sin recursos en caso de que las condiciones de propagación o interferencias hagan que sea imposible satisfacer simultáneamente a todos los usuarios de voz activos. A continuación, se dará servicio a los usuarios de vídeo conferencia, asignándoles la tasa máxima establecida para ese servicio.

Respecto a los servicios NRT, se han considerado dos opciones. En la primera no se diferencia entre navegación web y FTP, ya que se suponen ambos servicios interactivos. En la segunda se sirve primero a los usuarios web y posteriormente a los FTP.

### ***Weighted Round Robin (WRR)***

El algoritmo RR asigna la misma tasa binaria al servicio web y al servicio FTP. El algoritmo WRR [32] es similar al RR salvo por el hecho de que se introducen unos pesos para priorizar los distintos tipos de servicios. En última instancia los pesos indican qué tasa binaria hay que asignar al siguiente usuario de la lista. Gracias al uso del WRR se asigna una tasa binaria a los usuarios web,  $R_{web}$ , distinta a la asignada a

los usuarios FTP,  $R_{FTP}$ . Además de asignar distintas tasas binarias a web y FTP, se puede diferenciar completamente ambos tipos de servicios, tal y como se hacía en RR, dando primero recursos a los usuarios web.

### ***Optimum Bit Rate (OBR)***

Este algoritmo, definido en [33], tiene en cuenta el retardo en la transmisión de paquetes, una importante métrica de calidad de servicio de un usuario.

En el algoritmo OBR se intenta asignar a todos los usuarios la tasa mínima,  $R_{\min,i}$ , que garantice su *delay* máximo. Tras establecer una lista de usuarios activos, se va seleccionando de manera aleatoria al siguiente usuario al que se le asignará exactamente su tasa binaria  $R_{\min,i}$ . Si llegado un determinado usuario se exceden los recursos máximos del sistema, entonces ese usuario se descarta y se intenta dar servicio al resto de usuarios en cola aún no asignados hasta finalizar la lista de usuarios o agotar los recursos disponibles. Como es posible que  $R_{\min,i} \notin \mathfrak{R}_i$ , el algoritmo asigna la tasa binaria más cercana que esté incluida dentro del subconjunto  $\mathfrak{R}_i$ .

Cuando se trabaja con múltiples servicios, igual que en los anteriores algoritmos, primero se sirve a los usuarios de voz, después a los de vídeo conferencia y finalmente a los NRT con o sin diferenciación entre web y FTP.

### ***Descend Bit Rate (DBR)***

Este nuevo algoritmo, llamado DBR, es una variante orientada a *delay* del algoritmo MaxCIR [6]. El DBR asigna inicialmente la máxima tasa binaria a todos los usuarios e iteración a iteración va reduciendo la tasa binaria. El objetivo es, como siempre, alcanzar una asignación de recursos que no sobrepase los recursos máximos disponibles. Cada iteración se reducirá la tasa binaria asignada al usuario que consuma más recursos radio. El proceso de reducción de tasa se divide en dos fases. En la primera, el DBR nunca asigna tasas binarias menores a la  $R_{\min,i}$  de cada usuario, de manera que si no hay recursos suficientes todos los usuarios acaben teniendo asignada

su  $R_{\min,i}$  correspondiente. En la segunda fase, si la asignación de recursos aún sobrepasa el máximo, el algoritmo sigue reduciendo la tasa binaria pero ahora sin un límite inferior definido. De esta manera se garantiza, siempre que sea posible, la asignación de la  $R_{\min,i}$  y, por tanto, no sobrepasar el retardo máximo de los paquetes.

Con múltiple servicio, el procedimiento de diferenciación de servicios será el mismo que para el resto de algoritmos.

### ***Priorized Earliest Deadline First (PEDF)***

En este algoritmo se prioriza a los usuarios por *deadline* [34]. Primero, el algoritmo encuentra al usuario cuyo objeto a transmitir esté más cerca de exceder su *deadline*. Después, se asigna a este usuario su máxima tasa binaria asignable. Este proceso de búsqueda se realiza priorizando por tipo de servicio de manera que, una vez acabado de asignar a todos los usuarios de una clase de servicio, se pasa al siguiente. El proceso se continúa hasta que no quedan más recursos disponibles, en cuyo caso se baja la tasa binaria asignada al último usuario hasta que la asignación sea factible.

Después de la ejecución de todos los algoritmos, la asignación resultante se optimiza en términos de uso de recursos utilizando un procedimiento adicional, el *Minimum Noise Rise* (MNR). Este ajuste se basa en subir la tasa binaria al usuario que menor nivel de interferencia esté introduciendo en el sistema. Esta subida de tasa binaria se realiza siempre que dicha asignación sea factible y que no se sobrepasen los recursos máximos. El MNR tiene como objetivo maximizar la cantidad total de recursos asignados.

### **4.5.3.3 Escenario de simulación**

A continuación, se resumen algunos detalles importantes en cuanto a las simulaciones llevadas a cabo en este apartado.

El radio celular se ha definido de 500m. La potencia máxima disponible es de 43 dBm y la potencia transmitida por las celdas interferentes es de 40dBm. Se ha considerado un modelo sencillo de *shadowing* lineal (ver Capítulo 6 para más detalles), con una desviación estándar de 6 dB. Los usuarios se mueven a una velocidad constante de 50 km/h. La potencia de ruido es de -102 dBm. El ancho de banda de transmisión,  $W$ , es de 3.84 Mchips/s. El factor de ortogonalidad,  $\rho$ , se a definido igual a 0.5. Los algoritmos de asignación de recursos se ejecutan cada 0.1 segundos [19].

El valor de  $\eta_{\max}$  es de 0.7 y, por tanto, la potencia máxima asignable será de 14W. Para el cálculo de  $R_{\text{System,max}}$  se ha seguido el estándar en cuanto a que una rama de SF 256 se utiliza para la señalización de cada usuario DSCH. Además, cinco códigos adicionales de SF 256 se reservan para los canales comunes y de *broadcast*.

El conjunto de tasas binarias disponibles y los requerimientos de  $E_b/N_0$  se han extraído del Capítulo 3. Para la transmisión de vídeo conferencia se ha supuesto una fuente de tráfico medio de 32 kbps, limitando la tasa máxima asignable a este servicio a 64 kbps, tal y como indican las especificaciones. El resto de servicios no están acotados en cuanto a su tasa máxima, que será de 256 kbps para web y FTP y de 12.2 kbps para voz. El *delay* máximo es distinto en función del servicio: 0ms para voz, es decir, máxima prioridad, 100ms para una trama H.263, 30s para una descarga completa de página web y 30 minutos o  $\infty$  para la descarga de FTP, en función de si se considera a este servicio como interactivo o *background*. Respecto al algoritmo WRR,  $R_{\text{web}} = 256$  kbps y  $R_{\text{FTP}} = 64$  kbps.

Respecto a la red Hopfield,  $\Delta t = 10^{-4}$  y  $\alpha = 1$ . El valor de las constantes es el calculado en el apartado anterior mientras que la tolerancia considerada para finalizar las iteraciones del algoritmo será  $tol = 10^{-5}$ . El número máximo de iteraciones se ha establecido en 10.000.

#### 4.5.3.4 Discusión de los resultados de simulación

Se ha evaluado la calidad de servicio experimentada por los usuarios en el escenario multi-servicio planteado, aumentando progresivamente la carga de la red. Además, se ha tratado de evaluar separadamente los distintos tipos de servicio para así diferenciar el comportamiento de los algoritmos por servicio. En una primera serie de simulaciones se ha estudiado únicamente el funcionamiento de los algoritmos con servicios de RT. A continuación, se ha procedido con un análisis similar, pero esta vez evaluando servicios NRT. Finalmente, se ha estudiado el comportamiento conjunto de ambos tipos de servicios.

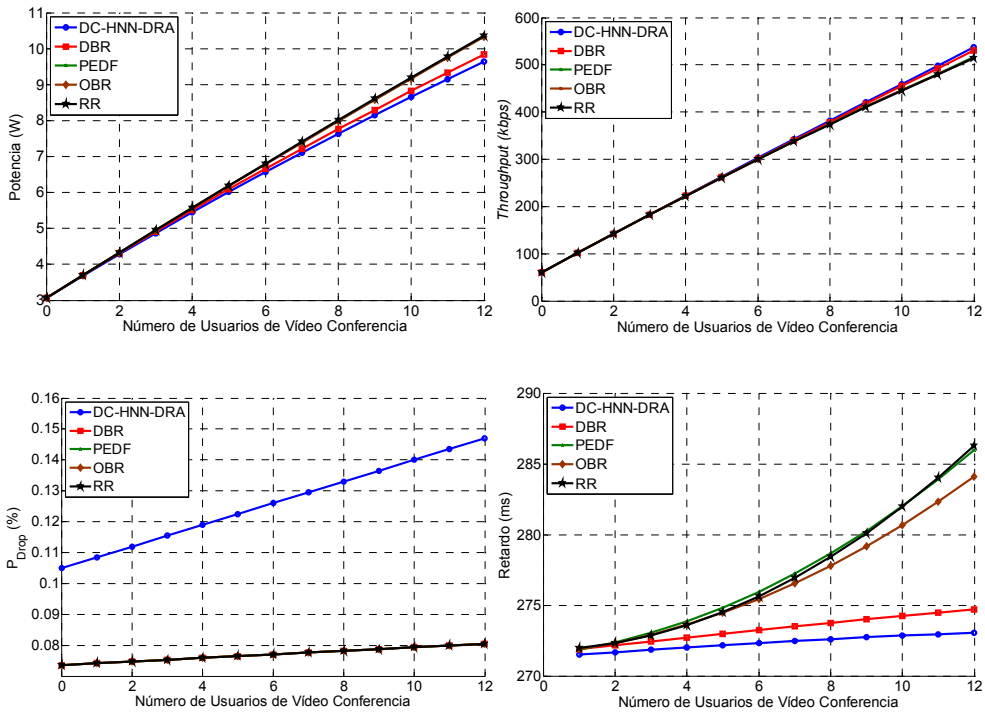
A lo largo del resto de la sección, la mejora en el *delay* conseguida por el algoritmo DC-HNN-DRA se ha calculado utilizando la siguiente expresión:

$$\text{Mejora (\%)} = 100 \cdot (1 - \text{delay}_{\text{DC-HNN-DRA}} / \text{delay}_{\text{Referencia}}) \quad (4.79)$$

#### *Comportamiento con servicios RT*

En primer lugar se han estudiado las prestaciones del algoritmo DC-HNN-DRA cuando se sirve conjuntamente a usuarios de voz y de vídeo conferencia. Se han introducido en el sistema 5 usuarios de voz junto con un número creciente de usuarios de vídeo conferencia, desde 0 hasta 12. En la Figura 4.14 se muestran gráficamente los principales indicadores de comportamiento del sistema, comparando las prestaciones de todos los algoritmos considerados, salvo el WRR, cuyo comportamiento es exactamente el mismo que el mostrado por el RR ya que sólo se diferencian para servicios NRT.

En primer lugar, el algoritmo DC-HNN-DRA es el que presenta un menor consumo de potencia, a la vez que es también quien maximiza más la tasa binaria asignada en el sistema. Sin embargo, la diferencia respecto al resto de algoritmos no es muy significativa. De entre todos los algoritmos comparados, el que más se asemeja al DC-HNN-DRA es el DBR, basado en MaxCIR. Se demuestra, por tanto, que la minimización del consumo de recursos es fundamental en UMTS.



**Figura 4.14** Prestaciones de los algoritmos de DRA para servicios RT. Arriba izquierda, potencia consumida por el Nodo-B. Arriba derecha, tasa binaria total asignada en el sistema. Abajo izquierda, probabilidad de pérdida de tramas de voz. Abajo derecha, retardo medio de transmisión de tramas de video conferencia.

Excepto DC-HNN-DRA, el resto de algoritmos presentan las mismas prestaciones en cuanto a la probabilidad de pérdida de trama de voz, o FER. Esto es debido a que el algoritmo utilizado en todos estos algoritmos para asignar recursos a la voz es el mismo: darle la máxima calidad a la voz siempre que sea posible y si no reducir a aquellos que consuman más recursos. Es interesante destacar que la FER crece levemente al aumentar el número de usuarios de video conferencia. Esto es debido a que al haber más nivel de ruido se cometerán más errores en el control de potencia rápido, por lo que se descartarán paquetes con una mayor probabilidad aun teniendo la misma tasa asignada.

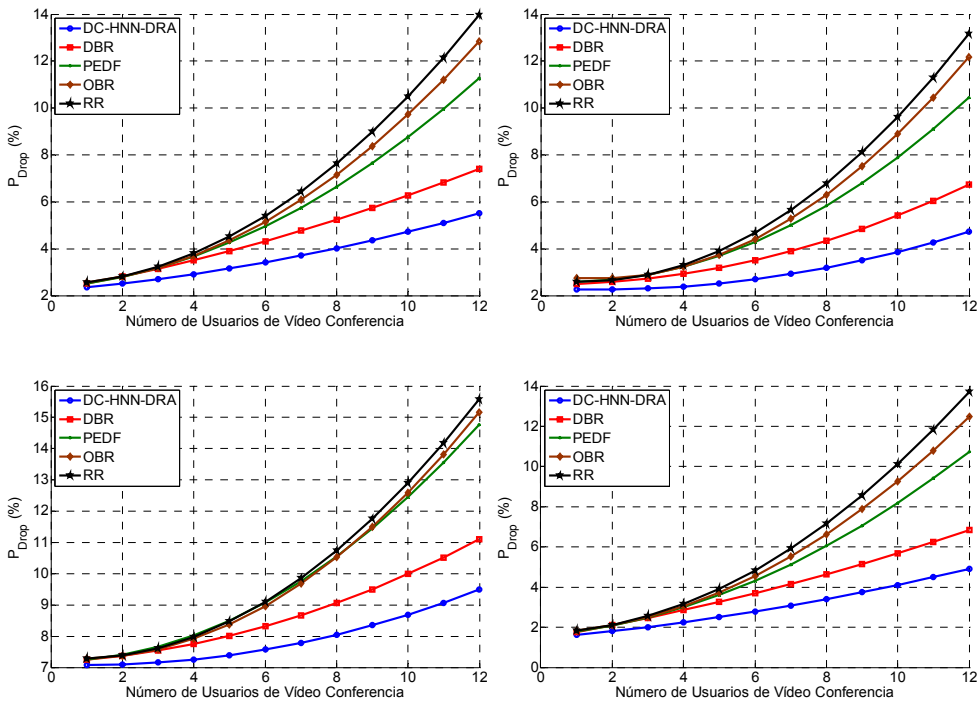
El FER medio del servicio de voz para estos algoritmos está alrededor del 0.08%, mientras que el valor mostrado por DC-HNN-DRA es ligeramente superior. Además, como el algoritmo neuronal no separa totalmente ambos servicios RT, salvo por los distintos requerimientos de *delay* máximo, es normal observar que la FER crece con mayor rapidez para DC-HNN-DRA. A pesar de que el nuevo algoritmo no es capaz de alcanzar la misma calidad que el resto, los valores registrados son bastante satisfactorios y siempre inferiores al 0.15%. Este valor está muy por debajo del típico valor objetivo del 1% [1], con lo que se puede afirmar que el algoritmo DC-HNN-DRA asegura también una buena satisfacción de los usuarios de voz.

Respecto al *delay* percibido por los usuarios de vídeo conferencia, de nuevo el algoritmo DC-HNN-DRA presenta las mejores prestaciones, aunque en porcentajes no se aprecia una mejora especialmente significativa. Es importante destacar que este *delay* se ha medido únicamente para aquellas tramas que se han transmitido con éxito, no considerando las tramas descartadas.

Es precisamente al estudiar la probabilidad de descarte de tramas de vídeo cuando más se aprecia la enorme mejora que supone utilizar el nuevo algoritmo propuesto en esta Tesis Doctoral. La Figura 4.15 muestra gráficamente el valor de probabilidad de pérdida de trama, tanto en valor medio como distinguiendo los distintos tipos de trama que se transmiten en H.263, es decir, tramas tipo I, P y PB.

Considerando todas las tramas de vídeo, el algoritmo DC-HNN-DRA consigue una mejora del 25% respecto al mejor del resto de algoritmos, el DBR, y de hasta un 60% respecto a RR, para el caso de máxima carga.

Esta mejora es muy significativa, sobre todo teniendo en cuenta que el valor objetivo de referencia de pérdida de tramas se encuentra en torno al 5%. Solamente con el algoritmo DC-HNN-DRA es posible garantizar una calidad de servicio aceptable para un número relativamente alto de usuarios de vídeo conferencia.



**Figura 4.15** Probabilidad de descarte de tramas de vídeo. Arriba izquierda, valor medio considerando todos los tipos de trama. Arriba derecha, tramas tipo I. Abajo izquierda, tramas tipo P. Abajo derecha, tramas tipo PB.

Analizando los distintos tipos de trama, son las tramas tipo P las que más sufren pérdidas. Esto es debido a que este tipo de tramas son las que tienen una duración menor, por lo que son las más sensibles al retardo en la transmisión. También hay que resaltar que la mejora obtenida gracias al uso del nuevo algoritmo es mucho más significativa para las tramas tipo I y tipo PB. Esto es debido a que estas tramas son las que tienen un mayor tamaño, por lo que será en las que la red Hopfield ponga un mayor esfuerzo para su transmisión a tiempo. Respecto al resto de algoritmos comparados, de nuevo el DBR es el único que consigue aproximarse a las prestaciones del DC-HNN-DRA. El RR es el que presenta las peores prestaciones puesto que su funcionamiento no está basado en *delay*.

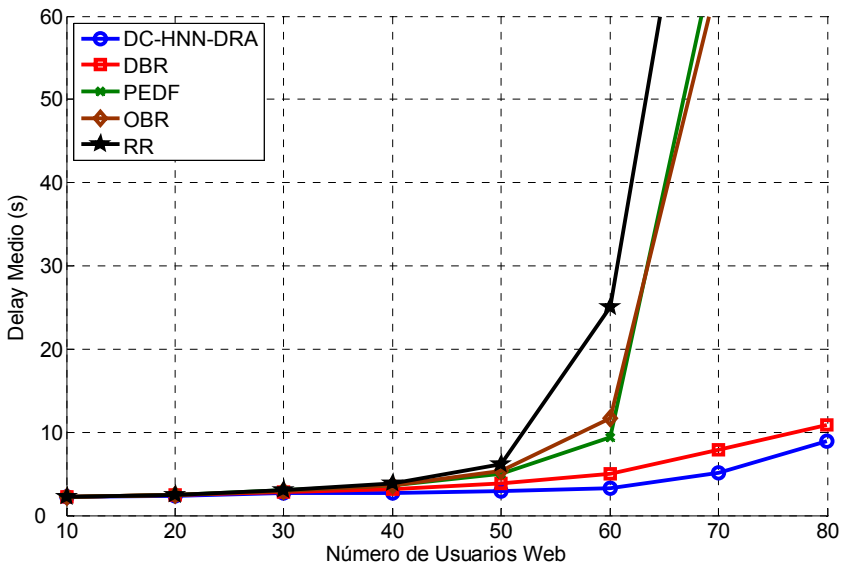


En general, la conclusión alcanzada tras el análisis de los servicios de tipo RT es que el algoritmo DC-HNN-DRA es capaz de mejorar considerablemente el comportamiento de la vídeo conferencia a costa de un ligero empeoramiento del servicio de voz. Por lo tanto, a través de la diferenciación por *delay*, el algoritmo DC-HNN-DRA es capaz de satisfacer los requerimientos de los servicios RT.

### ***Comportamiento con servicios NRT***

En este caso solamente se va a considerar el acceso de usuarios web y FTP. En un primer estudio se ha introducido un número creciente de usuarios web en el sistema, específicamente  $\{10,20,30,\dots,80\}$ . En la Figura 4.16 se representa el tiempo medio necesario para transmitir una página web en función del número de usuarios activos en el sistema. La gráfica muestra como el algoritmo DBR y el DC-HNN-DRA mejoran considerablemente el tiempo de respuesta del servicio, siendo el DC-HNN-DRA el algoritmo que presenta las mejores prestaciones. Al igual que para los servicios RT, asignar más recursos a aquellos usuarios con mejor estado del canal favorece la reducción de las interferencias, lo que en una tecnología limitada por interferencias, como UMTS, redundará en una mejora significativa de la satisfacción global de los usuarios.

Inicialmente, con pocos usuarios, el comportamiento de todos los algoritmos de gestión de recursos es bastante similar, pero cuando el número de usuarios aumenta DC-HNN-DRA es capaz de hacer una mejor distribución de los recursos. Con ochenta usuarios el DC-HNN-DRA mejora el resto de algoritmos hasta en un 92%. Respecto al DBR la mejora en el *delay* medio experimentado por los usuarios es del 35%, valor mucho más significativo que para los servicios RT. Salvo el DBR, el resto de algoritmos de referencia son incapaces de dar un servicio de calidad a más de 60 usuarios.



**Figura 4.16** Tiempo de descarga de página web con un número creciente de usuarios

La Figura 4.17 muestra la potencia media consumida por la estación base con un número creciente de usuarios web. El mejor comportamiento de la red hopfield en cuanto a *delay* está refrendado, además, por una importante reducción en la potencia consumida. La minimización del nivel de interferencias permite al algoritmo DC-HNN-DRA servir a más usuarios y satisfacer sus requerimientos de calidad de servicio. La potencia siempre está por debajo de 14 W, valor máximo de potencia establecido en la fase de diseño, de manera que la potencia asignada tiende asintóticamente a alcanzar este valor. Será precisamente cuando se sature en potencia cuando se produzca el incremento más significativo del *delay* en la transmisión, ya que no se dispone de más recursos para satisfacer simultáneamente a todos los usuarios.

La Figura 4.18 muestra el número medio de usuarios en cola, que representa cuántos usuarios tienen información pendiente de transmitir en sus buffers de salida.

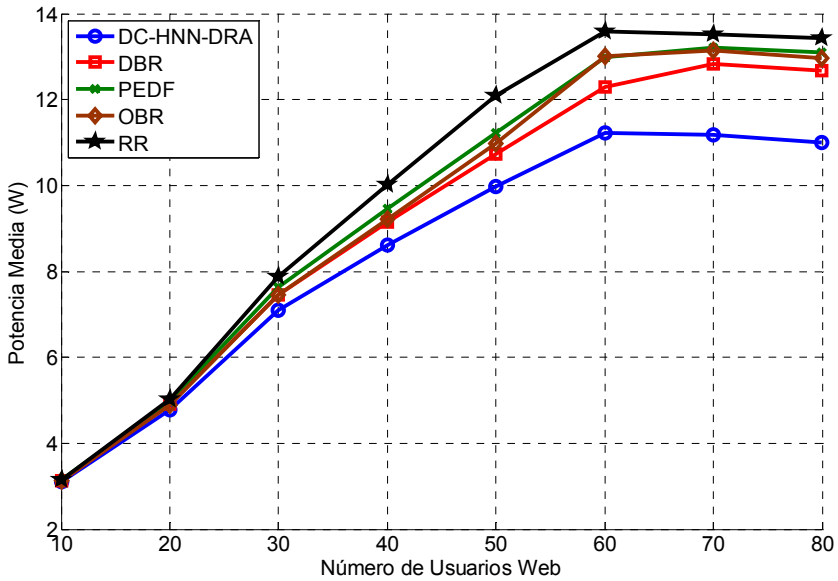


Figura 4.17 Potencia media consumida con un número creciente de usuarios web

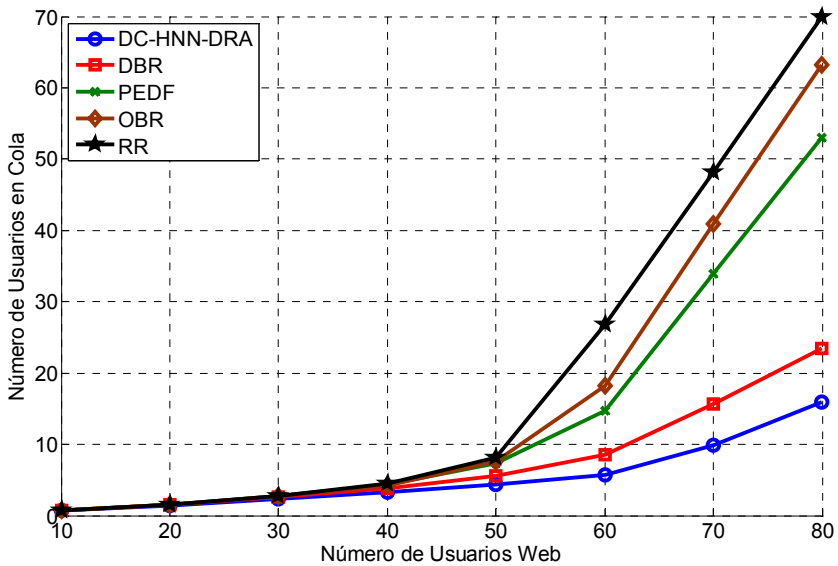


Figura 4.18 Número de usuarios en cola con un número creciente de usuarios web

Para todos los algoritmos, cuanto mayor es el número de usuarios activos en el sistema, mayor es el número de usuarios en cola, lo que resulta bastante obvio. Dado que el algoritmo DC-HNN-DRA presenta el más rápido tiempo de descarga de las páginas web, también es el algoritmo que menor número de usuarios en cola tiene.

La Figura 4.19 representa, con 80 usuarios activos, la CDF de la tasa binaria total asignada, normalizada por la tasa binaria máxima disponible en el sistema. En primer lugar, es importante mencionar que el aspecto en escalera que muestra esta gráfica es debido a que existe un conjunto discreto de tasas binarias asignables y a que algunas de esas tasas binarias tienen una alta probabilidad de suceso. Observando la gráfica, el RR presenta el peor comportamiento en términos de maximización del *throughput*, tal y como era de esperar. Por su parte, el OBR, PEDF y DBR van mejorando en este orden la eficiencia en el reparto de los recursos, pero aun muy lejos del comportamiento del DC-HNN-DRA. De nuevo las redes Hopfield proveen las mejores prestaciones, asignando más recursos incluso con un número medio de usuarios en cola inferior.

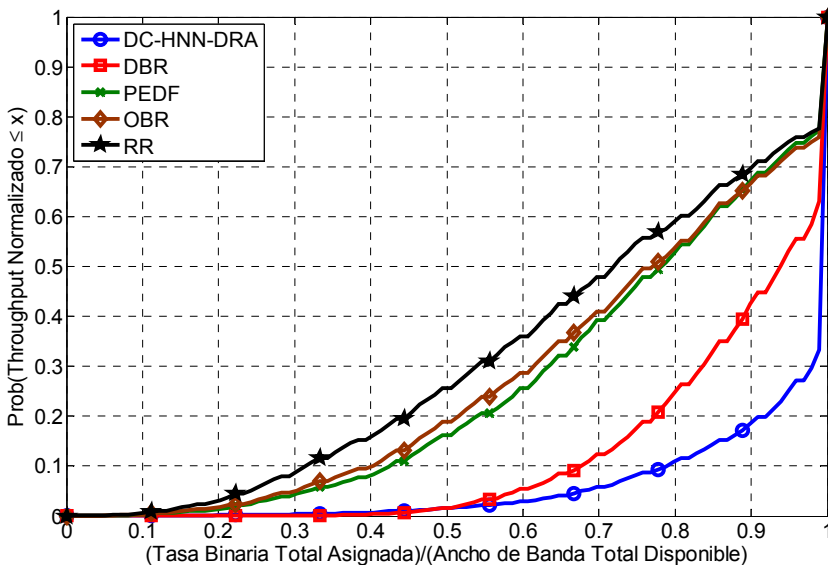
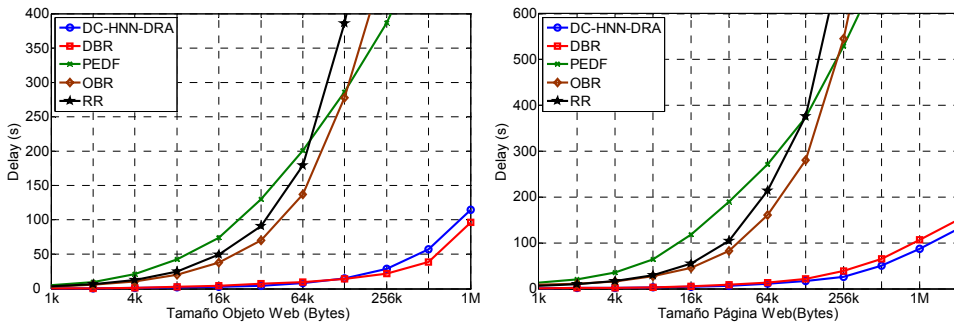


Figura 4.19 CDF del *throughput* normalizado del sistema para 80 usuarios web

Es importante destacar que, en muchos casos, los recursos asignados están acotados por el máximo número de códigos disponibles. Por ejemplo, si se utiliza DC-HNN-DRA, cerca de un 70% del tiempo se estará asignando el máximo ancho de banda disponible. En resumidas cuentas, el algoritmo DC-HNN-DRA es capaz de mejorar simultáneamente el consumo de potencia, el uso del ancho de banda disponible y el tiempo de respuesta del servicio web.

Centrando la atención de nuevo en la Figura 4.17, hay que destacar que al final de la gráfica la potencia consumida por ochenta usuarios es incluso menor que la consumida por 70. Para explicar este extraño comportamiento es necesario tener en cuenta el modo en que se usa el ancho de banda disponible en el sistema. Si el número de usuarios en cola aumenta, los algoritmos de *scheduling* basados en la minimización del consumo de potencia tendrán más diversidad de usuarios para seleccionar los que consuman menos recursos. Como todos los algoritmos, salvo el DC-HNN-DRA, implementan MNR, este efecto se aprecia claramente en todos ellos. Sin embargo, el efecto observado es más evidente en el caso del OBR ya que este algoritmo inicialmente asigna únicamente la mínima tasa binaria necesaria, obteniendo por tanto una asignación muy lejos de la máxima tasa binaria. El OBR ejecuta después MNR, por lo que existe un mayor margen para que esta técnica minimice el consumo de potencia. En el caso del DC-HNN-DRA, el efecto de la reducción de potencia también es bastante evidente y se explica por la introducción del factor  $O_i$  en la función de energía.

Centrando de nuevo la atención en el escenario más cargado de 80 usuarios, la Figura 4.20 muestra el *delay* en la transmisión de objetos y páginas web en función del tamaño de estos. De nuevo, el algoritmo DC-HNN-DRA exhibe el mejor comportamiento salvo cuando el tamaño de los objetos es muy grande, en cuyo caso el DBR mejora al DC-HNN-DRA. Esto es debido al hecho de que DC-HNN-DRA trata de minimizar el *delay*, al igual que el DBR, pero, además, pretende maximizar el uso de los recursos. El balance entre la satisfacción del *delay* y la maximización del *throughput* depende de la relación entre los dos primeros pesos de la función de energía,  $\mu_1$  y  $\mu_2$ .

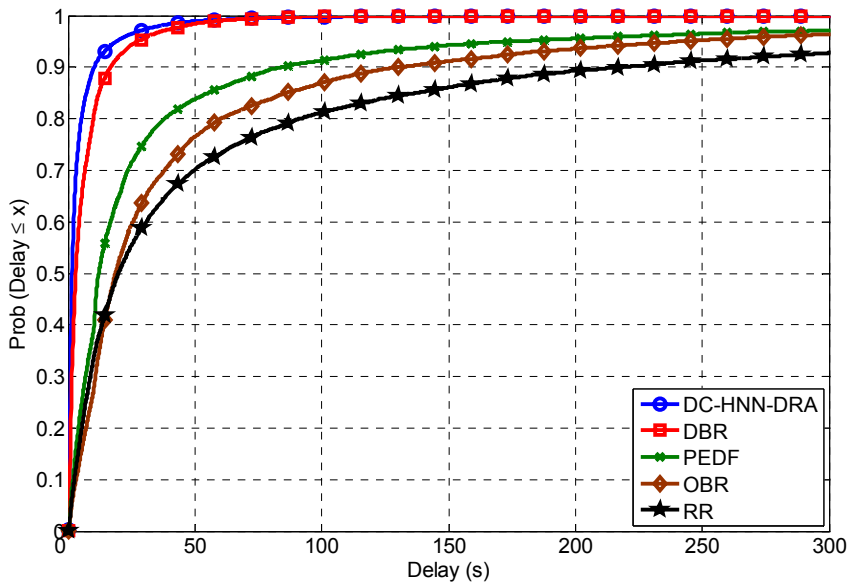


**Figura 4.20** *Delay* medio de descarga con 80 usuarios web en función del tamaño de objeto (izquierda) o de página (derecha)

Cuando se produce la transmisión de una página web de grandes dimensiones, el algoritmo DBR le da absoluta prioridad a esta descarga mientras que DC-HNN-DRA reconsidera esta decisión para satisfacer también los requisitos del resto de usuarios. Además, como la probabilidad de transmitir un objeto muy grande es muy pequeña, DC-HNN-DRA es el algoritmo que mejor comportamiento tiene en cuanto a la transmisión de la página web completa, tal y como se puede observar en la gráfica derecha de la Figura 4.20. Es también destacable que el algoritmo PEDF mejora su tendencia respecto al RR o al OBR cuando el tamaño el objeto aumenta. Este hecho se puede entender teniendo en cuenta que el *scheduler* PEDF da más prioridad a aquellos usuarios con restricciones temporales más acuciantes, de manera que favorece la transmisión de objetos grandes.

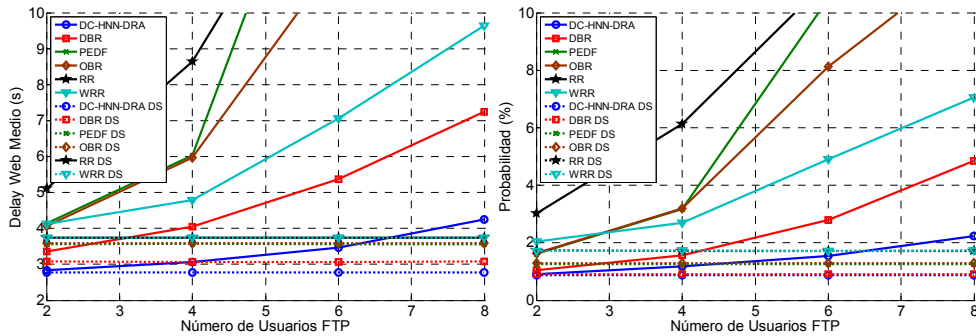
Por último, para no solamente estudiar el valor medio, la Figura 4.21 muestra la CDF del tiempo de respuesta del servicio web para 80 usuarios. De nuevo, se puede observar que el algoritmo DC-HNN-DRA es el que ofrece unas mejores prestaciones seguido por el DBR. Las peores prestaciones en cuanto a tiempo de respuesta las muestra el algoritmo RR, ya que es el único cuyo funcionamiento es independiente de este factor.

En el siguiente estudio realizado se sirve simultáneamente a usuarios web y FTP. En concreto, se han considerado 40 usuarios web y un número creciente de usuarios FTP, desde 1 hasta 8.



**Figura 4.21 CDF del tiempo de descarga de página web con 80 usuarios**

En primer lugar, se ha considerado que los algoritmos de gestión de recursos sirven simultáneamente a los usuarios web y FTP como usuarios interactivos, diferenciándolos únicamente por el valor de su *delay* máximo. A continuación, se ha simulado considerando que los *schedulers* aplican diferenciación de servicios y sirven en primer lugar a los usuarios web. En la Figura 4.22 (a) se ha dibujado el tiempo medio de descarga de página web, comparando ambas aproximaciones con un número creciente de usuarios FTP. Se utiliza la línea discontinua y el acrónimo DS (*DiffServ*) para diferenciar el método de diferenciar web y FTP. Observando al comportamiento de los algoritmos sin diferenciación, la primera conclusión que se alcanza es que el tráfico FTP afecta a los usuarios web que ven incrementado su tiempo medio de descarga. Además, las peores prestaciones las ofrece el algoritmo de RR, lo que parece lógico porque este *scheduler* no está orientado a *delay* por lo que todos los usuarios se sirven por igual. Por otro lado, el algoritmo WRR da menos tasa binaria a los usuarios FTP lo que reduce considerablemente el *delay* de los usuarios web.

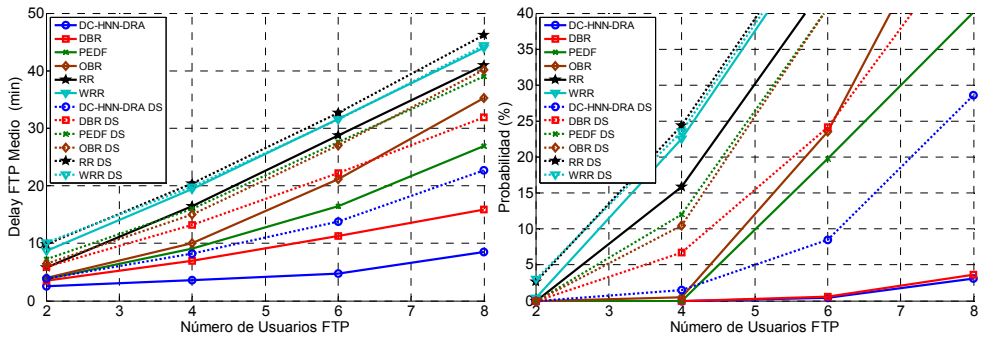


**Figura 4.22** Comportamiento del servicio de navegación web con un número creciente de usuarios FTP. Tiempo medio de descarga de página (izquierda) y Probabilidad de exceder el máximo *delay* (derecha)

Respecto al algoritmo OBR, éste asigna la correspondiente tasa mínima a todos los usuarios, lo que generalmente supone asignar bajas tasas binarias. Una vez concluida esta fase, muchos recursos estarán aun disponibles por lo que el proceso MNR aumentará la tasa binaria asignada por menor consumo de potencia, sin distinguir entre los tipos de servicio o las restricciones temporales. Por lo tanto, cuando haya un número mayor de usuarios FTP, será más probable que un usuario FTP consuma menos potencia que cualquier otro usuario web, lo que evita que tasas binarias altas se asignen a los usuarios web, lo que finalmente implica un aumento del *delay* web. El algoritmo PEDF también se ve afectado por el aumento de usuario FTP ya que, como se explicó con antes, este algoritmo tiende a dar mayores tasas binarias a los usuarios con más datos por transmitir, lo que claramente favorece a los usuarios FTP. Por su parte, el algoritmo DC-HNN-DRA consigue el menor *delay* medio, mejorando al DBR, que es el que mejores prestaciones ofrece del resto de algoritmos, hasta en un 41%.

Cuando se diferencian ambos servicios, representado en la Figura 4.22 con línea discontinua, el servicio web queda protegido del FTP y por lo tanto su comportamiento es independiente del número de usuarios FTP. Todos los algoritmos mejoran considerablemente las prestaciones de los usuarios web pero a costa de un importante deterioro del *delay* FTP, tal y como se observa en la Figura 4.23.





**Figura 4.23 Comportamiento del servicio FTP con un número creciente de usuarios FTP. Tiempo medio de descarga de archivo (izquierda) y Probabilidad de exceder el máximo *delay* (derecha)**

Cuando se diferencian ambos servicios, RR y WRR se comportan exactamente igual para los usuarios web. Este efecto es razonable ya que las diferencias entre ambos se debían a la diferenciación hecha por el WRR. Entonces, si ahora ambos algoritmos realizan la misma diferenciación, el comportamiento debe ser el mismo. De nuevo, DC-HNN-DRA consigue el servicio óptimo, mejorando un 10% respecto a DBR.

La Figura 4.22 (b) analiza la probabilidad de exceder el *delay* máximo de web. El algoritmo DBR se aproxima al DC-HNN-DRA en este indicador puesto que, tal y como se mostró en la Figura 4.20, reduce el retardo de los objetos más grandes y por tanto la probabilidad de exceder el máximo *delay*.

Este estudio no se puede dar por concluido sin antes analizar el comportamiento del servicio FTP con y sin diferenciación de servicio. La Figura 4.23 (a) representa el tiempo medio de descarga de un archivo FTP. Es importante destacar que el DC-HNN-DRA no solo reduce el retardo de la navegación web sino que, además, mejora considerablemente las prestaciones del servicio FTP. Para 8 usuarios FTP, el DC-HNN-DRA presenta una mejora del 46% con respecto al tiempo requerido por el algoritmo DBR, 68% respecto al PEDF, 75% con OBR, 79% con RR y 80% con WRR. Esta mejora se ha conseguido gracias a la mejor maximización del uso de los recursos disponibles y al menor consumo de potencia. Por otro lado, la diferenciación entre web

y FTP mejora las prestaciones del servicio web pero por el contrario aumenta considerablemente el retardo medio de descarga de fichero FTP. Observando la probabilidad de exceder el máximo *delay* en la Figura 4.23 (b), se puede ver claramente como la diferenciación penaliza mucho esta probabilidad. Por ejemplo, para DC-HNN-DRA la probabilidad aumenta del 3% hasta el 28%. Por lo tanto, la diferenciación entre ambos servicios no parece recomendable, sobre todo teniendo en cuenta que el servicio FTP no es estrictamente un servicio *background*. Además, comparando ambas probabilidades de exceder el máximo *delay*, el único algoritmo capaz de satisfacer al mismo nivel, cerca del 3%, ambos servicios es DC-HNN-DRA.

La Figura 4.24 y la Figura 4.25 comparan el comportamiento de los algoritmos de gestión de recursos cuando el servicio FTP se considera como servicio de *background* (el acrónimo BG se utiliza para su identificación). En este caso  $R_{\min,i} = 0$  para todos los usuarios FTP. La Figura 4.24 muestra cómo todos los algoritmos basados en *delay* mejoran el comportamiento de los usuarios web al considerar el servicio FTP como *background*.

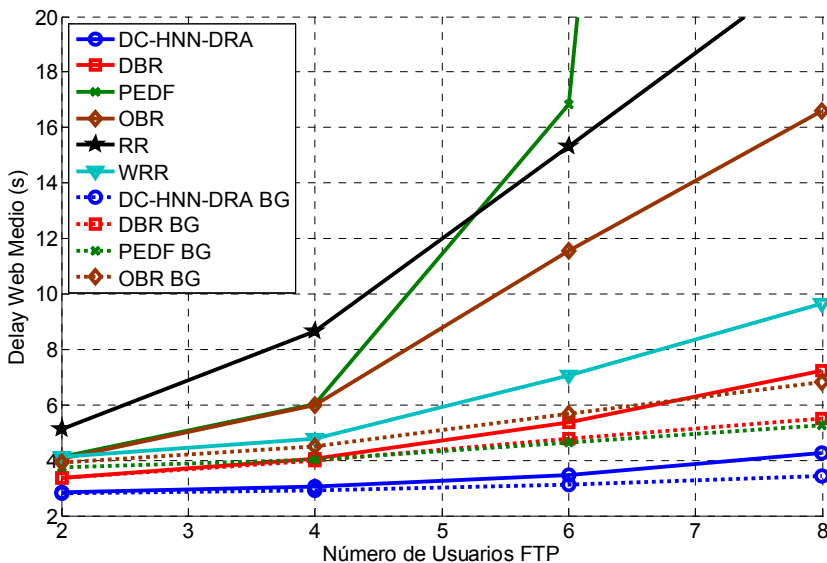
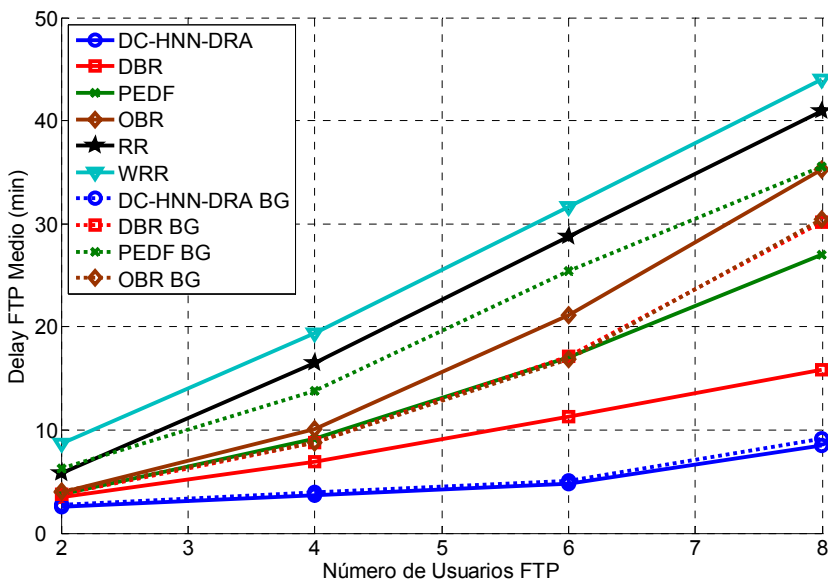


Figura 4.24 Tiempo medio de descarga de página web con número creciente de usuarios FTP definidos como interactivos o como *background*

Únicamente RR y WRR se comportan exactamente de la misma manera ya que su funcionamiento no depende del valor de  $R_{\min,i}$ . El algoritmo PEDF mejora en gran medida su funcionamiento para web ya que los usuarios FTP dejarán de acaparar gran parte de los recursos por disponer de objetos más grandes. El algoritmo DC-HNN-DRA es el único *scheduler* capaz de mantener prácticamente el mismo valor de *delay* independientemente del número de usuarios FTP en el sistema. Si este valor de retardo no se mantiene exactamente constante es porque la función de energía no está únicamente diseñada para minimizar el *delay* sino que también tratará de maximizar el uso de los recursos. Dependiendo del balance entre  $\mu_1$  y  $\mu_2$  el valor medio de retardo web será más o menos constante.

Respecto a los usuarios FTP, tal y como se muestra en la Figura 4.25, su comportamiento apenas empeora utilizando DC-HNN-DRA. Parece que, para sacar el máximo partido al nuevo algoritmo propuesto en esta Tesis, la asignación de recursos conjunta entre web y FTP definido como un servicio *background* es la mejor opción.



**Figura 4.25** Tiempo medio de descarga de archivo FTP con número creciente de usuarios FTP definidos como interactivos o como *background*

Es muy significativo ver cuál es el comportamiento del algoritmo OBR al definir el servicio FTP como *background*. Aunque para el resto de los algoritmos la tendencia es aumentar el tiempo medio de respuesta del servicio FTP, el OBR mejora las prestaciones, comportándose de manera muy similar al DBR. Esto es debido a que, al asignar tasa binaria mínima 0, toda la asignación de recursos para FTP se realiza mediante el proceso MNR. Éste proceso asigna más recursos a los que menos consumen, de manera muy similar a cómo actúa el DBR. Este mecanismo es el que mejor rendimiento da por la minimización de las interferencias conseguida, lo que justifica la mejora en las prestaciones del OBR.

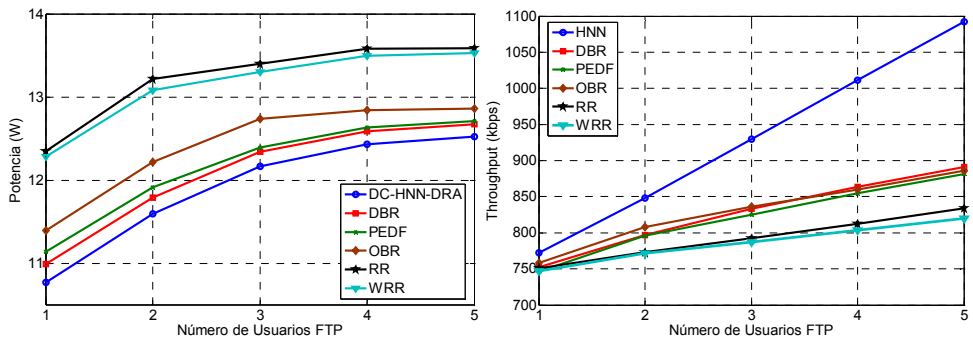
### ***Comportamiento con tráfico mixto***

Los anteriores estudios han demostrado que, aunque el algoritmo DC-HNN-DRA tiende a dar más recursos a los servicios más restrictivos en términos de *delay*, introducir usuarios menos prioritarios afecta en cierta medida a las prestaciones del resto. Cuando se desea trabajar conjuntamente con servicios de tiempo real y de no tiempo real no se puede permitir ese tipo de efectos perniciosos. Es por ello que al trabajar con tráfico mixto es preferible repartir en primer lugar los recursos entre los usuarios RT para posteriormente dar servicio a los usuarios NRT. Dentro del algoritmo DC-HNN-DRA este hecho se traducirá en ejecutar dos veces la red neuronal. En la primera ejecución solamente se excitarán las neuronas de los usuarios RT. La configuración a la que converja esta red se introducirá de nuevo en la red neuronal para una segunda ejecución, pero esta vez sólo se modificarán las neuronas correspondientes a los usuarios NRT, mientras que las otras permanecerán constantes.

Para conocer cuál es el comportamiento de la red en un escenario mixto, se han introducido 5 usuarios de voz, 5 usuarios de vídeo conferencia, 30 usuarios web y desde 1 hasta 5 usuarios FTP.

El comportamiento de los usuarios de voz y de vídeo conferencia es el mismo que el mostrado en el análisis de tráfico RT y no depende del número de usuarios NRT introducidos en el sistema, lo que valida la independencia conseguida entre ambos tipos de servicios. Para recalcar el menor consumo de recursos y la mayor

maximización del *throughput* conseguida mediante el uso de DC-HNN-DRA, se ha evaluado la potencia media consumida y la tasa binaria media del sistema para todos los algoritmos considerados. La Figura 4.26 muestra ambos parámetros para un número creciente de usuarios FTP. Se puede observar de nuevo que el algoritmo DC-HNN-DRA es el que menor consumo de potencia realiza. Además, es especialmente significativo observar la gran diferencia existente en cuanto a la maximización del uso de los recursos. No sólo se reduce el tiempo medio de descarga de página web y fichero FTP, sino que además, debido al segundo término de la función de energía, se hace un gran esfuerzo por maximizar el uso de los recursos que al final redunda en un mejor servicio para los usuarios.



**Figura 4.26** Potencia media consumida por la estación base (izquierda) y *throughput* del sistema (derecha)

## 4.6 Conclusiones

En este capítulo se han analizado las técnicas más importantes de gestión de recursos en UMTS. Fruto de este análisis se ha propuesto un algoritmo centrado en *delay* implementado mediante redes neuronales de Hopfield. Este algoritmo no solamente ejerce las funciones de *scheduler* sino que a su vez actúa como gestor del control de admisión y de control de la congestión. A través de simulaciones en el enlace descendente, se ha demostrado que el nuevo algoritmo, denominado DC-HNN-DRA, ofrece un excelente comportamiento para la gestión de servicios de tiempo real y de no tiempo real en un escenario multi-servicio.

Ambos servicios RT y NRT se gestionan independientemente debido a las estrictas restricciones temporales del tráfico RT. En DC-HNN-DRA, los servicios RT se gestionan simultáneamente en contraposición con otras técnicas que utilizan políticas de diferenciación de servicio para proteger a los usuarios de voz. Los resultados obtenidos han demostrado que, con esta diferenciación, la voz mejora ligeramente su calidad de servicio en comparación con el algoritmo DC-HNN-DRA. Sin embargo, mezclando los usuarios de voz y de vídeo conferencia, el DC-HNN-DRA tiene más flexibilidad para sacar el máximo partido de los recursos disponibles. Así pues, a costa de un ligero empeoramiento del servicio de voz, la mejora de la calidad de servicio de los usuarios de vídeo conferencia es muy significativa. En conclusión, el algoritmo DC-HNN-DRA es capaz de priorizar perfectamente a los usuarios de voz a través únicamente de la diferencia existente en el *delay* máximo permitido. Esto abre las puertas a su uso en redes exclusivamente de conmutación de paquetes, sin necesidad de establecer circuitos dedicados ni siquiera para los usuarios de voz.

Una vez servidos los usuarios RT, el algoritmo DC-HNN-DRA distribuye simultáneamente los recursos restantes entre los usuarios NRT interactivos y los usuarios *background*. En general, el DC-HNN-DRA favorece a los usuarios con mejores condiciones de propagación, reduciendo el consumo de potencia. El mejor uso de la potencia junto con una máxima utilización del *throughput*, hace que el algoritmo DC-HNN-DRA mejore considerablemente respecto al resto de algoritmos utilizados

como referencia. Así, las prestaciones de los usuarios FTP se mejoran claramente, únicamente a costa de un ligero aumento en el tiempo de respuesta de los usuarios interactivos. Sin embargo, este incremento en el retardo no implica una pérdida significativa de calidad de servicio.

Así pues, el mismo efecto observado con los usuarios de voz y vídeo conferencia se puede apreciar con los usuarios interactivos y de *background*, es decir, una asignación conjunta de recursos mejora ostensiblemente los servicios menos prioritarios perjudicando a los más prioritarios. Sin embargo, el efecto dañino en el caso del DC-HNN-DRA es casi despreciable, mientras que la mejora es muy significativa, justificando la toma de esta decisión. Por todo, el algoritmo DC-HNN-DRA con asignación conjunta por tipo de servicio, RT y NRT, es la solución más adecuada para la gestión dinámica de recursos en UMTS, ya que se ha demostrado que su uso mejora las prestaciones globales del sistema.

Las mejores prestaciones del DC-HNN-DRA se reflejan asimismo en una importante reducción del número medio de usuarios activos. Por lo tanto, la mayor eficiencia en la distribución de recursos es doblemente útil, en primer lugar permite que los usuarios transmitan más rápido, y además reduce la carga media del sistema, permitiendo que se admitan más usuarios.

## Referencias

- [1] 3GPP TS 23.107 v 3.9.0., “QoS Concept and Architecture”, 2002.
- [2] B.M. Epstein, M. Schwartz, “Reservation Strategies for Multi-Media Traffic in a Wireless Environment”, in Proceedings of the IEEE *Vehicular* Technology Conference (VTC’95), vol. 1, pp.165 – 169, 1995.
- [3] Y. Xiao, P. Chen, Y. Wang, “Optimal Admission Control for Multi-Class of Wireless Adaptive Multimedia Services”, IEICE Transactions on Communications, vol. E84-B, no. 4, pp. 795 – 804, 2001.
- [4] M. Ei-Kadi, S. Olariu, H. Abdel-Wahab, “Rate-based Borrowing Scheme for QoS Provisioning in Multimedia Wireless Networks”, IEEE Transactions on Parallel and Distributed Systems, vol. 13, no. 2, pp. 156 – 166, 2002.
- [5] C. Mihailescu, X. Lagrange, P. Godlewski, “Performance Evaluation of a Dynamic Resource Allocation Algorithm for UMTS-TDD Systems”, in Proceedings of the IEEE *Vehicular* Technology Conference (VTC’00), vol 3, pp. 2339 – 2343, 2000.
- [6] L. Forkel, T. Kriengchaiyapruk, B. Wegmann, E. Schulz, “Dynamic Channel Allocation in UMTS Terrestrial Radio Access TDD Systems”, in Proceedings of the IEEE *Vehicular* Technology Conference (VTC’01), vol. 2, pp. 1032 – 1036, 2001.
- [7] A. Hernández Solana, A. Valdovinos Bardaji, F. Casadevall Palacio, “*Scheduling* with Quality of Service Constraints for Real-Time and Non-Real-Time Traffic in WCDMA”, in Proceedings of the Wireless Personal Multimedia Communications, pp. 1543 – 1548, 2001.
- [8] L. Almajano, J. Pérez-Romero, “Packet *Scheduling* Algorithms for Interactive and Streaming Services under QoS Guarantee in a CDMA System”, in Proceedings of the IEEE *Vehicular* Technology Conference (VTC’02), vol. 3, pp. 1657 – 1661, 2002.
- [9] M.R. Sherif, I. Habib, M. Nagshineh, P. Kermani, “Adaptive Allocation of Resources and Call Admission Control for Wireless ATM Using Genetic Algorithms”, IEEE Journal on Selected Areas in Communications, vol. 18, pp. 268–282, 2000.
- [10] S. Papavassiliou, A. Puliafito, O. Tomarchio, J. Ye, “Integration of Mobile Agents and Genetic Algorithms for Efficient Dynamic Network Resource Allocation”, in Proceedings of the IEEE Symposium on Computers and Communications, 2001.



- 
- [11] O. Lázaro, “Dynamic Radio Resource Management Algorithms and Traffic Models for Emerging Mobile Communication Systems”, Ph.D. dissertation, University of Strathclyde, Glasgow, 2001.
- [12] C. Ahn, R. Ramakrishna, “QoS Provisioning Dynamic Connection-Admission Control for Multimedia Wireless Networks Using Hopfield Neural Networks”, IEEE Transactions on *Vehicular* Technology, vol. 53, no. 1, pp. 106-117, 2004.
- [13] N. García, R. Agustí, J. Pérez-Romero, “A User-Centric Approach for Dynamic Resource Allocation in CDMA Systems based on Hopfield Neural Networks”, in Proceedings of the IST Summit, 2005.
- [14] D. Calabuig, J. F. Monserrat, D. Gómez-Barquero, N. Cardona, “Hopfield Neural Network Algorithm for Dynamic Resource Allocation in WCDMA Systems”, in Proceedings of the IEEE International Symposium On Wireless Communication Systems, 2006.
- [15] D. Calabuig, J.F. Monserrat, D. Gómez-Barquero, O. Lázaro, “User Bandwidth Usage-Driven HNN Neuron Excitation Method for Maximum Resource Utilization within Packet-Switched Communication Networks”, IEEE Communication Letters, vol. 10, no. 11, pp. 766 – 768, 2006.
- [16] S. Haykin, “Neural Networks, A Comprehensive Foundation”, MacMillan, New York, 1994.
- [17] M. Forti, S. Manetti, M. Marini, “A Condition for Global Convergence of a Class of Symmetric Neural Circuits”, IEEE Transactions on Circuits and Systems I, vol. 39, pp. 480 – 483, 1992.
- [18] K. Smith, M. Palaniswami, M. Krishnamoorthy, “Neural Techniques for Combinatorial Optimization with Applications”, IEEE Transaction on Neural Networks, vol. 9, no. 6, pp 1301 – 1318, 1998.
- [19] H. Holma, A. Toskala, “WCDMA for UMTS”, John Wiley & Sons, 3rd edition, 2004.
- [20] 3GPP TS 25.211 V3.12.0, “Physical channels and mapping of transport channels onto physical channels (FDD)”, 2002.
- [21] T Minn et al. “Dynamic Assignment of Orthogonal Variable Spreading Factor Codes in WCDMA”, IEEE Journal on Selected Areas in Communications, pp. 1429-1440, 2000.
- [22] 3GPP TS 25.212 V3.11.0, “Multiplexing and channel coding (FDD)”, 2002.

- [23] T.K. Moon and W.C. Stirling, "Mathematical Methods and Algorithms for Signal Processing", Ed. Prentice-Hall, pp. 632–642, New Jersey 2000.
- [24] J.J. Hopfield and D. Tank, "Neural Computation of Decisions in Optimization Problems", *Biological Cybernetics*, vol. 52, pp. 141-152, July 1985.
- [25] J.J. Hopfield, "Neurons with graded response have collective computational properties like those of two-state neurons", in *Proceedings of National Academy of Sciences*, vol. 81, pp. 3088-3092, May 1984.
- [26] N. Funabiki and Y. Takefuji, "A Neural Network Parallel Algorithm for Channel Assignment Problems in Cellular Radio Networks", *IEEE Transactions On Vehicular Technology*, vol. 41, pp. 430–437, November 1992.
- [27] O. Lazaro and D. Girma, "A Hopfield Neural Network-based Dynamic Channel Allocation with Handoff Channel Reservation Control", *IEEE Transactions On Vehicular Technology*, vol. 49, pp. 1578–1587, September 2000.
- [28] H. G. Sandalidis, P. P. Stavroulakis, and J. Rodriguez-Tellez, "Borrowing Channel Assignment Strategies Based on Heuristic Techniques for Cellular Systems", *IEEE transactions On Neural Networks*, vol. 10, no. 1, pp. 176–181, January 1999.
- [29] N. Funabiki and S. Nishikawa, "A Binary Hopfield Neural-Network Approach for Satellite Broadcast *Scheduling* Problems", *IEICE Transactions On Neural Networks*, vol. 8, no. 2, pp. 441–445, March 1997.
- [30] M. Forti, S. Manetti and M. Marini, "Necessary and sufficient condition for absolute stability of neural networks", *IEEE Transactions on Circuits and Systems - I*, vol. 41, pp. 491-494, July 1994.
- [31] H. Jonathan Chao, Xiaolei Guo, "Quality of Service Control in High-Speed Networks", John Wiley & Sons, 2002.
- [32] M. Katavenis, et al., "Weighted round-robin cell multiplexing in a general-purpose ATM switch *chip*," *IEEE Journal on Selected Areas in Communication*, vol. 9, no. 8, pp 1265-1279, 1991.
- [33] N. García et al., "A Novel *Scheduling* Algorithm for *Delay*-Oriented Services Based on Hopfield Neural Networks Methodology", in *Proc. Of the IEEE Wireless Communications and Networking Conference*, 2006.

- [34] V. Sivaraman and F. Chiussi, “End-to-End Statistical *Delay* Guarantees Using Earliest Deadline First (EDF) Packet *Scheduling*”, in Proc. of GLOBECOM, pp. 1307-1312, 1999.



# Capítulo 5

## *Gestión de Recursos en Redes MNB*

*Lo importante es no dejar de hacerse preguntas.*

*Albert Einstein*

### **5.1 Introducción**

Desde sus inicios, la red UMTS fue concebida como una red celular flexible que permitiera el despliegue de redes jerárquicas o HCS (*Hierarchical Cell Structure*) [1]. Se entiende que una red jerárquica está compuesta por, al menos, dos tipos de celdas, como por ejemplo, celdas macro-celulares y celdas micro-celulares. El uso de celdas micro-celulares reduce en gran medida los niveles de interferencia experimentados en enlace ascendente y, además, aumenta la frecuencia de reutilización del espectro, lo que implica un incremento importante en la capacidad del sistema. Por el contrario, si se utilizaran únicamente redes micro-celulares existiría irremediablemente muchas zonas sin cobertura debido al efecto esquina, es decir, un usuario sufriría una importante reducción de cobertura al perder la visión directa con el transmisor micro-celular. Además, los usuarios a alta velocidad sufrirían muchos *handovers* e incluso se podría dar el caso en el que, iniciado un procedimiento de *handover*, al concluirlo la mejor celda hubiera cambiado, lo que provocaría la terminación forzosa de la llamada.

Para evitar estos dos efectos, la cobertura micro-celular se complementa con la de una celda macro-celular. De esta forma, la macro-celda dará servicio en aquellos puntos carentes de cobertura micro-celular y, además, se encargará de los usuarios de alta velocidad, evitando así tiempos de permanencia en la celda muy pequeños que

ocasionen cortes en las llamadas. En general, las micro-celdas servirán a usuarios de poca movilidad y de alta capacidad, mientras que las macro-celdas se usarán para los usuarios de alta movilidad y baja capacidad [2]-[11].

Existen dos posibles modos de configurar la red HCS en UMTS. La primera consiste en utilizar una frecuencia distinta para los distintos niveles de jerarquía. De esta manera, las celdas macro y micro-celulares están totalmente aisladas entre sí, por lo que no existe interferencia mutua. Si se desea pasar de un nivel a otro de jerarquía es necesario realizar un procedimiento de traspaso inter-frecuencia, traspaso que puede conllevar un retardo de entre 1 y 3 segundos, dependiendo del fabricante. La decisión sobre en qué tipo de celda se va a introducir un nuevo usuario se hace normalmente utilizando un umbral de velocidad [2], estimando ésta por medio de alguno de los múltiples métodos descritos en la literatura (véase como ejemplo [3]). A partir de este tipo de algoritmos, otros autores han diseñado métodos algo más complejos, introduciendo información adicional, como por ejemplo el grado de congestión de las celdas [4] o incluso el tipo de tráfico servido [5]. En [6] y [7] se proponen distintos métodos para variar de manera adaptativa los umbrales utilizados en la asignación del tipo de celda, de manera que se obtenga un balanceo adecuado de la carga entre los distintos tipos de celda. Por otro lado, y a pesar del elevado retardo que supone hacer un traspaso inter-frecuencia, son muchos los autores que proponen realizar una asignación dinámica del nivel de jerarquía utilizado en la transmisión. En [8], se resumen algunos de los mecanismos más utilizados de reasignación de tipo de celda.

El segundo modo de configurar la red HCS no fue contemplado en el estándar y consiste en que todos los niveles de jerarquía compartan la misma frecuencia y sólo se diferencien en el código de aleatorización utilizado [9]-[10]. Los usuarios utilizan una capa u otra siguiendo los mecanismos convencionales de traspaso, traspaso realizado normalmente en función del nivel de potencia recibido de los canales piloto. Cuando un usuario esté en situación de *soft handover* disfrutará de la transmisión sincronizada desde ambas antenas y de la consiguiente ganancia por macro y micro-diversidad experimentada en el terminal.

El principal problema de esta propuesta radica en el hecho de que UMTS es un sistema limitado por interferencias. El estándar de UMTS recoge un conjunto de métodos de control de potencia dedicados precisamente a reducir la interferencia del sistema y, por tanto, a aumentar su capacidad. Sin embargo, estas técnicas están diseñadas considerando únicamente una celda, por lo que no se ajustan a la HCS que sufrirá una interferencia inter-celular muy alta en las micro-celdas [9]. Algunos trabajos, como [10], intentan determinar la potencia óptima del canal piloto para minimizar el efecto de las interferencias inter-celulares, pero aun así no parece factible aislar ambos niveles de jerarquía, salvo que las micro-celdas se sitúen en los bordes de la cobertura macro-celular [11].

Otra limitación de la HCS con separación por código es que su principal virtud, que es la posibilidad de hacer uso del *soft handover* entre macro y micro-celdas, no tiene cabida para los canales compartidos, ya que para ellos el estándar no prevé su uso.

En esta Tesis Doctoral se ha optado por dar un paso más en la integración de las redes celulares jerárquicas, de manera que las macro y micro-celdas compartan no solamente la frecuencia, sino también el código de aleatorización. De este modo, será posible conseguir un total balanceo de recursos entre ambos niveles, solución que ha demostrado ser óptima para el funcionamiento de la HCS [12]. Además, como se utilizan códigos ortogonales para separar los distintos canales dentro del MNB, la interferencia inter-celular se reducirá, al igual que la intra-celular, ya que la potencia transmitida dentro de la micro-celda será quasi-ortogonal para todos los usuarios incluidos en la misma y estará muy atenuada para el resto del MNB. Además, y gracias al procesado en *array* introducido en el MNB, se podrá emular un *soft handover* virtual dentro del MNB que no sólo será útil para los usuarios de canales dedicados, sino también para los de canales compartidos, aumentando las prestaciones de estos últimos.

En el Capítulo 3 de esta Tesis Doctoral se demuestra que se puede obtener una importante ganancia en el sistema al transmitir sincronizadamente desde las antenas del MNB. Además, en el Capítulo 4 se recoge un nuevo algoritmo de gestión de recursos capaz de maximizar el uso de recursos y optimizar la calidad de servicio de todo tipo de usuarios UMTS. Este capítulo tiene como objetivo extender esta formulación para

incluir la distribución de recursos entre los transmisores presentes en el MNB y evaluar las prestaciones finales de la nueva arquitectura radio propuesta. Para ello, el apartado 5.2 describe las principales modificaciones realizadas en los algoritmos de gestión de recursos dentro del MNB sin incluir la cronoexecución, aspecto que se analizará en el apartado 5.3. Finalmente, en el apartado 5.4 se analizan las prestaciones de las distintas arquitecturas MNB consideradas, así como los algoritmos propuestos para cada una de ellas, comparándolos y extrayendo las conclusiones oportunas.

## **5.2 Gestión de Recursos en el MNB**

Tal y como se comentó en el Capítulo 4, los algoritmos de gestión de recursos radio son fundamentales para el buen funcionamiento del sistema UMTS. Con la nueva arquitectura radio Multinodo B se deberá asegurar la validez de los algoritmos utilizados en UMTS o adaptarlos convenientemente. Además, la selección de desde qué antena o antenas se transmite a cada usuario es una decisión fundamental. Una buena elección aumentará la diversidad, disminuirá el nivel de interferencias del sistema y, por tanto, incrementará su capacidad.

En este apartado se analizan los principales cambios introducidos en los algoritmos de gestión de recursos en el MNB respecto al control de potencia y la selección de celda. La cronoexecución junto con el control de admisión y de congestión se resolverá mediante redes Hopfield, reservando el apartado 5.3 para su descripción. Por último, la implementación de los algoritmos de traspaso no se ha considerado en esta Tesis, mientras que la gestión de códigos será la misma que la explicada en el Capítulo 4, por lo que no se repetirá su estudio.



## 5.2.1 Control de potencia

En UMTS se definen tres procedimientos de control de potencia: el control de potencia en bucle abierto, el control de potencia en bucle cerrado rápido y el control de potencia en bucle cerrado externo. A continuación, se describen las particularidades e implementación de cada uno de éstos en la nueva arquitectura radio Multinodo B.

### 5.2.1.1 Control de potencia en bucle abierto

En el Anexo B se demuestra que la potencia total transmitida por el MNB,  $P_T$ , respondía a la siguiente expresión:

$$P_T = \frac{\sum_{i=1}^N \frac{P_N + \chi_i}{D_{ijk}} + \sum_{p=1}^P P_{C,p}}{1 - \sum_{i=1}^N \frac{\sum_{p=1}^P \frac{\rho_p Q_p}{L_{p,i}}}{D_{ijk}}} \quad (5.1)$$

$$D_{ijk} = \sum_{p=1}^P \left( \frac{q_{pik}}{L_{p,i}} \cdot \left( \rho_p + \frac{W}{\left( \frac{E_b}{N_0} \right)_{ijk} R_{ij}} \right) \right) \quad (5.2)$$

donde  $P_N$  es la potencia de ruido térmico,  $\chi_i$  la potencia interferente recibida por el usuario  $i$ -ésimo,  $q_{pik}$  es la proporción concreta de potencia útil transmitida al usuario  $i$ -ésimo desde la antena  $p$ -ésima y cuyo valor estará determinado por el algoritmo utilizado de gestión de recursos radio,  $L_{p,i}$  representa las pérdidas de propagación desde la antena  $p$  hasta el usuario  $i$ ,  $\rho_p$  es el factor de ortogonalidad propio de la antena  $p$ ,  $W$  es el ancho de banda de UMTS,  $E_b / N_0|_{ijk}$  es la relación objetivo de energía de bit respecto a densidad espectral de ruido más interferencia para el usuario  $i$  y la tasa binaria  $j$  en las condiciones de propagación y diversidad correspondientes,  $R_{ij}$  es la tasa

binaria  $j$ -ésima asignada al usuario  $i$ ,  $Q_p$  es la proporción de potencia total transmitida por la antena  $p$  y, finalmente,  $P_{C,p}$  es la potencia transmitida en los canales de control común por la antena  $p$ .

A partir del valor calculado de  $P_T$  se puede calcular la potencia que se tiene que asignar al usuario  $i$ -ésimo para asignarle la tasa binaria  $j$ -ésima con la distribución de potencia  $k$ -ésima, como:

$$P_{ijk} = \frac{P_N + \chi_i + P_T \sum_{p=1}^P \frac{\rho_p Q_p}{L_{p,i}}}{D_{ijk}} \quad (5.3)$$

Como el valor de  $P_T$  depende de la relación de potencia entre las antenas del MNB y este valor depende a su vez de la potencia transmitida a todos los usuarios y de  $q_{pik}$ , para calcular definitivamente el valor de  $P_{ijk}$  se deberá repetir iterativamente el cálculo de  $P_T$  y de todas las  $P_{ijk}$  hasta alcanzar la convergencia en el cálculo.

El control de potencia en lazo abierto se ejecuta cada vez que el *scheduler* reasigna los recursos, ya que en ese momento se debe iniciar un nuevo procedimiento de comunicación. El algoritmo de gestión de recursos debe estimar la potencia inicial transmitida a cada usuario, estimación que se realiza a través de la información que los usuarios envían a la estación base en enlace ascendente y utilizando las ecuaciones (5.1) y (5.3). La principal cuestión ahora es cómo puede el MNB obtener los datos necesarios para resolver dichas ecuaciones y calcular así la potencia consumida por cada usuario.

En UMTS, y a través del protocolo RRC, el RNC es capaz de configurar al terminal para que realice una serie de medidas o bien de manera puntual o periódicamente o tras ocurrir un determinado evento. Una vez establecido el procedimiento de medida, que dentro del MNB se configurará periódico, el terminal envía lo que se conoce como *Measurement Report* o informe de medidas [13]. Es mucha la información que se puede solicitar al terminal [14] pero, desde el punto de vista del cálculo de la potencia transmitida en el MNB, los elementos de medida fundamentales son los siguientes:

- RSSI (*Received Signal Strength Indicator*): el móvil mide toda la potencia recibida en el ancho de banda de su filtro de recepción justo tras la antena.

En esta potencia se incluye tanto la potencia recibida de la estación que presta servicio al usuario como la potencia interferente y el ruido térmico.

- RSCP (*Received Signal Code Power*): según las especificaciones, el móvil debe ser capaz de monitorizar la potencia recibida del canal CPICH transmitido desde las dos antenas que transmiten en diversidad.
- Potencia transmitida por el UE en enlace ascendente.

Además, se puede solicitar al terminal que envíe estas medidas al RNC de manera periódica, siendo el periodo mínimo de envío de 200 ms. En el periodo entre envíos el móvil irá realizando las medidas solicitadas *slot a slot*, promediando los resultados obtenidos para obtener un valor más fiable. Todo este procedimiento de medidas se ha emulado en el simulador de sistema implementado. De este modo, al principio de la ejecución del *scheduler*, el sistema dispondrá de los últimos datos enviados por el móvil y trabajará con ellos para asignar convenientemente los recursos.

Como la potencia transmitida por el canal piloto es fija, a partir de la medida RSCP se puede estimar  $L_{p,i}$ , las pérdidas de propagación desde la antena  $p$ -ésima hasta el terminal  $i$ -ésimo, para las antenas que hayan transmitido al usuario en el periodo anterior. Para estimar las pérdidas de propagación del resto de antenas que no estén transmitiendo al móvil se ha de hacer uso de la potencia transmitida por el usuario para, a partir de la potencia recibida en cada antena del MNB, calcular las pérdidas de propagación en enlace ascendente, extrapolando este dato para las pérdidas en enlace descendente. Para calcular el nivel de interferencias inter-celulares más la potencia de ruido de cada usuario,  $\chi_i + P_N$ , el sistema dispone de la información de RSSI. Como se conoce la potencia transmitida por las antenas del MNB y las pérdidas de propagación desde cada antena, se puede calcular la potencia intra-celular recibida por el terminal. Si al valor medido de RSSI se le resta esta potencia intracelular se obtiene el valor de  $\chi_i + P_N$ . Finalmente, el valor de potencia total transmitida por el MNB,  $P_T$ , o la distribución de potencias,  $q_{pik}$  y  $Q_p$ , dependerá de la asignación de recursos realizada, por lo que se ajustará de manera iterativa y en paralelo para todos los usuarios. Cuando se alcance el equilibrio se llegará a la solución deseada.

El resto de términos de la ecuación (5.3) se deben definir en el proceso de planificación y despliegue, como por ejemplo el nivel de  $E_b/N_0$  objetivo para cada servicio y circunstancia de propagación o el factor de ortogonalidad,  $\rho_p$ . La Figura 5.1 resume el proceso seguido por el MNB para obtener los parámetros necesarios para determinar la potencia transmitida a cada usuario.

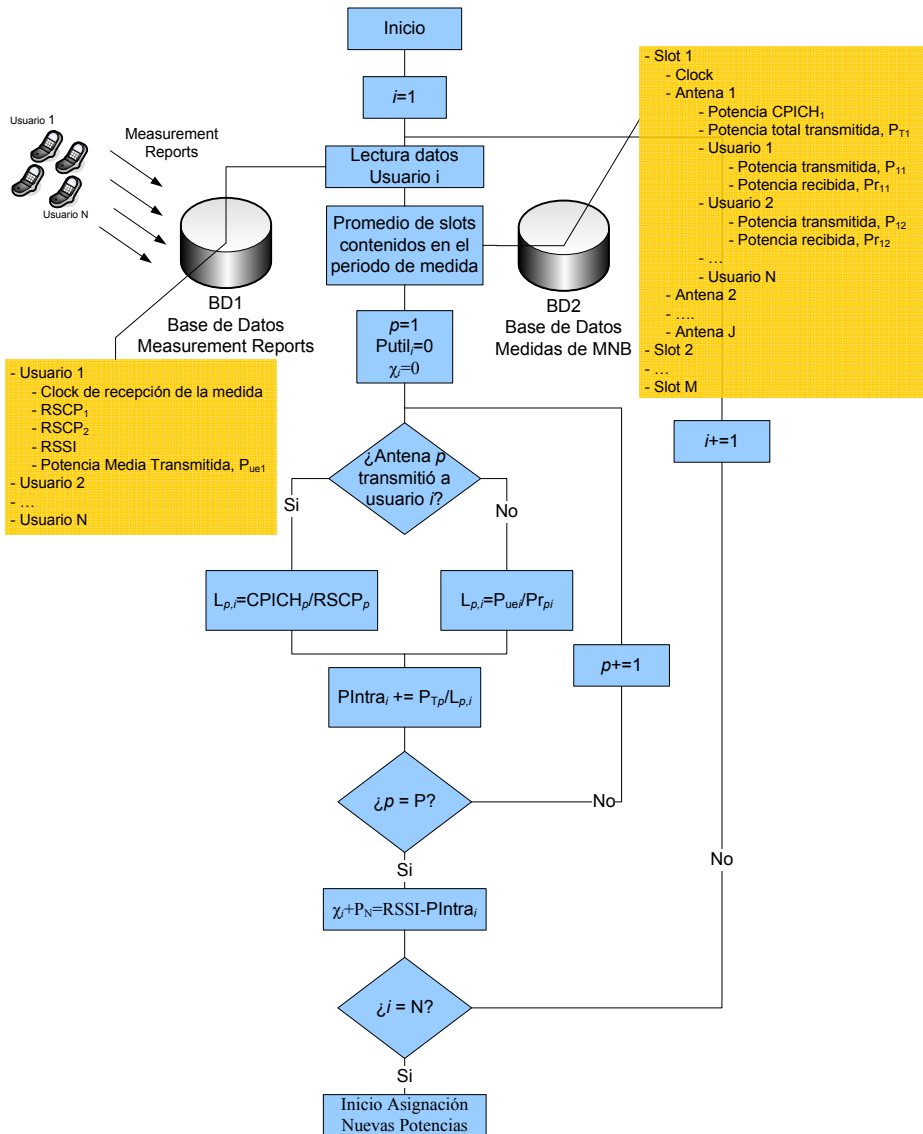


Figura 5.1 Cálculo de los parámetros del control de potencia en bucle abierto

### 5.2.1.2 Control de potencia en bucle cerrado

Cada cierto tiempo el algoritmo de *scheduling* establece la asignación de recursos para el siguiente periodo. A partir de este momento será el control de potencia en bucle cerrado rápido el que solucionará los posibles cambios en las condiciones de propagación. Como se explicó en el Capítulo 4, el control de potencia en bucle cerrado rápido aumenta o reduce en 1 dB la potencia asignada al usuario si no se satisface la  $E_b/N_0$  objetivo del usuario o viceversa. Cuando se trabaja con el Multinodo B, la distribución de potencias,  $q_{pik}$ , asignada por el *scheduler* se mantiene constante hasta la siguiente ejecución del gestor de recursos. Si no se satisfacen los niveles de calidad deseados se aumentará en 1dB por *slot* la potencia total asignada al usuario,  $P_i$ , de manera que en función de los valores de  $q_{pik}$  se aumentará más o menos la potencia asignada a cada antena.

Por su parte, el control de potencia en bucle cerrado externo se implementará en el simulador al considerar un valor de  $E_b/N_0$  objetivo diferente para cada una de las configuraciones de transmisión, es decir un valor distinto de  $E_b/N_0$  en función de si transmite una o dos antenas y del nivel de potencia relativo entre las contribuciones recibidas. Ante la movilidad de los usuarios el control de potencia en bucle externo se supone lo suficientemente rápido (100 Hz) como para cambiar convenientemente el valor de  $E_b/N_0$  para ajustarse a la nueva posición. Esta aproximación es razonable, puesto que en 10 ms a una velocidad de 120 km/h el espacio recorrido por un terminal es de apenas 0.3 metros, por lo que las condiciones de propagación serán relativamente constantes (claro está sin considerar el fenómeno de *fast fading*). De esta manera, puede suponerse que el control de potencia en bucle externo puede seguir perfectamente los cambios en las condiciones del canal.

## 5.2.2 Selección de celda

Uno de los procesos fundamentales en un sistema de comunicaciones móviles es determinar la celda o celdas a las que debe conectarse un terminal móvil en cada momento. En el simulador utilizado en esta Tesis Doctoral sólo se ha considerado una celda MNB, por lo que la selección de celda no consiste tanto determinar qué Nodo-B da servicio al usuario, sino qué antenas van a transmitir simultáneamente al usuario.

Dada una determinada relación entre las pérdidas de propagación experimentadas por las señales recibidas del MNB, existe una proporción óptima de potencia transmitida desde cada antena para minimizar la potencia consumida por el sistema global. En la arquitectura MNB-activo es la unidad de control la encargada de determinar la asignación de potencias (ADP), mientras que en la MNB-pasivo es la unidad de procesado en *array* del Multi Nodo-B la que realiza esa función.

En esta Tesis Doctoral se proponen tres variantes de selección de antena. Las dos primeras se corresponden con la arquitectura MNB-pasivo, mientras que la última se ajusta de forma clara a la arquitectura MNB-activo.

En la primera variante, denominada de Asignación de Recursos por Mínima Potencia (ARMIP), es el RNC quien asigna los códigos a todos los usuarios, considerando como única antena transmisora del Nodo-B la antena principal. Posteriormente, la unidad de procesado en *array* ajustará la potencia transmitida desde cada una de las antenas desplegadas del MNB, garantizando el consumo mínimo de potencia.

Para transmitir la mínima potencia posible a un usuario se ha de maximizar el valor de  $D_{ijk}$ , expresado en la ecuación (5.2). Para ello, lo primero que tiene que hacer el sistema es calcular las pérdidas de propagación experimentadas por la señal transmitida desde cada antena hasta el terminal. El método utilizado para este cálculo es el mismo que se resume en la Figura 5.1 y que utiliza las medias enviadas por el propio equipo. A partir de aquí, y puesto que sólo se va a transmitir a un usuario desde dos antenas como máximo, se seleccionan los dos transmisores que presenten mejores características de propagación. Posteriormente, es necesario discretizar los valores de

$q_{pik}$  de manera que para cada uno de esos valores se pueda extraer el valor de  $E_b/N_0$  objetivo asociado a la diferencia de potencia experimentada en el móvil entre ambas contribuciones y al tipo de antenas involucradas en la transmisión. Con cada  $q_{pik}$  y su valor de  $E_b/N_0$  asociado, se calcula el valor de  $D_{ijk}$ , utilizando para ello la ecuación (5.2). Finalmente, se selecciona como  $q_{pik}$  óptimo aquel que se corresponda con un valor de  $D_{ijk}$  máximo.

En la segunda variante, si la relación de potencia recibida por el móvil desde ambas antenas es inferior a un determinado margen, entonces se fuerza al sistema a transmitir simultáneamente por ambas antenas, a pesar de que esa opción pueda no resultar óptima en cuanto a consumo global de recursos. A esta segunda variante, muy similar al clásico *soft handover*, se le ha denominado técnica de Asignación de Recursos por Márgenes de Diversidad (ARMAD). Tal y como se ve en el Anexo B, el algoritmo ARMIP coincidirá con el ARMAD siempre que el margen considerado sea inferior al característico del escenario concreto de propagación. Para valores superiores a este margen, el protocolo ARMAD necesariamente consumirá más recursos de potencia, a costa de un mayor balanceo de carga entre las antenas del MNB.

El tercer algoritmo de selección de celda se corresponde con la arquitectura de red de acceso radio MNB-activo y considera simultáneamente la asignación de potencia y de códigos. Ya que el sistema UMTS está limitado por interferencias, reducir la potencia transmitida a un usuario tiene un efecto directo en el nivel de interferencia global y, por tanto, permite aumentar la capacidad de la celda. De esta manera, asignar los recursos considerando una potencia transmitida por usuario, cuando, posteriormente, ésta se va a reducir, hace que se infrutilicen los recursos realmente disponibles. El algoritmo de selección de celda denominado de Asignación Combinada de Recursos y Potencias, ACORP, se basa en considerar directamente la potencia final consumida por el usuario tras la asignación de antenas transmisoras, para así distribuir los recursos de manera óptima. El algoritmo ACORP está integrado dentro del *scheduler* basado en redes Hopfield propuesto en esta Tesis Doctoral para el MNB, cuya descripción completa se realiza a continuación.

## 5.3 Cronoejecución mediante Redes Hopfield

Uno de los principales objetivos de esta Tesis Doctoral es diseñar un algoritmo de gestión de recursos que permita obtener las mejores prestaciones de la nueva arquitectura radio Multinodo B. En este capítulo se exponen un conjunto de algoritmos alternativos de gestión de recursos, donde cada uno de ellos estará adaptado a una de las dos arquitecturas MNB descritas en el Capítulo 2.

En el Capítulo 4 se demostró el enorme potencial de las redes neuronales de Hopfield para la resolución del problema de asignación dinámica de recursos en UMTS, problema de optimización sumamente complejo. Tras plantear una función de energía, se evaluaron las prestaciones de la red Hopfield comparando los resultados obtenidos con otros algoritmos también basados en la minimización del *delay* de los paquetes. Los resultados de la comparación fueron aplastantes: el algoritmo desarrollado basado en redes de Hopfield reducía la potencia consumida, maximizaba el uso de los recursos y optimizaba el *delay* de los paquetes.

De las dos arquitecturas básicas MNB, la primera, denominada MNB-pasivo, se caracteriza porque la distribución de recursos entre los usuarios se decide en el RNC y, por lo tanto, el MNB no tiene más remedio que asumir dicha distribución de códigos. El MNB sólo tiene que distribuir las potencias entre las distintas antenas, con el objetivo de minimizar el nivel de interferencia introducido en la interfaz aire. En esta arquitectura, el algoritmo utilizado para la distribución de recursos entre los usuarios puede ser cualquiera de los comentados en el Capítulo 4: o utilizando redes neuronales o bien cualquiera de los disponibles en la literatura. A partir de aquí, y tal y como se ha mostrado en el apartado 5.2.2, se puede seleccionar el algoritmo de asignación de potencia entre las antenas que minimice el nivel de interferencia (ARMIP) o asegure un mayor balanceo de la carga entre las antenas del MNB (ARMAD).

La segunda arquitectura, la MNB-activa, se caracteriza porque es el propio MNB el que decide simultáneamente qué recursos se van a asignar a cada usuario y qué distribución de potencia se va a llevar a cabo entre las antenas. De esta manera, se puede aprovechar la reducción en los niveles de interferencia conseguida con la



distribución de antenas para aumentar la capacidad de la celda y así mejorar la calidad de servicio final experimentada por el usuario.

De nuevo, el algoritmos de gestión de recursos propuesto para esta segunda arquitectura está basado en redes neuronales Hopfield. Lo que se propone en esta Tesis Doctoral es extender la función de energía propuesta en el Capítulo 4 para incluir, asimismo, la asignación de antenas y potencia. Gracias a la toma de decisión simultánea, se podrá asignar de una manera mucho más eficiente los recursos, como se demostrará en el apartado de resultados. A continuación, se describe la función de energía que caracteriza al algoritmo que se ha denominado ACORP.

### 5.3.1 Formulación del problema

Antes de pasar a definir la función de energía, primero se describe la formulación empleada. Para el MNB se ha considerado una agrupación lógica de las neuronas en forma de matriz tridimensional. Al igual que antes, la implementación física de la red neuronal no depende de esta agrupación lógica, que sólo pretende simplificar la identificación de la funcionalidad de las distintas neuronas. Se define la matriz de salida de las neuronas,  $\overline{V}$ , de dimensiones  $N \times M \times K$  donde  $N$  es el número de usuarios del sistema,  $M$  es el máximo número de tasas binarias disponibles, definido de la misma manera que en el caso de redes convencionales, y  $K$  es el número de elementos del conjunto de proporciones de potencia que se puede asignar a la antena principal. Por ejemplo, en las simulaciones llevadas a cabo se ha considerado el siguiente conjunto de proporciones de potencia,  $\mathfrak{K} = \{1, 0.6, 0.5, 0.4, 0\}$ , de manera que  $K=5$ . Si, por ejemplo, el algoritmo de gestión de recursos decidiera que para el usuario  $i$  la relación de potencia óptima es  $q_i=0.6$ , eso quiere decir que el 60% de la potencia asignada al usuario se transmitirá por la antena principal, mientras que el 40% restante se transmitirá por la antena secundaria más próxima al equipo.

El estado final de las neuronas representa la asignación de recursos llevada a cabo. De esta manera, si la neurona  $V_{ijk}$  está en estado ON, eso quiere decir que al usuario  $i$ -

ésimo se le ha asignado la tasa binaria  $j$ -ésima con la  $k$ -ésima distribución de potencia, mientras que el resto de neuronas del mismo usuario, que coincide con el resto de elementos de la misma fila de la matriz  $\overline{\overline{V}}$ , debe estar en estado OFF. La Figura 5.2 representa esta idea.

La ecuación que describe la dinámica de la red Hopfield debe modificarse para adecuarse a la nueva formulación de la siguiente manera:

$$U_{ijk}^{[n+1]} = U_{ijk}^{[n]} - \Delta t \left( \frac{\partial E}{\partial V_{ijk}} \right) = U_{ijk}^{[n]} + \Delta t \left( \sum_{r=1}^N \sum_{s=1}^M \sum_{t=1}^K T_{ijk,rst} V_{rst}^{[n]} + I_{ijk} \right) \quad (5.4)$$

La matriz  $\overline{\overline{T}}$  de interconexión entre neuronas sigue siendo una matriz de dos dimensiones. La  $MK$  primeras entradas representan al primer usuario, las  $MK$  siguientes con el segundo y así sucesivamente. Para el primer usuario las  $K$  primeras entradas se corresponden con la asignación del primer elemento del conjunto  $\mathfrak{X}_i$  para cada una de las distribuciones de potencia del conjunto  $\mathfrak{K}$ , las  $K$  siguientes se corresponden con la segunda tasa binaria y así sucesivamente. De esta manera, el indicador de fila  $ijk$  se corresponden con la siguiente fila de la matriz  $\overline{\overline{T}}$ :

$$(i-1) \cdot M \cdot K + (j-1) \cdot K + k \quad (5.5)$$

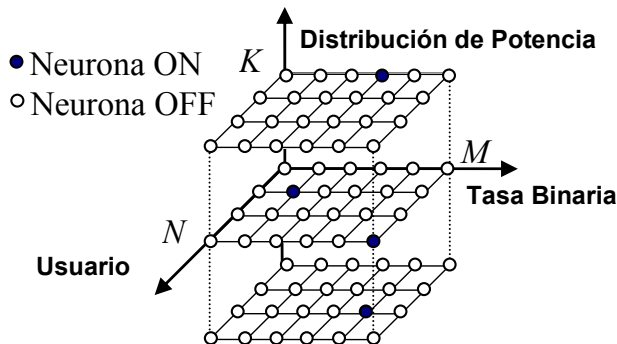


Figura 5.2 Red Hopfield 3-D para la asignación de recursos

y de igual forma se procede para calcular la columna.

Una vez descrita la formulación empleada, es necesario definir una función de energía de manera que su mínimo se corresponda con la mejor asignación posible de recursos en el sistema MNB. La función de energía del algoritmo ACORP propuesto en esta Tesis Doctoral para la red MNB es la siguiente:

$$\begin{aligned}
 E = & -\mu_1 \sum_{i=1}^N O_i \sum_{j=1}^M C_{ij} \sum_{k=1}^K V_{ijk} - \mu_2 \sum_{i=1}^N O_i \sum_{j=1}^M \frac{R_j}{R_{\max,i}} \sum_{k=1}^K V_{ijk} + \\
 & + \frac{\mu_3}{2} \sum_{i=1}^N \sum_{j=1}^M \frac{R_j}{R_{\max,i}} \sum_{k=1}^K \xi_{ijk} V_{ijk} + \frac{\mu_4}{2} \sum_{i=1}^N \sum_{j=1}^M \psi_{ij} \sum_{k=1}^K V_{ijk} + \frac{\mu_5}{2} \sum_{i=1}^N \sum_{j=1}^M \left( \frac{R_j}{R_{\max,i}} \cdot \sum_{k=1}^K \zeta_{ijk} \cdot V_{ijk} \right) \\
 & + \frac{\mu_6}{2} \sum_{i=1}^N \sum_{j=1}^M \sum_{k=1}^K V_{ijk} (1 - V_{ijk}) + \frac{\mu_7}{2} \sum_{i=1}^N \left( 1 - \sum_{j=1}^M \sum_{k=1}^K V_{ijk} \right)^2
 \end{aligned} \quad (5.6)$$

Las constantes  $\mu_1$  a  $\mu_7$  se han de ajustar convenientemente para garantizar la rápida convergencia a la solución deseada. A continuación, se va a proceder a justificar uno a uno todos los términos que conforman la ecuación, analizando el efecto introducido por cada uno de ellos. En los términos que sean idénticos a los del algoritmo DC-HNN-DRA no se entrará en detalle.

### ***Primer Término de la Función de Energía***

Este término actuará de igual manera que en el caso del sistema convencional. De hecho, la función de coste no depende de la distribución de potencias sino únicamente de la tasa binaria asignada a cada usuario.

### ***Segundo Término de la Función de Energía***

Al igual que el término anterior, éste no se modifica respecto a la formulación original puesto que tiene en cuenta únicamente la tasa asignada.

**Tercer Término de la Función de Energía**

Este término penaliza cualquier asignación de recursos que provoque que se superen los máximos recursos disponibles en el sistema. Para ello se hace uso de la función  $\xi_{ijk}$  que tomará únicamente los valores 0 ó 1. El desarrollo completo del cálculo de la potencia total transmitida por el MNB se analiza en el Anexo B, aunque la ecuación (5.1) recoge su expresión matemática final. Teniendo en cuenta esta expresión, se puede calcular  $P_T$  a partir de los estados de las neuronas como:

$$P_T = \frac{\sum_{i=1}^N \sum_{j=1}^M \sum_{k=1}^K \frac{P_N + \chi_i}{D_{ijk}} V_{ijk} + \sum_{p=1}^P P_{C,p}}{1 - \sum_{i=1}^N \sum_{j=1}^M \sum_{k=1}^K \frac{\sum_{p=1}^P \frac{\rho_p Q_p}{L_{p,i}}}{D_{ijk}} V_{ijk}} \tag{5.7}$$

A partir de esta expresión, el tercer término de la energía se define como:

$$\xi_{ijk} = u \left( \frac{H_{ijk}}{\eta_{\max} P_{T \max}} - 1 \right) \parallel u \left( \frac{\sum_{\substack{r=1 \\ r \neq i}}^N \sum_{s=1}^M \sum_{t=1}^K R_s V_{rst} + R_j}{R_{\text{System,max}}} - 1 \right) \tag{5.8}$$

donde

$$H_{ijk} = \frac{\sum_{\substack{r=1 \\ r \neq i}}^N \sum_{s=1}^M \sum_{t=1}^K \frac{P_N + \chi_r}{D_{rst}} V_{rst} + \frac{P_N + \chi_i}{D_{ijk}} + \sum_{p=1}^P P_{C,p}}{1 - \sum_{\substack{r=1 \\ r \neq i}}^N \sum_{s=1}^M \sum_{t=1}^K \frac{\sum_{p=1}^P \frac{\rho_p Q_p}{L_{p,r}}}{D_{rst}} V_{rst} - \frac{\sum_{p=1}^P \frac{\rho_p Q_p}{L_{p,i}}}{D_{ijk}}} \tag{5.9}$$

**Cuarto Término de la Función de Energía**

Este término se encarga de evitar el uso de tasas binarias no permitidas para un determinado tipo de servicio o para un usuario particular. La definición de la matriz  $\psi_{ij}$  será la misma que la utilizada para redes UMTS convencionales.

### Quinto Término de la Función de Energía

Este término favorecerá que se seleccione el factor  $q_{ijk}$  que menor consumo de potencia suponga al sistema. Para tener en cuenta esto, existe la matriz  $\zeta_{ijk}$ . Para construir esta matriz se debe estudiar qué relación de potencia entre ambas estaciones base es más beneficiosa para cada usuario y cada tasa binaria  $(i,j)$ . Lo que se pretende es ordenar los elementos del conjunto  $\mathfrak{K}$ , de manera que si una determinada distribución de potencias es la que consumiría menos recursos, la matriz  $\zeta_{ijk} = 0$ , mientras que la relación más perjudicial, es decir, la que requiriese de un mayor consumo de potencia por parte de la estación base, tendría un valor de  $\zeta_{ijk} = K - 1$ .

A continuación se vuelve a mostrar la expresión de la potencia transmitida al usuario  $i$  para asignarle la tasa binaria  $j$  con una distribución de potencias específica:

$$P_{ijk} = \frac{P_N + \chi_i + P_T \sum_{p=1}^P \frac{\rho_p Q_p}{L_{p,i}}}{D_{ijk}} \quad (5.10)$$

Teniendo en cuenta que el numerador es indiferente de la distribución de potencias entre las antenas del MNB,  $q_{ijk}$ , minimizar (5.10) es lo mismo que maximizar  $D_{ijk}$ .

A partir de la expresión de  $D_{ijk}$ , y como solamente se considera la transmisión desde dos antenas simultáneamente, es sencillo comprobar qué elemento del conjunto  $\mathfrak{K}$  maximiza esta expresión. Por ejemplo, dado el conjunto  $\mathfrak{K} = \{1, 0.6, 0.5, 0.4, 0\}$ , si se transmitiera desde la primera estación y la tercera, el elemento  $k = 2$  tendría la siguiente expresión:

$$D_{ij2} = \frac{0.6}{L_{1,i}} \left( \rho_1 + \frac{W}{\left( \frac{E_b}{N_0} \right)_{ij2} R_j} \right) + \frac{0.4}{L_{3,i}} \left( \rho_3 + \frac{W}{\left( \frac{E_b}{N_0} \right)_{ij2} R_j} \right) \quad (5.11)$$

Tras calcular los  $K$  elementos  $D_{ijk}$  y ordenarlos por orden decreciente, se construirá la matriz  $\zeta_{ijk}$ , teniendo en cuenta que el elemento igual a cero será el que maximice  $D_{ijk}$  y, por tanto, minimice la potencia transmitida al usuario. Si dos asignaciones distintas conllevan un valor idéntico de  $D_{ijk}$ , se considerará siempre aquella que transmita más potencia desde la antena principal. Con esta medida más conservadora se pretende reducir al mínimo la potencia transmitida por las antenas remotas.

Por último, el quinto término también está normalizado respecto a la tasa binaria, asegurando así ser más restrictivo para las tasas binarias más altas, ya que estas consumen muchos más recursos. Para las tasas binarias bajas se admitirá con más facilidad que no se asigne la distribución de antenas óptima, si así se optimiza el funcionamiento del resto del sistema.

#### ***Sexto Término de la Función de Energía***

El sexto término fuerza a las neuronas a irse a los extremos, o bien  $V_{ijk} = 0$  o bien  $V_{ijk} = 1$ , ya que para estos casos este término es igual a 0, mientras que para el resto este término perjudica, o sea, aumenta la función de energía.

#### ***Séptimo Término de la Función de Energía***

Al igual que la anterior formulación, este término garantiza, para cada usuario, que solamente se asigne una tasa binaria.

### **5.3.2 Dinámica del sistema y parámetros *hardware***

Como antes, el algoritmo de Hopfield se inicia con todas las neuronas inactivas, es decir,  $V_{ijk} = 0 + \varepsilon$ . El gradiente de  $E$  resulta ser:

$$\begin{aligned} \frac{\partial E}{\partial V_{ijk}} = & -\mu_1 O_i C_{ij} - \mu_2 \frac{O_i R_j}{R_{\max,i}} + \frac{\mu_3}{2} \cdot \frac{R_j}{R_{\max,i}} \xi_{ijk} + \frac{\mu_4}{2} \psi_{ij} + \\ & + \frac{\mu_5}{2} \cdot \frac{R_j}{R_{\max,i}} \zeta_{ijk} + \frac{\mu_6}{2} (1 - 2V_{ijk}) - \mu_7 \left( 1 - \sum_{l=1}^M \sum_{m=1}^K V_{ilm} \right) \end{aligned} \quad (5.12)$$

Considerando la estructura matricial de las neuronas, la nueva formulación del gradiente de la función de energía resulta ser:

$$\frac{\partial E}{\partial V_{ijk}} = - \sum_{r=1}^N \sum_{s=1}^M \sum_{t=1}^K T_{ijk,rst} V_{rst} - I_{ijk} \quad (5.13)$$

Ahora solo resta identificar los parámetros *hardware* de la red Hopfield mediante comparación directa entre (5.12) y (5.13). La matriz  $\overline{\overline{T}}$  y el vector de corrientes son:

$$\begin{aligned} T_{ijk,rst} &= \overline{\mu_6} \delta_{ir} \delta_{js} \delta_{kt} - \mu_7 \delta_{ir} \\ I_{ijk} &= \mu_1 O_i C_{ij} + \mu_2 \frac{O_i R_j}{R_{\max,i}} - \frac{\mu_3}{2} \frac{R_j}{R_{\max,i}} \xi_{ijk} - \frac{\mu_4}{2} \psi_{ij} - \frac{\mu_5}{2} \frac{R_j}{R_{\max,i}} \zeta_{ijk} - \frac{\mu_6}{2} + \mu_7 \end{aligned} \quad (5.14)$$

Como se puede apreciar, de nuevo la matriz  $\overline{\overline{T}}$  es simétrica y el valor de las resistencias constante, mientras que las corrientes variarán dinámicamente en función del estado de la red y de las características de transmisión de cada usuario.

### 5.3.3 Cálculo de las constantes

El cálculo de las constantes realizado en el Capítulo 4 es válido sea cual sea la proporción de potencia asignada a cada una de las antenas. La evolución de las tasas binarias asignadas a un usuario se rige por el *delay* y por los recursos disponibles, por lo que la evolución de las neuronas será la misma en la segunda dimensión independientemente de lo que ocurra en la tercera.

Es por esto que el proceso descrito para calcular las constantes en el Capítulo 4 se puede también emplear para la formulación utilizada en el MNB. Así pues, solo resta

obtener el valor por el que se debe ponderar el quinto término para garantizar la mejor asignación posible de recursos entre las antenas.

**Quinto Término**

Uno de los objetivos principales de la red neuronal es no exceder nunca los recursos máximos disponibles del sistema. En caso de que una determinada asignación de potencia entre antenas supere los recursos (situación diferenciada por el acrónimo *exc*) y otra no (*fav*), se deberá asegurar que la segunda tiene un gradiente más pequeño que favorezca su asignación, incluso cuando el quinto término esté forzando su no utilización por suponer un mayor consumo de potencia para el usuario particular. Suponiendo que el usuario puede hacer uso de todas las tasas binarias disponibles, en el caso de la neurona favorecida (*i,j,fav*) el gradiente de la energía es:

$$\frac{\partial E}{\partial V_{i,j,fav}} = -\mu_1 O_i C_{ij} - \mu_2 \frac{O_i R_j}{R_{max,i}} + \frac{\mu_5}{2} \frac{R_j}{R_{max,i}} \zeta_{i,j,fav} + \frac{\mu_6}{2} (1 - 2V_{i,j,fav}) - \mu_7 \left( 1 - \sum_{l=1}^M \sum_{m=1}^K V_{ilm} \right) \quad (5.15)$$

Mientras que el gradiente de la energía de la neurona que sí que excede los recursos máximos del sistema (*i,j,exc*) es:

$$\begin{aligned} \frac{\partial E}{\partial V_{i,j,exc}} = & -\mu_1 O_i C_{ij} - \mu_2 \frac{O_i R_j}{R_{max,i}} + \frac{\mu_3}{2} \frac{R_j}{R_{max,i}} + \frac{\mu_5}{2} \frac{R_j}{R_{max,i}} \zeta_{i,j,exc} + \\ & + \frac{\mu_6}{2} (1 - 2V_{i,j,exc}) - \mu_7 \left( 1 - \sum_{l=1}^M \sum_{m=1}^K V_{ilm} \right) \end{aligned} \quad (5.16)$$

La condición necesaria para garantizar que se realice una asignación correcta de recursos es que

$$\frac{\partial E}{\partial V_{i,j,fav}} < \frac{\partial E}{\partial V_{i,j,exc}} \quad (5.17)$$

con lo que



$$\mu_5 < \frac{\mu_3 + 2\mu_6 \frac{R_{\max,i}}{R_j} (V_{i,j,\text{fav}} - V_{i,j,\text{exc}})}{\zeta_{i,j,\text{fav}} - \zeta_{i,j,\text{exc}}} \quad (5.18)$$

El caso peor se encuentra cuando la  $R_j$  es mínima,  $\zeta_{i,j,\text{exc}} = 0$ ,  $\zeta_{i,j,\text{fav}} = K - 1$ ,  $V_{i,\text{exc}} = 1$  y  $V_{i,\text{fav}} = 0$ . En este caso

$$\mu_5 < \frac{\mu_3 - 2\mu_6 \frac{R_{\max,i}}{R_j}}{K - 1} \quad (5.19)$$

Así pues, el valor de  $\mu_5$  será siempre el mayor posible para evitar consumir más potencia innecesariamente, pero siempre cumpliendo (5.19). Dado el caso particular de las tasas binarias definidas en esta Tesis Doctoral,  $\mu_5$  deberá cumplir

$$\mu_5 < \frac{\mu_3 - 32\mu_6}{K - 1} \quad (5.20)$$

En concreto, en las simulaciones realizadas se ha considerado un valor de  $\mu_1 = \mu_2 = 500$ ,  $\delta = 0.1$  y  $K = 5$ , por lo que el resto de constantes tienen los siguientes valores:  $\mu_6 = 15$ ,  $\mu_3 = 17500$ ,  $\mu_7 = 10000$ ,  $\mu_4 = 22000$  y  $\mu_5 = 4200$ .

## 5.4 Análisis de Resultados y Discusión

En el Capítulo 4 se demostró que la función de energía propuesta para la red neuronal de Hopfield mejoraba considerablemente las prestaciones del sistema UMTS para distintos tipos de servicio. Por esta razón se ha seleccionado también el uso de redes neuronales para la gestión de recursos en el MNB. Como se acaba de ver, la función de energía propuesta para el MNB es exactamente la misma que en el capítulo anterior, salvo que se ha añadido un término adicional para realizar la distribución de recursos entre las antenas del MNB. El estudio realizado en este apartado no pretende evaluar el comportamiento del *scheduler* basado en *delay*, sino, más bien, se centra en comparar las distintas arquitecturas MNB consideradas y sus diferentes alternativas de distribución de potencia entre las antenas.

Para simplificar el análisis de los resultados se ha considerado un único tipo de tráfico, el servicio web, por ser éste el servicio más completo y significativo. Además, se ha seleccionado un número fijo de usuarios, en concreto 70. Con este número de usuarios el sistema UMTS convencional está ya congestionado, tal y como se demostró en el Capítulo 4. Lo que se pretende es evaluar hasta qué punto se consigue mejorar el consumo de potencia y la calidad de servicio para este escenario concreto, utilizando para ello la nueva arquitectura radio Multinodo B.

A continuación, se describirán los distintos algoritmos utilizados para distribuir los recursos entre los usuarios y las antenas del MNB. Posteriormente se analizarán todos los escenarios de simulación considerados, detallando las características de cada uno de ellos. Finalmente se evaluarán y discutirán los resultados obtenidos.

### 5.4.1 Algoritmos de distribución de recursos utilizados

El objetivo fundamental del análisis de resultados es comparar el nuevo algoritmo de cronoejecución, ACORP, desarrollado mediante redes neuronales de Hopfield, con otras alternativas más simples. Dentro de los otros algoritmos de referencia se han

incluido no sólo algoritmos para el MNB sino también otras técnicas más clásicas con las que comparar el funcionamiento global del MNB.

### ***Asignación de Recursos UMTS Convencional (Sin MNB)***

En primer lugar, es fundamental comparar la nueva solución MNB con la arquitectura UMTS convencional. Para ello se utilizarán los resultados que se obtuvieron en el Capítulo 4 utilizando el algoritmo DC-HNN-DRA.

### ***Asignación de Macro o Micro-celda sin Diversidad (Sin Diversidad)***

Según este algoritmo, cada vez que se ejecuta el algoritmo de gestión de recursos se selecciona desde qué antena el usuario recibe mejor calidad de señal. La transmisión al usuario se hará únicamente considerando dicha antena. El algoritmo utilizado para la distribución de recursos es el basado en redes Hopfield explicado en este capítulo, con la salvedad de que el vector de proporciones de tasas binarias será  $\mathbf{x} = \{1, 0\}$ , es decir, o se transmite desde la antena principal o desde la secundaria. Este algoritmo implementa la idea propuesta por Dong y Wha en [12] y representa una aproximación idélica a las estructuras híbridas celulares clásicas, ya que supone que no hay ningún *delay* en el *handover* intra-frecuencial.

### ***Transmisión con Diversidad desde un Único Emplazamiento (CLMTD)***

Otra solución planteada en el estándar es transmitir con diversidad desde un único emplazamiento para mejorar la calidad del enlace [15]. Este mecanismo de transmisión también se ha evaluado en esta Tesis Doctoral, de manera que se consideran dos transmisores localizados en el mismo emplazamiento transmitiendo con diversidad en bucle cerrado. Respecto a la diversidad introducida, los algoritmos de realimentación y

estimación de canal utilizados en este algoritmo son los propuestos en esta Tesis Doctoral para el MNB, por lo que la ganancia por diversidad macro-celular calculada en el Capítulo 3 será la empleada en este algoritmo. De nuevo, para asignar los recursos se va a considerar el algoritmo Hopfield explicado en este capítulo, pero con la salvedad de que los dos transmisores están situados en el mismo emplazamiento, compartiendo por tanto pérdidas de propagación y *shadowing* hasta cualquier punto del escenario de simulación.

### ***Asignación Combinada de Recursos y Potencias (ACORP)***

Esta es la primera de las soluciones MNB propuesta en esta Tesis Doctoral y se corresponde con la arquitectura MNB-activa, es decir, la toma de decisión sobre asignación de códigos y distribución de potencia se realiza simultáneamente. Para dicha asignación se utiliza la red Hopfield descrita en el apartado 5.3, considerando un vector de proporciones de potencia entre antenas  $\mathfrak{K} = \{1, 0.6, 0.5, 0.4, 0\}$ .

### ***Asignación de Recursos por Mínima Potencia (ARMIP)***

Esta técnica se corresponde con una arquitectura MNB-pasivo. La distribución de códigos se realizará utilizando el algoritmo DC-HNN-DRA y tras esta asignación se distribuirá la potencia entre las antenas de manera que se minimice la potencia transmitida al usuario. Para ello se seleccionará el elemento del conjunto  $\mathfrak{K}$  que maximice (5.2).

### ***Asignación de Recursos por Márgenes de Diversidad (ARMAD)***

En esta técnica la distribución de códigos también se realiza mediante el algoritmo DC-HNN-DRA. Por su parte, la distribución de potencia entre las antenas se ejecuta de

manera muy similar al *soft handover* de UMTS, es decir, se define un margen en la diferencia relativa del nivel de potencia recibido desde ambas antenas, de forma que, dentro de ese margen, se transmite simultáneamente desde ambas. Cuanto mayor sea el margen mayor será el balanceo de la red pero mayor también el consumo de recursos.

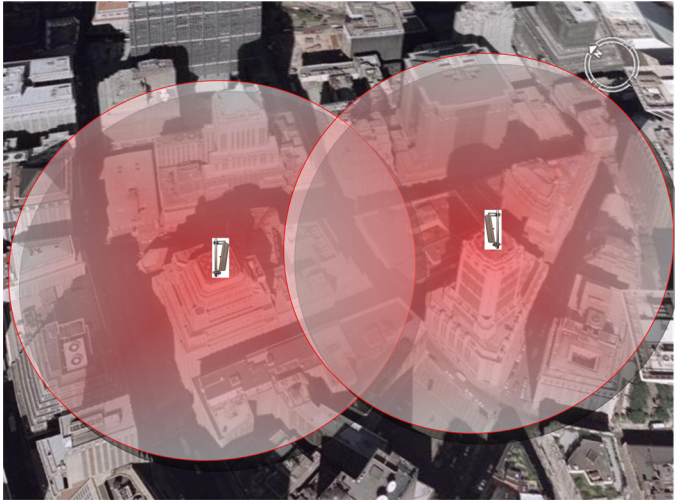
## 5.4.2 Escenarios de simulación

Las condiciones generales de simulación son exactamente las mismas que se utilizaron en el capítulo anterior. Es necesario considerar el mismo escenario ya que se pretende comparar los resultados obtenidos con el MNB con la arquitectura radio UMTS convencional.

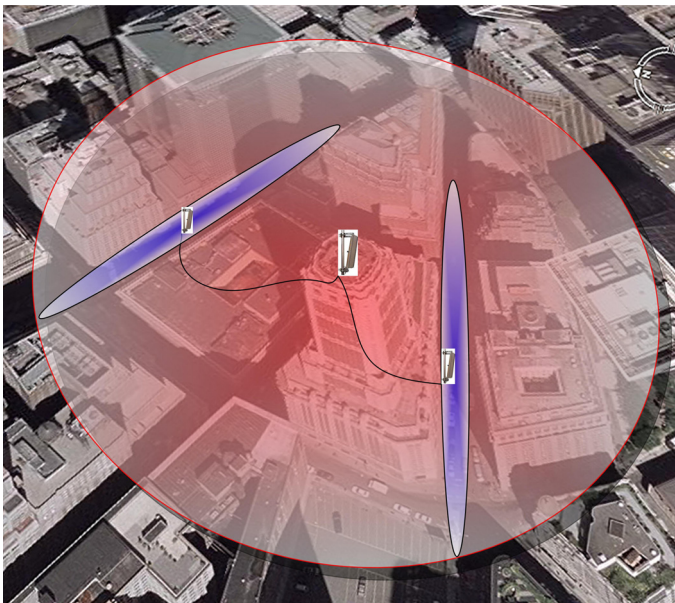
La única diferencia respecto a las anteriores simulaciones es que ahora se dispone de un conjunto adicional de antenas distribuidas en la celda, o micro o macro-celulares. Cabe destacar que el factor de correlación de las micro-celdas es de 0.9.

Tal y como se explicó en el Capítulo 2, el MNB hará uso de un método de transmisión con diversidad espacial estandarizado, conocido como CLMTD. Para garantizar el correcto funcionamiento de este sistema, y tal y como se ve en el Anexo B, todas las antenas del MNB deberán transmitir un canal piloto, o bien el primario o el secundario, mientras que el terminal en cualquier posición sólo debe recibir información de, como máximo, dos pilotos de la misma celda. Esto obliga a que en la zona de cobertura del MNB solamente se pueda percibir señal piloto de dos nodos simultáneamente, es decir, no puede existir más de dos coberturas solapadas. Si se utiliza una antena adicional macro-celular, como el escenario mostrado en la Figura 5.3, no se podrá añadir ninguna antena más, ya que si no podrían existir más de dos zonas de cobertura solapadas.

En cambio en un escenario mixto, en el que se añaden celdas micro-celulares, como la señal está confinada en las calles y presenta una gran atenuación, si el diseño se hace correctamente se evitará tener solapamiento entre los transmisores. Esta idea queda reflejada en la Figura 5.4.



**Figura 5.3 Escenario macro-celular**



**Figura 5.4 Escenario mixto**

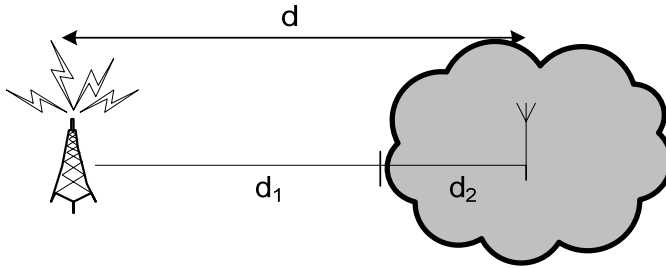
Cuando se trabaja con un escenario sintético con antenas omnidireccionales es necesario controlar que la señal transmitida por las antenas micro-celulares no se recibe en las áreas de cobertura de otras estaciones micro-celulares.

Tal y como viene determinado en la especificación 3GPP TS 25.133 [16], el terminal móvil debe ser capaz de medir el canal piloto, o *Common Pilot Channel* (CPICH), mientras reciba un nivel de potencia de éste de hasta -114 dBm. El radio de cobertura máximo dependerá por tanto de la potencia transmitida en el canal piloto, potencia comprendida entre 27 y 33 dBm. El valor más típico de potencia de P-CPICH para celdas situadas sobre azotea es de 30 dBm, tal y como se ha constatado hablando con diversos operadores.

Si se considera este valor típico para el transmisor macro-celular y aplicando la ecuación de pérdidas de propagación macro-celulares, descrita en el Anexo B, se obtiene un radio de cobertura de 1.54 km. Dado que el radio celular considerado en los estudios realizados es de 500 m, un valor de 20mW (13dBm) sería suficiente para garantizar la cobertura celular. Aún así, se ha preferido mantener 1W (30dBm) para la potencia de CPICH transmitida por el emplazamiento macro-celular.

Al repetir el cálculo de cobertura considerando 30 dBm de potencia transmitida en el canal piloto y el modelo de pérdidas por propagación micro-celular, el valor del radio de cobertura resulta ser de 1.12 km, un valor demasiado elevado para desplegar múltiples antenas en el escenario mixto. Resulta por tanto fundamental determinar el valor de la potencia de CPICH transmitida por las antenas micro-celulares para limitar el radio de cobertura al que más interese, de manera que se evite solapamientos indeseados.

Tal y como se demuestra en [17], las contribuciones obtenidas en el receptor *Rake* con niveles de potencia que disten más de 10 dB en potencia respecto a la componente principal se pueden despreciar, puesto que no afectan al comportamiento del receptor. De esta manera, dada una determinada distancia entre el transmisor macro-celular y el micro-celular,  $d$ , y teniendo en cuenta que, a priori, la potencia transmitida por ambas antenas es la misma, se puede calcular el radio máximo de la celda micro-celular,  $d_2$ , a partir del cual la diferencia de potencia respecto a la macro-celular es superior a 10 dB y por tanto no será útil.



**Figura 5.5 Esquema del cálculo de radio máximo**

Por lo tanto, la potencia de CPICH se calculará, en primera instancia, de forma que se garantice ese radio micro-celular de cobertura. La Figura 5.5 muestra un esquema del problema considerado.

El sistema de ecuaciones a resolver, teniendo en cuenta los modelos de propagación macro y micro-celulares, es el siguiente:

$$d_1 + d_2 = d$$

$$142 + 40 \log(d_2) = 137.4 + 35.2 \log(d_1) + 10 \quad (5.21)$$

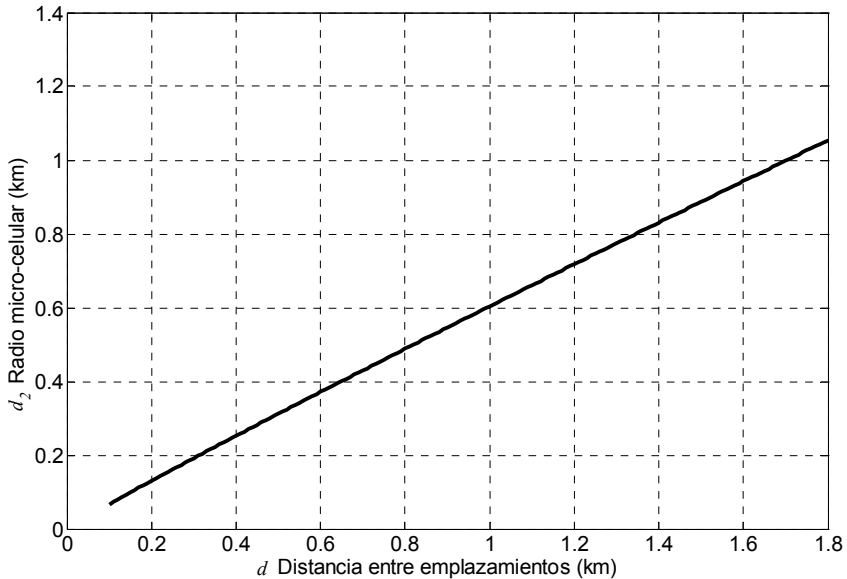
Operando convenientemente:

$$d_2 + 0.702 \cdot d_2^{1.1364} = d \quad (5.22)$$

La Figura 5.6 muestra el radio micro-celular resultante, obtenido tras resolver la ecuación (5.22) y considerando, además, distintos valores de distancia entre transmisores.

Utilizando la Figura 5.6 y conociendo la distancia entre la antena micro-celular y la central, se podría calcular el radio máximo de cobertura micro-celular y despejar de ahí la potencia necesaria del canal piloto. Sin embargo, este razonamiento es válido siempre que las áreas de cobertura de los transmisores micro-celulares no se solapen, ya que, en ese caso, tal y como se describió en el Capítulo 3, las contribuciones podrían ser destructivas y empeorarían gravemente la calidad de servicio experimentada por el usuario. Así pues, el radio micro-celular deberá ser el mínimo entre la mitad de la distancia entre emplazamientos micro-celulares y el valor de  $d_2$ .

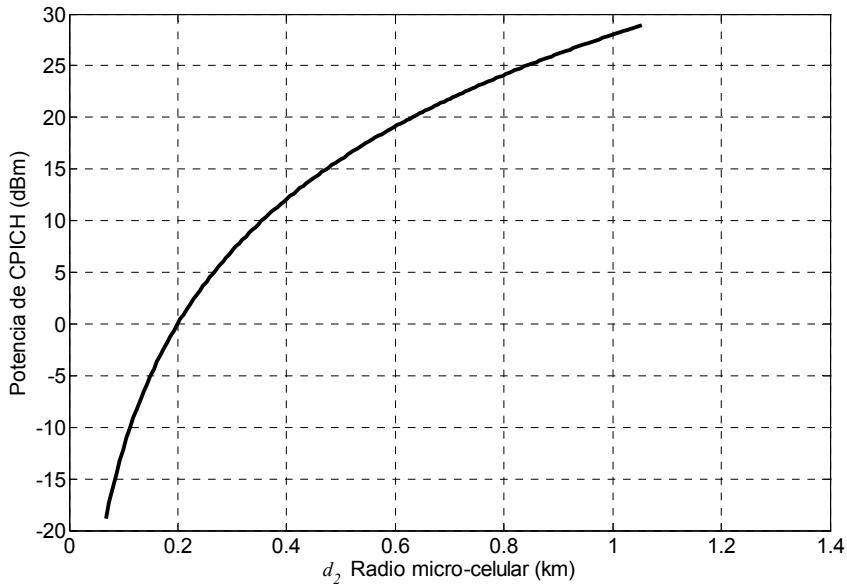




**Figura 5.6 Relación entre distancia entre emplazamientos y radio micro-celular máximo**

En resumen, el procedimiento a seguir para calcular la potencia de CPICH transmitida será el siguiente:

1. En primer lugar se definirá los emplazamientos de las antenas micro-celulares.
2. Para cada una de ellas se calculará la distancia entre el emplazamiento principal y el micro-celular y, utilizando la Figura 5.6, se obtendrá el valor de  $d_2$ .
3. El radio óptimo será el mínimo entre el valor de  $d_2$  y la mitad de la mínima distancia con otro emplazamiento micro-celular.
4. A partir del radio óptimo y la ecuación de pérdidas se determinará el nivel de potencia del canal piloto necesario para garantizar la cobertura micro-celular deseada. La Figura 5.7 muestra este valor de potencia de CPICH en dBm necesario para garantizar el radio de cobertura deseado.

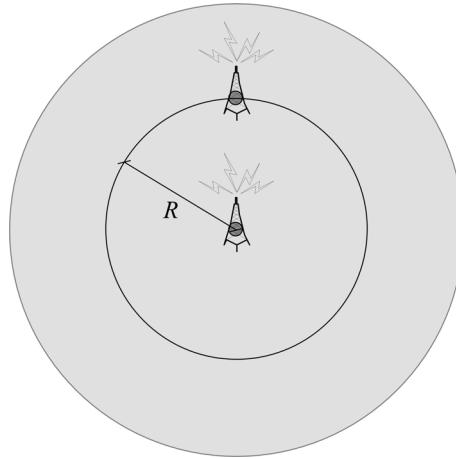


**Figura 5.7 Potencia de CPICH en función del radio de cobertura deseado**

Para poder evaluar el número de emplazamientos adicionales que interesaría introducir en la arquitectura MNB, se han considerado 8 escenarios distintos de distribución de antenas. A continuación, se analizarán los detalles más importantes que permiten conocer las peculiaridades de cada uno de estos escenarios.

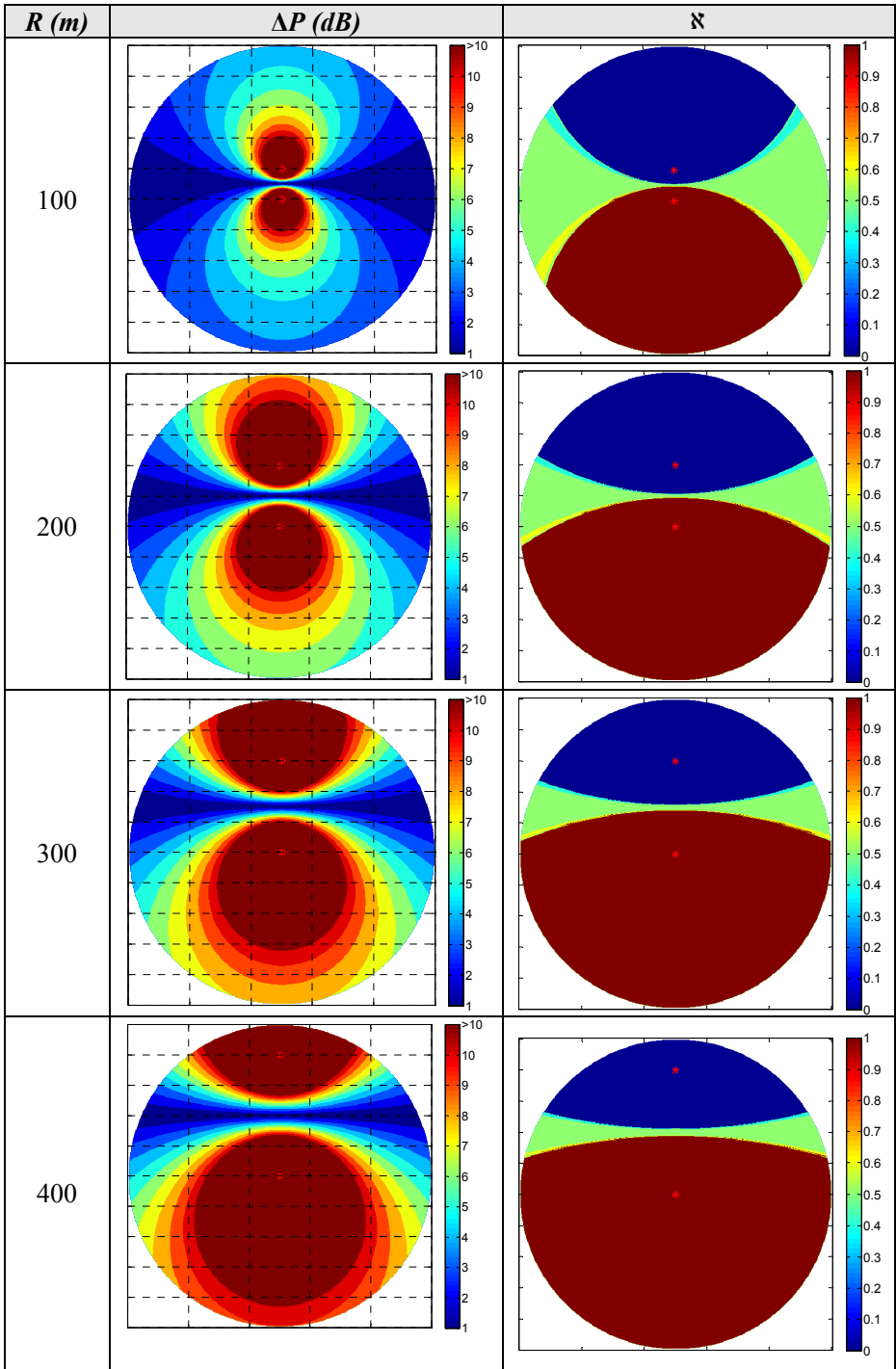
#### 5.4.2.1 Escenario 1

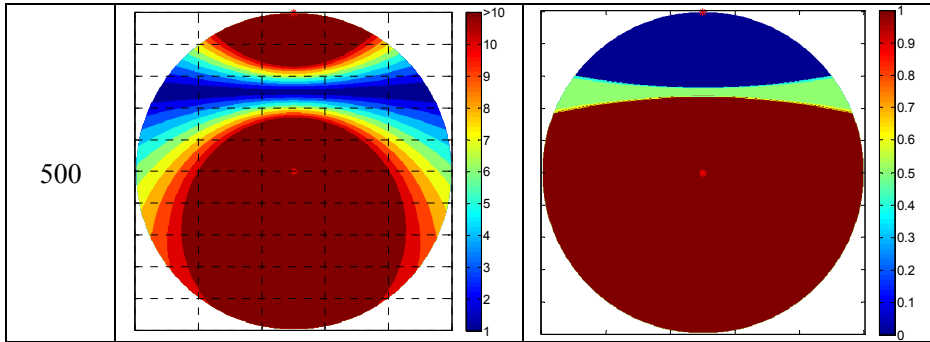
El escenario 1 es un escenario MNB macro-celular con dos transmisores. El primero de ellos se localiza en el punto central de la celda, garantizando una cobertura uniforme de toda la zona de interés. El segundo transmisor se puede colocar en cualquier punto de la celda. Como el escenario es simétrico, el comportamiento del sistema será el mismo para cualquier emplazamiento de la celda auxiliar en una circunferencia concéntrica. Identificando el valor del radio de dicha circunferencia,  $R$ , se identifica unívocamente el escenario concreto que se pretende evaluar.



**Figura 5.8 Estructura del escenario 1**

El comportamiento en diversidad del MNB, explicado en el Capítulo 3, depende en gran medida de la diferencia de potencia relativa entre las dos contribuciones recibidas simultáneamente desde dos antenas,  $\Delta P$ . En la columna izquierda de la Tabla 5.1 se muestra las regiones donde se experimenta un determinado valor de  $\Delta P$  cuando las antenas transmiten con igual potencia. Además, no se ha considerado el efecto del *shadowing*. La región marcada como 1dB incluye todos los puntos geográficos del área de simulación cuya diferencia relativa de potencia es menor de 1dB. Tal y como se muestra en el Anexo B, para la transmisión macro-celular existe un margen  $\Delta P$  máximo de 2.3 dB a partir del cual no resulta interesante transmitir con diversidad, puesto que el consumo necesario de potencia extra es mayor que el beneficio obtenido por la diversidad. Además, los valores óptimos de distribución de potencia rondan valores próximos a 0.5, no excediendo normalmente el valor de 0.6 ni de 0.4. Es por ello que se ha cuantificado los posibles valores de asignación de potencia a los siguientes:  $\mathfrak{K} = \{1, 0.6, 0.5, 0.4, 0\}$ . Cuando se selecciona  $\mathfrak{K} = 1$  quiere decir que toda la potencia se transmite por la antena principal. Cuando  $\mathfrak{K} = 0.6$  un 60% se transmite por la principal y el 40% restante por la secundaria y así sucesivamente. La columna de la derecha de la Tabla 5.1 muestra qué valores óptimos de  $\mathfrak{K}$  se asignaría en todos los puntos geográficos de una celda, variando el radio de la antena secundaria.



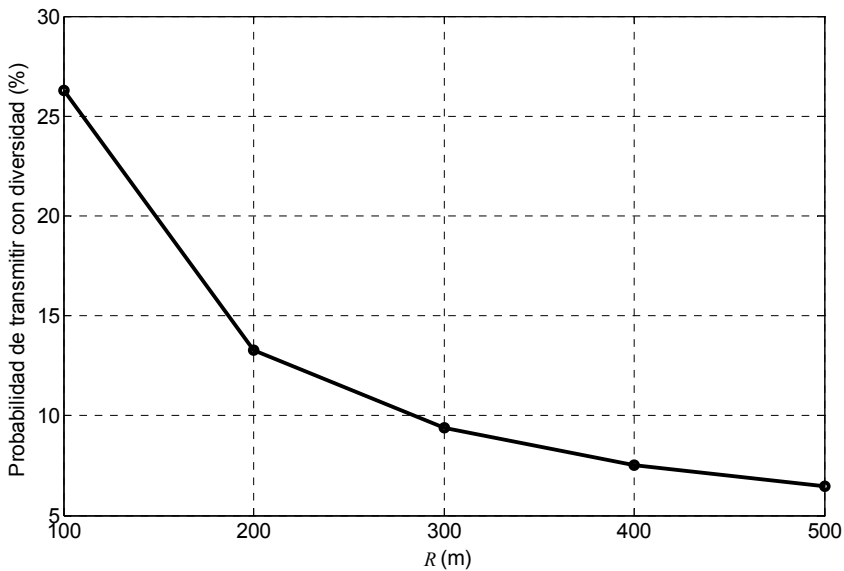


**Tabla 5.1 Relación de potencia recibida y distribución óptima de potencia para radios crecientes del escenario 1**

Como ya se ha explicado anteriormente, para el cálculo del valor de  $\alpha$  óptimo se busca cuál de los elementos del conjunto maximiza la ecuación (5.2). Analizando dicha ecuación se observa que el valor de  $D_{ijk}$  es específico para un determinado servicio y un valor de  $\Delta P$  que determinará el valor de la  $E_b/N_0$  objetivo. En concreto, para obtener los valores mostrados en la Tabla 5.1, se ha seleccionado el servicio de transmisión de paquetes mediante DSCH a una tasa binaria de 32 kbps.

Comparando la columna de la izquierda y de la derecha se puede observar como, efectivamente, para valores de  $\Delta P$  menores de 2 dB siempre se transmite con diversidad, mientras que si la diferencia de potencia está por encima de 3dB entonces la antena más cercana será la encargada de transmitir. Además, en el escenario 1 separar más las antenas implica irremediablemente ver reducida el área en el que se transmite con diversidad. Por ejemplo, si las antenas estuvieras físicamente emplazadas en el mismo punto se transmitiría con un 100% de probabilidad con diversidad, puesto que en cualquier punto la  $\Delta P$  sería de 0dB.

En la Figura 5.9 se muestra la evolución de la probabilidad de transmitir con diversidad respecto al radio de posicionamiento de la antena secundaria. Tal y como se intuía a partir de los mapas de la Tabla 5.1, separar más la antena secundaria reduce la probabilidad de transmitir con diversidad. Para decidir qué posición es óptima habrá que estudiar, además, cuál es la potencia transmitida por usuario en cada escenario.

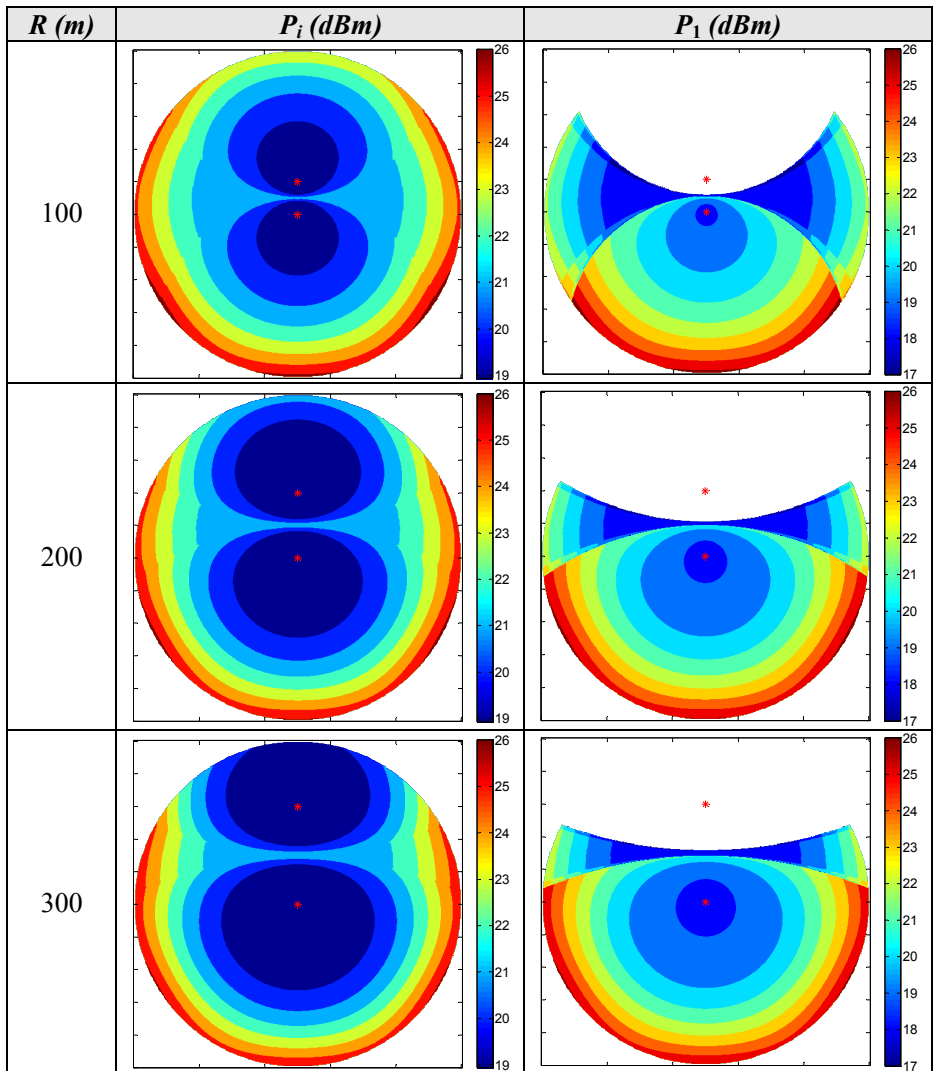


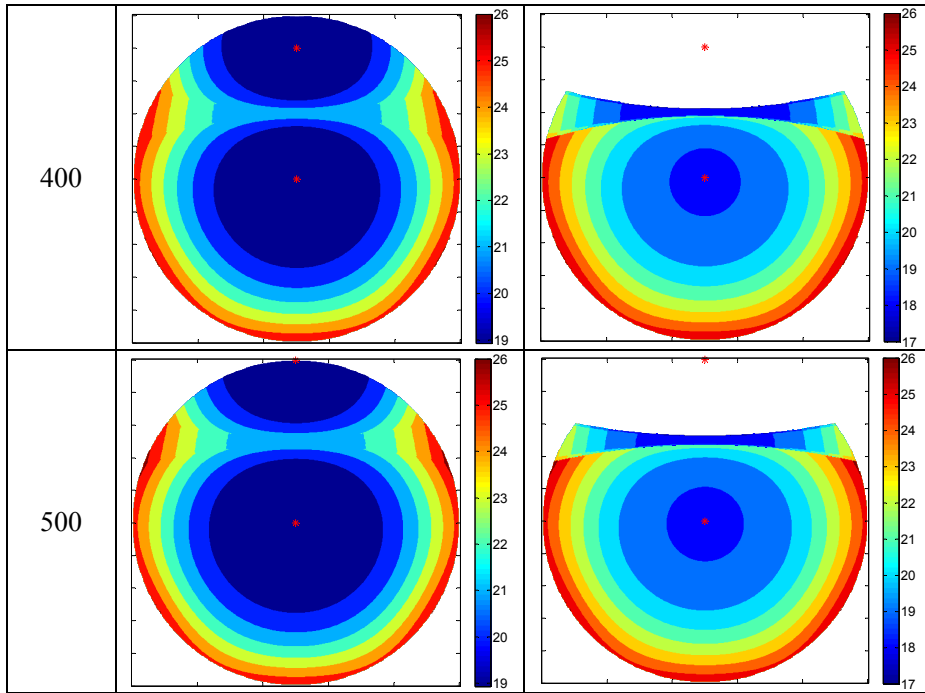
**Figura 5.9 Probabilidad de transmitir simultáneamente desde las dos antenas del MNB para radios crecientes del escenario 1**

Para evaluar la potencia consumida por un usuario de datos a 32 kbps se ha considerado que el MNB está transmitiendo a media carga, es decir, a 10 W, al igual que el resto de interferentes. Además, la distribución global de potencia entre las antenas es del 70% transmitido desde la antena principal y el 30% restante desde las secundarias, que en este caso incluye únicamente a la otra antena macro-celular. Esta asignación estática independiente del radio de la segunda antena no es realista pero permite, al menos, evaluar de manera global el comportamiento del sistema. En el apartado dedicado específicamente a analizar los resultados de simulación se comprobará si, efectivamente, el sistema se comporta de manera estática o balancea la carga de las antenas en función de la posición de las mismas.

En la columna izquierda de la Tabla 5.2 se muestra la potencia requerida por un usuario a 32 kbps localizado en cada punto geográfico de la celda. Por otro lado, la columna de la derecha recoge la potencia que transmite la antena principal a dicho usuario. La ventaja de tener las antenas cerca es que la zona con diversidad es mayor,

lo que hace que la potencia transmitida por la antena principal sea más baja, es decir, se reparte más equitativamente la potencia entre ambas antenas. En el límite, si las dos antenas estuvieran situadas en el mismo punto la potencia transmitida por la antena principal sería menos de la mitad de la transmitida si estuviera ella sola. Si no existiera diversidad en la transmisión sería exactamente la mitad, pero gracias a la ganancia por diversidad adicional introducida, esta potencia se verá reducida.

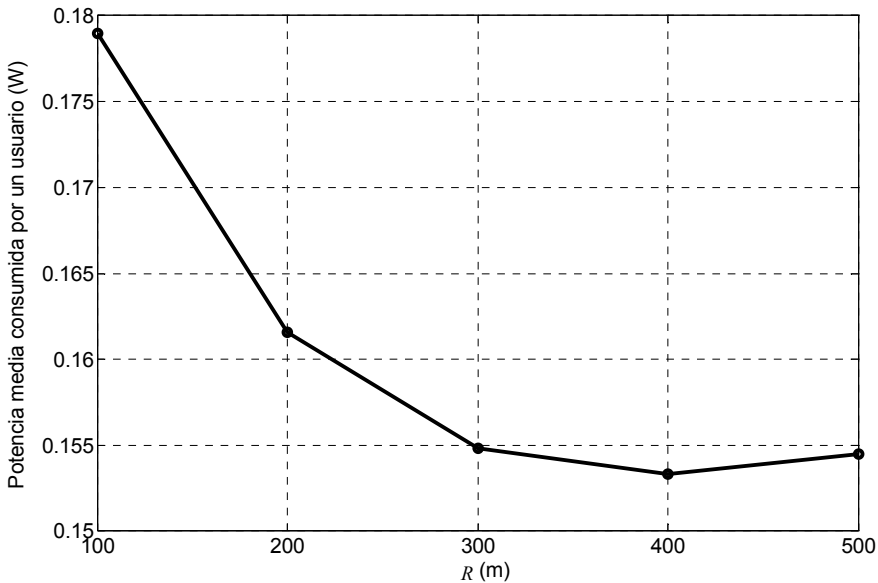




**Tabla 5.2 Potencia transmitida por el MNB y por la antena principal a un usuario de datos a 32 kbps para radios crecientes del escenario 1**

Por otro lado, como se puede observar en la potencia transmitida al usuario, al separar más la antena secundaria se consigue consumir menos recursos para los usuarios más alejados del centro, usuarios que precisamente son los que se encuentran en una situación más crítica por las mayores pérdidas de propagación experimentadas. Además, al separar parte de las interferencias intra-celulares, la zona de menor consumo de cada antena también se hace más grande. Se ha conseguido, separando las antenas, reducir el nivel de interferencia en el sistema. El problema es que si se aumenta mucho el radio considerado la zona de buena cobertura se saldrá de la celda de interés, lo cual no es deseable. Por lo tanto, existirá un valor óptimo del radio,  $R$ , para el que se minimizará la potencia consumida en el sistema.





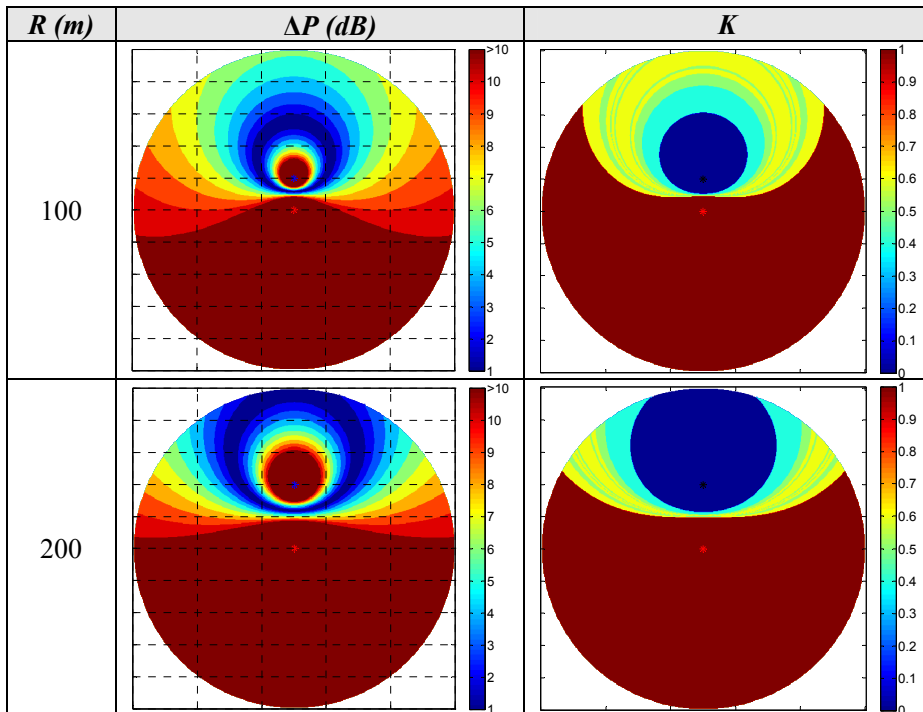
**Figura 5.10 Potencia media consumida por un usuario de datos a 32 kbps en todo el área celular para radios crecientes del escenario 1**

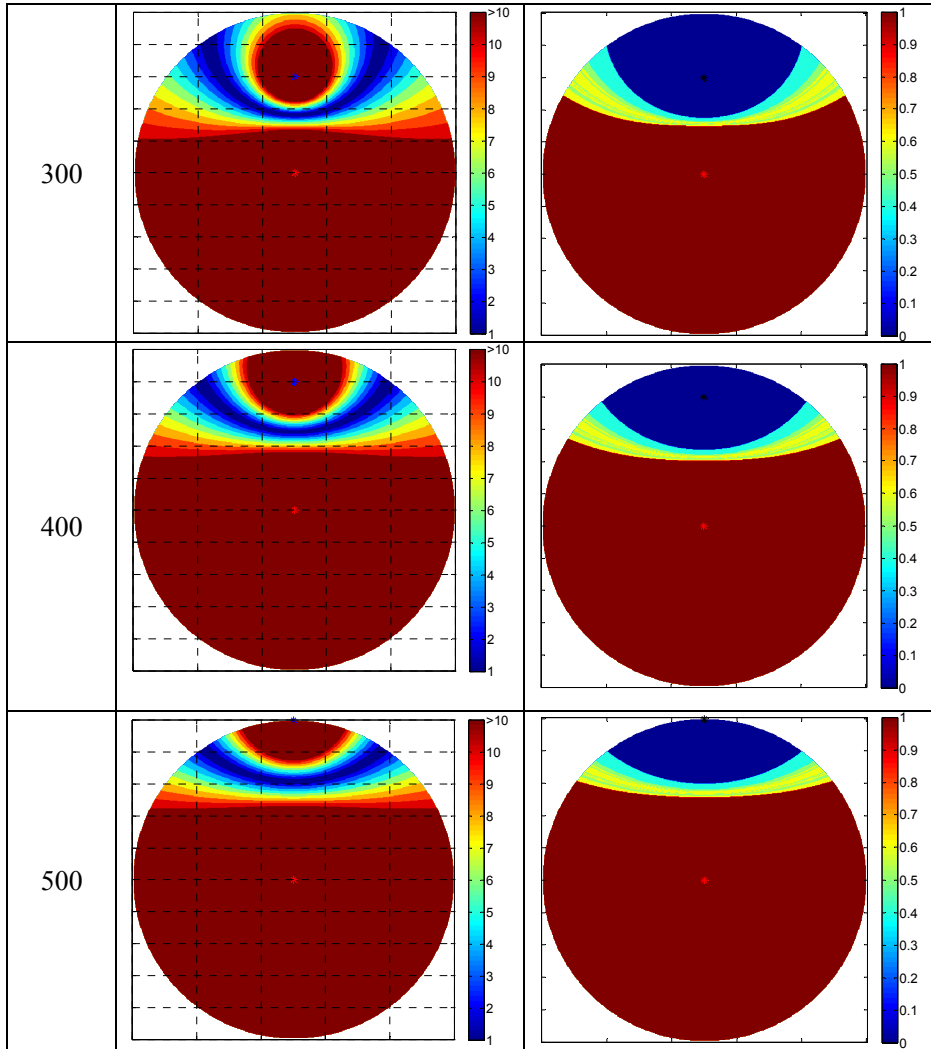
En la Figura 5.10 se recoge el valor medio de la potencia transmitida a un usuario de datos a 32 kbps situado en cualquier punto de la celda. Como se puede observar, separar las antenas reduce de manera muy significativa la potencia consumida, sobre todo porque se mejoran las condiciones de propagación de una parte de los usuarios más alejados y porque las interferencias intracelulares se distribuyen mejor en la celda. Sin embargo, a partir de 400m se pierde parte de este beneficio puesto que las antenas están demasiado separadas. Parece pues que este valor de 400m es el radio óptimo de emplazamiento de la antena secundaria macro-celular del escenario 1. Sin embargo, hay que tener en cuenta que se han hecho unas suposiciones muy estrictas. Por tanto, el estudio realizado no permite afirmar rotundamente que éste sea el radio óptimo, pero, al menos, sí que constata que existirá una mejor aproximación al usuario si se separan las antenas secundarias del centro de la celda bajo estudio. En este aspecto, la conclusión alcanzada coincide con lo expuesto en [11].

### 5.4.2.2 Escenario 2

El escenario 2 es el mismo que el anterior salvo que, en esta ocasión, la antena secundaria conforma una celda micro-celular. Las celdas micro-celulares se caracterizan por tener unas pérdidas de propagación más altas y también por carecer de diversidad, lo que conlleva tener un nivel de  $E_b/N_0$  objetivo más alto para un mismo servicio que una celda macro-celular. Sin embargo, una de las principales ventajas de las celdas micro-celulares es precisamente que el radio de cobertura es más reducido. Esto permite confinar interferencias a la vez que mejora considerablemente la correlación de los códigos ortogonales, lo que redonda en una menor interferencia intra-celular.

El estudio realizado de este escenario es parejo al acometido anteriormente con el escenario 1, lo que permite comparar ambas soluciones y ver las ventajas e inconvenientes de las celdas macro y micro-celulares. La Tabla 5.3 muestra las regiones con una determinada  $\Delta P$  y el valor de  $\aleph$  óptimo.





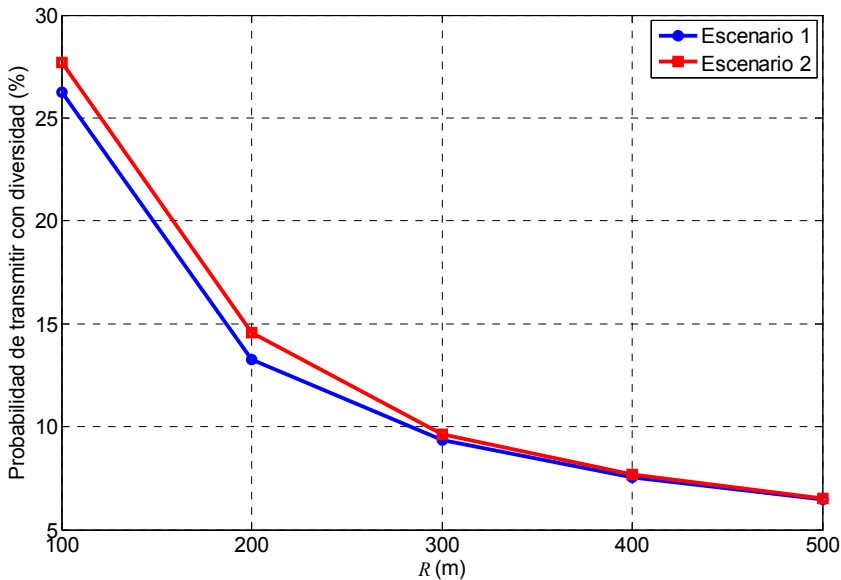
**Tabla 5.3 Relación de potencia recibida y distribución óptima de potencia para radios crecientes del escenario 2**

Lo primero que se observa en la Tabla 5.3 es que, debido a las mayores pérdidas de propagación, la diferencia de potencia entre las contribuciones recibidas es, en valor medio, más alta. A pesar de que este hecho podría llevar a pensar que la probabilidad de transmitir con diversidad se reduciría, esto no se produce, sino que incluso con 7 dB

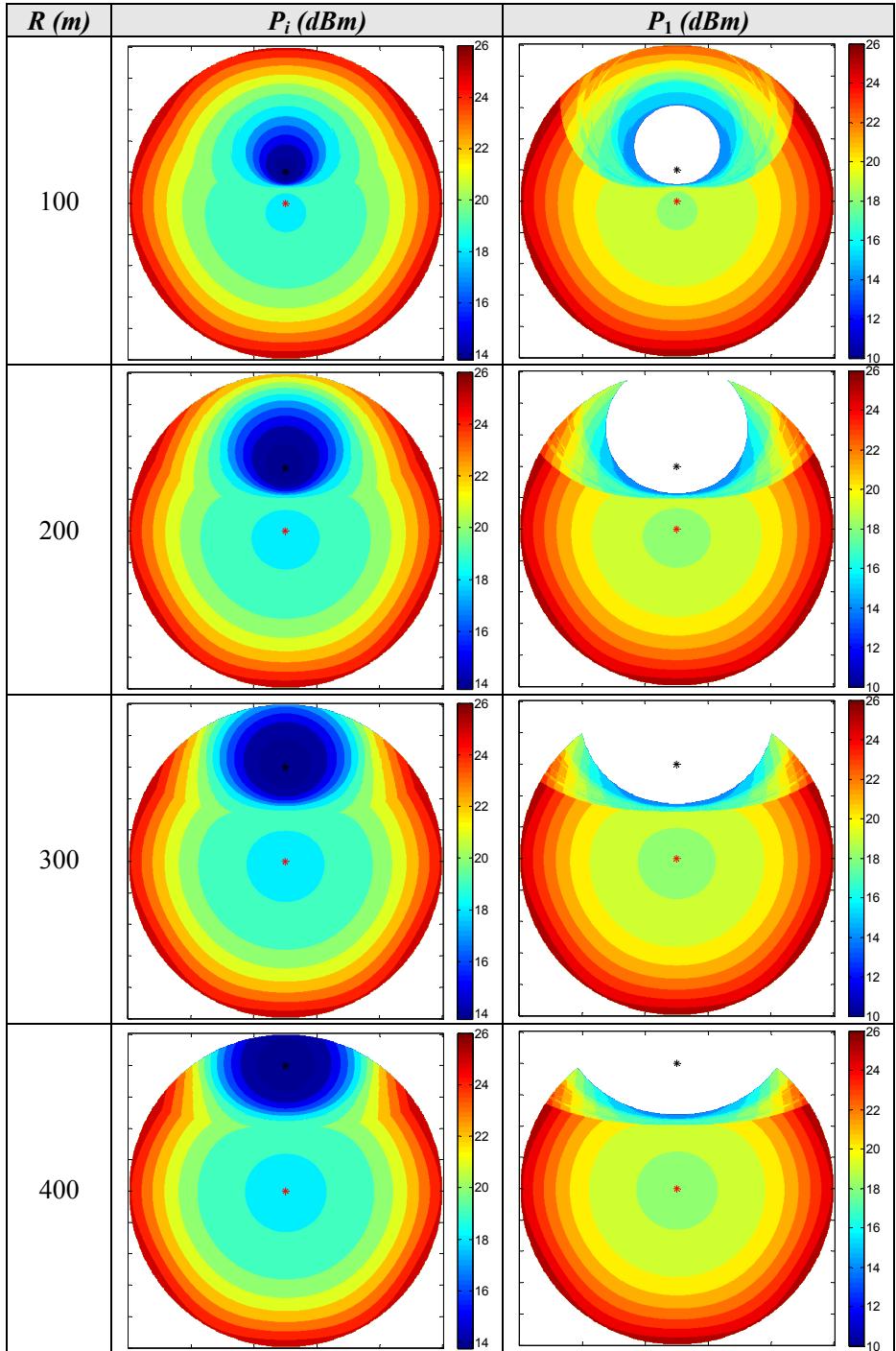
de diferencia de potencia se sigue transmitiendo con diversidad. Este hecho corrobora lo mostrado en el Anexo B, donde se concluye que el margen máximo de  $\Delta P$  para un escenario mixto se sitúa cerca de los 7 dB.

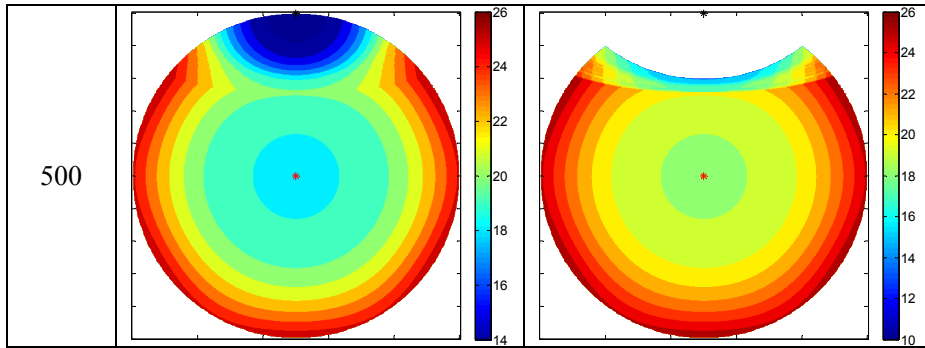
Al igual que ocurría en el escenario 1, al aumentar  $R$  se reduce el área donde resulta óptimo transmitir simultáneamente desde ambas antenas. La Figura 5.11 compara la probabilidad de transmitir con diversidad en el escenario 1 y el escenario 2. Debido a que la micro-celda se ve muy beneficiada por la diversidad adicional que consigue gracias al apoyo de la celda principal, la mayor diferencia de potencia se compensa, llegando incluso a experimentarse una mayor diversidad en el escenario 2. Cuando el valor de  $R$  aumenta, ambos escenarios reducen la probabilidad de transmitir con diversidad, convergiendo hasta un mismo punto.

Por último, la Tabla 5.4 representa la evolución de la potencia asignada en dBm a un usuario de 32 kbps y la potencia transmitida a ese usuario por la antena principal.



**Figura 5.11 Probabilidad de transmitir simultáneamente desde las dos antenas del MNB para radios crecientes. Comparación de los escenarios 1 y 2**



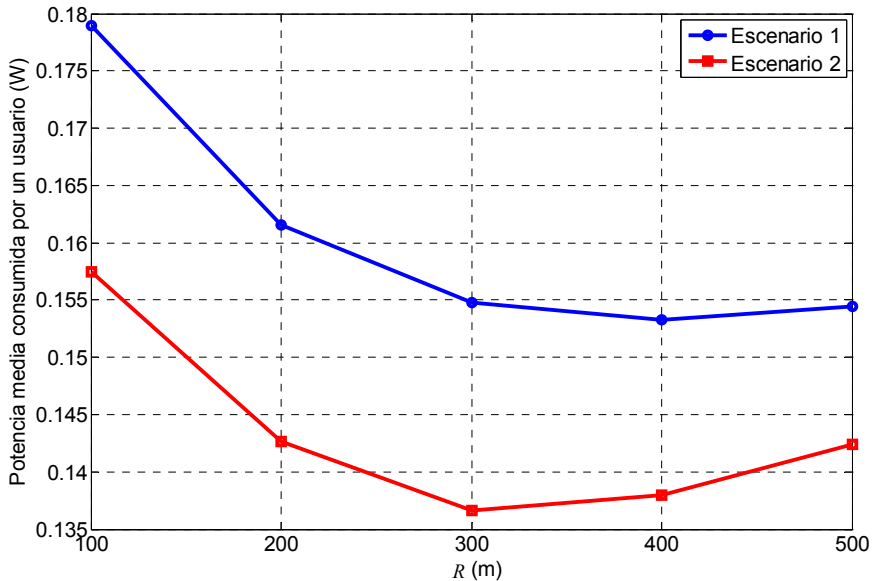


**Tabla 5.4 Potencia transmitida por el MNB y por la antena principal a un usuario de datos a 32kbps para radios crecientes del escenario 2**

Lo primero que destaca de la Tabla 5.4 es que el nivel de potencia requerido dentro del área de influencia de la micro-celda es mucho más bajo que para la macro-celda. Esto corrobora que el aumento de la correlación entre los códigos es mucho más importante que las mayores pérdidas de propagación experimentadas. Además, el área cubierto por la micro-celda es mucho menor que para la macro-celda del escenario 1 y la antena principal debe transmitir potencia a usuarios ubicados en muchos más puntos de la celda. De esta manera, la falta de diversidad de la micro-celda se compensa con la diversidad adicional del MNB, lo que reduce la potencia consumida.

Como ocurría antes, al separar la micro-celda del transmisor principal su área de influencia aumenta y la zona de poca potencia se hace mayor, sobre todo debido a la mayor reducción en las interferencias intra-celulares. Además, los usuarios más alejados se aprovecharán de las mejores condiciones de transmisión de la micro-celda, lo que mejorará el comportamiento medio del sistema.

La carga entre las antenas se reparte más equitativamente en el escenario 1 que en el escenario 2. La antena principal en el escenario 2 transmite prácticamente en toda la celda, mientras que la antena secundaria sirve para mejorar altamente la cobertura en su región más cercana. Aunque más desequilibrado el reparto de potencias, la potencia media del sistema se reduce en este escenario, tal y como se confirma en la Figura 5.12. El radio óptimo en el caso del escenario 2 es más pequeño, cercano a los 300m.

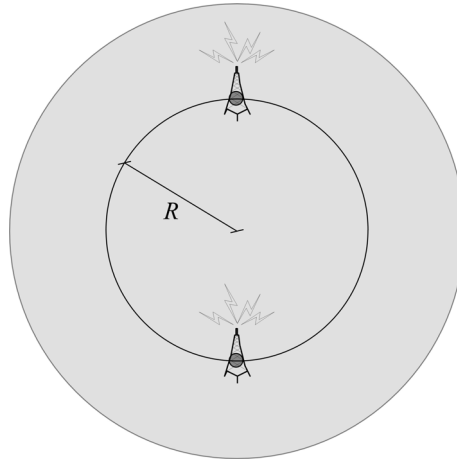


**Figura 5.12** Potencia media consumida por un usuario de datos a 32 kbps en todo el área celular para radios crecientes. Comparación de los escenarios 1 y 2

### 5.4.2.3 Escenario 3

Con los resultados mostrados hasta ahora, parece obvio pensar que separar las antenas resulta altamente beneficioso para el comportamiento del sistema por tres razones: en primer lugar, se aumenta la probabilidad de estar cerca de una antena, en segundo lugar, se reduce el nivel de interferencias intra-celulares experimentadas por los usuarios y, finalmente, se da mayor cobertura a los usuarios más distantes de la celda, los más críticos.

En el escenario 3 se ha considerado dos antenas macro-celulares pero, a diferencia del escenario 1, ninguna de ellas ocupa el centro de la celda sino que se localizan lo más distantes posibles en una circunferencia de radio  $R$ . La Figura 5.13 muestra gráficamente la configuración empleada en este escenario.



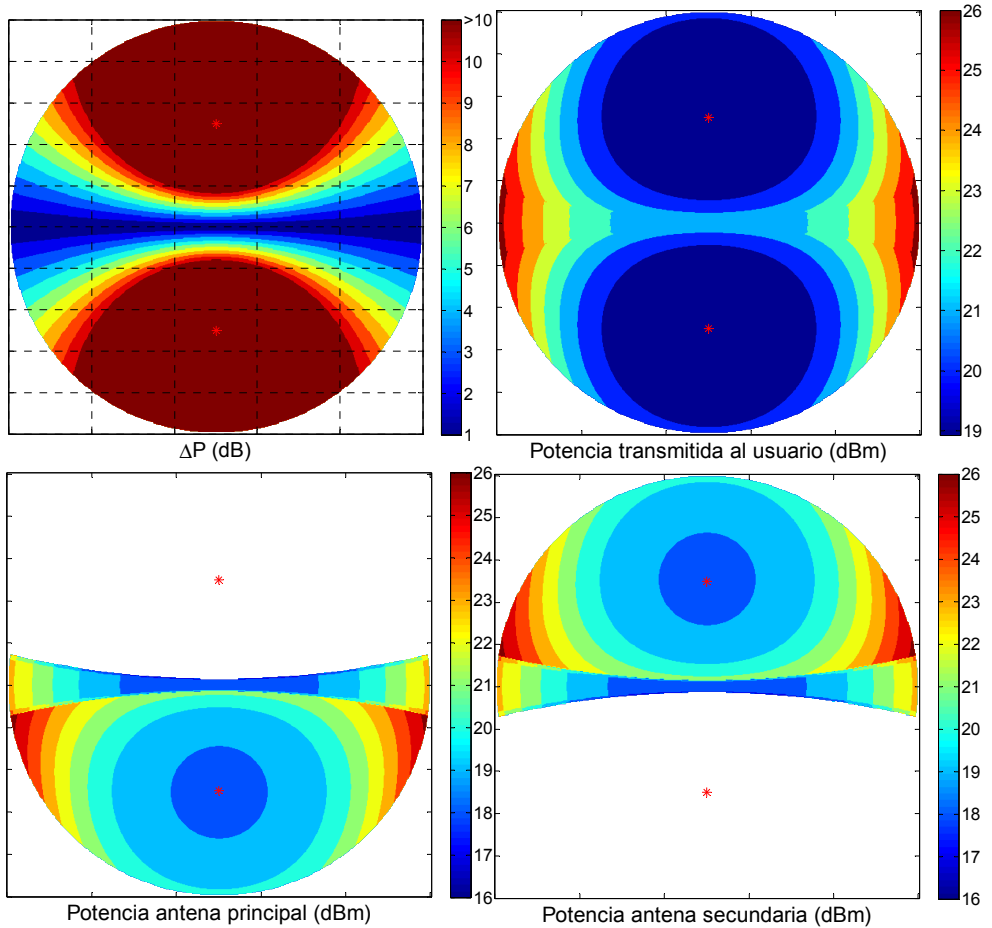
**Figura 5.13 Estructura del escenario 3**

Una de las principales peculiaridades de este escenario es la enorme simetría que existe entre las dos antenas que componen el MNB. Debido a esta simetría, en la evaluación del escenario 3 se ha considerado un reparto equitativo de recursos entre las antenas. La Figura 5.14 recoge un caso concreto del escenario 3 con radio  $R=250$  m. Como se puede apreciar en la figura, la potencia transmitida por ambos transmisores es exactamente la misma, lo que corrobora el perfecto balanceo de carga y la simetría antes mencionada.

Observando la gráfica de  $\Delta P$ , es importante destacar que la región con diferencia de potencia menor de 3 dB se sitúa precisamente en el ecuador de la celda, precisamente la región que más dista de las dos antenas transmisoras.

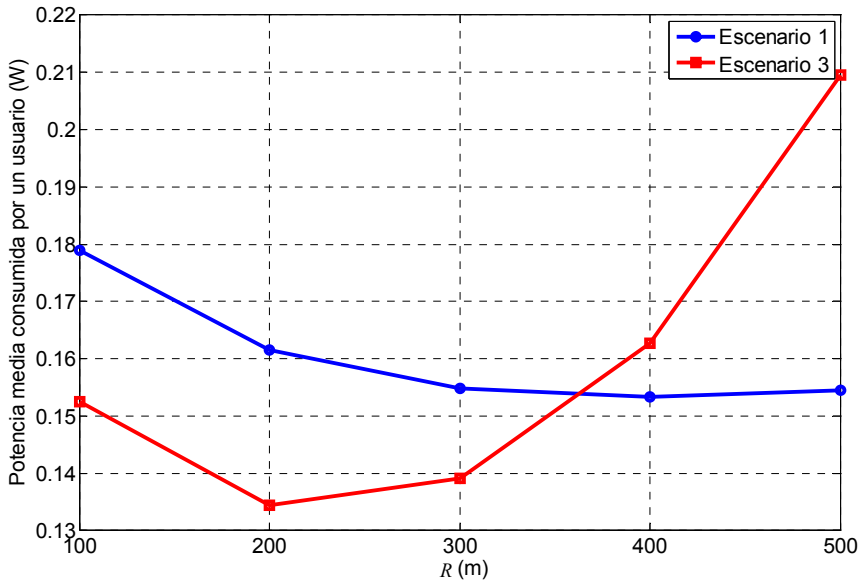
Otro aspecto importante es que la región de peor cobertura ya no es toda la periferia circular, sino que se concentra en los dos polos del escenario, zona con mayor diversidad en transmisión. Todo ello hará que el comportamiento de este escenario sea mejor que el mostrado por el escenario 1. Para demostrarlo, se ha analizado comparativamente la potencia media consumida por un usuario de datos en toda la celda. La Figura 5.15 recoge los resultados obtenidos. Como se muestra en la figura, en el punto óptimo de separación entre los transmisores existe una importante reducción de la potencia requerida por el usuario.





**Figura 5.14** Comportamiento del escenario 3. Arriba izquierda, diferencia relativa de potencia recibida. Arriba derecha, potencia asignada a un usuario de datos a 32 kbps. Abajo izquierda, potencia transmitida por la antena principal a dicho usuario. Abajo derecha, potencia transmitida por la antena secundaria.

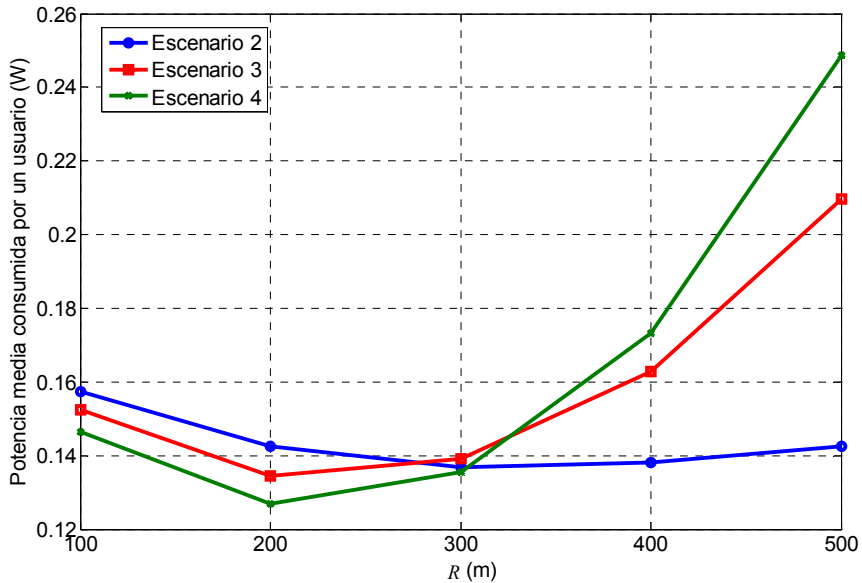
Sin embargo, una separación excesiva de las antenas provoca que la zona central sea precisamente la que se quede sin cobertura. Por eso, la potencia media por usuario aumenta de manera más significativa que en el caso del escenario 1 cuando las antenas se separan más de 350m.



**Figura 5.15** Potencia media consumida por un usuario de datos a 32 kbps en todo el área celular para radios crecientes. Comparación de los escenarios 1 y 3

#### 5.4.2.4 Escenario 4

El escenario 4 es idéntico al 3 salvo por el hecho de que uno de los transmisores en vez de ser macro-celular es micro-celular. Por una parte, este escenario tiene como ventaja respecto al escenario 2 que la separación entre los transmisores es mayor, lo que favorece la separación entre interferencias y mejora el área de influencia del transmisor micro-celular. El problema es que el funcionamiento de la micro-celda depende mucho de la diversidad adicional introducida por la antena macro-celular, por lo que si la separación es grande el funcionamiento tenderá a empeorar rápidamente. Además, si la antena macro-celular se escora demasiado habrá una gran parte de la celda que carezca de su necesaria cobertura. El efecto será el mismo que el observado en el escenario 3 pero con el agravio adicional de que la antena micro-celular tiene menos capacidad de cobertura que la macro-celular.



**Figura 5.16** Potencia media consumida por un usuario de datos a 32 kbps en todo el área celular para radios crecientes. Comparación de los escenarios 2, 3 y 4

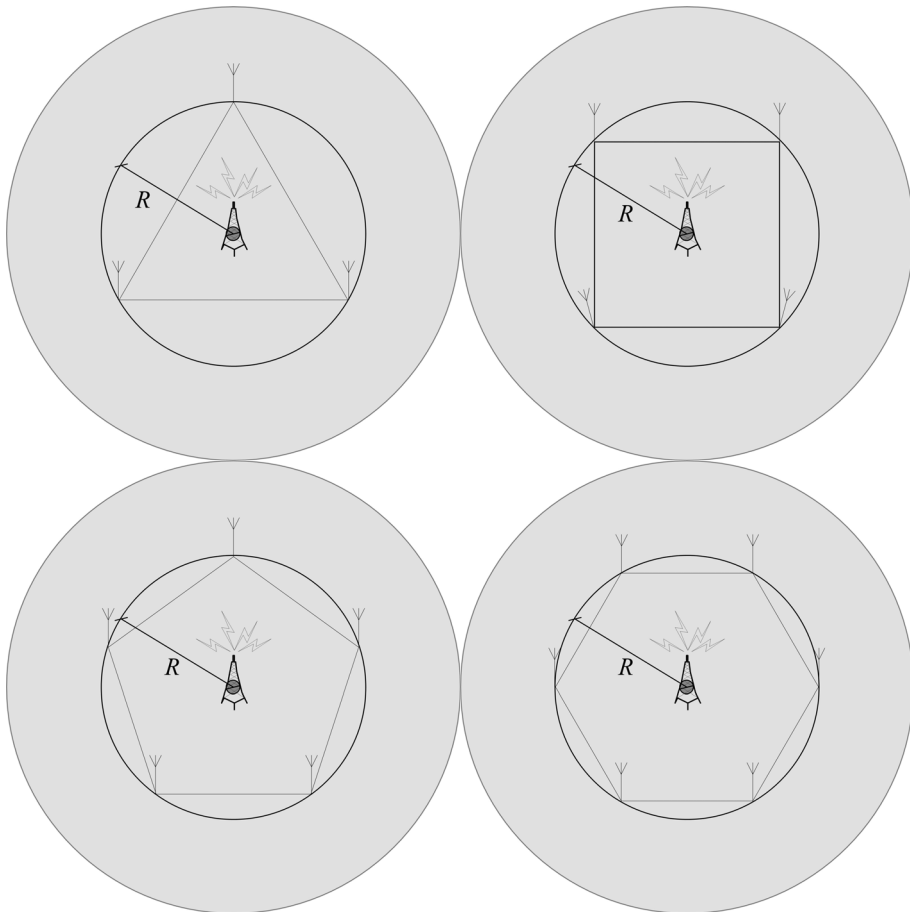
Todo este razonamiento intuitivo queda patente al analizar la evolución de la potencia asignada a un usuario conforme aumenta el valor de  $R$ . La Figura 5.16 muestra este análisis. Como se puede ver, si la distancia es pequeña el funcionamiento de este escenario es el óptimo, mejorando todos los escenarios anteriores. Sin embargo, si se separa demasiado la macro-celda de su posición central el empeoramiento es el más abrupto. La micro-celda se queda sin diversidad y existen regiones con mucha peor cobertura.

#### 5.4.2.5 Escenario 5-8

Los resultados obtenidos hasta ahora demuestran que, si se consigue distribuir convenientemente las potencias entre las antenas, las mejores prestaciones se obtienen haciendo uso de emplazamientos micro-celulares. Además, resulta interesante

distanciar las micro-celdas de manera que se de mejor servicio a los usuarios más alejados del centro de la celda. Por último, la cobertura macro-celular ha de estar presente en todo el Multinodo B ya que si no es así se perderá la diversidad adicional que las micro-celdas necesitan para proporcionar el mejor servicio.

Por todo esto, los escenarios restantes considerados son combinaciones de una macro-celda situada en el centro del MNB y de un número creciente de celdas micro-celulares separadas una distancia  $R$  del centro. Las estructuras de los escenarios 5, 6, 7 y 8 se recogen en la Figura 5.17.



**Figura 5.17 Estructura del escenario 5 (arriba izquierda), 6 (arriba derecha), 7 (abajo izquierda) y 8 (abajo derecha)**

Como se puede observar, todos estos escenarios tienen una estructura simétrica respecto a la celda, de manera que se pueda dotar de cobertura micro-celular a todo el MNB, separando lo máximo posible las antenas entre sí para evitar solapamientos.

En la Tabla 5.5 se muestra un análisis de estos escenarios mixtos. Lo primero que se puede destacar es que, gracias a la utilización de un mayor número de celdas micro-celulares, la antena central deja de transmitir a una importante porción de la celda Multinodo. Cabe recordar que en el cálculo de la potencia se ha considerado en todo momento que el 70% de la potencia es transmitida por la celda principal y el resto se distribuye entre las secundarias. Quizá este resultado invalida esta suposición, pero tampoco son importantes los valores absolutos obtenidos, que se reservan para el apartado 5.4.3.2, sino más bien comparar el comportamiento de los distintos escenarios.

Respecto a la diversidad en transmisión, al aumentar el número de antenas esta diversidad también disminuye, como era de esperar. Existen muchas más regiones con una antena cercana, regiones en las que no tiene sentido transmitir con diversidad puesto que la propia antena por sí misma puede satisfacer la calidad de servicio del usuario. Respecto al radio  $R$ , conforme éste aumente se reducirá la zona de posible transmisión por diversidad ya que las micro-celdas coparán la transmisión en toda la periferia.

Sin embargo, y a pesar de transmitir con menor diversidad, conforme se aumenta el número de celdas micro-celulares existe una clara tendencia a reducir la potencia consumida por los usuarios. Este hecho se debe básicamente a la mayor probabilidad de tener una antena cerca. Además, como cada una de estas antenas presenta una menor potencia total, también se reduce la interferencia intra-celular.

Para comprobar que efectivamente se produce una reducción en la potencia consumida al aumentar el número de antenas, se ha evaluado la potencia media consumida por un usuario para distintos valores de  $R$  en cada uno de los escenarios. La Figura 5.18 muestra los resultados de esta evaluación.

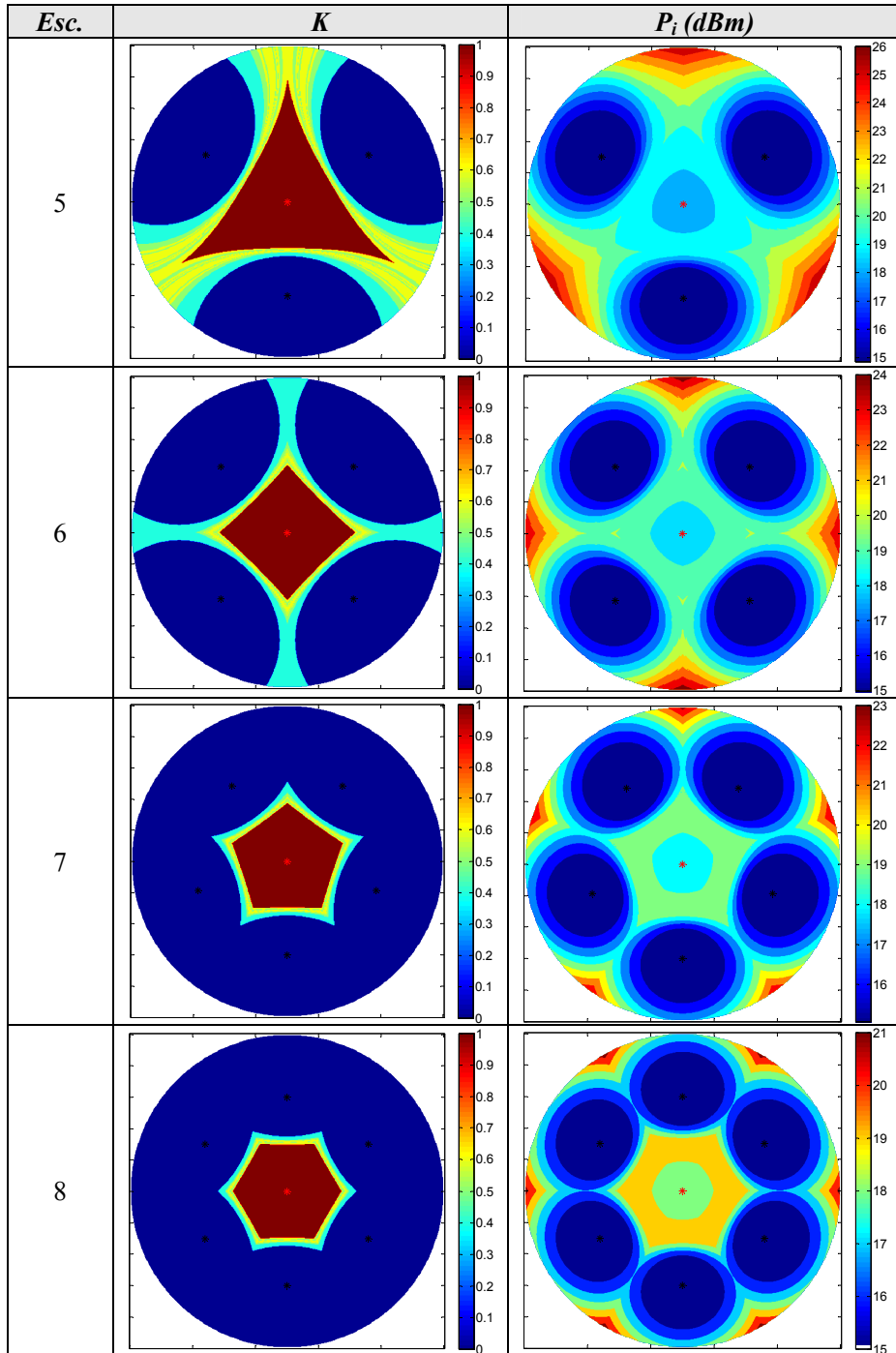
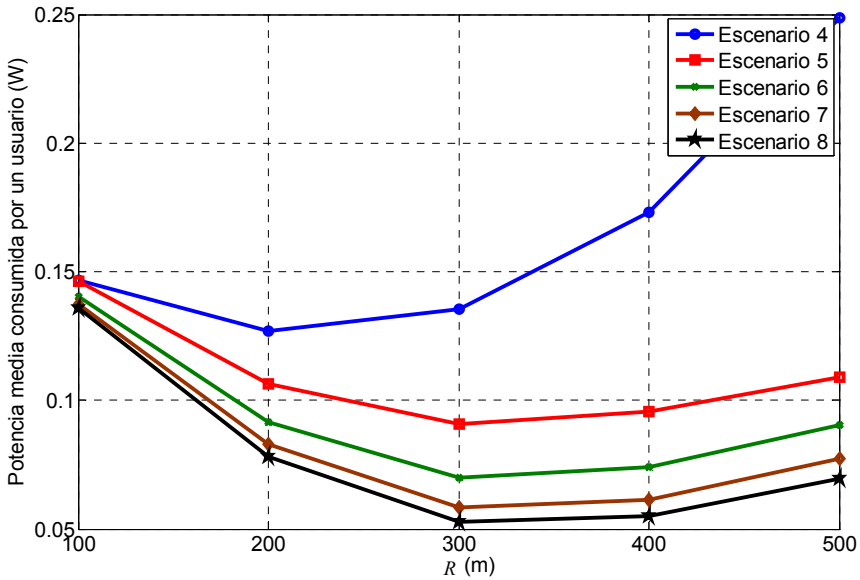


Tabla 5.5 Distribución óptima de potencia y potencia transmitida por el MNB a un usuario de datos a 32kbps para los escenarios 5, 6, 7 y 8 con  $R=300m$



**Figura 5.18 Potencia media consumida por un usuario de datos a 32kbps en todo el área celular para radios crecientes. Comparación del escenario 4, 5, 6, 7 y 8**

Efectivamente, el escenario que menor potencia media consume es el 8. Además, cabe destacar que los cuatro escenarios presentan su mínimo de potencia para un radio  $R=300$  m. A esta distancia las antenas micro-celulares ya se han separado lo suficiente como para no causarse mutua interferencia y, además, así dan suficiente buena cobertura a los usuarios más alejados de la celda principal.

### 5.4.3 Discusión de los resultados de simulación

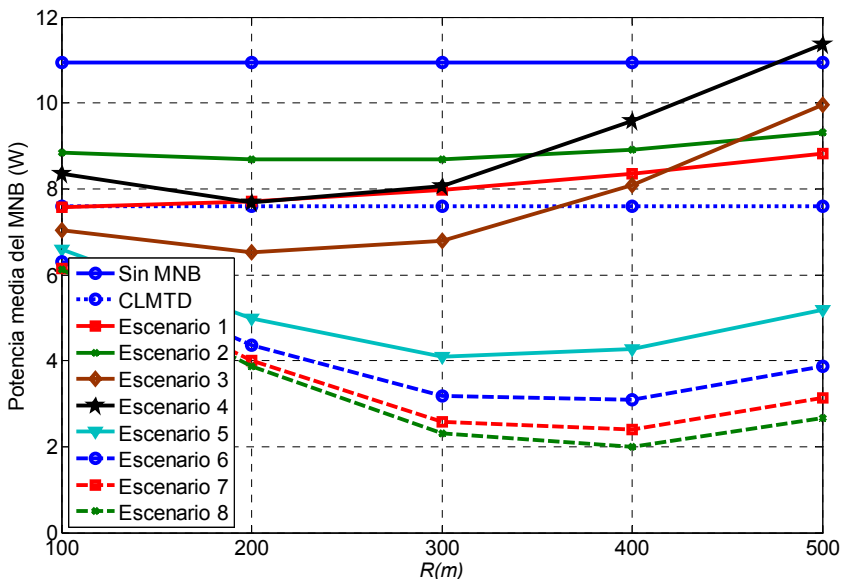
En este apartado se evaluará, en primer lugar, la calidad de servicio experimentada por los usuarios web en todos los escenarios de simulación descritos. En esta primera evaluación se comparará el algoritmo más potente desarrollado en la Tesis, el ACORP, con otros más convencionales. Posteriormente, se seleccionará un escenario concreto,

el escenario 5, para analizar las distintas técnicas de distribución de recursos contempladas y el efecto de considerar una arquitectura u otra del MNB.

### 5.4.3.1 Comparación de los distintos escenarios de simulación

El primer aspecto que se debe analizar de los distintos escenarios considerados es, sin duda, la potencia media consumida por el MNB. Además de ejecutar el algoritmo neuronal ACORP en los 8 escenarios, se ha comparado estos resultados con el escenario sin MNB y el CLMTD. En este último se sigue ejecutando el ACORP pero con las dos antenas macro-celulares localizadas en el mismo emplazamiento. La Figura 5.19 muestra los resultados obtenidos.

En primer lugar, cabe destacar que las técnicas de MNB propuestas en esta Tesis Doctoral reducen considerablemente el consumo global de potencia del sistema. Esta reducción tiene un gran impacto en las prestaciones del sistema, sobre todo en UMTS.



**Figura 5.19** Potencia media consumida en el MNB en función del valor de  $R$  y comparando los distintos escenarios considerados



La Tabla 5.6 recoge la ganancia obtenida por el MNB respecto a la arquitectura convencional en términos de potencia total consumida para los distintos escenarios. Dicha ganancia se ha calculado considerando el radio óptimo de cada escenario. La ganancia llega a ser de hasta el 81.73 % para el escenario 8, lo que supone dividir la potencia consumida por un factor mayor de 5.

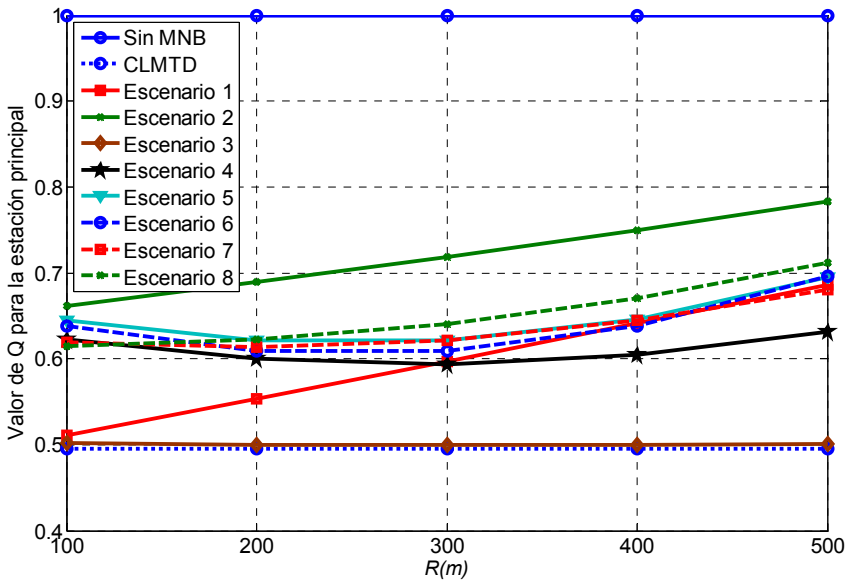
Escenario	1	2	3	4	5	6	7	8
Mejora (%)	31.14	19.26	40.45	28.58	62.19	71.84	78.26	81.73

**Tabla 5.6 Mejora conseguida por el MNB en los distintos escenarios**

Utilizar únicamente dos antenas en transmisión reduce la potencia consumida, pero no supone una ganancia tan importante como la conseguida con los escenarios mixtos con más antenas desplegadas. Por otro lado, considerando la transmisión desde únicamente dos antenas, parece que transmitir desde un único emplazamiento (CLMTD) no es una mala opción en absoluto. El escenario 1 tiende a presentar su mejor comportamiento para un valor de  $R=0$ , convergiendo hacia el valor mostrado por CLMTD. Únicamente el escenario 3 obtiene unas mejores prestaciones que transmitir desde un único emplazamiento con diversidad, pero esta ganancia es de apenas un 14% respecto a éste.

Analizando el comportamiento del escenario 1, se observa que los resultados obtenidos no se corresponden con los mostrados en el apartado de descripción de los escenarios. Para entender mejor este hecho, la Figura 5.20 muestra el valor de  $Q_1$ , es decir, la proporción de potencia transmitida por la antena principal respecto al total.

Según se observa en la Figura 5.19, el escenario 1 presenta el mínimo de consumo de potencia para el radio más pequeño analizado, con una clara tendencia a presentar el mínimo precisamente para  $R=0$ , que se corresponde con el CLMTD. Sin embargo, en el apartado anterior se obtuvo que el óptimo se alcanzaba para un radio de 400 m. Esto era así suponiendo que  $Q_1=0.7$  para cualquier radio, lo que no se corresponde con lo experimentado en el sistema, como se ve en la Figura 5.20.



**Figura 5.20 Proporción de potencia transmitida por la antena principal del MNB en función del valor de  $R$  y comparando los distintos escenarios considerados**

Cuando  $R=0$  la potencia se reparte por igual entre ambas antenas y  $Q_1=0.5$ . Conforme se van separando las antenas, la antena principal va asumiendo mayor carga ya que da cobertura a una región mayor y, además, contribuye aportando diversidad a la segunda antena. Una tendencia creciente con el radio se observa también para el escenario 2 pero, en este caso, la antena principal asume una mayor carga puesto que la cobertura de la antena micro-celular es menor y, además, requiere de más apoyo en diversidad.

Considerando un valor igual de  $Q_1$ , en el apartado anterior se llegó a la conclusión de que en el escenario 2 se consumía menos potencia que el escenario 1. Sin embargo, como un mayor balance de potencia entre las antenas supone una importante reducción de potencia consumida y el escenario 1 consigue, precisamente, este mejor reparto de carga, en el análisis global del sistema se observa que es el escenario 1 el que realiza un consumo medio de potencia menor.

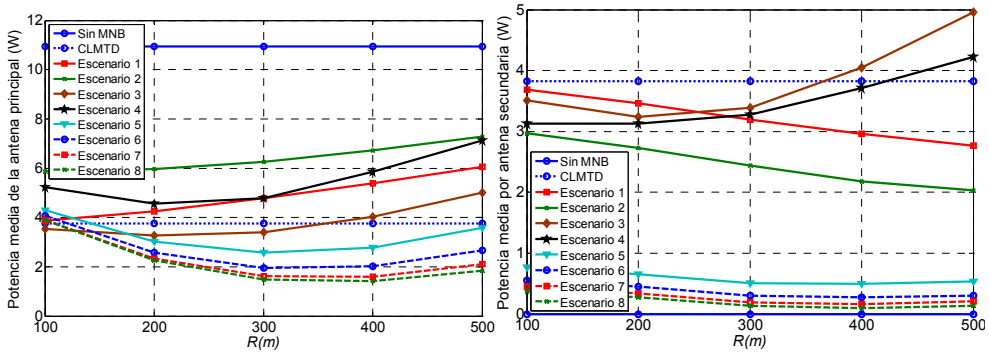
El análisis del escenario 3 corrobora la simetría que le caracteriza. El valor de  $Q_1$  es muy próximo a 0.5, como era de esperar. Además, y tal y como se comentó anteriormente, considerando únicamente dos transmisores es preferible equiespaciadas respecto al centro de la celda pero hasta un determinado radio a partir del cual se empieza a perder cobertura. Es por esto que el escenario 3 consume menos potencia que el 1, obteniéndose el mínimo en  $R=200$  m, mientras que para valores de  $R$  grandes, el comportamiento del escenario 3 es peor que el mostrado por el escenario 1.

Este análisis es igualmente válido entre el escenario 2 y su equivalente, el escenario 4. De nuevo, con dos antenas es mejor transmitir con emplazamientos macro-celulares.

Respecto a los escenarios mixtos con más de dos antenas, aumentar el número de transmisores supone reducir la potencia media consumida por el MNB. Esto es lógico puesto que al existir más puntos de transmisión es más probable que un usuario tenga una antena cercana desde la que se le pueda transmitir con baja potencia. Además, transmitir desde emplazamientos micro-celulares con diversidad adicional aportada por la celda central reduce las interferencias intra-celulares del sistema, por el importante aumento del factor de ortogonalidad.

Conforme se aumenta el número de antenas micro-celulares el radio óptimo de localización de las antenas tiende a ser mayor. Esto es debido a que si las antenas están demasiado juntas sufren una mayor interferencia entre ellas, efecto que se reduce al aumentar  $R$ . Como siempre, existe un límite a partir del cual aumentar más esta distancia tiene la contrapartida de que se pierde cobertura dentro del MNB.

El análisis del comportamiento en términos de potencia de los distintos escenarios se completa con el estudio de la potencia media consumida por la antena principal y por cada una de las secundarias. En la Figura 5.21 se muestran los resultados obtenidos. Respecto a la potencia consumida por la antena principal, es importante destacar que en cualquier caso el MNB reduce considerablemente la potencia consumida, lo cual resulta especialmente interesante sobre todo teniendo en cuenta la creciente alarma social generada en torno a los emplazamientos de comunicaciones móviles. Para el mejor de los casos (escenario 8) la potencia consumida por la macro-celda será de sólo 1.3 W para un escenario con 70 usuarios web activos.



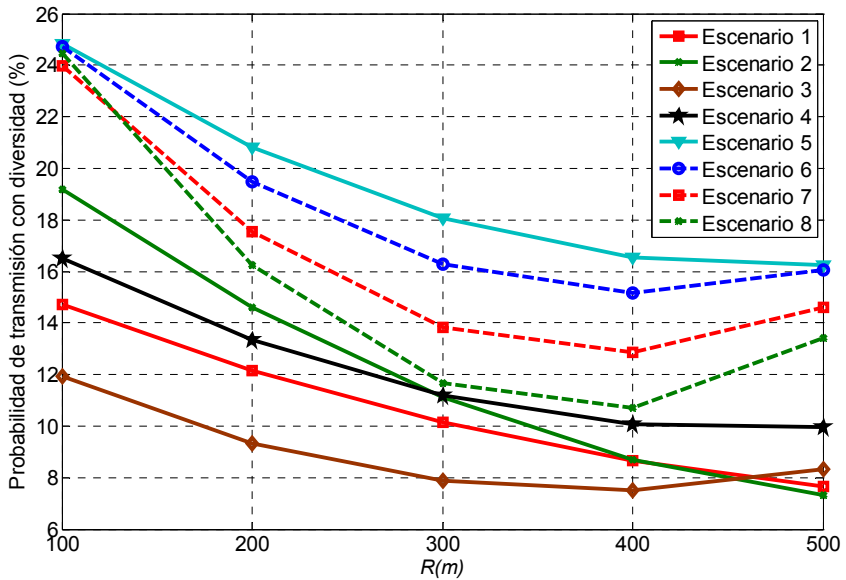
**Figura 5.21** Potencia media consumida por la antena principal (izquierda) y por cada uno de los emplazamientos secundarios (derecha) en función del valor de  $R$

En los escenarios 1 y 2, aumentar el valor de  $R$  supone siempre incrementar la potencia consumida por el emplazamiento central, ya que éste asume una mayor zona de cobertura, mientras que la potencia consumida por el emplazamiento secundario tenderá a bajar, pues dará servicio a un menor número de usuarios.

Para el resto de escenarios, existe siempre un valor de  $R$  óptimo para tanto la antena principal como para la secundaria, que coincide con el punto de mínimo consumo de potencia medio en el MNB. Es significativo recalcar de nuevo que, en el caso del escenario 3, el consumo de potencia es el mismo para ambas antenas.

Respecto a la potencia consumida por cada antena secundaria, resulta especialmente significativo estudiar cuál es la potencia consumida por las celdas micro-celulares ya que, al estar situadas a menor altura, deberán transmitir con la menor potencia posible. En el caso del escenario 3, el nivel de potencia consumida es siempre superior a 3 W lo que no es recomendable para un escenario de estas características. Sin embargo, al aumentar el número de emplazamientos micro-celulares, el consumo medio por antena es muy reducido y siempre inferior a 1 W, que es el valor recomendado por las autoridades.

A continuación, se ha procedido a analizar la probabilidad de transmisión con diversidad dentro del MNB para los distintos escenarios. La Figura 5.22 muestra los resultados obtenidos.



**Figura 5.22 Probabilidad de transmitir con diversidad en el MNB en función del valor de  $R$  y comparando los distintos escenarios considerados**

Los valores de probabilidad de transmisión con diversidad del CLMTD y del escenario sin MNB no se han representado gráficamente puesto que su valor es trivial: 100% y 0% respectivamente.

Como era de esperar, existe una menor probabilidad de diversidad para los escenarios macro-celulares, puesto que en ese caso la transmisión simultánea sólo es eficaz para valores de diferencia de potencia de 2.3 dB. Tal y como se demostró en la descripción de los escenarios de simulación, aumentar el valor de  $R$  hace que se reduzca la probabilidad de transmitir con diversidad. Al comparar el escenario 1 con el 3 (o alternativamente el 2 con el 4), se aprecia como para valores de  $R$  bajos la diversidad en el escenario 3 es menor puesto que los transmisores están más separados. Sin embargo, para valores altos la tendencia se invierte debido a que en el escenario 1 se pierde la simetría.

Respecto al resto de escenarios mixtos, el escenario 5 es el que presenta una mayor probabilidad de transmisión con diversidad. Como las micro-celdas están más distantes entre sí, existen más regiones en las que se requiere la diversidad adicional aportada por la macro-celda para asegurar una transmisión de calidad. Conforme se aumenta el número de emplazamientos existen más zonas solapadas entre las micro-celdas y, por tanto, las regiones de interés con diferencia de potencia pequeña se van reduciendo, reduciéndose también la diversidad en transmisión.

A continuación, se procede a evaluar el comportamiento del MNB en términos de la tasa binaria o *throughput* del sistema. La Figura 5.23 muestra el *throughput* efectivo, es decir, descontando la información no recibida correctamente por el usuario.

El MNB es capaz de optimizar la tasa binaria asignada en el sistema en comparación con la arquitectura convencional UMTS. Como el MNB reduce considerablemente la potencia transmitida, el factor limitante pasará a ser el número de códigos disponibles y, por tanto, se podrá asignar recursos a más usuarios simultáneamente. Es interesante decir que en el escenario considerado el número de códigos limita la tasa a 1280 kbps.

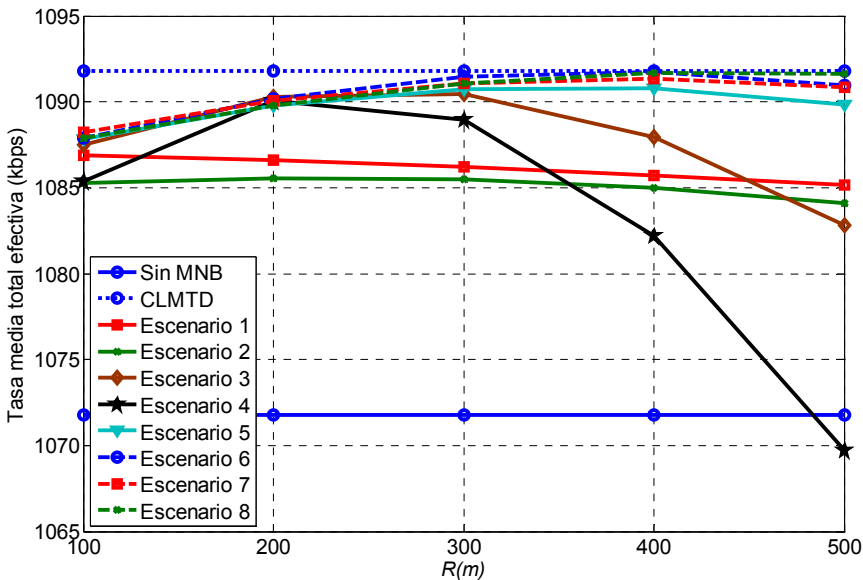


Figura 5.23 *Throughput* efectivo del sistema en función del valor de  $R$

Como era de esperar, los escenarios que más reducen el consumo de potencia son también los que maximizan en mayor media los recursos asignados. Además, los escenarios 3, 5, 6, 7 y 8 tienen un comportamiento similar, por encima del resto de escenarios y de la arquitectura UMTS convencional.

La Figura 5.24 muestra el *throughput* medio asignado a los usuarios web activos, es decir, con información pendiente de transmitir. Analizando este indicador de calidad, se puede concluir que la mejora ofrecida por el Multinodo B es mucho mayor en términos de maximización del *throughput* de lo que parecía al analizar el valor medio total de la celda. Gracias al MNB se consigue hasta un 35% de incremento en la tasa binaria asignada al usuario, lo que sin duda alguna redundará en una mayor calidad de servicio de los usuarios. De nuevo, cuanto mayor sea el número de micro-celdas desplegadas mejor se comportará el sistema, aunque ya con el escenario mixto con 3 micro-celdas se consiguen unos resultados muy satisfactorios.

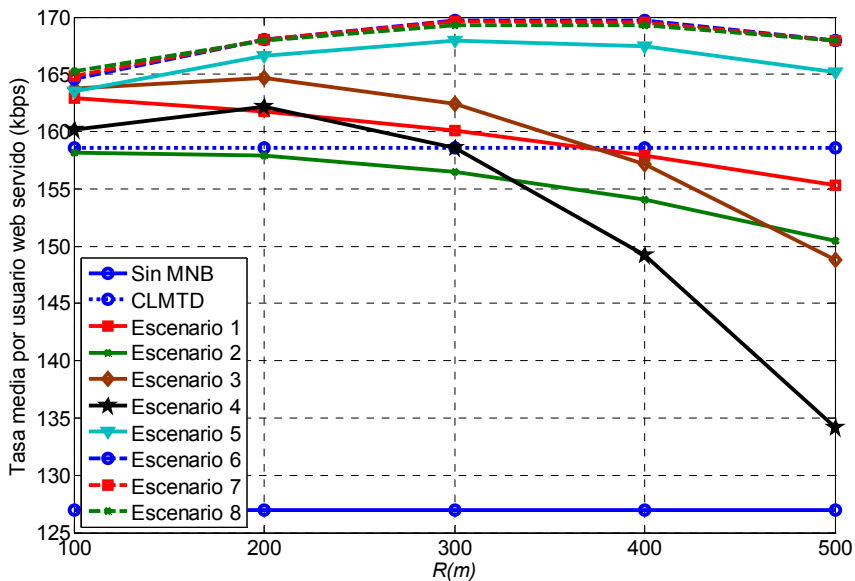


Figura 5.24 *Throughput* medio efectivo por usuario activo en función del valor de  $R$

El incremento en diversidad es un factor muy importante en esta ganancia. Como se puede observar, con el CLMTD diseñado en esta Tesis Doctoral y sin emplazamientos remotos es posible mejorar hasta un 27% las prestaciones del sistema, pero solamente combinando esta diversidad con las antenas remotas es posible obtener las mejores prestaciones.

Para corroborar estos resultados, se ha analizado el número medio de usuarios con información pendiente de transmitir. En la Figura 5.25 se comparan los resultados obtenidos para los distintos escenarios. Efectivamente, gracias al MNB se pasa de un número medio de 9.6 usuarios en cola a menos de 7. Es muy significativo destacar que del quinto al octavo escenario apenas se aprecia diferencia en los resultados. Así pues, el menor número de antenas del escenario 5 se compensa con la mayor probabilidad de transmisión en diversidad observada en la Figura 5.22.

Resultaría interesante estudiar otros parámetros de calidad de servicio de los usuarios para concluir si efectivamente no es necesario introducir en el MNB más de 3 micro-celdas. Por eso, a continuación se ha analizado el *delay* en la descarga de página web.

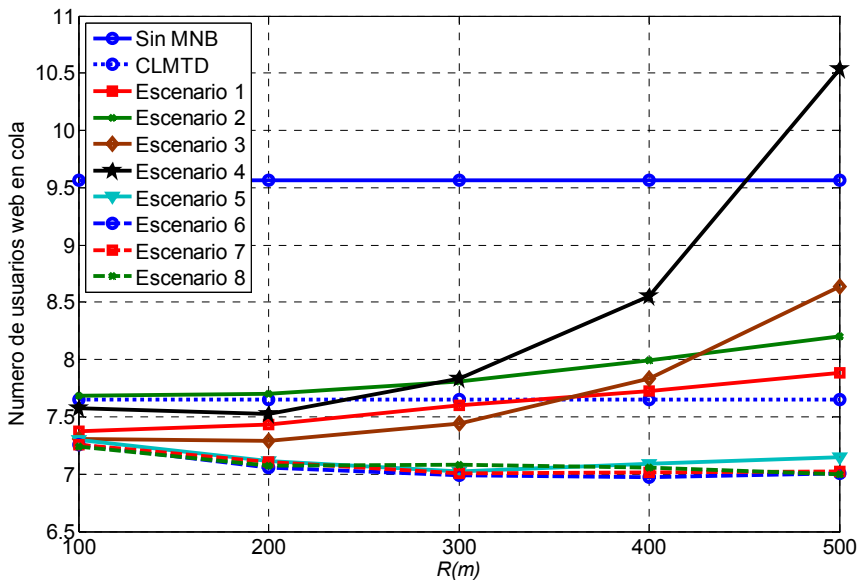
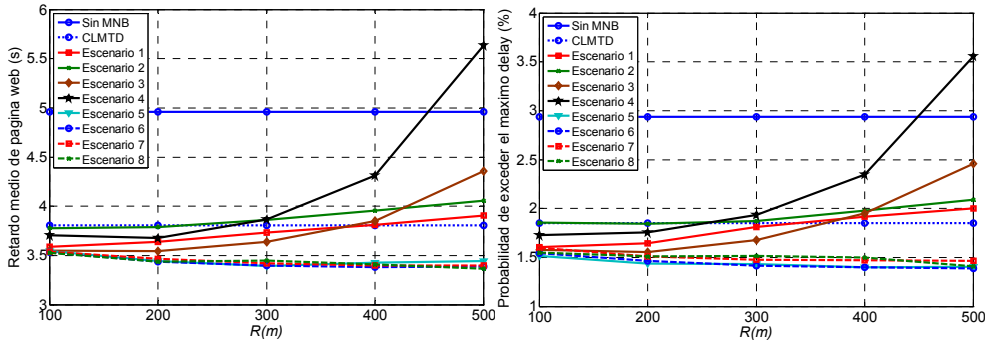


Figura 5.25 Número de usuarios web en cola en función del valor de  $R$





**Figura 5.26** Análisis de la descarga de página web en función del valor de  $R$ . Izquierda, tiempo de descarga de página web. Derecha, probabilidad de exceder el máximo delay

Como se puede observar en la Figura 5.26, entre los escenarios 5 y 8 apenas hay diferencias ni en el tiempo medio de descarga de una página web ni en la probabilidad de exceder el máximo delay. Una vez que se ha reducido la potencia consumida en el sistema hasta un límite en el que este factor deja de ser limitante para la asignación de recursos, todos los escenarios tendrán un comportamiento similar pues en todos ellos se utiliza el mismo algoritmo de gestión de recursos mediante redes Hopfield. Por ejemplo, simplemente utilizando CLMTD se reduce en 3W la potencia consumida, lo que al final se traduce en una calidad de servicio similar a la conseguida mediante otros despliegues más sofisticados.

En cualquier caso, cualquiera de los escenarios MNB consigue una importantísima reducción en el tiempo medio de descarga de página web y sobre todo en la probabilidad de exceder el máximo delay que se ve mejorada en hasta un 50%.

Así pues, gracias al MNB se da mejor servicio sobre todo a los usuarios más alejados de la celda central, aquellos que con mayor probabilidad superaban el retardo máximo establecido. Los nuevos emplazamientos y sobre todo la mayor diversidad en transmisión favorecen enormemente la calidad de servicio de los usuarios del sistema.

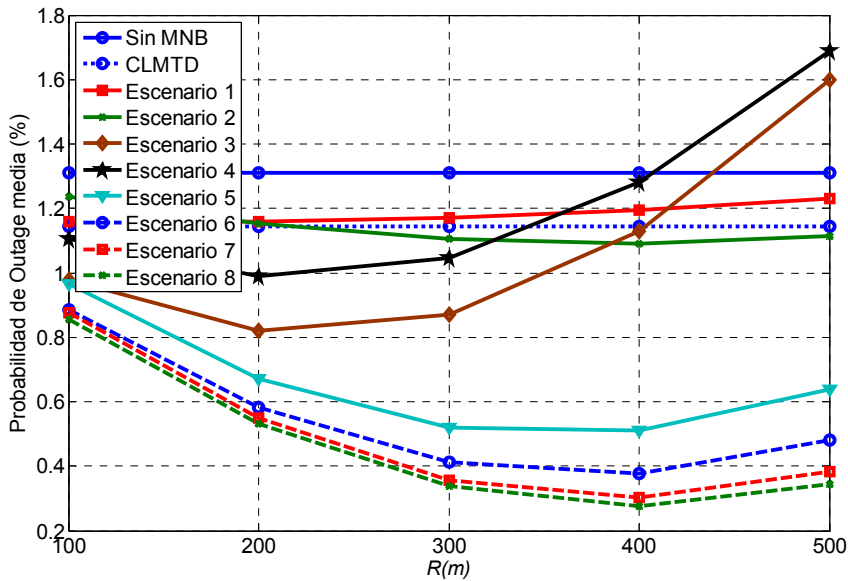


Figura 5.27 Probabilidad de *Outage* medio por usuario en función del valor de  $R$

Por último, se ha evaluado la probabilidad de *outage* en todos los escenarios, mostrando los resultados obtenidos en la Figura 5.27. En la nueva arquitectura MNB, además de minimizarse la potencia, maximizarse el uso de los recursos y reducirse el tiempo de descarga de las páginas web, una vez asignados los códigos y la potencia, la proporción de *slots* en los que un usuario activo presenta mala calidad disminuye. Esto es así por dos motivos. En primer lugar, al transmitirse en general con un nivel de potencia medio más bajo, la probabilidad de quedarse sin recursos es menor. Además, como se dispone de una diversidad adicional, en caso de que un enlace falle siempre se dispone de la posibilidad de transmitir desde la otra antena, reduciéndose la probabilidad de quedarse de repente sin un nivel adecuado de señal.

Cabe destacar que, al transmitir con una tasa binaria más alta, el sistema MNB debería caracterizarse por una mayor probabilidad de *outage*, pero los dos factores mencionados hacen que la tendencia sea la inversa.

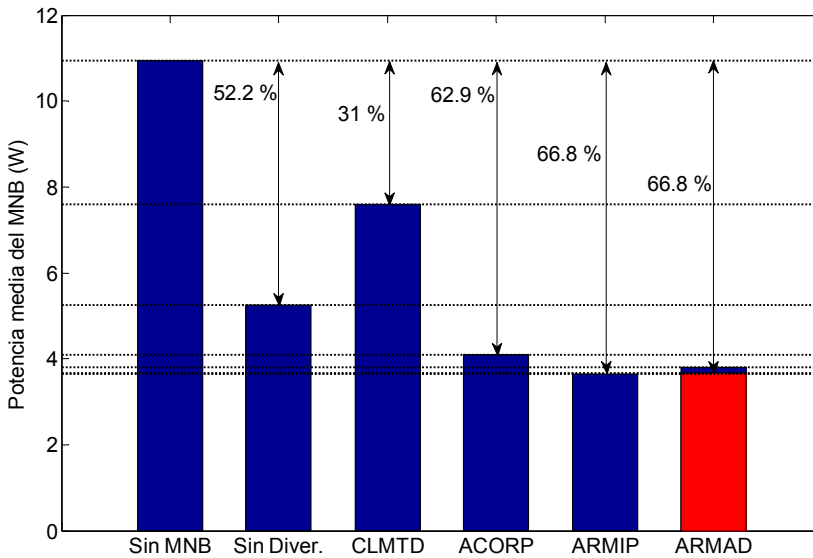
Tras analizar los parámetros de calidad de servicios de los usuarios, se puede concluir que los escenarios 5, 6, 7 y 8 presentan unos resultados muy similares. La única diferencia entre ellos es la potencia media transmitida por el MNB que será menor cuantas más antenas se consideren en el sistema. Sin embargo, y como se observa en la Figura 5.21, la potencia transmitida por las antenas micro-celulares es muy similar y la mayor diferencia de potencia se presenta en la antena principal. Por todo esto, parece lógico pensar que el escenario idóneo es el quinto, ya que un menor número de emplazamientos supone un ahorro importante en el despliegue para las operadoras. Considerando el quinto escenario, el radio óptimo de despliegue de las antenas micro-celulares es de 300m.

El siguiente apartado se centra en este escenario específico para comparar el comportamiento de los distintos algoritmos de distribución de recursos utilizados.

#### **5.4.3.2 Comparación de las distintas técnicas de reparto de recursos**

En el apartado 5.4.1 se describieron tres algoritmos distintos de asignación de recursos dentro del MNB. Dos de ellos son específicos de la arquitectura MNB-pasivo, el ARMIP y ARMAD, mientras que el tercero es específico de la arquitectura MNB-activo, el ACORP. Además de estas tres técnicas de asignación de recursos, en este capítulo también se van a comparar otras dos soluciones: o bien transmitir con diversidad desde un único emplazamiento (CLMTD), o bien no utilizar la diversidad y sólo transmitir simultáneamente o desde el emplazamiento macro-celular o desde el micro-celular (Sin Diversidad). Estos dos últimos algoritmos se han implementado particularizando el algoritmo ACORP descrito en este capítulo, por lo que son capaces de asignar códigos y potencia simultáneamente.

En primer lugar se ha evaluado el efecto de las distintas técnicas de gestión del MNB en términos de la potencia media consumida en el sistema. La Figura 5.28 muestra los resultados obtenidos, representando, además, la mejora relativa respecto a la arquitectura UMTS convencional.



**Figura 5.28 Potencia media consumida en el MNB para los distintos algoritmos de asignación de recursos. Mejora relativa a la arquitectura UMTS convencional**

Transmitir sin diversidad aprovechando el despliegue micro-celular adicional reduce considerablemente la potencia consumida, al igual que transmitir con diversidad por bucle cerrado desde el emplazamiento macro-celular. Comparando la técnica Sin Diversidad y el CLMTD, el despliegue micro-celular es mucho más efectivo reduciendo la potencia consumida en el sistema. Al tener antenas más cercanas y menos interferencia intra-celular, en el despliegue micro-celular una comunicación requiere mucha menos potencia transmitida. Sin embargo, mediante la combinación de ambas técnicas, es decir, lo que se realiza en las técnicas ACORP, ARMIP y ARMAD desarrolladas en esta tesis, es posible reducir aún más la potencia consumida, pasando de cerca de los 11 vatios a menos de 4, prácticamente una tercera parte.

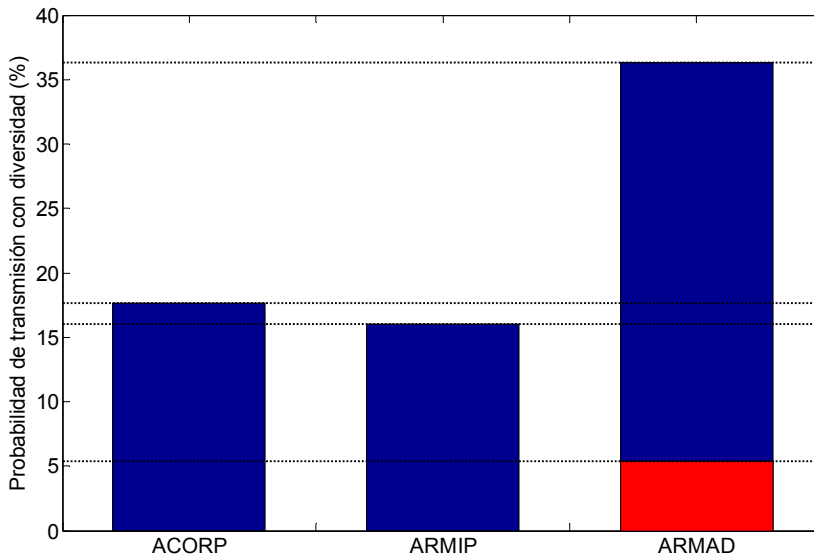
Entre las técnicas ARMAD y ARMIP, propias del MNB-pasivo, y ACORP parece que las primeras consiguen una mayor reducción en la potencia. Esto es debido a que como la asignación de códigos se realiza sin considerar el despliegue micro-celular

existente y la posterior diversidad por transmisión, se asigna recursos a muchos menos usuarios que en el caso del MNB-activo. Al posteriormente reducirse la potencia por la presencia del MNB, el MNB-pasivo consume menor potencia pero porque básicamente asigna recursos a menos usuarios y, por tanto, está desaprovechando la capacidad del sistema.

El algoritmo de asignación de recursos por márgenes de diversidad se ha evaluado para distintos umbrales, desde 1dB hasta 7, con saltos de 1dB. En los diagramas de barras utilizados se ha representado para el algoritmo ARMAD el valor mínimo y máximo del parámetro evaluado en cuestión. Respecto a la potencia consumida, la técnica ARMAD se comporta exactamente igual a la ARMIP cuando se utiliza un margen de 3dB. Si se utiliza un margen inferior se está infrutilizando la diversidad disponible en el sistema, mientras que si es superior el mayor balanceo de carga tiene como contrapartida un mayor consumo de potencia. El problema que tiene decidir transmitir con diversidad a partir de la diferencia de potencia entre las contribuciones recibidas desde ambas antenas, es que el comportamiento de la técnica de diversidad depende de qué enlace es el dominante. Cuando el escenario micro-celular es el que domina sobre el macro-celular, hasta con una diferencia de potencia relativa de 7 dB es preferible transmitir desde ambas antenas, mientras que en el caso contrario por encima de 1.5dB no resulta rentable transmitir con diversidad. Como el algoritmo ARMAD no diferencia entre ambos casos, el valor de 3dB resulta ser el valor intermedio óptimo entre ambos márgenes.

Siguiendo el mismo análisis realizado en el apartado anterior, a continuación se ha evaluado la probabilidad de transmitir con diversidad. No se han representado los valores para el escenario Sin MNB, Sin Diversidad y CLMTD, puesto que, en estos casos, se sabe a priori cuál va a ser dicha probabilidad: 0, 0 y 100% respectivamente.

El algoritmo ACORP presenta una probabilidad de transmisión en diversidad ligeramente superior al ARMIP. Esto quiere decir que, debido al quinto término de la función de energía, la red Hopfield transmite más veces simultáneamente desde ambas antenas, incluso en ocasiones en las que no resulta óptima esta decisión en términos de potencia consumida.

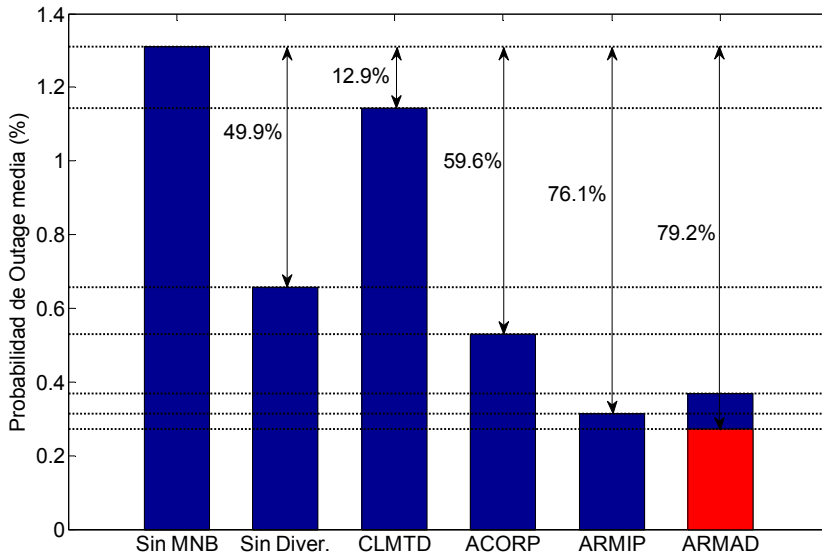


**Figura 5.29 Probabilidad de transmitir con diversidad en el MNB para los distintos algoritmos de asignación de recursos**

El mayor balanceo de carga puede estar en estos casos aconsejado por cuestiones de interferencias, ya que puede ser rentable descargar por ejemplo una micro-celda si casualmente muchos usuarios coinciden temporalmente alrededor de la misma.

Respecto al ARMAD, con un margen de 1dB se consigue la menor probabilidad de transmisión con diversidad, mientras que con 7 dB se consigue el máximo, como era de esperar. El cruce entre ARMAD y ARMIP se produce de nuevo para 3 dB, consiguiéndose en ese caso una probabilidad del 16.02 %.

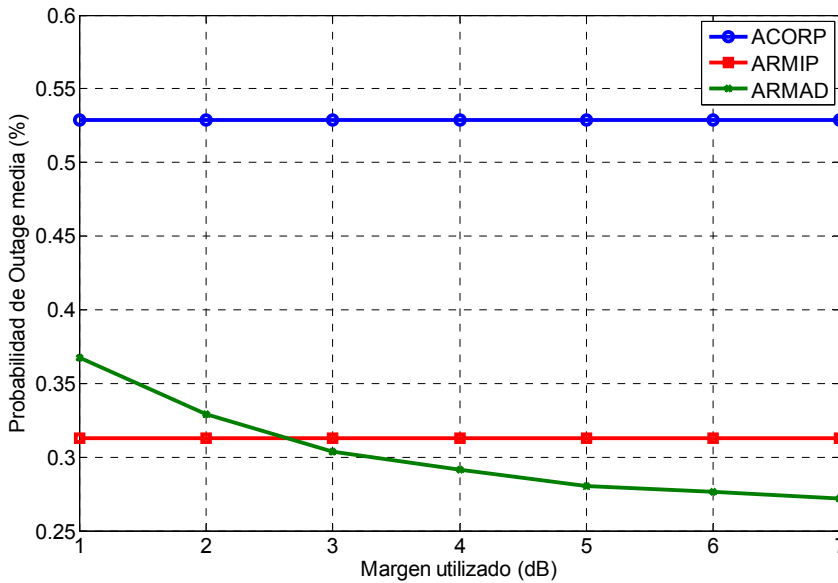
Una vez concluido el estudio de la probabilidad de transmisión en diversidad, se ha analizado la probabilidad de *outage*, ya que este parámetro afecta en gran medida a la calidad de servicio proporcionada por las distintas técnicas utilizadas. En la Figura 5.30 se puede apreciar como la mayor probabilidad de mala calidad la presenta la arquitectura convencional UMTS. Como no dispone de diversidad en transmisión ni de un despliegue micro-celular adicional, la calidad de los enlaces establecidos es menor.



**Figura 5.30 Probabilidad de *Outage* de los distintos algoritmos**

Introducir diversidad en transmisión mejora la probabilidad de *outage*, como se puede apreciar para el CLMTD, pero aún así los usuarios alejados del centro de la celda siguen sufriendo de mala calidad, por lo que la mejora es de tan solo un 12.9%. El cambio es más significativo (49.9%) al introducir la estructura micro-celular del MNB. Ahora los usuarios más alejados tienen otras antenas más cercanas capaces de satisfacer sus necesidades de potencia. Combinando ambas técnicas, como realizan ACORP, ARMIP y ARMAD, se puede conseguir una reducción aun mayor de la probabilidad de *outage*. De nuevo, los algoritmos ARMIP y ARMAD parece que consiguen un mejor funcionamiento que el ACORP, pero como antes eso es debido a la peor utilización de los recursos.

Es interesante analizar el comportamiento específico del algoritmo de asignación de recursos por márgenes de diversidad. La Figura 5.31 muestra la variación de la probabilidad de *outage* para ese algoritmo con diferentes márgenes utilizados. En general, conforme aumenta el margen disminuye la probabilidad de *outage*.

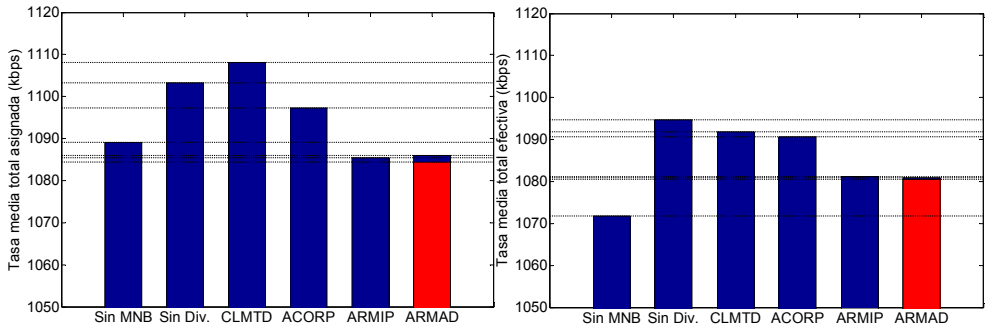


**Figura 5.31 Variación de la probabilidad de *Outage* del algoritmo ARMAD en función del margen de diversidad utilizado**

Es importante recordar que el comportamiento de la técnica ARMAD en términos de potencia consumida no presentaba esta tendencia decreciente sino que, por el contrario, la potencia aumentaba a partir de 3 dB. Esto reafirma el hecho de que la diversidad espacial adicional incorpora al sistema una mayor protección frente a los desvanecimientos en la calidad del canal. Cuanto mayor es el margen utilizado, mayor es la probabilidad de transmitir con diversidad y menor es la probabilidad de sufrir mala calidad en el enlace radio.

A continuación se ha analizado el *throughput* del sistema, tanto en términos de *throughput* asignado, es decir, de cantidad de recursos repartidos entre los usuarios, como en términos de *throughput* efectivamente cursado en la celda.



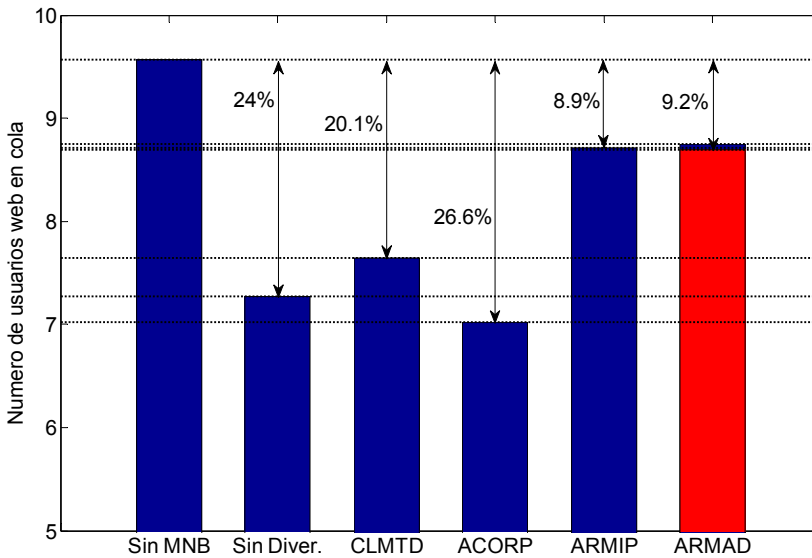


**Figura 5.32** *Throughput* del sistema para los distintos algoritmos de asignación de recursos. Izquierda, *throughput* asignado. Derecha, *throughput* efectivo

Observando el *throughput* asignado se puede distinguir claramente entre dos grupos: los algoritmos que consideran simultáneamente la asignación de códigos y las capacidades de transmisión adicionales del sistema (Sin Diversidad, CLMTD y ACORP) y, por otro lado, los algoritmos que asignan recursos considerando únicamente el emplazamiento principal (Sin MNB, ARMIP y ARMAD). Como los del primer grupo, los basados en un MNB-activo, saben que los usuarios consumen realmente menos potencia, pueden asignar más recursos simultáneamente, por lo que la tasa binaria asignada total es mayor.

Por otra parte, el hecho de que la arquitectura Sin MNB asigne más recursos que el ARMIP y ARMAD es únicamente debido a que existe un mayor número de usuarios en cola, lo que permite tener una mayor diversidad para elegir aquellos usuarios que consuman menos potencia.

Es muy destacable apreciar las diferencias entre la tasa binaria asignada y la efectiva. Como era de esperar, el *throughput* efectivo es siempre inferior al asignado, puesto que si no se dispone de suficiente calidad en el canal no se transmite correctamente la información. Los algoritmos que presentaban una menor probabilidad de *outage* son los que menos ven reducida su tasa binaria por culpa de la mala calidad. Por ejemplo, es significativo ver como ahora los algoritmos ARMAD y ARMIP tienen mucho mayor *throughput* efectivo que Sin MNB, cambiando la tendencia observada antes.



**Figura 5.33** Número de usuarios web en cola para los distintos algoritmos de asignación de recursos

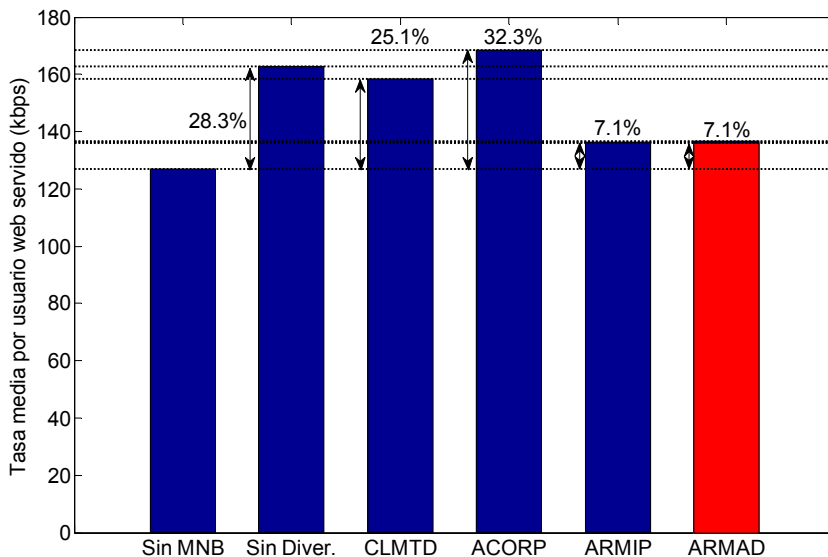
Como se ha comentado antes, el *throughput* medio del sistema tiene una clara dependencia con el número de usuarios activos en el sistema, cuyo valor medio se representa en la Figura 5.33. Como se puede ver, el algoritmo ACORP es el que mejor administra los usuarios activos y, por tanto, el que menor número de ellos tiene con información pendiente de transmitir.

Es interesante notar que el algoritmo Sin MNB tiene más usuarios en media que el ARMIP y ARMAD, siendo que ambos utilizan el mismo algoritmo de asignación de códigos entre los usuarios. El MNB-pasivo no sólo reduce la potencia consumida sino que, por esa misma reducción y la diversidad adicional introducida, también mejoran la calidad del enlace radio, reduciendo la probabilidad de *outage*. Como se desperdician menos bits asignados, al final el número de usuarios tiende a ser más pequeño. Como antes, tener una mayor diversidad incrementa la potencia consumida pero, por el

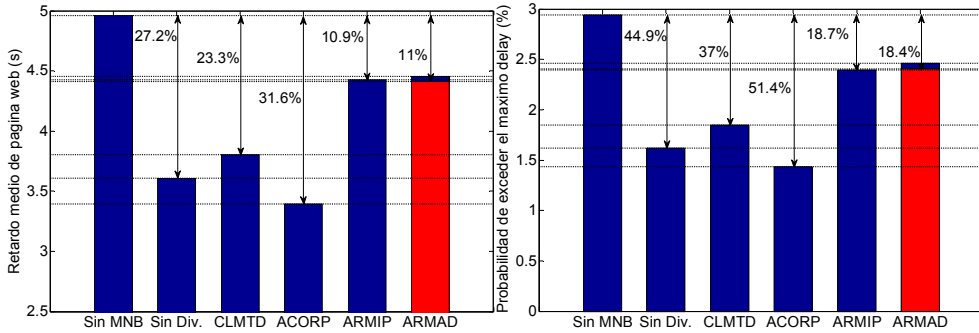
contrario, reduce la probabilidad de *outage* y también el número de usuarios medio en el sistema.

Respecto a los algoritmos de MNB-activo, existe una clara relación entre la potencia consumida, la probabilidad de *outage* y el número de usuarios en cola, puesto que las tendencias observadas son las mismas en todos los resultados mostrados. En cualquier caso, el algoritmo propuesto en esta Tesis Doctoral consigue las mejores prestaciones, reduciendo el número de usuarios en cola hasta en un 26.6%.

El diferente número de usuarios tiene un claro efecto sobre la tasa binaria efectiva por usuario activo en el sistema. En la Figura 5.34 se puede observar la comparación entre los distintos algoritmos considerados. El MNB-activo con asignación de recursos óptima mediante redes Hopfield es capaz de asignar un 32.3% más de recursos a cada usuario activo en el sistema. La optimización del uso de los recursos es más que patente a la vista de estos resultados y tendrá unas claras consecuencias en la calidad del servicio web, tal y como se puede ver en la Figura 5.35.



**Figura 5.34** *Throughput* medio efectivo por usuario activo para los distintos algoritmos de asignación de recursos



**Figura 5.35** Análisis de la descarga de página web para los distintos algoritmos de asignación de recursos. Izquierda, tiempo de descarga de página web. Derecha, probabilidad de exceder el máximo *delay*

Gracias al MNB-activo no sólo se reduce la potencia consumida en el sistema sino que, además, se mejora hasta en un 31.6% el tiempo medio de descarga de una página web. Además, el número de páginas cuyo tiempo de descarga supera los 30s se ve también reducido en un 51.4%.

Comparando los resultados de la arquitectura MNB-activo y MNB-pasivo se corrobora el hecho de que efectivamente el MNB-pasivo consume menos potencia media a costa de una peor calidad de servicio de los usuarios del sistema, aunque en cualquier caso los resultados son mejores que los mostrados por la arquitectura UMTS convencional.

## 5.5 Conclusiones

En este capítulo se ha abordado con éxito la gestión de recursos en las redes Multinodo B. Para ello, se han considerado todas las posibilidades en cuanto a la arquitectura de la nueva red de acceso radio propuesta en esta Tesis Doctoral.

En primer lugar, se ha explicado de manera realista y práctica cuál sería la implementación del control de potencia en el MNB. Gracias al envío de cierta información por parte del móvil, el sistema es capaz de estimar perfectamente la potencia requerida por los usuarios, controlando así el consumo total de potencia y por tanto la carga de la red. Los resultados obtenidos han demostrado que el procedimiento utilizado funciona y que el periodo que permite el estándar UMTS de envío de los informes es suficiente para garantizar el buen funcionamiento del MNB.

La selección de los códigos de ensanchado, es decir, la asignación de tasas binarias, se ha llevado a cabo mediante redes neuronales de Hopfield, ya que su buen funcionamiento quedó ya demostrado en el capítulo anterior. Por otro lado, en este capítulo se han propuesto tres técnicas distintas de selección de celda. Las dos primeras son independientes de la asignación de tasa binaria y, por tanto, se corresponden con la arquitectura MNB-pasivo. La tercera, se basa en una extensión de la red neuronal utilizada para el sistema UMTS convencional y permite que la decisión de código y antena se realice simultáneamente. De las dos técnicas de selección de celda del MNB-pasivo, la primera se basa en minimizar el consumo de potencia del usuario, mientras que la segunda establece un umbral de selección de antena, tal y como funciona el mecanismo de traspaso convencional de UMTS.

Para evaluar cuál es la mejor configuración de la red MNB y comprobar hasta qué punto el MNB ofrece una ganancia respecto a la red convencional, se han evaluado hasta ocho escenarios distintos de despliegue, modificando en cada uno de ellos la posición específica de las antenas.

Tras la evaluación exhaustiva realizada, se puede concluir que el MNB, en sus dos variantes activo y pasivo, es capaz de reducir en gran medida la potencia consumida en

el sistema. Además, fruto de este menor consumo y de la diversidad adicional introducida por la transmisión desde dos emplazamientos, la calidad de servicio se ve también mejorada en gran medida.

Comparando las dos soluciones de arquitectura MNB propuestas en esta Tesis, la mayor flexibilidad del MNB-activo permite obtener las mejores prestaciones en cuanto a calidad de servicio, a costa de un ligero incremento en el consumo de potencia respecto al MNB-pasivo.

Dentro de las dos técnicas de selección de celda en el MNB-pasivo, la basada en márgenes de diversidad es la que consigue una mejor calidad, siempre que los márgenes sean suficientemente altos. Este mejor comportamiento se debe a que esta técnica es capaz de dotar de más diversidad al sistema y, por tanto, de mayor protección frente a situaciones de mala calidad del canal. En contra, consume más potencia que la técnica basada en la minimización de la potencia consumida por el usuario.

Respecto a los distintos escenarios de despliegue de los transmisores, la conclusión alcanzada tras el análisis de los resultados es que las mejores prestaciones se obtienen aumentando el número de emplazamientos celulares, siempre que estos no incurran en interferencia mutua. A partir de tres micro-celdas adicionales no se consigue mejorar la calidad de servicio, pero sí reducir ligeramente el consumo de potencia del sistema. Además, cuantas más micro-celdas se desplieguen más interesa separarlas del centro de la celda para precisamente evitar esta interferencia indeseable. Finalmente, y de cara a reducir los costes de despliegue, parece razonable considerar que el escenario con tres micro-celdas es el óptimo, puesto que el pequeño mayor consumo de potencia se justifica sin duda por el hecho de no tener que alquilar más emplazamientos para las antenas.

Finalmente, es importante recalcar de nuevo que en este capítulo se ha demostrado la enorme mejora que conlleva el uso del MNB, corroborando así las hipótesis planteadas al inicio de la Tesis Doctoral.

---

## Referencias

- [1] E. Dahlman, B. Gudmundson, M. Nilsson and J. Skold, "UMTS/IMT-2000 Based on Wideband CDMA", IEEE Communications Magazine, pp.70-80, September 1998.
- [2] W. Xiaoxin, B. Murherjee, D. Ghosal, "Hierarchical Architectures in the Third-generation Cellular Network", IEEE Wireless Communications, vol. 11, no. 3, pp. 62-71, June 2004.
- [3] M. Hellebrandt, R. Mathar and M. Scheibenbogen "Estimating Position and Velocity of Mobiles in a Cellular Radio Network", IEEE Transactions on *Vehicular* Technology, vol. 46, no. 1, February 1997.
- [4] S.A. Ghorashi, H.K. Cheung, F. Said and A.H. Aghvami, "Performance of a CDMA-based HCS Network with Hybrid Speed/overflow-sensitive *handover* Strategy", IEE Proceedings- Communications, vol. 150, no. 4, pp. 293-297, Aug. 2003.
- [5] Y. Hwang and Y. Han, "Multimedia Traffic Management Schemes under Hierarchical Cell Structure", in Proc. IEEE TENCON Conference, vol. 2, pp. 1494-1497, Sept. 1999.
- [6] C. Hartmann and O. Schlegelmilch, "Hierarchical Cell Structures with Adaptive Radio Resource Management", in Proc. IEEE *Vehicular* Technology Conference, vol. 4, pp. 1764-1771, Sept. 2000.
- [7] Taesoo Kwon and Dong-Ho Cho, "Adaptive Radio Resource Management Based on Cell Load in CDMA Based Hierarchical Cell Structure", in Proc. IEEE *Vehicular* Technology Conference, vol. 4, pp. 2337-2341, Sept. 2002.
- [8] J.R.B. de Marca, I.F e Silva, "Further Simulation Results on Hierarchical Cell Structures", in Proc. IEEE Wireless Communications and Networking, vol. 3, pp. 1705-1708, March 2003.
- [9] S. Hoon Shin and K. Sup Kwak, "Power Control for CDMA Macro-Micro Cellular System", in Proc. IEEE *Vehicular* Technology Conference, vol. 3, pp. 2133-2136, May 2000.
- [10] I. Jami and H. Tao, "Micro-cell Planning within Macro-cells in UMTS: Downlink Analysis", in Proc. International Conference on 3G Mobile Communication Technologies, pp. 211-215, May 2002.

- [11] Sangbum Kim, Daehyoung Hong and Jaeweon Cho, "Hierarchical Cell Deployment for High Speed Data CDMA Systems", in Proc. IEEE Wireless Communications and Networking Conference, vol. 1, pp. 7-10, 2002.
- [12] Dong Geun Jeong and Wha Sook Jeon, "Load Sharing in Hierarchical Cell Structure for High Speed Downlink Packet Transmission", in Proc. IEEE Global Telecommunications Conference, vol. 2, pp. 1815-1819, 2002.
- [13] 3GPP TS 25.331 V6.0.1, "Radio Resource Control (RRC); protocol specification", 2004.
- [14] 3GPP TS 25.215 V6.4.0, "Physical layer - Measurements (FDD)", 2005.
- [15] 3GPP TS 25.214 V5.9.0, "Physical Layer Procedures (FDD)", 2004.
- [16] 3GPP TS 25.133 V3.20.0, "Requirements for support of radio resource management (FDD)", 2005.
- [17] J.F. Monserrat, D. Gomez-Barquero, D. Calabuig, L. Rubio and N. Cardona, "Evaluation of soft *handover* Micro Diversity Gain on the UMTS System Capacity and QoS", in Proc. IEEE 17th International Symposium on Personal, Indoor and Mobile Radio Communications, 2006.



# Capítulo 6

## *Modelado del Shadowing y su Influencia sobre la Evaluación de Sistemas*

*La verdadera grandeza de la ciencia  
acaba valorándose por su utilidad*

*Gregorio Marañón*

### **6.1 Introducción**

La propagación radio a través de un canal móvil se suele caracterizar por medio de tres fenómenos: las pérdidas básicas de propagación o *path loss*, un término de pérdidas a largo plazo conocido como *shadowing*, y un término de pérdidas a corto plazo que recibe el nombre de *fast fading*.

El modelo de pérdidas de propagación básicas se describe en el apartado B.1.5 mientras que el modelo de pérdidas a corto plazo se analiza en el apartado A.1.11. Este capítulo se centra en el estudio de los modelos de pérdidas a largo plazo, puesto que este aspecto del modelado del canal radio ha sido objeto de un meticuloso análisis en la presente Tesis Doctoral. Además de proponer un conjunto de modelos de creciente realismo y complejidad, este capítulo tiene como objetivo analizar el efecto del uso de los distintos modelos sobre la evaluación realizada de las prestaciones de un sistema, particularizando este estudio para el escenario de simulación propuesto en esta Tesis Doctoral. A partir de este análisis se podrá concluir la utilidad de los modelos complejos de *shadowing* en función del tipo de estudio que se pretenda realizar.

El capítulo comienza con un análisis de las características del fenómeno de propagación conocido como *shadowing*. A partir de este análisis se expondrán un

conjunto de modelos desarrollados en el seno de esta Tesis Doctoral, modelos que irán considerando sucesivamente cada una de las características del *shadowing*.

Por otro lado, en el sistema implementado, el organizador o *scheduler* realiza una toma de decisión que conlleva una determinada asignación de recursos cada cierto tiempo. Esta distribución de los recursos se mantiene entre dos tomas de decisiones del *scheduler* lo que implica que una decisión que se ha tomado con unas determinadas condiciones de propagación de los usuarios puede no ser válida si estos se alejan demasiado de la posición en la que se encontraban. Una mayor correlación de las muestras de *shadowing* hará que la asignación sea más exitosa, por lo que se obtendrán unas mejores prestaciones del algoritmo. Si, por el contrario, se carece de tal correlación, la asignación, que en un determinado instante pudiera ser ideal, puede tener fatales consecuencias en las siguientes tramas.

Es por esto por lo que en esta Tesis Doctoral no solamente se han propuesto distintos modelos de *shadowing* evaluando sus propiedades, sino que también se ha caracterizado el efecto que los distintos modelos tienen en las prestaciones finales del sistema. Así pues, se analizarán los resultados obtenidos en un mismo escenario de simulación utilizando los distintos modelos propuestos. A partir de estos resultados se podrá concluir qué nivel de complejidad es necesario alcanzar para tener un buen equilibrio entre coste computacional y fiabilidad del análisis.

## 6.2 Modelos de *Shadowing*

El *shadowing* representa un término de pérdidas aleatorio que se añade a las pérdidas básicas obtenidas de forma determinista mediante el modelo de propagación.

Campañas de medidas efectuadas en entornos reales [1][2] han demostrado que las pérdidas de *shadowing* siguen una distribución log-normal (distribución normal en unidades logarítmicas) de media 0 y desviación típica  $\sigma$ , cuyo valor depende de la diversidad del entorno de propagación. La desviación típica implementada en el simulador es de 6 dB al ser este un valor típico en entornos urbanos [1][3].

Con las consideraciones anteriores, el efecto del *shadowing* entre una estación base y un terminal móvil puede modelarse añadiendo a las pérdidas básicas, expresadas en dB, una variable aleatoria con la siguiente función de densidad de probabilidad:

$$f_s(s) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{s^2}{2\sigma^2}} \quad (6.1)$$

No obstante, para modelar adecuadamente el *shadowing* entre una estación base y un terminal móvil, no basta con caracterizar su distribución estadística, es necesario tener en cuenta, además, otras propiedades.

En primer lugar, en un determinado punto geográfico, el *shadowing* correspondiente a las distintas estaciones base del escenario no es independiente, sino que presenta una correlación distinta de cero. En otras palabras, supóngase que  $\{S_1, S_2, \dots, S_N\}$  representan los valores de *shadowing* entre un conjunto de  $N$  estaciones base y un terminal situado en una determinada posición, entonces para cualquier par de valores  $(S_i, S_j)$  debe cumplirse que el coeficiente de correlación entre ellos debe ser igual a un valor  $\rho_{ij}$  distinto de cero. Como consecuencia, para calcular el *shadowing* entre un terminal situado en una posición determinada, y un conjunto de estaciones base, no basta con generar una variable independiente para cada una de las estaciones, puesto que el *shadowing* presenta unas determinadas características de correlación cruzada.

Por otra parte, el *shadowing* es un proceso correlado espacialmente, lo que se denomina autocorrelación espacial, lo cual quiere decir que el *shadowing* entre una estación base y un terminal móvil en una determinada posición, esta correlado con el observado en las posiciones cercanas. En [2] y [4] se refleja que la función de autocorrelación espacial del *shadowing* se puede modelar con suficiente precisión mediante una función exponencial. Así pues, la función de autocorrelación entre dos puntos separados una distancia  $\Delta r$  viene dada por:

$$R_s(\Delta r) = e^{-\ln 2 \frac{\Delta r}{d_{corr}}} = 2^{-\frac{\Delta r}{d_{corr}}} \quad (6.2)$$

En la expresión anterior  $d_{corr}$  es un parámetro importante conocido como distancia de decorrelación. Este parámetro representa la distancia mínima que tiene que haber

entre dos puntos para que la correlación entre las muestras de *shadowing* en esas posiciones sea menor o igual que 0.5. La distancia de decorrelación depende del entorno de propagación. En [2] se propone una distancia de decorrelación de 20 m para el entorno de prueba *vehicular*.

Como se puede observar, para modelar correctamente el proceso de desvanecimiento lento, es decir el *shadowing*, sería necesario considerar de manera conjunta la autocorrelación espacial y la correlación cruzada. Sin embargo, un modelo de tales características es bastante complejo y requiere una caracterización meticulosa de los procesos estadísticos de segundo orden de un gran número de muestras de pérdidas de propagación. En el desarrollo de esta Tesis Doctoral se han propuesto diversos modelos de *shadowing* abordando paso a paso la caracterización de los diversos procesos de correlación inherentes a este fenómeno de propagación.

A continuación, se describen todos los modelos de *shadowing* considerados por orden de creciente detalle de caracterización. Los dos primeros son ampliamente utilizados en la literatura mientras que los cuatro últimos, más sofisticados, han sido propuestos en el seno de esta Tesis Doctoral. El estudio del efecto producido por los distintos modelos de *shadowing* sobre las prestaciones del sistema se deja para su posterior análisis en la sección de resultados (Apartado 6.3).

### **6.2.1 Modelo 1: Unidimensional Incorrelado**

Según este modelo, el *shadowing* sólo se caracteriza por su distribución estadística log-normal. Cada vez que se requiera calcular el valor del desvanecimiento lento se elegirá aleatoriamente un valor de una distribución gaussiana de media cero y desviación típica  $\sigma$ . Este valor representa por tanto el valor de *shadowing* en unidades logarítmicas. Si transcurrida una trama se necesita de nuevo una muestra de *shadowing* se repetirá la operación sin tener en cuenta el valor anterior asignado. Además, según este modelo, no existe ningún tipo de correlación cruzada entre el *shadowing* experimentado por las distintas estaciones base consideradas.

A pesar de que pueda parecer un modelo extremadamente simple es bastante común su utilización, sobre todo en simuladores estáticos. Véase como ejemplo [5].

## 6.2.2 Modelo 2: Unidimensional con Autocorrelación Espacial

Este modelo se corresponde con el propuesto por Gudmundson en [2]. La correlación cruzada no se considera pero sí la autocorrelación espacial. Por lo tanto, se debe garantizar que la correlación formulada en (6.2) se está implementando correctamente. Según este modelo el valor de *shadowing* se calcula para todos los enlaces de un móvil (los de señal útil e interferentes) cada unidad de tiempo, que en UMTS se corresponderá con una trama de 10ms. Como, además, la velocidad de movimiento de los usuarios es conocida, en ese intervalo de tiempo se habrá recorrido una cierta distancia. El modelo de Gudmundson permite conocer la correlación de dos muestras separadas una determinada distancia, por lo que directamente se puede calcular el valor de esa correlación utilizando la expresión (6.2).

Kim y Han [6] propusieron un modelo sencillo para implementar esta correlación que se enuncia a continuación. Se denomina  $S_{ik}$  a la muestra  $k$ -ésima de *shadowing* de un determinado enlace  $i$  y  $R_S$  al valor de correlación entre dos muestras consecutivas dada una determinada distancia  $\Delta r$  entre ellas de manera que:

$$R_S(\Delta r) = 2^{-\frac{v \cdot \Delta t}{d_{corr}}} \quad (6.3)$$

siendo  $v$  la velocidad del móvil en  $m/s$  y  $\Delta t$  el periodo de trama (10ms).

El valor de la siguiente muestra de *shadowing* para el enlace  $i$ -ésimo se calculará como:

$$S_{ik+1} = R_S \cdot S_{ik} + \sqrt{1 - R_S^2} Z_{ik} \quad (6.4)$$

donde  $Z_{ik}$  es la muestra  $k$ -ésima de una variable aleatoria gaussiana,  $Z_i$ , independiente, de media cero y desviación típica  $\sigma$ .

A continuación se demostrará que, mediante la utilización de (6.4), se introduce la correlación deseada en las muestras, manteniendo los estadísticos de primer orden de la variable de *shadowing*.

Las variables aleatorias  $S_i$  y  $Z_i$  cumplen las siguientes propiedades:

$$\begin{aligned} E[S_i] &= 0, E[S_i^2] = \sigma^2 \\ E[Z_i] &= 0, E[Z_i^2] = \sigma^2 \end{aligned} \quad (6.5)$$

Además, como son incorreladas se cumple que:

$$E[S_i \cdot Z_i] = 0 \quad (6.6)$$

Ahora se define una nueva variable aleatoria,  $T_i$ , de manera que satisfaga (6.4):

$$T_i = R_s \cdot S_i + \sqrt{1 - R_s^2} Z_i \quad (6.7)$$

La media y varianza de  $T_i$  será:

$$\begin{aligned} E[T_i] &= E[R_s \cdot S_i + \sqrt{1 - R_s^2} \cdot Z_i] = E[R_s \cdot S_i] + E[\sqrt{1 - R_s^2} \cdot Z_i] = \\ &= R_s \cdot E[S_i] + \sqrt{1 - R_s^2} \cdot E[Z_i] = 0 \end{aligned} \quad (6.8)$$

$$\begin{aligned} E[T_i^2] &= E[R_s^2 \cdot S_i^2 + (1 - \rho^2) \cdot Z_i^2 + 2 \cdot R_s \cdot \sqrt{1 - R_s^2} \cdot S_i \cdot Z_i] = \\ &= E[R_s^2 \cdot S_i^2] + E[(1 - R_s^2) \cdot Z_i^2] = R_s^2 \cdot \sigma^2 + (1 - R_s^2) \cdot \sigma^2 = \sigma^2 \end{aligned}$$

Mientras que la correlación entre  $T_i$  y  $S_i$  resulta ser:

$$\rho = \frac{E[S_i T_i]}{\sqrt{E[S_i^2] E[T_i^2]}} = \frac{E[R_s \cdot S_i^2 + \sqrt{1 - \rho^2} \cdot T_i \cdot S_i]}{\sigma^2} = \frac{R_s \cdot \sigma^2}{\sigma^2} = R_s \quad (6.9)$$

De manera que se cumplen todos los requisitos que se buscaba conseguir.

Finalmente, es importante recalcar que el modelo descrito en esta sección es ampliamente utilizado en la literatura dada su sencilla implementación (véanse como ejemplo [6] y [7] aunque en la literatura se pueden encontrar muchos más). Además, el consumo de memoria es muy reducido ya que solamente es necesario que, por cada

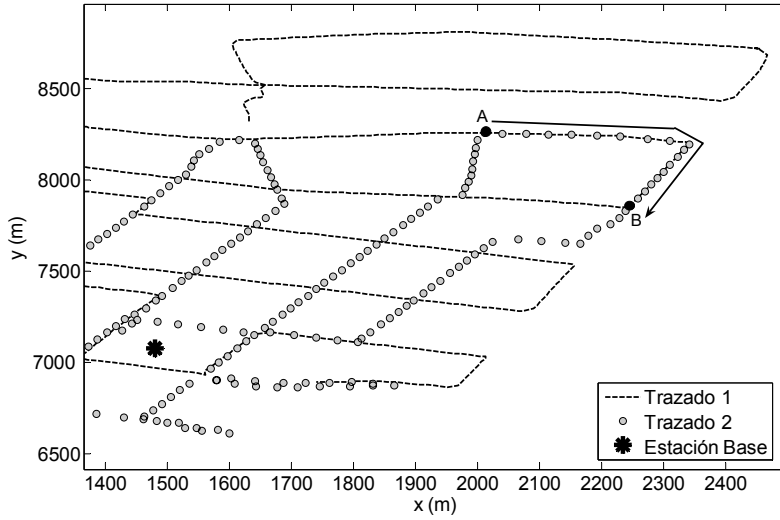
terminal simulado, se almacene el valor de *shadowing* asignado a cada uno de los enlaces con las estaciones base implicadas en su transmisión y/o interferencia.

### 6.2.3 Modelo 3: Bidimensional con Autocorrelación Espacial

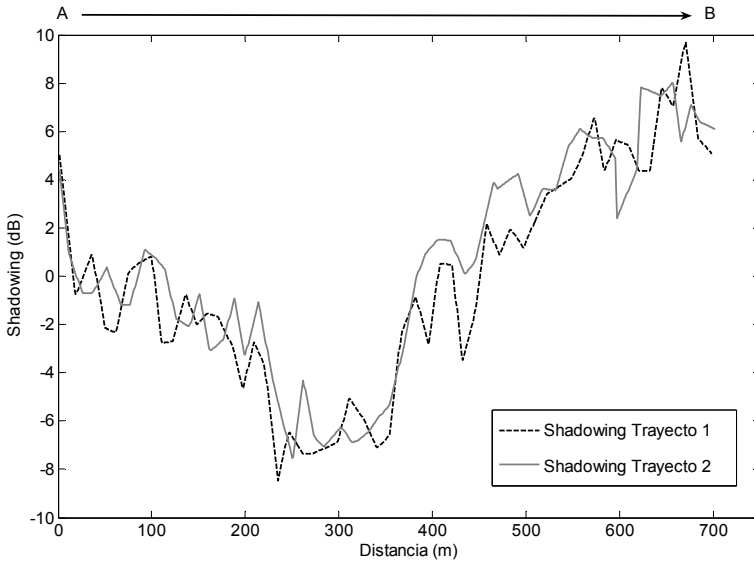
El modelo propuesto en el anterior apartado presenta una clara limitación dado que el *shadowing* se calcula independientemente para cada usuario. Este hecho tiene como consecuencia que el *shadowing* experimentado por usuarios que se encuentren muy próximos físicamente puede ser radicalmente distinto, siendo las condiciones de ocultamiento las mismas. Además, el modelo anteriormente expuesto no tiene en cuenta la trayectoria del usuario. Si éste, por ejemplo, describe un círculo y vuelve a una posición inicial, puede darse el caso de que el *shadowing* para una misma posición sea totalmente dispar.

Intuitivamente parece que un modelo más preciso debería asegurar que el *shadowing* experimentado en un punto fuera el mismo independientemente del usuario y del momento en que se mida ese nivel de ocultamiento. Para corroborar esta idea intuitiva se ha analizado los resultados obtenidos en una campaña de medidas realizada en Munich y cuyos datos fueron compartidos dentro de la iniciativa europea COST231 [8]. En esta campaña de medidas se analizaron las pérdidas de propagación respecto de una estación base cercana en una serie de trayectos. La Figura 6.1 recoge dos trazados recorridos en la campaña de medida. Las medidas realizadas en el trazado 1 y trazado 2 se realizaron por un equipo de medidas distinto en momentos distintos. Para comprobar la invariabilidad temporal del *shadowing* se ha analizado un recorrido en el que se dispone de medidas de ambos trazados, en concreto se ha estudiado el trayecto entre A y B mostrado en la Figura 6.1. A partir de las medidas de pérdidas de propagación se puede calcular fácilmente la desviación de estas pérdidas respecto a la regresión lineal. Esta desviación se corresponde con el valor de *shadowing* en cada punto. La Figura 6.2 representa el valor obtenido de *shadowing* en ambos trayectos. Como se puede observar, dicho valor es claramente similar y las diferencias obtenidas se corresponden

con posibles variaciones del punto de medida o con el efecto propio de *fast fading* que no queda totalmente eliminado tras el promediado.



**Figura 6.1** Recorridos seguidos en la campaña de medidas



**Figura 6.2** Valor de *shadowing* en los dos trayectos



Tras este análisis se puede concluir que, efectivamente, la hipótesis postulada anteriormente es cierta y el *shadowing* no varía temporalmente sino que solamente depende de las condiciones de ocultamiento que hay en un determinado enlace.

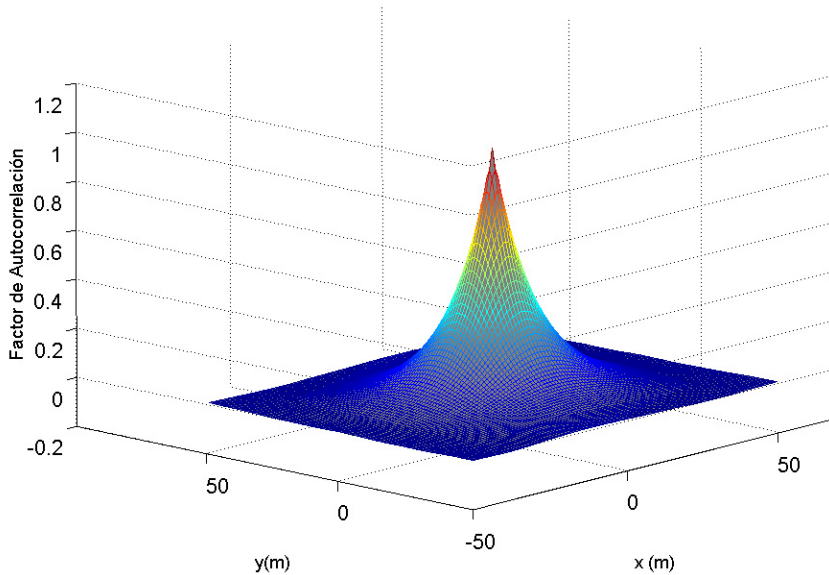
Es por tanto fundamental desarrollar un modelo que permita caracterizar el *shadowing* atendiendo a este hecho real. Para modelar el efecto del *shadowing* teniendo en cuenta las propiedades de autocorrelación espacial y de invariabilidad temporal, se ha empleado un nuevo procedimiento desarrollado en esta Tesis Doctoral contando con la colaboración del Dr. Rubén Fraile y publicado en [4].

Este método tiene como objetivo generar un conjunto de  $N$  matrices o mapas cubriendo la zona bajo estudio, siendo  $N$  el número de emplazamientos del escenario y, por tanto, el número de enlaces de propagación caracterizados en cada punto. Cada uno de estos mapas contendrá las muestras de *shadowing* correspondientes a un emplazamiento, para un conjunto de puntos del área de simulación. El conjunto de puntos estará determinado por los nodos de una rejilla cuadrada que representa la matriz. Estos mapas deben presentar la propiedad de autocorrelación deseada, es decir, los valores de un mismo mapa correspondientes a puntos cercanos deben mostrar la autocorrelación espacial propuesta en [2]. A continuación, se describe el procedimiento seguido para generar dichas matrices de *shadowing* con correlación espacial.

En primer lugar se generarán  $N$  matrices con muestras aleatorias de una distribución log-normal con la media y desviación típica deseadas. A partir de estos mapas, la autocorrelación espacial se puede introducir filtrando cada uno de ellos mediante un filtro bidimensional, diseñado de forma que, al final del proceso, las muestras de los mapas presenten la siguiente función de autocorrelación bidimensional:

$$R_s(\Delta x, \Delta y) = e^{-\ln 2 \cdot \frac{\sqrt{\Delta x^2 + \Delta y^2}}{d_{corr}}} = 2^{-\frac{\sqrt{\Delta x^2 + \Delta y^2}}{d_{corr}}} \quad (6.10)$$

siendo  $\Delta x$  y  $\Delta y$  la distancia entre dos puntos del mapa en las coordenadas horizontal y vertical. La Figura 6.3 representa esta función de autocorrelación teórica.



**Figura 6.3 Función de autocorrelación del *shadowing***

Aplicando las propiedades de la transformada de Fourier, es sencillo obtener, a partir de la autocorrelación, la función de transferencia del filtro deseado:

$$\left| F \left[ R_s (\Delta x, \Delta y) \right] \right| = \left| H (f_x, f_y) \right|^2 \quad (6.11)$$

Una posible solución para  $H (f_x, f_y)$  viene dada por:

$$\begin{aligned} \left| H (f_x, f_y) \right| &= \sqrt{\left| F \left[ R_s (\Delta x, \Delta y) \right] \right|} \\ \angle H (f_x, f_y) &= \angle F \left[ R_s (\Delta x, \Delta y) \right] \end{aligned} \quad (6.12)$$

Los cálculos involucrados en las ecuaciones anteriores son complejos de realizar analíticamente, por lo que es necesario obtener la solución mediante métodos numéricos.

Se debe tener en cuenta, además, que el filtro no debe alterar la varianza del *shadowing*. Esto se consigue normalizando los coeficientes del filtro, esto es, debe verificarse:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h^2(x, y) \cdot dx \cdot dy = 1 \quad (6.13)$$

Una vez realizado el proceso descrito, se obtendrán los mapas que contendrán las muestras de *shadowing* de cada una de las estaciones del escenario, presentando dichas muestras las propiedades requeridas de correlación espacial.

Con la utilización de los mapas descritos anteriormente, el cálculo de las pérdidas de propagación totales entre una estación base y un terminal móvil se reduce al siguiente procedimiento:

1. A partir de la distancia, en valor absoluto, entre la estación base y la estación móvil se obtiene el valor de las pérdidas medias de propagación utilizando la expresión correspondiente.
2. A partir de la posición de la estación móvil sobre el área de simulación en coordenadas  $x$  e  $y$ , se obtiene el valor de las pérdidas asociadas al *shadowing* accediendo a la posición adecuada de la matriz de *shadowing* correspondiente al emplazamiento al que pertenece la estación base.
3. Se suman los dos valores anteriores para obtener las pérdidas totales de propagación en dB.

## Validación del modelo

Para comprobar la validez del modelo propuesto para obtener los desvanecimientos causados por el *shadowing* es necesario verificar los siguientes aspectos:

- Las muestras de *shadowing* de cada mapa siguen la distribución propuesta en el modelo.
- Las muestras de *shadowing* de cada mapa exhiben las propiedades de autocorrelación espacial esperadas, coincidiendo la distancia de decorrelación con la establecida.

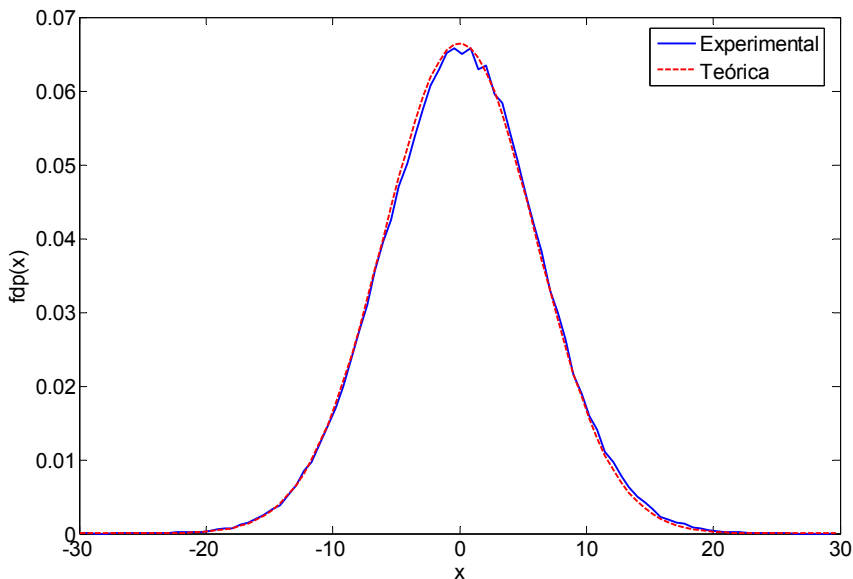
Debido a la imposibilidad de generar mapas de *shadowing* genéricos que puedan ser utilizados para cualquier definición del escenario de simulación, se ha optado por

realizar las comprobaciones anteriores sobre los mapas correspondientes al escenario de simulación planteado en la presente Tesis Doctoral. Las principales características de los mapas de *shadowing* correspondientes al escenario descrito se resumen en la Tabla 6.1.

Número de mapas, $N$	7
Tamaño del mapa en x	1000 m
Tamaño del mapa en y	1000 m
Número de puntos por mapa	200 x 200
Distancia entre puntos	5 m
Correlación cruzada entre estaciones base	0
Distancia de decorrelación $d_{corr}$	20 m

**Tabla 6.1 Parámetros de los mapas de *shadowing* analizados**

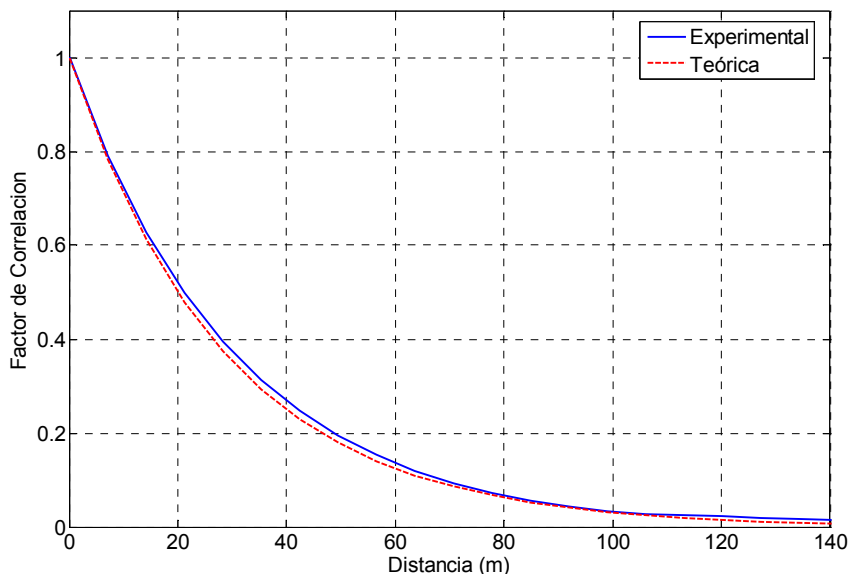
La Figura 6.4 muestra la media de las funciones de densidad de probabilidad de las muestras de *shadowing* de los 7 mapas considerados. Junto a ella se representa la distribución log-normal teórica propuesta en el modelo de *shadowing*. Como se puede observar, el ajuste entre las dos trazas es bastante satisfactorio.



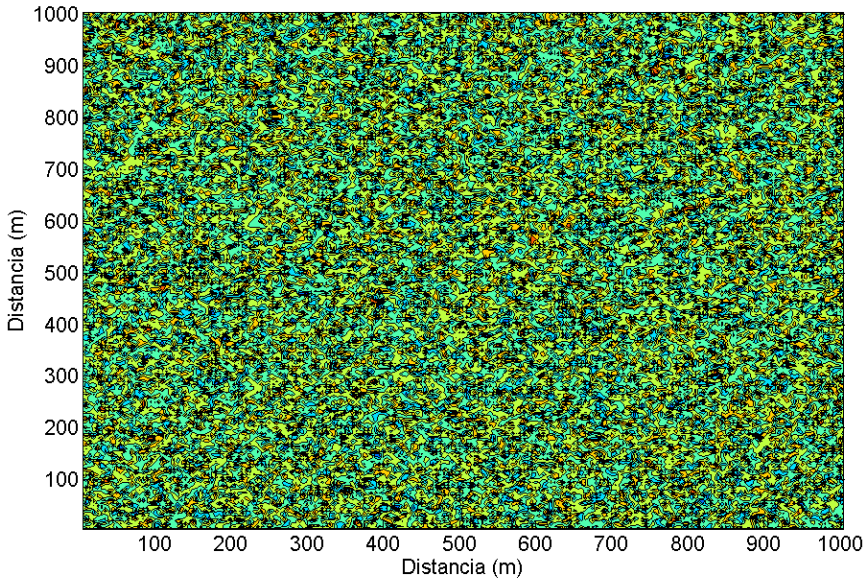
**Figura 6.4 Función densidad de probabilidad del *shadowing***

La Figura 6.5 muestra, en color azul, la media de las funciones de autocorrelación espacial correspondientes a cada uno de los 7 mapas de *shadowing* y, en color rojo, la función de autocorrelación teórica. La gráfica permite comprobar que la función de autocorrelación de los mapas se ajusta considerablemente bien a la función de autocorrelación teórica. En particular, la autocorrelación entre muestras de los mapas correspondientes a puntos geográficos separados una distancia de 20m es igual a 0.5, tal y como se esperaba del modelo de *shadowing*.

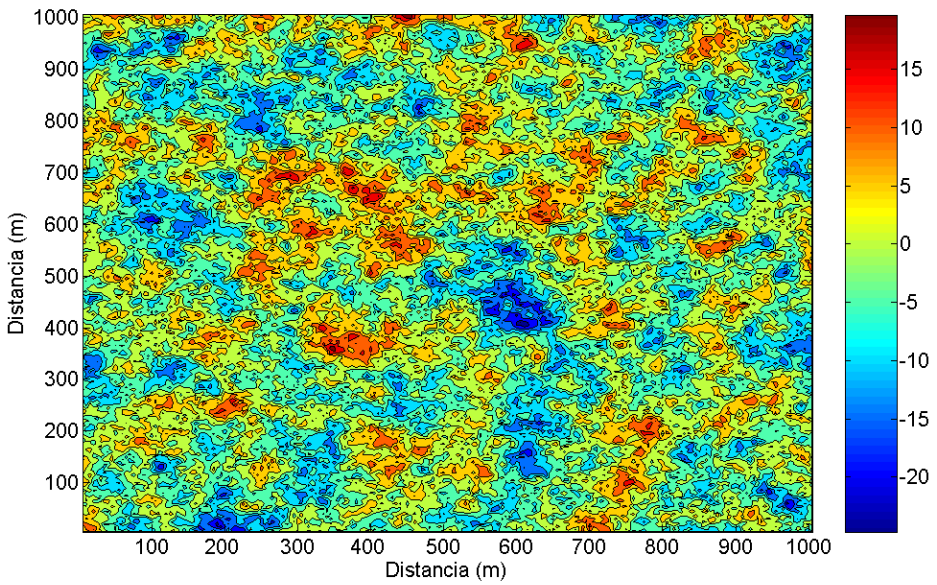
Para finalizar el análisis de los mapas de *shadowing*, en la Figura 6.6 se representan dos mapas. El primero de ellos, Mapa A, contiene muestras de *shadowing* incorreladas espacialmente, mientras que el segundo, Mapa B, es el mapa resultante de filtrar las muestras del primero con el filtro de correlación. Comparando ambos mapas se puede observar visualmente el efecto de este filtrado.



**Figura 6.5** Autocorrelación espacial de las muestras de los mapas de *shadowing*



*Mapa A. Shadowing incorrelado espacialmente*



*Mapa B. Shadowing correlado espacialmente*

**Figura 6.6** Mapas de *Shadowing* correlados e incorrelados espacialmente

## 6.2.4 Modelo 4: Bidimensional con Correlación Cruzada Constante

Para introducir en los mapas la correlación entre estaciones base se ha empleado el siguiente procedimiento, publicado asimismo en [4]. Según este método, si se desea obtener un conjunto de  $N$  variables aleatorias gaussianas  $\{S_1, S_2, \dots, S_N\}$ , de forma que cualquier par de estas variables presente un coeficiente de correlación cruzada constante,  $\rho$ , basta con generar  $N+1$  variables aleatorias gaussianas independientes  $\{G, G_1, G_2, \dots, G_N\}$  y posteriormente calcular cada  $S_i$  mediante la siguiente expresión:

$$S_i = \sqrt{\rho} \cdot G + \sqrt{1-\rho} \cdot G_i \quad (6.14)$$

De acuerdo con este método, la variable aleatoria gaussiana de *shadowing* en un determinado punto geográfico correspondiente a cada una de las estaciones base puede obtenerse como combinación de dos variables aleatorias gaussianas independientes, una común a todas las estaciones base,  $G$ , y otra particular para cada estación,  $G_i$ . Utilizando este método en cada uno de los puntos del escenario se obtienen  $N$  mapas de *shadowing*, siendo  $N$  el número de emplazamientos del escenario, de forma que los valores de los distintos mapas para el mismo punto presentan la correlación deseada.

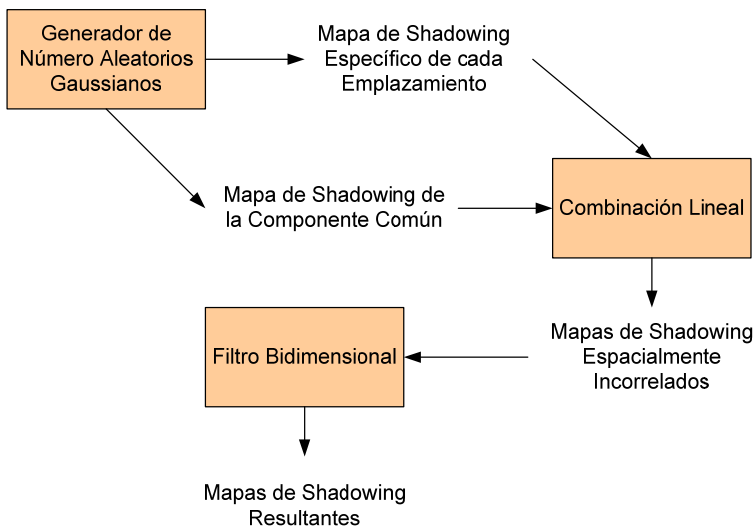
La comparativa de las ecuaciones (6.14) y (6.4) puede llevar a confusión. Mediante (6.4) se pretendía correlar dos muestras consecutivas de una misma variable aleatoria mientras que (6.14) lo que hace es introducir una determinada correlación constante entre cualquier combinación de dos variables aleatorias distintas de un conjunto de variables. A continuación, se propone una sencilla demostración de que la correlación cruzada entre dos variables aleatorias cualquiera,  $S_i$  y  $S_j$ , generadas mediante la expresión (6.14), se corresponde con lo esperado.

$$\begin{aligned}
 R_S(S_i, S_j) &= \frac{E[S_i S_j]}{\sqrt{E[S_i^2] E[S_j^2]}} = \frac{E[(\sqrt{\rho} \cdot G + \sqrt{1-\rho} \cdot G_i)(\sqrt{\rho} \cdot G + \sqrt{1-\rho} \cdot G_j)]}{\sqrt{E[(\sqrt{\rho} \cdot G + \sqrt{1-\rho} \cdot G_i)^2] E[(\sqrt{\rho} \cdot G + \sqrt{1-\rho} \cdot G_j)^2]}} = \\
 &= \frac{E[\rho G^2 + \sqrt{\rho-\rho^2} \cdot GG_j + \sqrt{\rho-\rho^2} \cdot GG_i + (1-\rho)G_i G_j]}{\sqrt{E[\rho G^2 + (1-\rho)G_i^2 + 2\sqrt{\rho-\rho^2} \cdot GG_i] E[\rho G^2 + (1-\rho)G_j^2 + 2\sqrt{\rho-\rho^2} \cdot GG_j]}} = \\
 &= \frac{\rho \cdot \sigma^2}{\sqrt{(\rho\sigma^2 + (1-\rho)\sigma^2)(\rho\sigma^2 + (1-\rho)\sigma^2)}} = \rho
 \end{aligned}$$

De manera similar se podría proceder para comprobar que, efectivamente, la media y la desviación típica son las esperadas.

Una vez que se han generado los mapas con correlación entre estaciones base, se introducirá la correlación espacial tal y como se propuso en el apartado anterior.

La Figura 6.7 resume el proceso seguido para la generación de los mapas de *shadowing* bidimensionales con correlación cruzada y espacial.



**Figura 6.7** Proceso de generación de los mapas de *Shadowing* con correlación cruzada y espacial



## Validación del modelo

Para comprobar la validez del procedimiento empleado es necesario verificar los siguientes aspectos:

- La correlación cruzada entre las estaciones base es la esperada.
- La inclusión de la correlación cruzada no altera las propiedades de correlación espacial conseguidas con la utilización del filtro bidimensional diseñado en el apartado anterior.

Para realizar este estudio se ha seleccionado un escenario muy parecido al utilizado en el apartado anterior (ver Tabla 6.1) con la única salvedad de que en esta ocasión se ha considerado una correlación cruzada  $\rho = 0.5$  [9].

La Tabla 6.2 muestra el coeficiente de correlación entre las muestras de cada pareja de mapas, promediando los resultados obtenidos con mapas generados con 10 semillas aleatorias distintas. Como se puede observar, en todos los casos el valor obtenido está muy próximo a 0.5, que es el valor de correlación entre estaciones base que se pretendía obtener al generar los mapas.

	Mapa 1	Mapa 2	Mapa 3	Mapa 4	Mapa 5	Mapa 6	Mapa 7
Mapa 1	1.0000	0.5029	0.5107	0.5039	0.5222	0.4972	0.4862
Mapa 2	0.5029	1.0000	0.4867	0.5119	0.4947	0.5124	0.5098
Mapa 3	0.5107	0.4867	1.0000	0.4942	0.4998	0.4945	0.4814
Mapa 4	0.5039	0.5119	0.4942	1.0000	0.4801	0.5058	0.4858
Mapa 5	0.5222	0.4947	0.4998	0.4801	1.0000	0.4841	0.4833
Mapa 6	0.4972	0.5124	0.4945	0.5058	0.4841	1.0000	0.4988
Mapa 7	0.4862	0.5098	0.4814	0.4858	0.4833	0.4988	1.0000

**Tabla 6.2 Correlación cruzada entre las muestras de los mapas de *shadowing***

Finalmente, se ha evaluado la función de correlación espacial de un mapa, comprobando que la introducción de correlación cruzada no afecta al filtrado. La Figura 6.8 representa tal correlación, muy similar a la mostrada anteriormente.

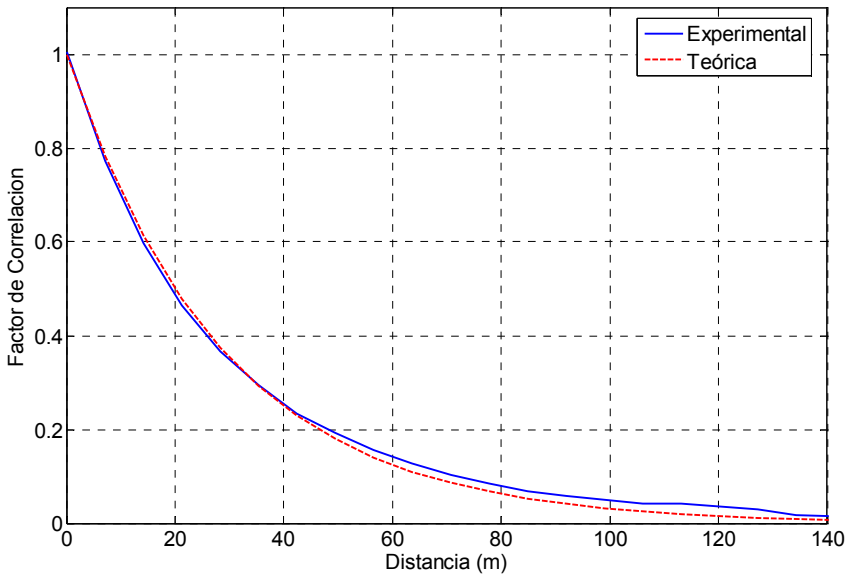


Figura 6.8 Autocorrelación espacial de las muestras de los mapas de *shadowing*

## 6.2.5 Modelo 5: Bidimensional con Correlación Cruzada Basada en Medidas

El efecto de la correlación cruzada, corroborado en distintas campañas de medidas [1][10][11], está provocado por el hecho de que el desvanecimiento aleatorio a largo plazo de las pérdidas de propagación está compuesto por dos componentes: la primera está originada por los obstáculos situados en los alrededores del equipo receptor, mientras que la segunda es propia del escenario de propagación específico de la estación base. Por lo tanto, como la primera componente es igual para todas las estaciones, el fenómeno del *shadowing* presenta cierta correlación.

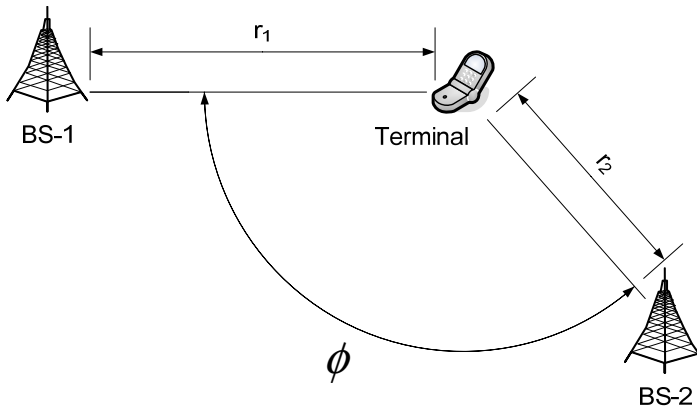
Se ha demostrado que el efecto de esta correlación cruzada es muy importante sobre todo cuando se evalúan ciertos aspectos del diseño de sistemas como por ejemplo [12]:

- Selección óptima del ancho de haz de las antenas para la sectorización.
- Prestaciones de la macro-diversidad producida por *soft handover*, ya que estos esquemas proveen de máxima ganancia cuando la correlación es baja.

- Prestaciones de los algoritmos de *handover*. Estos mecanismos realizan la actualización del conjunto de estaciones activas en base a los niveles relativos de potencia recibidos de cada una de las estaciones. Para poder optimizar correctamente las ventanas de actualización es necesario un conocimiento exacto de la correlación entre las señales recibidas por el terminal.
- Planificación óptima de frecuencias para GSM.
- Cálculo de las prestaciones de antenas adaptativas.
- Probabilidad de *outage* de la calidad de la señal recibida, ya que la distribución de los valores de SIR tendrá menor variabilidad si aumenta la correlación del *shadowing*.

Algunos autores, como Viterbi [9], proponen que esta correlación se puede considerar como un valor constante de entre 0.3 y 0.5 y esta idea es la que se ha implementado en el modelo anterior. Sin embargo, todos los autores asumen que cuanto mayor sea el camino de propagación común entre dos estaciones base hasta el terminal mayor será el coeficiente de correlación. Dada esta afirmación, el coeficiente de correlación del *shadowing* se modela normalmente como una función de dos variables [13]: la diferencia del ángulo de llegada de la señal al terminal o *Angle-of-Arrival Difference* (AAD) y, opcionalmente, la diferencia relativa de distancia,  $\Delta R$ . A partir de la campaña de medidas de Graziano [1], Saunders [12] propuso un modelo que consideraba simultáneamente ambas variables y que lograba un ajuste casi perfecto a las medidas reales. Otros autores, como Zayana y Guisnet [14] o Sorensen [10], también propusieron modelos basados en medidas, pero estos no han alcanzado el nivel de aceptación del modelo de Saunders.

El modelo propuesto por Saunders considera el siguiente escenario



**Figura 6.9 Modelo físico de correlación cruzada propuesto por Saunders [12]**

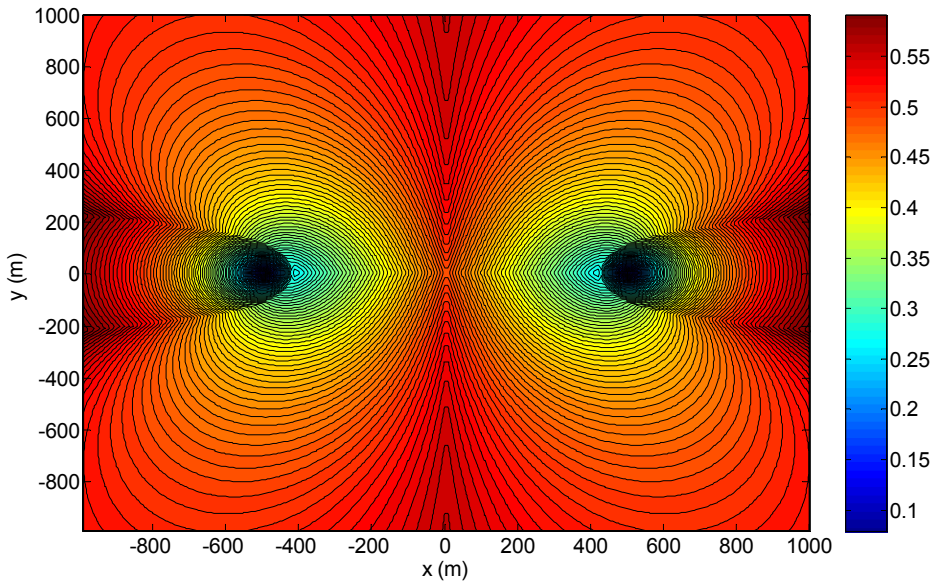
El valor de AAD se representa mediante  $\phi$  mientras que la diferencia relativa de distancia será la relación entre  $r_1$  y  $r_2$ . A partir de este escenario el factor de correlación cruzada entre ambas estaciones es [12]:

$$\rho_c = \begin{cases} \sqrt{\frac{r_2}{r_1}} & , 0 \leq \phi < \phi_T \\ \left(\frac{\phi_T}{\phi}\right)^\gamma \sqrt{\frac{r_2}{r_1}} & , \phi_T \leq \phi < \pi \end{cases} \quad (6.15)$$

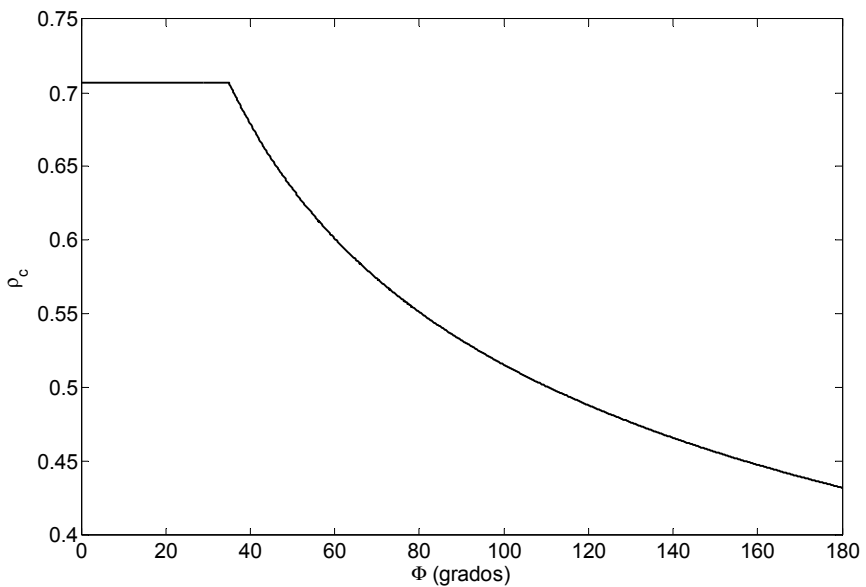
siendo  $r_2$  la menor de las distancias entre ambas estaciones base y  $\gamma$  un parámetro que permite el ajuste del modelo a una secuencia de datos obtenidos experimentalmente. Por su parte  $\phi_T$  depende del ancho de los obstáculos encontrados. Saunders demostró que existe una relación entre este parámetro y la distancia de decorrelación de manera que

$$\phi_T = 2 \sin^{-1} \frac{d_{corr}}{2r_2} \quad (6.16)$$

La Figura 6.10 muestra la correlación cruzada alrededor de dos estaciones base utilizando este modelo. Se puede observar como el valor medio de la correlación está en torno a 0.5. Por su parte, la Figura 6.11 muestra el comportamiento del modelo en función de la AAD.



**Figura 6.10** Ejemplo de cálculo de la correlación cruzada utilizando la ecuación (6.15).  
 Parámetros  $\gamma = 0.3$  y  $d_{corr} = 20$ . Las estaciones base están situadas en  $(\pm 500 \text{ m})$



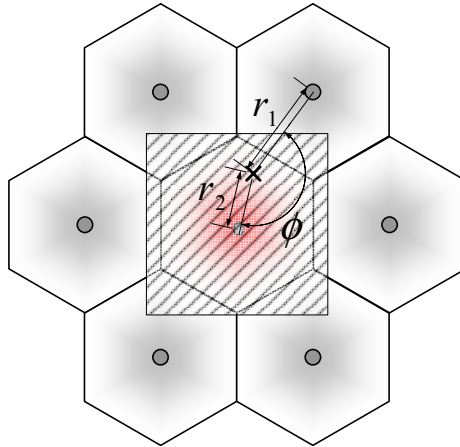
**Figura 6.11** Efecto del AAD. Parámetros  $\gamma = 0.3$  y  $d_{corr} = 20$ . Las estaciones base están situadas a una distancia de 500 y 1000 m respectivamente.

Una vez conocido el modelo de correlación basado en medidas que se va a implementar, hay que decir qué variables se van a correlar y de qué manera. Una primera aproximación podría garantizar únicamente la correlación cruzada entre el conjunto de estaciones base interferentes respecto a la contribución principal. Este modelo tiene como limitación que las interferentes entre sí no presentan una correlación basada en medidas, lo que incrementa la variabilidad de la SIR y, además, no permitiría evaluar convenientemente los procesos de *soft handover*. Sin embargo, es un modelo muy sencillo de implementar como se verá a continuación. Un modelo más sofisticado calcularía para el conjunto de  $N$  estaciones base consideradas una matriz de correlación de dimensiones  $N \times N$  y aplicaría dicha matriz para la generación de las muestras de *shadowing*. El primer modelo se corresponde con el descrito en este apartado mientras que la segunda aproximación será desarrollada en el siguiente apartado.

En primer lugar, se selecciona el área de interés, que en el escenario planteado coincide con el área de cobertura de la celda bajo estudio (representado en la Figura 6.12 como un cuadrado de traza diagonal). Se generan  $N$  matrices,  $\{\mathbf{G}_1, \dots, \mathbf{G}_N\}$ , una por cada estación base considerada (según el esquema de la figura harían falta 7 matrices), cuyos elementos se corresponden con cada uno de los nodos de una determinada rejilla espacial que coincide con el cuadrado. Se selecciona la primera matriz, la de la celda central o de interés, como la matriz común,  $\mathbf{G}_1$ , que además coincidirá con la matriz resultante del proceso de correlación cruzada para la estación base de interés,  $\mathbf{S}_1$ . Considerando la estación base  $i$ -ésima se obtiene su matriz de correlación,  $\mathbf{R}_i$ , calculando cada uno de sus componentes de acuerdo a su posición espacial equivalente y utilizando la ecuación (6.15). Finalmente, la matriz de *shadowing* con correlación cruzada de la estación base  $i$ -ésima se obtiene utilizando la siguiente expresión:

$$\mathbf{S}_i = \mathbf{R}_i * \mathbf{G}_1 + \sqrt{1 - \mathbf{R}_i^2} * \mathbf{G}_i \quad (6.17)$$

donde el operador  $*$  se corresponde con una multiplicación término a término, mientras que las operaciones de potenciación también se ejecutan término a término.



**Figura 6.12 Esquema del modelo**

Como se demostró en el apartado 6.2.2, mediante la ecuación (6.17) se garantiza la correlación  $\mathbf{R}_i$  entre  $\mathbf{S}_1$  y  $\mathbf{S}_i$  sin modificar las propiedades estadísticas de  $\mathbf{G}_i$ .

Finalmente, se filtran las matrices  $\mathbf{S}_i$  utilizando el filtro bidimensional descrito con anterioridad para introducir la correlación espacial deseada.

### Validación del modelo

Matemáticamente ya se ha demostrado que el método utilizado para introducir la correlación es el adecuado y también se evaluó en apartados anteriores que, independientemente de la correlación cruzada, el filtrado bidimensional garantiza la correlación espacial. Un estudio que resulta interesante es analizar la correlación cruzada entre los distintos mapas de *shadowing* obtenidos mediante este método. La Tabla 6.3 recoge estos valores de correlación.

Como se puede observar, la correlación entre la primera matriz de *shadowing*, correspondiente a la matriz de interés, y el resto de matrices tiene un valor medio muy próximo a 0.5, mientras que la correlación entre el resto de mapas es menor, del orden de 0.25, lo que pone de manifiesto la incapacidad de este método para modelar la correlación cruzada real existente entre todos los caminos de propagación.

	Mapa 1	Mapa 2	Mapa 3	Mapa 4	Mapa 5	Mapa 6	Mapa 7
Mapa 1	1.0000	0.4939	0.5066	0.5077	0.5149	0.4998	0.4814
Mapa 2	0.4939	1.0000	0.2508	0.2573	0.2604	0.2355	0.2409
Mapa 3	0.5066	0.2508	1.0000	0.2667	0.2543	0.2553	0.2567
Mapa 4	0.5077	0.2573	0.2667	1.0000	0.2636	0.2593	0.2514
Mapa 5	0.5149	0.2604	0.2543	0.2636	1.0000	0.2365	0.2467
Mapa 6	0.4998	0.2355	0.2553	0.2593	0.2365	1.0000	0.2329
Mapa 7	0.4814	0.2409	0.2567	0.2514	0.2467	0.2329	1.0000

Tabla 6.3 Correlación cruzada entre las muestras de los mapas de *shadowing*

## 6.2.6 Modelo 6: Bidimensional con Correlación Cruzada Completa

Dado un escenario con  $N$  estaciones base, en cada posición del terminal se puede definir el valor del coeficiente de correlación cruzada,  $\rho$ , entre cada uno de los valores de *shadowing*  $x_i$  ( $i=1,2,\dots,N$ ) resultado una matriz de correlación,  $\mathbf{R}=\{\rho_{ij}\}$ , de dimensiones  $N \times N$ , de forma que, para cada par de variables  $\{x_i, x_j\}$ , se calcula  $\rho_{ij}$  utilizando (6.15). Una vez definida la matriz de correlación para cada punto, la dificultad estriba en generar un valor de *shadowing* aleatorio para cada estación,  $x_i$ , de manera que, efectivamente, se cumplan las condiciones de correlación. Este problema se puede resolver matemáticamente utilizando la descomposición de Cholesky [15] de la siguiente manera. Si la matriz de correlación,  $\mathbf{R}_{N \times N} = \{\rho_{ij}\}$ , es simétrica y definida positiva, entonces su descomposición de Cholesky es  $\mathbf{R} = \mathbf{C}\mathbf{C}^T$ , donde  $\mathbf{C}$  es una matriz triangular inferior y el superíndice T representa la operación de transposición. Después de generar un vector  $\mathbf{r}$  que contenga  $N$  valores aleatorios gaussianos estadísticamente independientes y que presenten los mismos estadísticos de primer orden que se espera del *shadowing*, es decir media cero y desviación típica  $\sigma$ , las  $N$  muestras de *shadowing* correladas se pueden calcular como:

$$\mathbf{x} = \mathbf{C}\mathbf{r} \quad (6.18)$$

La principal limitación de este método matemático es precisamente que  $\mathbf{R}$  tiene que ser definida positiva. Esta limitación fue observada en primer lugar por Graziosi [16] quien propuso elegir adecuadamente una función matemática de la correlación que



asegurara que la matriz de correlación fuera positiva. Sin embargo, un modelo matemáticamente adecuado no se tiene porqué corresponder con la correlación entre los enlaces de propagación en circunstancias reales y, por lo tanto, puede no resultar válido. Por otro lado, las matrices de correlación calculadas mediante modelos ajustados a medidas pueden no ser definidas positivas. Esto introduce valores propios negativos en la matriz  $\mathbf{R}$  y, como resultado, la matriz  $\mathbf{C}$  no siempre existe, lo que hace que el método de generación de valores correlados mediante la descomposición de Cholesky sea inutilizable. El problema es que, hasta la fecha, no existe otro método matemático para generar  $N$  variables que satisfagan una determinada matriz de correlación.

En esta Tesis Doctoral se ha demostrado que muchos modelos de la correlación cruzada del *shadowing* extraídos de datos reales, incluido el modelo de Saunders, carecen de la necesaria condición de positividad, sobre todo cuando las matrices de correlación son de gran tamaño, es decir, cuando el número de estaciones base consideradas aumenta. El problema es, por tanto, definir un nuevo procedimiento que sea capaz de encontrar la matriz simétrica y definida positiva más cercana a la obtenida por el modelo experimental, de modo que sea posible aplicar la descomposición de Cholesky.

### **6.2.6.1 Procedimiento de ajuste de las matrices de correlación cruzada**

La descomposición de Cholesky solamente se puede aplicar si y solo si la matriz en cuestión es simétrica y definida positiva. Una matriz de correlación es siempre simétrica con diagonal principal unidad, por lo que solamente se ha de comprobar su positividad. En esta Tesis Doctoral se ha particularizado el método de proyecciones alternativas AP (*Alternating Projection*) [17] definiendo los subespacios de interés y las reglas de proyección. Se define  $D$  como el subespacio de las matrices simétricas definidas positivas y  $U$  como el subespacio de las matrices simétricas con diagonal

unitaria. Dada una matriz de correlación experimental,  $\mathbf{R} \in U$ , el algoritmo busca la matriz  $\mathbf{R}^C \in \{D \cap U\}$  más próxima a  $\mathbf{R}$  en la norma Frobenius<sup>3</sup>.

Según el método AP, la búsqueda se puede hacer mediante proyecciones iterativas de la matriz  $\mathbf{R}$  sobre los subespacios  $D$  y  $U$  hasta que se alcance la convergencia. La proyección sobre el subespacio  $D$ ,  $P_D(\mathbf{R}, \varepsilon)$ , consiste en eliminar todos los valores propios negativos o iguales a cero, sustituyéndolos por un valor propio de valor  $\varepsilon$  positivo pero lo bastante pequeño. Para proyectar una matriz sobre el subespacio  $U$ ,  $P_U(\mathbf{R})$ , se fuerza a todos los elementos de la diagonal a ser exactamente iguales a 1.

La particularización del método de las proyecciones alternativas propuesto en esta Tesis Doctoral resulta ser:

$$\begin{aligned}
 & t = tol + 1, \Delta \mathbf{E}_0 = 0, \mathbf{R}_0^C = \mathbf{R} \\
 & \text{for } (k = 1, t \geq tol, k = k + 1) \{ \\
 & \quad \mathbf{X}_k = \mathbf{R}_{k-1}^C - \Delta \mathbf{E}_{k-1} \\
 & \quad \mathbf{Y}_k = P_D(\mathbf{X}_k, \varepsilon) \\
 & \quad \Delta \mathbf{E}_k = \mathbf{Y}_k - \mathbf{X}_k \\
 & \quad \mathbf{R}_k^C = P_U(\mathbf{Y}_k) \\
 & \quad t = \frac{\|\mathbf{R}_k^C - \mathbf{R}_{k-1}^C\|_{\infty}}{\|\mathbf{R}_k^C\|_{\infty}} \\
 & \}
 \end{aligned} \tag{6.19}$$

Este método iterativo continúa el proceso de acercamiento hasta que la distancia relativa entre dos matrices resultado consecutivas sea menor que una determinada tolerancia,  $tol$ . El trabajo de Boyle y Dykstra [17] demuestra que con el procedimiento propuesto  $\mathbf{R}_k^C$  converge a la matriz mas cercana a  $\mathbf{R}$  dentro del subespacio de intersección  $D \cap U$ .

---

<sup>3</sup>  $\|R\|_F = \sqrt{\sum_{i=1}^n \sum_{j=1}^n |\rho_{ij}|^2}$

### Simulación del procedimiento de ajuste y resultados

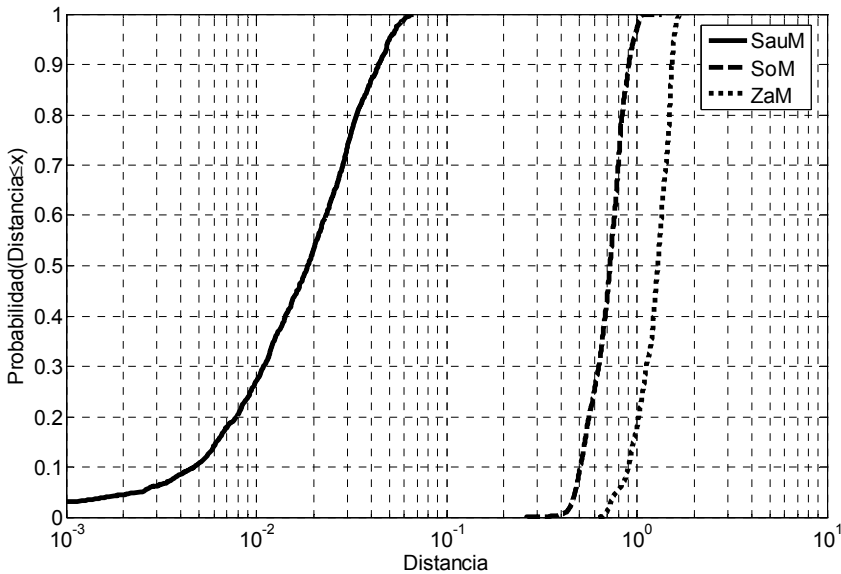
Se han implementado tres modelos de correlación cruzada del *shadowing* diferentes: el modelo de Saunders (SauM) [12], el modelo de Sorensen (SoM) [10] y el modelo de Zayana (ZaM) [14]. Se ha considerado un sistema celular de múltiples anillos con una celda central de radio 500m de manera que se ha ido aumentando paulatinamente el número de anillos celulares, desde 1 hasta 4. En cualquier caso, se define una rejilla de espaciado 10m cubriendo el área de simulación de manera que en cada posición de la rejilla se ha calculado la matriz de correlación cruzada del *shadowing* verificando si la matriz obtenida es efectivamente definida positiva. La Tabla 6.4 recoge la probabilidad, en tanto por uno, de que la matriz no sea definida positiva, dependiendo del modelo empleado y del número de anillos considerados.

	1-Anillo	2- Anillos	3- Anillos	4- Anillos
<i>Saunders</i>	0	0	0.43	0.91
<i>Sorensen</i>	0	0.97	1	1
<i>Zayana</i>	0.27	1	1	1

**Tabla 6.4 Probabilidad de que la matriz de correlación no sea definida positiva**

Se puede observar como al aumentar el número de estaciones base implicadas en el cálculo aumenta también la probabilidad de no ser definida positiva. El modelo de Saunders presenta el mejor comportamiento, seguramente porque su matriz de correlación no está definida a saltos como en el resto de modelos.

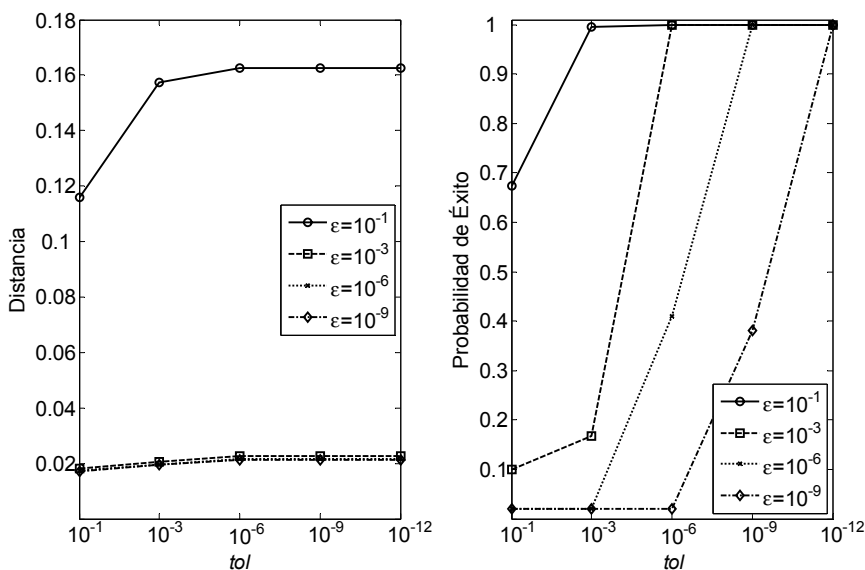
A partir de ahora se analiza con más profundidad el escenario de 3 anillos celulares. Se ha empleado el método AP particularizado propuesto en esta Tesis Doctoral para ajustar aquellas matrices que no sean definidas positivas. En la Figura 6.13 se muestra la función de distribución acumulativa (CDF) de la distancia, medida según la norma Frobenius, entre la matriz original y la matriz resultante. De nuevo el modelo de Saunders genera matrices de correlación que están más cerca de la matriz de correlación correcta en  $D \cap U$ .



**Figura 6.13 CDF de la distancia entre la matriz original y la ajustada**

A continuación, se investiga el efecto de  $\varepsilon$  y  $tol$  sobre las prestaciones del método AP en términos del éxito en la transformación, es decir, de la probabilidad de que la matriz resultante sea efectivamente simétrica, con matriz unidad y definida positiva. Además, se ha medido la distancia entre las matrices original y final, ya que se pretende encontrar siempre la matriz más cercana.

Inicialmente, se ha investigado la distancia entre las matrices. Para ello, a  $tol$  se le ha dado valores de  $10^{-1}$ ,  $10^{-3}$ ,  $10^{-6}$ ,  $10^{-9}$  y  $10^{-12}$ , mientras que  $\varepsilon$  se ha distribuido desde  $10^{-1}$  hasta  $10^{-9}$ . Los resultados mostrados en la Figura 6.14 muestran que no hay un efecto significativo en la distancia para valores de  $\varepsilon$  menores que  $10^{-6}$ . Sin embargo, tal y como era de esperar, un valor de  $\varepsilon$  igual a 0.1 hace que  $P_d(\mathbf{R}, \varepsilon)$  proyecte la matriz a una gran distancia ya que los valores propios sufren un cambio significativo, lo que justifica las distancias mayores experimentadas.



**Figura 6.14 Efecto de la tolerancia y  $\epsilon$  sobre las prestaciones del método AP**

Respecto a la probabilidad de éxito, se puede observar como ésta aumenta con valores menores que la tolerancia. Además, cuanto mayor sea  $\epsilon$  mayor será dicha probabilidad. Sin embargo, tal y como se discutió en el párrafo anterior, el uso de valores de  $\epsilon$  elevados conduce a la obtención de matrices más distantes de la original, lo que resulta indeseable. Por lo tanto, como  $\epsilon$  no debería ser mayor que  $10^{-6}$  la tolerancia debe seleccionarse menor de  $10^{-9}$  para garantizar la máxima probabilidad de éxito. Con esta configuración se consigue el ajuste óptimo de las matrices. Por último, es interesante resaltar que el número medio de iteraciones del método AP es de alrededor de 20 para los parámetros de  $\epsilon$  y tolerancia seleccionados.

### 6.2.6.2 Descripción del modelo

Una vez se dispone de un método para garantizar la aplicabilidad del método de generación de variables gaussianas correladas a partir de la descomposición de

Cholesky, se puede especificar el procedimiento de cálculo de las  $N$  matrices de *shadowing* bidimensionales con correlación completa.

En primer lugar, para cada punto geográfico  $\{x, y\}$  y para cada pareja de estaciones base del escenario  $\{i, j\}$ , se calcula, utilizando la expresión (6.15), el factor de correlación correspondiente,  $\rho_{ij}$ , de la matriz de correlación,  $\mathbf{R}_{x,y}$ . Esta matriz se podrá descomponer o no en función de si es definida positiva o no. Par comprobarlo se evalúa un menor<sup>4</sup> cualquiera de orden creciente desde 1 hasta  $N$  de la matriz  $\mathbf{R}_{x,y}$ . Si algún menor resulta ser negativo o cero, entonces la matriz no es definida positiva y, por tanto, no se puede descomponer utilizando el método de Cholesky. En este caso será necesario aplicar el procedimiento de ajuste basado en el método AP descrito.

Una vez se dispone de una matriz  $\mathbf{R}_{x,y}$  definida positiva se aplica la descomposición de Cholesky obteniendo la matriz triangular inferior  $\mathbf{C}_{x,y}$ . Posteriormente se genera un vector aleatorio gaussiano de media 0 y desviación típica  $\sigma$ ,  $\mathbf{r}$ , de  $N$  componentes, una para cada estación base considerada. Las  $N$  componentes de las  $N$  matrices de *shadowing* con correlación cruzada se calcularán a partir de la expresión (6.18).

Este procedimiento se repetirá para cada punto geográfico considerado, obteniéndose finalmente las matrices deseadas con correlación cruzada. Además, para introducir la correlación espacial se filtrarán las matrices con el filtro bidimensional diseñado anteriormente. La Figura 6.15 resume este procedimiento.

## Validación del modelo

Para validar el modelo, en primer lugar se debe comprobar que realmente al generar las muestras de *shadowing* conforme a (6.18) éstas cumplen la correlación deseada.

---

<sup>4</sup> Se define como menor de orden  $p$  de una matriz genérica  $\mathbf{A}$  al determinante que resulta de eliminar ciertas filas y columnas hasta quedar una matriz cuadrada de orden  $p$ . Es decir, al determinante de cualquier submatriz cuadrada de  $\mathbf{A}$ .

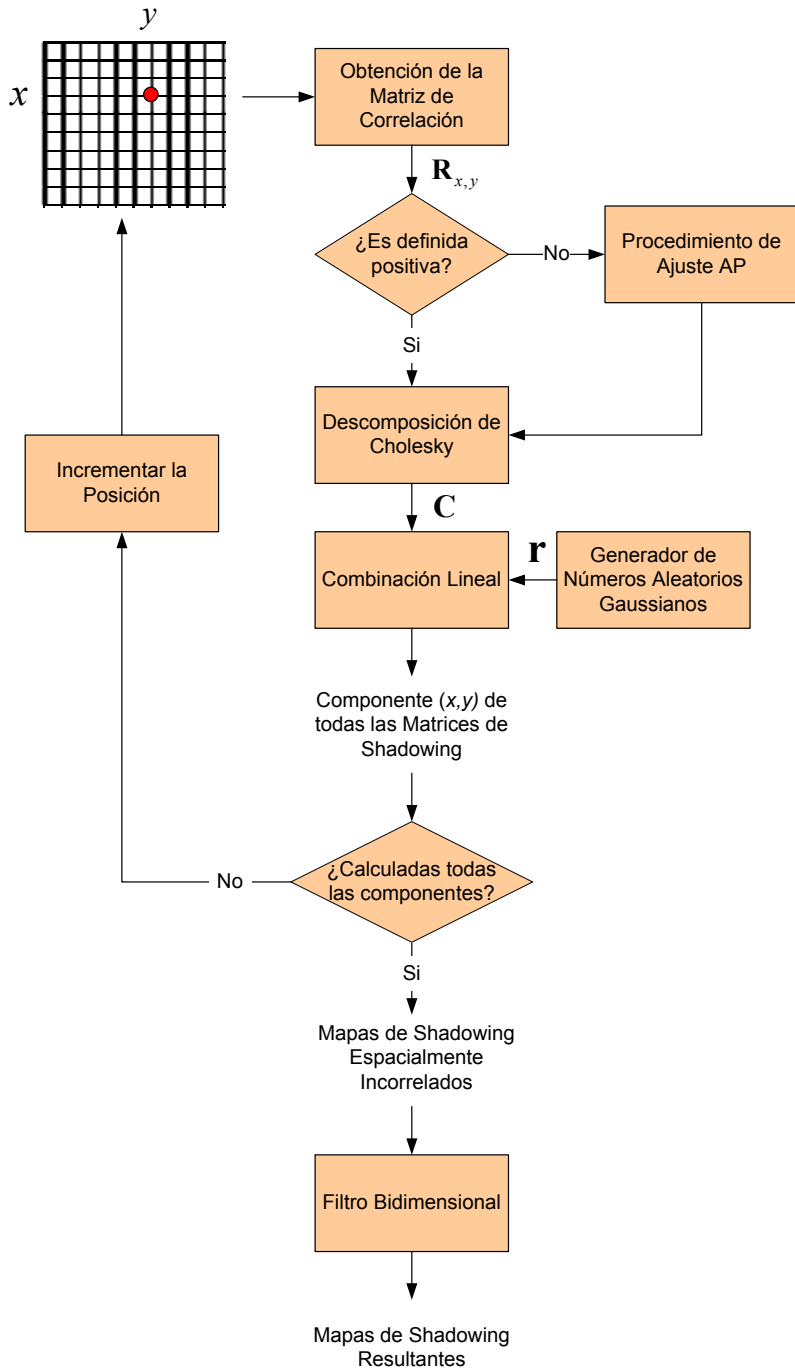
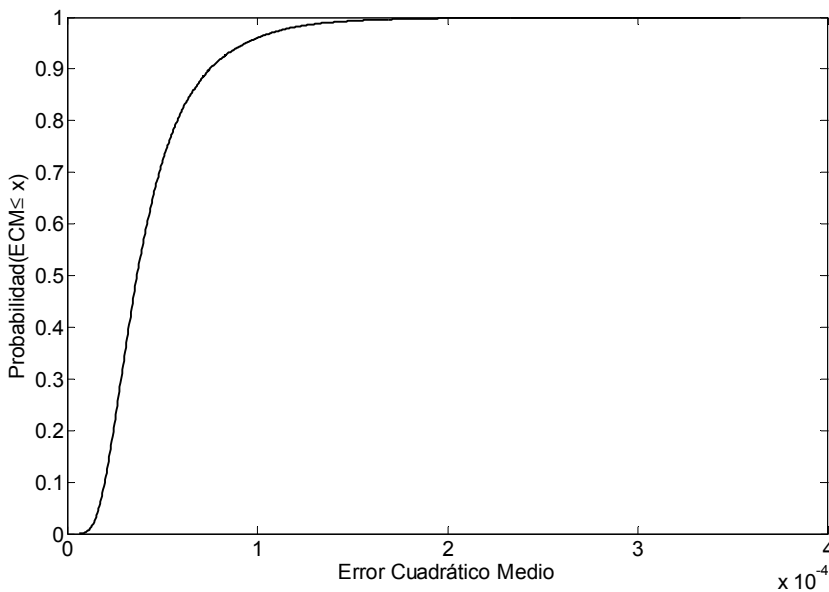


Figura 6.15 Proceso de generación de los mapas de *shadowing* con correlación cruzada

Para ello se ha seleccionado el mismo escenario de simulación descrito en el apartado 6.2.3, donde se consideran 7 estaciones base ( $N=7$ ). Para todos los puntos de la rejilla espacial se ha generado, utilizando el procedimiento descrito anteriormente, una secuencia suficientemente larga (100.000 muestras) de las 7 variables aleatorias correspondientes a las 7 estaciones base, de manera que se pueda calcular la correlación entre las variables aleatorias. Construida la matriz de correlación experimental se ha comparado con la matriz de correlación obtenida mediante el modelo de Saunders y se ha calculado el promedio del error cuadrático medio cometido considerando todos los puntos de la matriz de correlación. La Figura 6.16 muestra la CDF del error cometido.

Como se puede observar, este error es prácticamente despreciable (la media resulta ser de  $4.3 \cdot 10^{-5}$ ), lo que confirma la validez del método utilizado para generar las muestras con correlación cruzada.



**Figura 6.16 CDF del error cuadrático medio cometido en el generación de las muestras correladas para todos los puntos geográficos considerados**



Por último, se ha repetido el estudio realizado en los modelos anteriores y se ha analizado la correlación cruzada entre los distintos mapas de *shadowing* obtenidos mediante este método. La Tabla 6.5 recoge estos valores de correlación.

A diferencia del modelo descrito en el apartado 6.2.5, con el modelado completo se consigue una mayor correlación entre todos los mapas.

	Mapa 1	Mapa 2	Mapa 3	Mapa 4	Mapa 5	Mapa 6	Mapa 7
Mapa 1	1.0000	0.5249	0.5445	0.5449	0.5567	0.5326	0.5227
Mapa 2	0.5249	1.0000	0.6256	0.4811	0.4531	0.4759	0.6197
Mapa 3	0.5445	0.6256	1.0000	0.6451	0.5001	0.4442	0.4797
Mapa 4	0.5449	0.4811	0.6451	1.0000	0.6562	0.5133	0.4564
Mapa 5	0.5567	0.4531	0.5001	0.6562	1.0000	0.6455	0.4986
Mapa 6	0.5326	0.4759	0.4442	0.5133	0.6455	1.0000	0.6342
Mapa 7	0.5227	0.6197	0.4797	0.4564	0.4986	0.6342	1.0000

Tabla 6.5 Correlación cruzada entre las muestras de los mapas de *shadowing*

## 6.3 Influencia de los Distintos Modelos

### 6.3.1 Introducción

Hasta ahora en este capítulo de la Tesis Doctoral se han propuesto un conjunto de modelos de *shadowing* más o menos complejos que se pueden clasificar en tres grupos: los que no incluyen correlación de ningún tipo (modelo 1), los que modelan únicamente la correlación espacial (modelos 2 y 3) y los que incluyen además la correlación cruzada (modelos 4,5 y 6).

El efecto del modelado del *shadowing* dependerá en gran medida de si se está estudiando un proceso altamente dependiente del canal, como por ejemplo sistemas basados en la adaptación al canal (*link adaptation*) como GSM/GPRS y HSDPA, o si se está analizando un proceso dependiente de la macro-diversidad entre estaciones base, como los procesos de *handover*.

En esta sección se va a evaluar el efecto de los distintos modelos en los resultados obtenidos en el escenario de simulación utilizado en el Capítulo 4. Es importante

recordar que se está trabajando con un sistema no basado en *link adaptation* y donde no existe ningún tipo de macro-diversidad puesto que no se han implementado técnicas de *soft handover*. Aunque a priori parezca que en este escenario el modelado del *shadowing* no tendrá ningún efecto en los resultados medios, los resultados obtenidos confirman lo contrario. Para entender este hecho hay que considerar los siguientes factores.

1. El sistema de asignación de recursos (*scheduler*) ha de maximizar el uso de los recursos disponibles, es decir, potencia y códigos o tasa binaria. Así pues todos los algoritmos se basan en un conocimiento, al menos estimativo, de las condiciones de propagación de cada usuario. A partir de esta información se realiza una determinada asignación de recursos que se mantendrá estable hasta la siguiente ejecución del *scheduler*. Durante el periodo de tiempo transcurrido entre dos ejecuciones del *scheduler*, las condiciones del canal de los usuarios pueden variar fruto de su movilidad. Cuanto menor sea la correlación del *shadowing* mayor será la variabilidad experimentada y, por tanto, más errores se cometerán en la gestión de recursos.
2. El sistema UMTS cuenta con un control de potencia rápido que se ejecuta 1500 veces por segundo. Este procedimiento es capaz de seguir las variaciones a largo plazo del canal siempre que estas variaciones no sean muy grandes. Sin embargo, si las condiciones del canal cambian muy bruscamente o si el sistema entra en congestión, el control de potencia no podrá seguir las variaciones del canal y, por tanto, no se satisfarán las condiciones de relación señal a ruido necesarias para una correcta recepción. En estos casos se considera que el usuario se encuentra en un estado de mala calidad u *outage* en el que no puede recibir correctamente los paquetes transmitidos. Si los paquetes no se reciben correctamente se quedan almacenados en los *buffers* del Nodo B lo que aumentará las colas en la transmisión y por tanto aumentará los retardos.
3. La variabilidad de la relación señal a ruido no sólo depende de la correlación de la señal recibida con el espacio sino también de la variabilidad de la

interferencia. Si la señal interferente recibida de las celdas vecinas está más correlada entonces la relación señal a ruido será más estable y el sistema se comportará óptimamente.

Todo esto hace prever que no solamente la correlación espacial, sino también la correlación cruzada, tendrán un efecto sobre el simulador dinámico de sistema implementado. En la siguiente sección se evalúa hasta qué punto los distintos modelos afectan a la evaluación de las prestaciones del sistema.

### **6.3.2 Análisis de los resultados y discusión**

En las simulaciones llevadas a cabo se ha considerado un periodo de actualización de los recursos asignados de 0.5 s, una velocidad constante de los usuarios de 50 km/h y un radio celular de 500m. Para cada uno de los modelos de *shadowing* simulados se ha emulado el funcionamiento de la red durante una hora, replicando cada simulación con 10 semillas aleatorias distintas, tanto para los modelos de tráfico y movilidad como para la generación de los mapas o valores de *shadowing*, asegurando así la validez estadística de los resultados. Se ha considerado un escenario UMTS convencional siendo el gestor de recursos implementado el DBR, ya que éste se ha mostrado como el mecanismo de referencia más eficaz en la maximización del uso de los recursos (ver apartado 4.5.3). En el sistema se han introducido 70 usuarios web, suficientes como para llevar al sistema a un estado de carga razonablemente alto. No se han tenido en cuenta servicios de tiempo real puesto que estos necesitan una actualización de los recursos asignados mucho más rápida, lo que implicaría un cambio muy reducido en las condiciones de propagación. Los parámetros considerados en los modelos de *shadowing* se resumen en la Tabla 6.6.

Número de Nodos-B	7
Tamaño del mapa en x	1000 m
Tamaño del mapa en y	1000 m
Número de puntos por mapa	200 x 200
Distancia entre puntos	5 m
Distancia de decorrelación	20 m
Correlación entre estaciones base – Modelo 4	0.5
$\varepsilon$ – Modelo 6	$10^{-6}$
$tol$ – Modelo 6	$10^{-9}$

**Tabla 6.6 Parámetros de los mapas de *shadowing* generados**

La Tabla 6.7 recoge, para los seis modelos de *shadowing* implementados, un conjunto de parámetros de medida de las prestaciones del sistema. Todos esos parámetros, calculados en valor medio, se describen a continuación.

- $P_{Out}$  : probabilidad de *outage* media, es decir, probabilidad de tiempo en el que el usuario experimenta unas condiciones de propagación que no le permiten recibir correctamente la información transmitida.
- $P_T$  : representa la potencia media consumida en la celda central bajo estudio. Se mide en Vatios.
- $N_{User Act}$  : es el número medio de usuarios con información pendiente de transmisión.
- $N_{User Serv}$  : número medio de usuarios a los que se les asigna simultáneamente recursos.
- $R_T$  : representa la tasa binaria media total asignada tras el proceso de asignación de recursos. Se mide en kbps.
- $R_{TU}$  : parámetro similar al anterior con la diferencia de que éste recoge la tasa binaria efectiva incluyendo el efecto de pérdidas del canal. Se mide en kbps.
- $R_{User}$  : tasa binaria media asignada por usuario. Se mide en kbps.
- $D_{Obj}$  : tiempo medio transcurrido desde la llegada al Nodo B de un objeto web dirigido a un usuario, hasta su recepción completa en el destinatario. Se mide en segundos.

- $D_{Pag}$  : igual que el anterior pero en este caso recoge el retardo experimentado en la transmisión de la página web completa. Se mide en segundos.
- $P_{Exc D}$  : probabilidad media de exceder el retardo máximo de transmisión de una página web.

	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5	Modelo 6
$P_{Out}$ (%)	1.53	1.75	1.12	0.421	0.369	0.346
$P_T$ (W)	13.82	12.85	12.83	11.53	11.71	11.42
$N_{User Act}$	43.68	14.11	13.95	10.63	10.64	10.38
$N_{User Serv}$	30.45	13.56	13.38	10.58	10.61	10.35
$R_T$ (kbps)	945.25	1051.24	1043.87	1089.87	1088.41	1095.74
$R_{TU}$ (kbps)	930.60	1026.42	1028.48	1083.49	1083.28	1090.24
$R_{User}$ (kbps)	31.03	77.58	78.39	103.66	103.21	106.38
$D_{Obj}$ (s)	14.49	1.21	1.22	0.875	0.877	0.851
$D_{Pag}$ (s)	83.85	7.12	7.17	5.15	5.16	5.02
$P_{Exc D}$ (%)	36.06	4.38	4.47	2.45	2.46	2.33

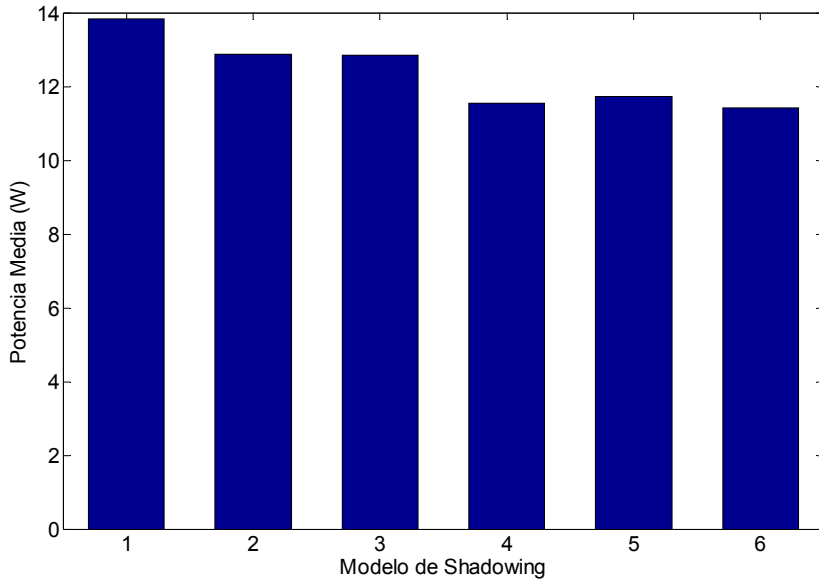
**Tabla 6.7 Resultados medios obtenidos**

De todos los parámetros analizados el que más información da sobre el efecto de la correlación es sin duda la probabilidad de *outage*. De manera intuitiva se espera que a medida que se aumente la correlación se reduzca la probabilidad de *outage* ya que el canal será más estable y el control de potencia será capaz de seguir los cambios experimentados en la relación señal a ruido. Así pues, la probabilidad de *outage* más baja se experimenta para el modelo de correlación más completo, el modelo 6, bidimensional y con correlación cruzada completa, en el que tanto interferentes como la señal útil están correladas conforme al modelo de Saunders. El modelo 5, que no incluye correlación entre las interferentes, aumenta su probabilidad de *outage* debido a la mayor variabilidad de la SIR. El modelo 4 de correlación cruzada constante, a pesar de tener una correlación mayor que el modelo 5 entre las interferentes, presenta una mayor probabilidad de *outage*. Los modelos que no incluyen la correlación cruzada aumentan considerablemente la  $P_{Out}$ , teniendo el modelo bidimensional una menor probabilidad de *outage* respecto al modelo lineal. Destaca el hecho de que el modelo sin correlación tenga menos  $P_{Out}$  que el modelo con correlación lineal. Este

comportamiento se explica porque con el modelo sin correlación se asigna una menor tasa binaria por usuario (ver resultados de  $R_{User}$ ). Las tasas binarias más bajas requieren un consumo de potencia menor y son más robustas a los cambios de SIR, lo que justifica el menor valor de probabilidad de *outage*.

Analizando la potencia media de forma gráfica (Figura 6.17) se puede observar claramente cómo los tres grupos identificados anteriormente presentan unos valores de potencia bastante similares. Así pues, no considerar ningún tipo de correlación supone estimar que el sistema va a hacer un 18% más de consumo de potencia que el modelo de correlación más completo implementado, mientras que considerar únicamente correlación espacial supone estimar un consumo un 9.5% mayor de potencia.

Dentro de los modelos que sólo consideran correlación espacial, es decir, los modelos 2 y 3, apenas se aprecia ninguna diferencia destacable entre ambos, lo que podría llevar a considerar que la correlación bidimensional no tiene un efecto real sobre las prestaciones del sistema. Aunque el primer análisis realizado entre las probabilidades de *outage* ya demostró que el modelo bidimensional aumenta la correlación, para refrendar la importancia de un modelado bidimensional, se ha evaluado las prestaciones de los modelos 2, 3 y 4 en un escenario GPRS basado en *link adaptation*, mucho más sensible a las variaciones del canal. Para realizar esta evaluación se utilizó una herramienta de simulación prestada por la Universidad Miguel Hernández y se evaluó el comportamiento de los tres modelos [1]. Para más detalles sobre el escenario de simulación empleado consultar [1].



**Figura 6.17 Potencia media en función del modelo de *shadowing* empleado**

La Figura 6.18 muestra la CDF de la tasa media asignada por el sistema utilizando el modelo 2 de *shadowing* lineal, representado con línea continua, el modelo 3 bidimensional sin correlación, dibujado con una línea de puntos, y el modelo 4 bidimensional con correlación constante, representado con una línea discontinua. Se puede observar claramente cómo el modelo bidimensional sin correlación cruzada aumenta considerablemente la tasa binaria asignada lo que implica que el *link adaptation* está funcionando mejor debido a la mayor estabilidad del canal. También se puede ver cómo la correlación cruzada aumenta aún más dicho valor de tasas binaria. Este aumento en la tasa binaria media se corresponde asimismo con una disminución del BLER, tal y como se observa en la Figura 6.19, es decir, el modelado bidimensional no sólo aumenta la tasa de transmisión, sino que además mejora la calidad de servicio de los usuarios.

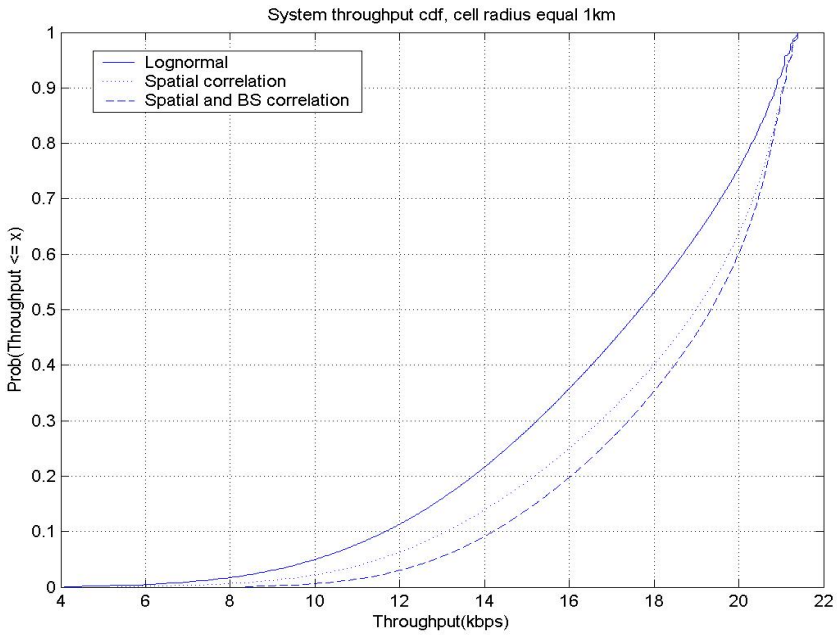


Figura 6.18 CDF de la tasa media del sistema actualizando el LA cada 20 ms [1]

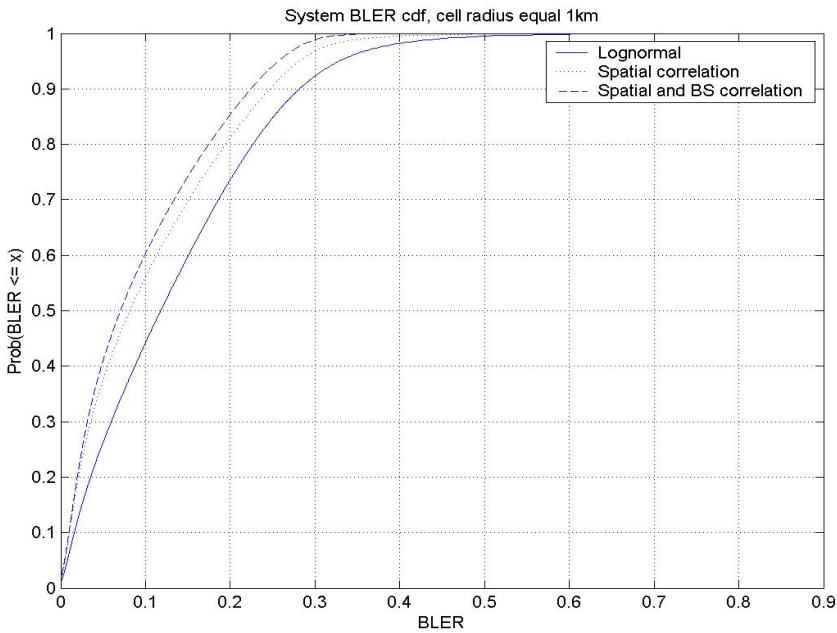
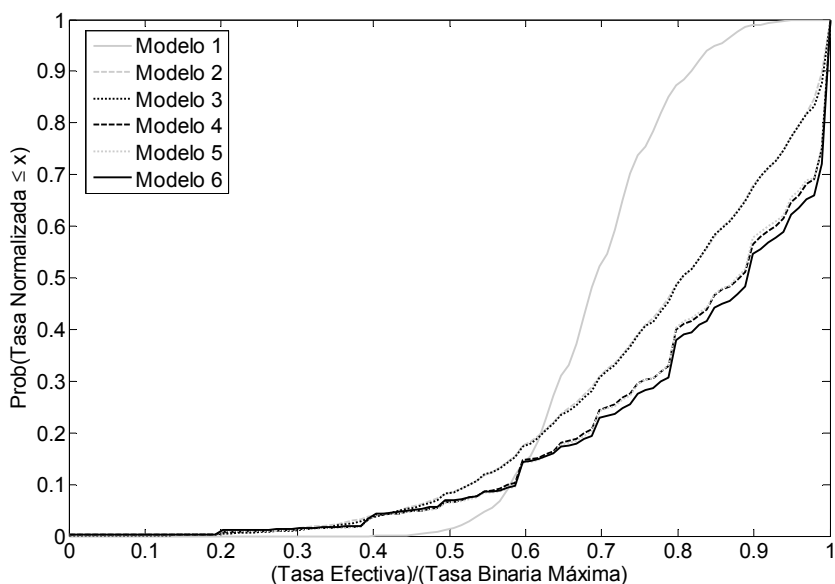


Figura 6.19 CDF del BLER actualizando el LA cada 20 ms [1]

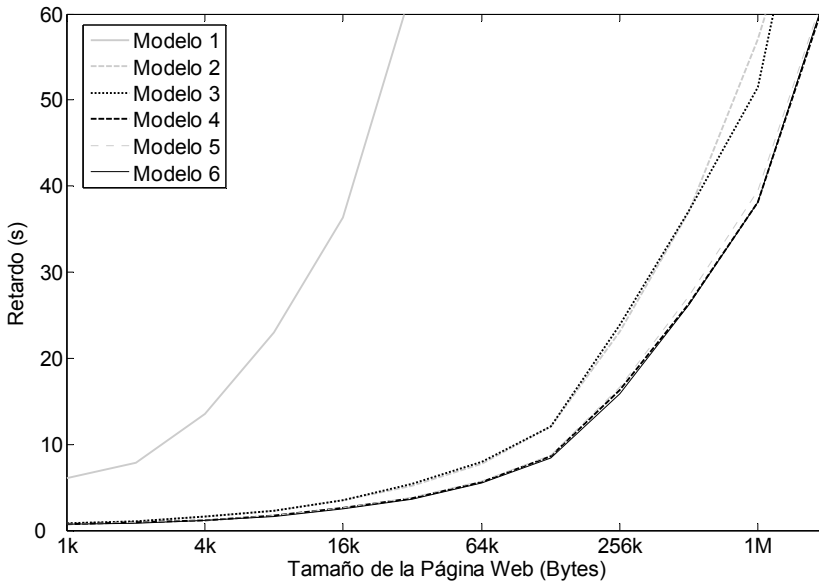


Volviendo de nuevo a los resultados obtenidos con el simulador de sistema utilizado en esta Tesis Doctoral, se puede observar en la Tabla 6.7 que la disminución de la probabilidad de *outage* trae consigo una sustancial mejora de los recursos asignados, lo cual implica un menor número de usuarios activos y una más que notable reducción en el retardo medio de objeto y de página. De nuevo, los resultados diferencian claramente tres grupos: el modelo sin correlación, los modelos sin correlación cruzada y los modelos que incluyen ambas correlaciones, espacial y cruzada. Además, dentro de cada grupo las diferencias observadas en los resultados son muy poco significativas aunque cumplen la tendencia observada.

La Figura 6.20 muestra la CDF de la tasa efectiva asignada en la celda normalizada respecto a la tasa binaria máxima. De nuevo, se puede observar como los modelos que no consideran la correlación cruzada y la autocorrelación ofrecen un análisis pesimista de los resultados de la red. Además, el modelo de *shadowing* más aproximado a la realidad, es decir, el bidimensional con correlación cruzada completa, ofrece los mejores resultados en cuanto a tasa binaria y retardo en la transmisión (Figura 6.21).



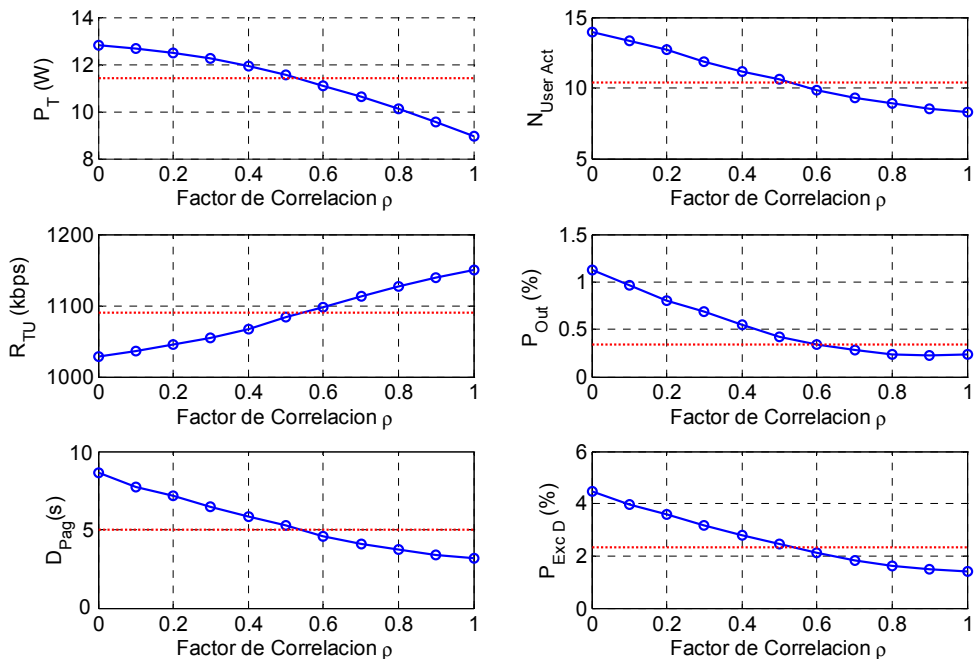
**Figura 6.20 CDF de la tasa normalizada del sistema**



**Figura 6.21 Retardo medio en función del tamaño de página web**

Tras analizar estos resultados se puede concluir que, en primer lugar, es necesario introducir la correlación bidimensional tal y como se mostró en el estudio realizado en [1]. Además, es también igualmente importante incluir los dos efectos de correlación característicos del *shadowing*, es decir, correlación espacial y correlación cruzada, ya que sin su modelado los resultados obtenidos son claramente pesimistas. Por último, es importante destacar que, al menos en las simulaciones de sistema sin macro-diversidad, un modelo más simple como es el 4 puede obtener resultados muy similares a modelos mucho más complejos como son el 5 y el 6. Al fin y al cabo, el modelo 4 se puede aproximar a los modelos más realistas sin más que ajustar convenientemente el factor de correlación constante utilizado.

Para corroborar esta idea se ha estudiado el efecto del factor de correlación cruzada en los resultados obtenidos con el modelo 4. La Figura 6.22 muestra el comportamiento del sistema para distintos valores del factor de correlación. Además, una línea roja representa el valor obtenido con el modelo más completo, el modelo 6.



**Figura 6.22 Comportamiento medio del sistema utilizando el modelo 4 de shadowing con distintos factores de correlación**

Lo primero que se observa es que, claramente, la correlación cruzada tiene un efecto muy importante en los resultados obtenidos, de manera que una mayor correlación mejora considerablemente las prestaciones del sistema, reduciendo el consumo de potencia, aumentando la tasa binaria asignada y mejorando la calidad de servicio de los usuarios. La única cuestión es ver si eligiendo adecuadamente un valor del factor de correlación sería posible emular totalmente el comportamiento del shadowing, obviando el uso de modelos más complejos para simulaciones de sistema. La Tabla 6.8 recoge el valor del factor de correlación que sería necesario introducir en el modelo 4 para obtener exactamente los mismos resultados que el modelo más realista, el modelo 6. Los resultados muestran cómo resulta imposible reproducir exactamente el funcionamiento del modelo 6, pero con un valor de  $\rho = 0.55$  sería factible obtener unas prestaciones muy similares a las deseadas.

	$P_T$ (W)	$N_{User Act}$	$R_{TU}$ (kbps)	$P_{Out}$ (%)	$D_{Pag}$ (s)	$P_{Exc D}$ (%)
$\rho$	0.527	0.533	0.548	0.594	0.531	0.535

**Tabla 6.8 Factor de correlación óptimo para cada parámetro**

## 6.4 Conclusiones

En este capítulo se ha analizado en profundidad el comportamiento del fenómeno de desvanecimiento lento de la señal radio, también denominado *shadowing*. En la literatura se puede encontrar distintos modelos de correlación espacial y de correlación cruzada entre estaciones base, modelos extraídos del análisis de distintas campañas de medidas. A partir de esta caracterización, en este capítulo se han propuesto un conjunto de modelos de *shadowing* que permiten introducir el efecto del desvanecimiento lento de la señal en simulaciones de Monte Carlo. Desde modelos sencillos hasta otros mucho más realistas y complejos, se ha validado el funcionamiento de cada uno de ellos, remarcando sus ventajas e inconvenientes.

Una vez formulados por orden de creciente realismo todos los modelos, se ha analizado el efecto que su uso tiene sobre la evaluación de las prestaciones de un sistema de comunicaciones móviles. Asumiendo que el último modelo descrito, el que más efectos integra, es el que refleja con más fidelidad la realidad, se ha comprobado que los otros modelos se aproximan más a menos a éste en función de si implementan o no los dos efectos de correlación más destacados del *shadowing*: la autocorrelación espacial y la correlación cruzada. Sin embargo, dentro de los modelos que incluyen ambas correlaciones, las diferencias apenas son significativas con lo que el modelo más simple es el más adecuado. Es importante destacar que esta conclusión está sujeta al tipo de estudio realizado ya que, en la evaluación llevada a cabo, la tecnología de transmisión no es sensible a las variaciones de canal y, además, no incluye ningún tipo de técnica basada en macro-diversidad. Sería por tanto necesario realizar otros estudios similares de evaluación de prestaciones sobre distintos escenarios, preferentemente utilizando tecnologías basadas en *link adaptation*, de manera que se pudiera evaluar si las conclusiones alcanzadas en este capítulo son igualmente extrapolables a otros

escenarios de simulación. Este trabajo queda fuera de los objetivos de la presente Tesis Doctoral y se enmarca dentro de las líneas futuras de investigación de la Tesis.

## Referencias

- [1] V. Graziano, "Propagation Correlations at 900 MHz", IEEE Trans. On *Vehicular Technology*, Vol. 27, No. 4, November 1978.
- [2] D. Gudmundson, "Correlation Model For Shadow Fading in Mobile Radio Systems", IEE Electronic Letters, Vol. 27, No. 23, November 1991.
- [3] M. Barbiroli, C. Carciofi, G. Falciasecca, M. Frullone and P. Grazioso, "A Measurement-based Methodology for the Determination of Validity Domains of Prediction Models in Urban Environment", IEEE Transactions on *Vehicular Technology*, Vol. 49, No. 5, pp. 1508-1515, Sept. 2000.
- [4] R. Fraile, J. F. Monserrat, J. Gozávez and N. Cardona, "Mobile Radio Bi-dimensional Large-Scale Fading Modelling with Site-to-Site Cross-Correlation", Wiley InterScience European Transactions On Telecommunications (in press), DOI 10.1002/ett.1179, 2007.
- [5] J. Reig, J. Larrey and N. Cardona, "Power allocation strategies in downlink WCDMA systems using soft *handover* and SSDT power control techniques", in Proceedings of the IEEE *Vehicular Technology* Conference, Vol. 2, pp. 817-821, Sept. 2004.
- [6] H. Kim and Y. Han, "Enhanced Correlated *Shadowing* Generation in Channel Simulation", IEEE Communications Letters, Vol. 6, pp 279-281, 2002.
- [7] J. Gozávez y J. Dunlop, "System Performance and Adaptive Configuration of Link Adaptation Techniques in Packet-Switched Cellular Radio Networks", Computer Networks Journal (Elsevier), pp. 404-426, 2006.
- [8] E. Damosso (ed.), "Digital Mobile Radio Towards Future Generation Systems", COST 231 Final Report, pp. 115-207, 1998.
- [9] A. Viterbi, E. Zehavi, "Other-cell Interference in Cellular Power-controlled CDMA". IEEE Transactions on Communications, Vol. 42, pp. 1501-1504, 1994.
- [10] T. B. Sorensen, "Slow fading Cross-correlation Against Azimuth Separation of Base Stations", IEE Electronic Letters, Vol. 35, No. 2, January 1999.
- [11] J. Van Rees, "Cochannel measurements for interference limited small cell Planning", Arch. Elek. Ubertragung, Vol. 41, pp. 318-320, 1987.
- [12] S.R. Saunders, "Antennas and Propagation for Wireless Communication Systems", John Wiley & Son, Chichester, 1999.

- [13] T. Klingenbrunn, P. Mogensen, “Modelling Cross-Correlated *Shadowing* in Network Simulations”, in Proc. of the IEEE *Vehicular* Technology Conference, Vol. 3, September 1999.
- [14] K. Zayana, B. Guisnet, “Measurements and Modelisation of *Shadowing* Cross-Correlations Between Two Base-Stations”, in Proc. of the IEEE International Conference on Universal Personal Communications, Vol. 1, October 1998.
- [15] S. Haykin, Adaptive Filter Theory (fourth edition). Prentice Hall, 2002.
- [16] F. Graziosi, M. Pratesi, M. Ruggieri, F. Santucci, “A Multicell Model of *handover* Initiation in Mobile Cellular Networks”, IEEE Transactions on *Vehicular* Technology, Vol. 48, No. 3, May 1999.
- [17] J. P. Boyle, R. L. Dykstra, “A method for finding projections onto the intersection of convex sets in Hilbert spaces”, Lecture Notes in Statistics, Vol. 37, 1986.
- [18] R. Fraile, J. Gozalvez, O. Lazaro, J. F. Monserrat and N Cardona, “Effect of a Two Dimensional *Shadowing* Model on System Level Performance Evaluation”, in Proc. Wireless Personal and Mobile Communications, 2004.





# Capítulo 7

## *Conclusiones y Líneas Futuras de Investigación*

*Sólo sé que no sé nada.*

*Sócrates*

### **7.1 Conclusiones**

Esta Tesis Doctoral tiene su punto de partida en el proyecto ministerial CYCIT TIC2002-02678 y su continuación en el seno del proyecto TEC2005-08211-C02. En el primero de estos proyectos se propuso el uso coordinado de un conjunto de antenas distribuidas sobre el área de cobertura celular para la tecnología móvil UMTS. Esta nueva arquitectura radio, denominada Multinodo B, planteaba una serie de interrogantes de partida que se plantearon en el anteproyecto de Tesis y a los que esta Tesis Doctoral ha tratado de dar respuesta. A continuación, se contestará todas estas cuestiones a partir de las conclusiones alcanzadas en esta Tesis Doctoral.

***¿La arquitectura de protocolos UMTS permite ampliar el concepto clásico de celda UMTS añadiendo antenas distribuidas?***

Uno de los primeros objetivos planteados en el proyecto era garantizar que la nueva arquitectura radio Multinodo B modificase de manera mínima la arquitectura de red de acceso UMTS estándar, de forma que fuera compatible su despliegue con el uso de celdas convencionales. Si esto fuera posible se reduciría el coste de su implementación, favoreciendo la transferencia tecnológica. En esta Tesis Doctoral se han estudiado en

profundidad las especificaciones del 3GPP respecto a la arquitectura radio y su interfaz de protocolos. A partir de este análisis se han propuesto dos arquitecturas de protocolos distintas para el MNB: la primera, denominada MNB-activo, es más compleja pero totalmente flexible, mientras que la segunda, MNB-pasivo, resulta de muy fácil implementación pero únicamente decide sobre qué antenas se transmite la información enviada por el RNC. En cualquier caso, la implementación de ambas arquitecturas sólo implica un pequeño ajuste de los enlaces ATM del RNC y un nuevo diseño *software* del Nodo-B convencional, añadiéndole mayor capacidad de procesado.

Además, toda la información que necesita el sistema MNB para realizar la asignación óptima de recursos se puede obtener a partir de los informes de estado que el terminal envía periódicamente al sistema UMTS. El procedimiento seguido para procesar toda la información recibida del usuario y obtener los parámetros requeridos también se ha analizado en esta Tesis Doctoral, demostrando su sencillez y posibilidades de implementación.

### ***¿Cómo analizar las prestaciones de la nueva arquitectura radio planteada?***

Para analizar las prestaciones de la nueva arquitectura radio se han desarrollado dos simuladores: un simulador de enlace destinado a evaluar la ganancia por diversidad espacial y un simulador de sistema encargado de determinar cuál es la mejora introducida por el MNB en términos de potencia consumida y calidad experimentada por los usuarios. Además, se ha desarrollado la interfaz entre ambos simuladores para permitir un modelado correcto del comportamiento del MNB.

### ***¿Es posible aprovechar el despliegue jerárquico basado en micro-celdas para dotar a los terminales móviles de una elevada ganancia por diversidad espacial en transmisión?***

Dada la creciente demanda de servicios multimedia en movilidad, las operadoras de telefonía móvil celular necesitan aumentar constantemente la capacidad de sus sistemas 3G. Aunque la red esté perfectamente optimizada, llega un punto en que la única

solución posible es incrementar el número de transmisores. Sin embargo, utilizar más emplazamientos macro-celulares no garantiza siempre la mayor capacidad del sistema, por los grandes problemas de interferencia mutua que sufren dos macro-celdas si se encuentran demasiado próximas. La única solución que se puede adoptar es hacer uso de redes jerárquicas, añadiendo micro-celdas que utilicen una frecuencia de portadora distinta.

En esta Tesis Doctoral se plantea la posibilidad de aprovechar el despliegue de las redes jerárquicas para mejorar las prestaciones de las macro-celdas ya activas. Así, las nuevas micro-celdas, además de dar servicio a los usuarios estáticos que se encuentran en su área de cobertura, pueden transmitir en otra portadora a los usuarios admitidos en el MNB a los que, a priori, debería dar servicio la macro-celda. Gracias a la transmisión conjunta desde macro y micro-celdas se añade una importante diversidad espacial en transmisión.

Considerando la transmisión simultánea de más de una antena, se ha evaluado si efectivamente el MNB es capaz de aportar una diversidad espacial en transmisión y hasta qué punto influye esta diversidad adicional en las prestaciones del sistema UMTS. Con este objetivo, se ha implementado un simulador de enlace en sentido descendente de la red de acceso radio UTRAN, en modo FDD. Dicha implementación se ha realizado siguiendo el estándar e intentando alcanzar un equilibrio entre simplicidad y validez de los resultados.

El simulador desarrollado, implementado en C++, constituye una herramienta extremadamente flexible y potente que podrá ser empleada en numerosos estudios de investigación relacionados con el desarrollo y evaluación de nuevas técnicas de capa física destinadas a mejorar las prestaciones de los sistemas de comunicaciones de tercera generación.

Tras un primer análisis de las prestaciones del MNB, sin realizar ningún procesado previo adicional de la señal transmitida, se ha demostrado que el MNB es capaz de ofrecer una mejora importante en el comportamiento del sistema, siempre que las señales recibidas no se solapen en el tiempo, ya que al solaparse se pierde la diversidad

deseada. La ganancia experimentada depende del escenario, del número de *fingers* del receptor *Rake* y de la diferencia relativa de potencia de las contribuciones recibidas.

A modo de ejemplo, para el servicio de voz a 12.2 kbps y considerando la transmisión desde dos antenas macro-celulares en un escenario vehicular, la ganancia puede alcanzar valores de hasta 1.61 dB. Aunque desde el punto de vista radio esta ganancia pueda parecer pequeña, a efectos de gestión de recursos es un valor muy significativo, sobre todo teniendo en cuenta que se está considerando una velocidad del terminal de 50 km/h.

Como se detallará más adelante, en esta Tesis Doctoral también se ha solucionado el problema del posible solapamiento de las señales transmitidas utilizando para ello mecanismos de diversidad en transmisión en bucle cerrado.

### ***¿De qué manera habría que modificar la transmisión desde el MNB y las técnicas de estimación de canal para optimizar el funcionamiento de la nueva arquitectura radio?***

El diseño del MNB permite, en el caso de que las dos señales no estén solapadas, sumar coherentemente las señales recibidas, ya que en recepción se pueden discriminar las señales, ajustándolas en fase por separado. Sin embargo, cuando las señales llegan simultáneamente, como se utiliza el mismo código de aleatorización, el receptor es incapaz de separar ambas contribuciones, por lo que no se puede alcanzar ninguna ganancia.

Como las señales transmitidas por las antenas del MNB están a priori sincronizadas para reducir la ISI, es necesario desarrollar algún tipo de técnica de procesado en *array* que permita aprovechar al máximo las capacidades de combinación coherente que ofrece el MNB, independientemente de que las señales estén o no solapadas en el tiempo. Con este objetivo en mente, en esta Tesis Doctoral se ha propuesto el uso de una técnica de procesado en *array* convencional de maximización de potencia junto con una modificación de la técnica utilizada de estimación de canal, que empleará los bits piloto enviados por el canal CPICH.

Utilizando conjuntamente la técnica de procesado en *array* y el mecanismo adaptado de estimación de canal, se ha evaluado el comportamiento del MNB cuando las señales se reciben temporalmente solapadas. La mejora conseguida por los nuevos cambios es más que notable. Los resultados obtenidos muestran que, gracias a los nuevos mecanismos, en cualquier caso el MNB es capaz de combinar coherentemente las señales recibidas desde ambas antenas.

### ***¿Es posible optimizar el funcionamiento de la arquitectura radio UMTS?***

Se ha evaluado el valor concreto de la ganancia aportada por la diversidad espacial adicional del MNB para un conjunto de servicios y escenarios específicos. Esta ganancia depende de la diferencia relativa de potencia entre las señales recibidas pero, en cualquier caso, se puede afirmar que el MNB es capaz de mejorar notablemente el comportamiento del sistema, sobre todo en entornos micro-celulares a los que se complementa con una transmisión macro-celular. En este caso la reducción en la relación  $E_b/N_0$  conseguida puede llegar a alcanzar hasta los 3.5 dB aproximadamente.

Considerando que en ningún caso el uso del MNB perjudica el funcionamiento del sistema UMTS convencional, se puede afirmar que la nueva arquitectura radio es capaz de optimizar las prestaciones de UMTS.

### ***¿Sería posible diseñar una nueva arquitectura capaz de adaptarse automáticamente a cambios en la distribución de tráfico?***

Una vez corroborada la ganancia por diversidad del sistema MNB, otro de los grandes objetivos de esta Tesis Doctoral ha sido desarrollar un algoritmo de reparto óptimo de los recursos disponibles en la red, de manera que se pueda aprovechar al máximo esta ganancia. El algoritmo de asignación desarrollado es capaz de distribuir los recursos entre los usuarios activos, determinando al mismo tiempo desde qué antena o qué combinación de antenas se transmitirá a cada usuario. Además, el mecanismo de *scheduling* toma las decisiones de manera dinámica, de forma que se pueda variar la configuración del sistema en función de la distribución de la carga.

Un problema tan complejo como éste, de estas dimensiones y con estas grandes restricciones temporales de respuesta, sólo se puede resolver mediante alguna técnica matemática avanzada. En esta Tesis Doctoral se han analizado, investigado y evaluado las prestaciones de las redes neuronales Hopfield (HNN). Una de las principales ventajas de las HNN es que son capaces de encontrar una solución semi-óptima en un tiempo muy reducido, gracias sobre todo al inherente paralelismo de la red neuronal. Los resultados obtenidos han demostrado que, con un periodo de envío de informes por parte de los usuarios de 200 ms y un periodo de reasignación de recursos de 100 ms, el sistema MNB es capaz de seguir perfectamente las variaciones experimentadas en los niveles de interferencia y en la carga de la red. Así, el MNB consigue reducir el valor de probabilidad de *outage* de un 1.3% hasta un 0.27%, lo que demuestra el buen comportamiento dinámico de la formulación empleada.

***¿Qué algoritmos adaptativos de gestión de recursos habría que utilizar para optimizar la utilización de los recursos disponibles?***

Tras analizar distintas alternativas, se ha concluido que las mejores prestaciones del sistema UMTS se obtienen utilizando el algoritmo específico de gestión de recursos desarrollado en esta Tesis Doctoral para el MNB. El mecanismo desarrollado no solamente distribuye de manera óptima los recursos entre los usuarios y la potencia entre las antenas sino que, además, es capaz de tomar medidas de control de congestión, de manera que se garantice el comportamiento controlado del sistema.

Respecto a la distribución de recursos, la formulación de la red neuronal desarrollada maximiza el uso de los recursos a la vez que minimiza el *delay* en la transmisión de los paquetes. Las prestaciones conseguidas mediante el uso de la HNN se han comparado con otros algoritmos orientados a la satisfacción del *delay*. Los resultados obtenidos han demostrado que el nuevo algoritmo de gestión de recursos desarrollado en esta Tesis Doctoral es mejor que el resto de algoritmos encontrados en la literatura. Esto es debido, sobre todo, por la gran ventaja que supone considerar simultáneamente la satisfacción de múltiples objetivos.

Además, la red Hopfield es capaz de distribuir de manera óptima la señal transmitida al usuario entre las antenas que conforman el MNB. Gracias a la diversidad espacial y esta gestión óptima de recursos, la potencia consumida se puede dividir hasta por un factor cinco. Además de esta reducción en la potencia consumida, la calidad de servicio también se ve mejorada, reduciéndose así el número de usuarios en cola y el tiempo medio de espera en la transmisión de los paquetes.

### ***¿Cuál es la configuración óptima de la distribución espacial de las antenas?***

Respecto a los distintos escenarios de despliegue de los transmisores, la conclusión alcanzada tras el análisis de los resultados es que las mejores prestaciones se obtienen aumentando el número de emplazamientos celulares, siempre que estos no incurran en interferencia mutua. A partir de tres micro-celdas adicionales no se consigue mejorar la calidad de servicio, pero sí reducir sensiblemente el consumo de potencia del sistema. Además, cuantas más micro-celdas se desplieguen más interesa separarlas del centro de la celda para precisamente evitar esta interferencia indeseable. Finalmente, y de cara a reducir los costes de despliegue, parece razonable considerar que el escenario con tres micro-celdas es el subóptimo, puesto que el pequeño mayor consumo de potencia se justifica sin duda por el hecho de no tener que alquilar más emplazamientos para las antenas.

### ***¿Qué ganancia en términos de consumo de potencia, capacidad y calidad de servicio sería capaz de obtener la nueva arquitectura Multinodo B?***

Gracias al despliegue micro-celular del MNB se consigue aumentar el factor de ortogonalidad. Además, el factor de interferencia intercelular también se verá reducido puesto que la señal queda confinada en las calles y no hay visión directa entre las estaciones micro-celulares. Por último, se ha demostrado que disponer de más de una estación transmitiendo simultáneamente a un usuario aumenta la diversidad en frecuencia y crea macro-diversidad por tener dos enlaces totalmente incorrelados en

cuanto a las pérdidas de propagación. Esta diversidad adicional lleva consigo una reducción importante de los requisitos de  $E_b/N_0$ .

Todos estos efectos se han estudiado en las simulaciones de sistema realizadas. Tras la evaluación exhaustiva, se puede concluir que el MNB, en sus dos variantes activo y pasivo, es capaz de reducir en gran medida la potencia consumida en el sistema. Además, fruto de este menor consumo y de la diversidad adicional introducida por la transmisión desde dos emplazamientos, la calidad de servicio se ve también mejorada en gran medida.

Comparando las dos soluciones de arquitectura MNB propuestas en esta Tesis, la mayor flexibilidad del MNB-activo permite obtener las mejores prestaciones en cuanto a calidad de servicio, a costa de un ligero incremento en el consumo de potencia respecto al MNB-pasivo.

Finalmente, comparando la nueva arquitectura radio MNB-activo respecto al sistema UMTS convencional, la mejora en el consumo de potencia es de un 82%, la probabilidad de exceder el máximo *delay* se reduce en un 50% y la probabilidad de *outage* en un 80.7%.

Gracias a esta reducción espectacular del nivel de interferencias se consigue maximizar la capacidad del sistema en el MNB, que estará limitada por el máximo número de códigos disponibles. Además, como se transmite mucha menos potencia se consigue una reducción del impacto ambiental de las radiaciones electromagnéticas, con lo que será más fácil satisfacer la normativa vigente. A modo de ejemplo, tal y como se ha demostrado, el sistema MNB es capaz de dar servicio a una red congestionada con apenas 2 W, mientras que un sistema de telefonía inalámbrica doméstica transmite a 250 mW en el interior de nuestros hogares, los radioaficionados transmiten entre 5-20 W, GSM transmite entre 20 a 80W, las antenas de televisión entre 1 y 5 kW, y las antenas de radio-difusión hasta 1 MW. La diferencia comparativa es muy significativa.



***¿Realmente es posible la transferencia tecnológica de esta nueva arquitectura?***

Tal y como se ha demostrado en esta Tesis Doctoral, la incorporación del MNB aprovechando el despliegue de estructuras jerárquicas celulares requiere algunas modificaciones *software* en el sistema UMTS convencional y conectar las antenas desplegadas mediante fibra óptica. Este último aspecto es el que más problemas podría ocasionar a las operadoras. Sin embargo, teniendo en cuenta que prácticamente todos los núcleos urbanos disponen de una importante estructura de fibra óptica sobredimensionada, se podría llevar a cabo esta operación sin mucha complejidad. Además, las micro-celdas no se pueden conectar al núcleo de red mediante radioenlaces puesto que carecen de visión directa, por lo que el esfuerzo de cableado se realizará de cualquier manera.

Las grandes ventajas que aporta el uso de MNB justifican, sin duda alguna, su transferencia tecnológica.

***¿Hasta qué punto influyen los modelos de shadowing en la evaluación de las prestaciones de un sistema celular UMTS?***

Por último, dada la importancia de los modelos de simulación empleados en la evaluación de las prestaciones de los sistemas, en esta Tesis Doctoral también se ha trabajado activamente en el modelado del *shadowing*. No solamente se han propuesto distintos modelos de *shadowing* evaluando sus propiedades, sino que también se ha caracterizado el efecto que los distintos modelos tienen en las prestaciones finales del sistema. Así pues, a partir de los resultados obtenidos en un mismo escenario de simulación utilizando los distintos modelos propuestos, se ha comprobado que para evaluar correctamente el sistema UMTS es necesario modelar los dos efectos de correlación más destacados del *shadowing*: la autocorrelación espacial y la correlación cruzada. Sin embargo, dentro de los modelos que incluyen ambas correlaciones, las diferencias apenas son significativas, con lo que el modelo más simple es el más adecuado.

También cabe destacar como una de las principales aportaciones de esta Tesis Doctoral el método matemático propuesto para asegurar la factorización de matrices de correlación experimentales grandes. Este método soluciona un problema bastante común en el análisis de los sistemas MIMO (*Multiple-Input Multiple-Output*) y en cualquier problema de ingeniería en el que sea necesario generar secuencias aleatorias correladas a partir de una matriz de correlación dada.

## 7.2 Líneas Futuras de Investigación

A continuación se exponen algunas actuaciones futuras como continuación del trabajo realizado:

- Mejorar el simulador de enlace incluyendo un modelo de error en la recepción de comandos del canal de control o teniendo en cuenta el efecto del ruido de fase o de las no linealidades de los convertidores y amplificadores.
- Introducir nuevas técnicas de estimación de canal del MNB. En este punto se podría contemplar la combinación del canal piloto CPICH con los bits asociados al DPCCH para obtener unas mejores prestaciones en la estimación de canal.
- Implementar nuevas técnicas de codificación más evolucionadas e incluir el *puncturing*, técnica muy utilizada en la práctica.
- También se debería estudiar otras técnicas más sofisticadas de procesado en *array* para la transmisión con diversidad en bucle cerrado. El método utilizado en esta Tesis Doctoral únicamente maximiza la potencia recibida, mientras que existen otras técnicas más complejas que maximizan la relación señal a interferencia. El problema es que para poder utilizar estas técnicas es necesario caracterizar estadísticamente las interferencias del sistema y eso en un sistema con tanta variabilidad de las interferentes como es UMTS no es tarea sencilla. Además, se podría hacer uso de técnicas de *beamforming* para conformación de haz.
- Falta también por evaluar el comportamiento del MNB en enlace ascendente. Obviamente, al disponer de más antenas y, por tanto, de más caminos de recepción de la señal, el receptor *Rake* dispondrá de más diversidad para optimizar la recepción de la señal. Sin embargo, sería interesante evaluar hasta qué punto existe ganancia por utilizar esta estructura distribuida de antenas en enlace ascendente.

- Como se ha comentado, las redes MNB están pensadas para desplegarse simultáneamente con las redes jerárquicas celulares. Un aspecto interesante por investigar es ver cuál sería la interacción entre la capa micro-celular y la celda MNB y hasta qué punto puede afectar el MNB a las técnicas propuestas en la literatura para optimizar el funcionamiento de las redes jerárquicas convencionales.
- Respecto a las redes neuronales Hopfield, su buen funcionamiento motiva su extensión a otras tecnologías como HSDPA o LTE, en las que la rápida respuesta del sistema es fundamental. También se ha empezado a trabajar en la extensión de la formulación empleada para redes Heterogéneas y *multi-homing*.
- Finalmente, hasta ahora la evaluación de las redes Hopfield se ha realizado únicamente mediante simulación. Una futura línea de investigación es abordar la digitalización e implementación *Hardware* de estas redes neuronales en FPGAs, para evaluar el tiempo de respuesta real del sistema.

# Anexo A

## *Descripción del Simulador de Enlace*

Para entender completamente el análisis de las prestaciones radio del sistema Multinodo B es necesario describir con minuciosidad la herramienta de simulación a nivel de enlace empleada. Sin embargo, una descripción detallada del simulador en medio del texto de la Tesis Doctoral podría hacer que el lector perdiera el hilo del razonamiento seguido en el análisis del sistema MNB, por lo que se ha preferido relegar la descripción a este anexo. A pesar de ser información anexada, este capítulo es de vital interés para el buen entendimiento de la Tesis Doctoral y pretende no sólo describir todas las unidades funcionales que componen el simulador, sino también reflejar la meticulosidad con la que se ha abordado su implementación.

El desarrollo de una herramienta de simulación tan potente y compleja sólo se ha acometido en grandes empresas del sector de las comunicaciones móviles, como Ericsson o Nokia, o en proyectos de investigación punteros en Europa. Gran parte del éxito de esta Tesis Doctoral es precisamente dotar al Grupo de Comunicaciones Móviles de la UPV de una herramienta de estas características y de la capacidad que permite poder abordar retos similares en otras tecnologías.

Este anexo esta compuesto básicamente de dos apartados: el primero describe todos los bloques funcionales del simulador, mientras que el segundo se centra en la validación de los mismos, validación a partir de la comparación directa con otros proyectos europeos.

## A.1 Bloques Funcionales del Simulador

El simulador de enlace desarrollado en la presente Tesis Doctoral se compone de los bloques funcionales mostrados en la Figura A.1. A continuación, se procede a su descripción.

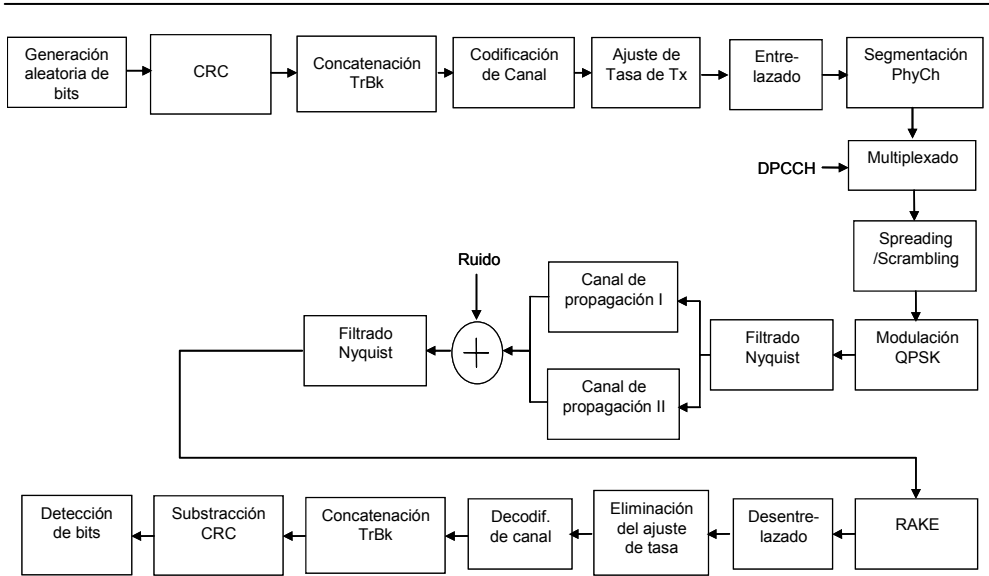
### A.1.1 Generación aleatoria de bits

La secuencia de bits de información en formato 0s y 1s se genera de manera aleatoria, almacenando dicha información para la posterior detección de errores.

### A.1.2 CRC

CRC, acrónimo de *Cyclic Redundancy Check* (comprobación de redundancia cíclica), es un procedimiento utilizado para detectar errores de transmisión de datos. Este procedimiento consiste en que, antes de enviar la información, el transmisor utiliza un algoritmo para generar una secuencia de bits calculada en función de los datos que se van a transmitir. Esta secuencia se añade tras el bloque de datos y se envía al receptor. A continuación, el receptor aplica el mismo algoritmo a los datos recibidos, y obtiene una secuencia que debe coincidir con la secuencia generada por el transmisor y enviada junto a los datos. Si coinciden, existe una elevada probabilidad de que la transmisión se haya completado con éxito.

Este bloque puede ser obviado en la implementación dada la naturaleza del simulador, en donde se conoce perfectamente la información transmitida. Así pues, los bits asociados al CRC serán rellenados con información útil y por tanto contribuirán al cálculo de la tasa de pérdidas. Sin embargo es importante tener en cuenta, a la hora de calcular la energía por bit de información útil  $E_b$ , que estos bits no contribuyen a la tasa de información de usuario.

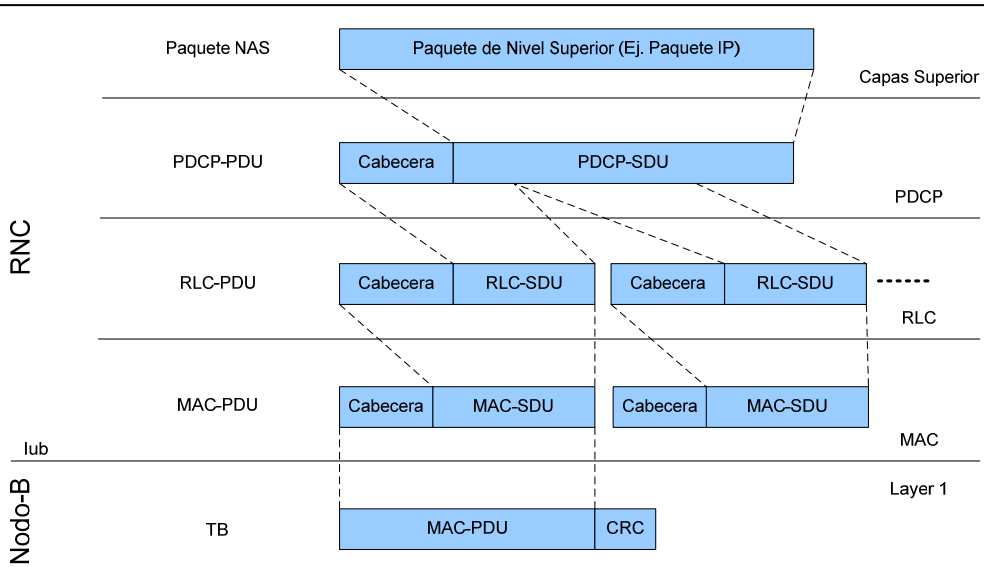


**Figura A.1 Diagrama de bloques del simulador de enlace**

### A.1.3 Concatenación de bloques de transporte

El proceso de formación del bloque de transporte viene especificado por la tabla RAB del servicio. Como se puede apreciar en la Figura A.2, al paquete de datos (PU) procedente del usuario se le anexa, si así corresponde, las cabeceras introducidas respectivamente por los niveles PDCP, RLC y MAC. Una vez añadido el CRC, ya se tiene perfectamente completado, a nivel físico, un bloque de transporte.

El formato de transporte (TF) define el número de bloques transmitidos en un intervalo temporal TTI. La concatenación de bloques de transporte lo único que hace es encadenar un bloque de transporte tras otro hasta completar el número de bloques especificados por el TF y pasarle esta secuencia de bits, o bloque código CB (*Code Block*), al módulo responsable de la codificación de canal.



**Figura A.2 Formación del TB**

Si el número de bits del CB excede el tamaño máximo del bloque a codificar, entonces el módulo de concatenación de bloques de transporte se encargará de segmentar apropiadamente la secuencia. El tamaño máximo del CB depende de la técnica de codificación empleada:

- Codificación convolucional: tamaño máximo 504 bits
- Turbo codificación: tamaño máximo 5114 bits.

## A.1.4 Codificación de canal

La finalidad de la transmisión reside en que el receptor reciba exactamente lo que se le envía desde un emisor dado. El principal problema se encuentra en que el canal añade ruido aleatorio y en consecuencia es necesario un proceso mediante el cual decidir qué mensaje, de los posibles, ha sido enviado. Si no existieran restricciones en potencia, bastaría con aumentar el nivel de potencia de las señales transmitidas para tener una demodulación más fiable, ya que la potencia de ruido no aumenta. Sin



embargo, en sistemas limitados por interferencias, como el sistema UMTS, esta solución no es válida y es necesario encontrar una opción alternativa.

En sus trabajos, *Shannon* concluyó que no hacía falta aumentar la potencia sino disminuir el número de mensajes, es decir, no enviar todos los posibles sino solo una fracción de ellos. Los mensajes posibles son  $2^n$  en el caso de transmisión binaria (siendo  $n$  el número de bits de la palabra codificada), lo que *Shannon* propuso es enviar  $2^{Rn}$  mensajes con  $R < 1$ . De esta forma, si  $R$  es menor que la capacidad del canal  $C$  (es una característica propia de cada canal), se puede crear un sistema de comunicaciones con una probabilidad de error característica muy baja, siempre que  $n$  sea lo suficientemente grande, es decir, cuando  $n$  tiende a infinito. El problema está en decidir qué  $2^{Rn}$  vectores se van a enviar. Al sistema que realiza dicha decisión se le llama codificador.

La codificación del canal consiste en asignar a cada secuencia de datos de entrada una secuencia de salida que será transmitida a través del canal radio. En recepción, con la información recibida del canal, se realiza una asignación inversa, obteniéndose una secuencia de datos de tal forma que los efectos del ruido sean mínimos. Además, en este proceso de asignación, también denominado mapeo por herencia de la expresión inglesa *mapping*, se introduce cierta información de redundancia en la secuencia de datos original, consiguiéndose de esta manera una mayor protección de los datos que supone obviamente una mejora en la fiabilidad de la transmisión.

Según el estándar de UMTS [1], un canal de transporte de datos, como por ejemplo el DCH (*Data Channel*), se puede codificar en el transmisor tanto mediante codificación convolucional como mediante turbo-codificación. Todos los posibles esquemas de codificación junto con las tasas de codificación asociadas, se resumen en la Tabla A.1.

Canal de Transporte	Esquema de Codificación	Tasa de Codificación
BCH	Codificación Convolutacional	1/2
PCH		
RACH		
CPCH, DCH, DSCH, FACH	Turbo Codificación	1/3

Tabla A.1 Esquemas de codificación [1]

### A.1.4.1 Codificación convolutacional

Los códigos convolucionales son códigos lineales que tienen la propiedad de que la suma de dos palabras código cualesquiera también es una palabra código. La codificación convolutacional es una manera de reducir el número de mensajes que enviamos por el canal, cumpliendo de esta forma la recomendación de *Shannon*.

Un código convolutacional queda especificado por tres parámetros ( $n$ ,  $k$ ,  $m$ ):

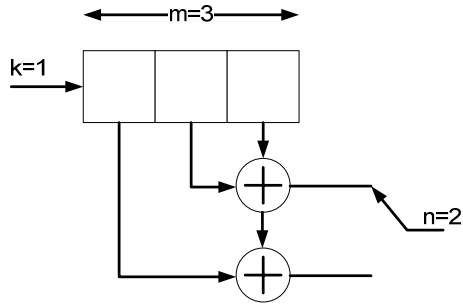
- $n$  es el número de bits de la palabra codificada.
- $k$  es el número de bits de la palabra de datos
- $m$  es la memoria del código o longitud restringida

El número de bits por palabra de datos,  $k$ , cumple:

$$\frac{k}{n} = R \quad (\text{A.1})$$

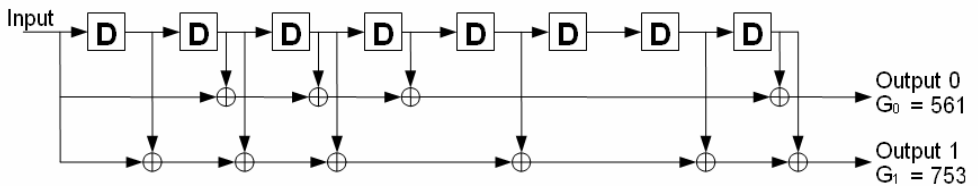
A este cociente se le denomina ratio del codificador. Como ejemplo, en la siguiente figura se muestra un codificador convolutacional (2, 1, 3), es decir: un bit para representar la palabra de datos, dos bits de palabra de codificación por cada bit de palabra de datos y tres bits de longitud de registro. En este caso el ratio es  $\frac{1}{2}$ .

Cuanto mayor sea  $n$ , más segura será la transmisión y menor será, por tanto, la probabilidad de error. Por el contrario, a mayor  $n$  más complejo será el diagrama de estados del codificador y más difícil su posterior decodificación. Hay que encontrar, como siempre, el compromiso idóneo entre calidad y complejidad.

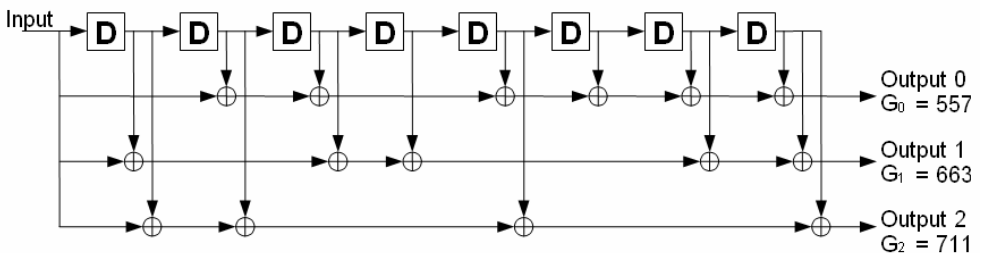


**Figura A.3 Codificador convolucional**

La estructura de los codificadores convolucionales que deben usarse en UMTS viene definida en la norma del 3GPP [1], y tal cual se han implementado en el simulador. En ella se especifica que la longitud de constricción del código convolucional (el número de símbolos de información necesarios para realizar la codificación) debe ser 9, y que su tasa de codificación puede ser 1/2 o 1/3. La estructura de los codificadores, tal y como se describe en el estándar, puede verse en las siguientes figuras.



**Figura A.4 Codificador convolucional UMTS de tasa 1/2 [1]**



**Figura A.5 Codificador convolucional UMTS de tasa 1/3 [1]**

Los valores de los polinomios generadores están expresados en notación octal.

La implementación de los codificadores convolucionales no reviste demasiada complejidad, puesto que solamente es necesario implementar una serie de multiplicadores, sumadores y registros de desplazamiento.

## Decodificación

A continuación, se explica el algoritmo de decodificación de máxima verosimilitud en sus dos versiones, la firme o *hard* y la indecisa o *soft*. El algoritmo de Viterbi está basado en la decodificación de máxima verosimilitud, por lo que es importante entender su fundamentos matemáticos.

El algoritmo de decodificación de máxima verosimilitud, pretende seleccionar la palabra código (secuencia de bits codificados asociados a un determinado bit de entrada del codificador) de manera que se minimice la probabilidad de detectar un mensaje (bit codificado) erróneo.

En el caso de la detección firme, la probabilidad de error en la decodificación viene dada por la siguiente expresión:

$$P_e(m_r, c_r) = P(m_r \neq m | c_r) = 1 - P(m_r = m | c_r) \quad (\text{A.2})$$

Donde  $m_r$  es el mensaje decodificado en el receptor,  $m$  es el mensaje codificado en el transmisor y  $c_r$  la palabra código recibida. La regla de decisión óptima es la definida por la siguiente ecuación:

$$\text{se decide } m_r = m^i \text{ si } P(m^i = m | c_r) \geq P(m^k = m | c_r) \text{ para } k \neq i \quad (\text{A.3})$$

siendo  $m^i$  el  $i$ -ésimo mensaje de la lista de posibles mensajes codificados. Aplicando el teorema de Bayes a la ecuación anterior, se puede escribir la regla de decisión óptima de la siguiente forma:

$$\text{se decide } m_r = m^i \text{ si}$$

$$\frac{P(m^i = m) f_{c_r}(c_r | m^i = m)}{f_{c_r}(c_r)} \geq \frac{P(m^k = m) f_{c_r}(c_r | m^k = m)}{f_{c_r}(c_r)} \quad (\text{A.4})$$

para  $k \neq i$

donde  $f_{c_r}(c_r)$  es la función densidad de probabilidad de recibir la palabra código  $c_r$ .

Considerando que todos los mensajes son equiprobables y que el mensaje  $m^i$  se corresponde uno a uno con el vector codificado,  $c^i$ , se obtiene el criterio de máxima verosimilitud como:

$$\text{se decide } m_r = m^i \text{ si para } k = i \text{ se maximiza } f_{c_r}(c_r | c^k = c) \quad (\text{A.5})$$

Si se asume un canal sin memoria, éste afecta de forma independiente a los bits que componen el vector codificado dando lugar a cada uno de los bits del vector binario recibido, por lo tanto, la función de verosimilitud del vector recibido es el producto de las funciones de verosimilitud de cada uno de los  $j$  bits recibidos, quedando finalmente:

$$f_{c_r}(c_r | c^k) = \prod_j f_{c_r^j}(c_r^j | c_k^j) \quad (\text{A.6})$$

Y expresado de manera logarítmica:

$$\ln(f_{c_r}(c_r | c^k)) = \sum_j \ln(f_{c_r^j}(c_r^j | c_k^j)) \quad (\text{A.7})$$

Maximizar la función de verosimilitud equivale a maximizar la función de máxima verosimilitud logarítmica. Por lo tanto, se denomina  $p$  a la probabilidad de que se produzca un error en un bit,  $d$  al número de bits en los que difiere el vector binario recibido  $c_r$  de una cierta palabra código  $c^k$  (se supone que cada código tiene  $N$  bits). La función de verosimilitud logarítmica será:

$$\ln(f_{c_r}(c_r | c^k)) = d \ln(p) + (N - d) \ln(1 - p) = d \ln\left(\frac{p}{1 - p}\right) + N \ln(1 - p) \quad (\text{A.8})$$

El último término de la ecuación es común a todos los vectores código, por lo que para maximizar la ecuación anterior, sabiendo que  $p < 1/2$ , basta con minimizar el término que esta multiplicado por  $d$ . Por lo tanto, dado que  $p$  es constante, el valor mínimo de  $d$  es el que minimiza la función anterior. Como conclusión, un

decodificador basado en el criterio ML (máxima verosimilitud) escoge, entre las posibles, la palabra código que presenta el menor número de bits diferentes respecto al vector binario recibido (mínima distancia *Hamming*).

En el caso de detección indecisa, el desarrollo es igual al anterior, lo único que varía son las funciones de probabilidad. Considerando que todos los mensajes son equiprobables y que el mensaje  $m^i$  se corresponde uno a uno con el vector codificado  $c^i$ , y en consecuencia con el vector de símbolos  $s^i$ , se obtiene el criterio de máxima verosimilitud como:

$$\text{se decide } m_r = m^i \text{ si para } k=i \text{ se maximiza } f_y(y | s^k = s) \quad (\text{A.9})$$

Donde  $y$  es el vector recibido directamente del canal, es decir en el espacio de la señal, y  $s$  es el vector de símbolos transmitidos. Como el canal afecta de forma independiente a los símbolos del vector transmitido para dar lugar a cada uno de los símbolos del vector recibido, la función de verosimilitud del vector recibido es el producto de verosimilitudes de cada uno de los símbolos recibidos

$$f_y(y | s^k) = \prod_j f_{y^j}(y^j | s_k^j) \quad (\text{A.10})$$

O en forma logarítmica

$$\ln(f_y(y | s^k)) = \sum_j \ln(f_{y^j}(y^j | s_k^j)) \quad (\text{A.11})$$

Como ocurría antes, maximizar la función de verosimilitud o la función de verosimilitud logarítmica es equivalente. Se supone que cada vector de símbolos tiene  $M$  elementos. La función de verosimilitud logarítmica será en este caso:

$$\ln(f_y(y | s^k)) = -\frac{2}{N_0^2} \sum_j (y^j - s_k^j)^2 + \frac{M}{2} \ln\left(\frac{2}{\pi N_0^2}\right) \quad (\text{A.12})$$

El último término de la ecuación es común a todos los vectores código. Por lo que, maximizar la ecuación equivale a maximizar el primer término, es decir, la distancia en el espacio de la señal al cuadrado de  $y$  y  $s^k$ . Como conclusión, un decodificador indeciso escoge la palabra código, de entre las posibles, cuya representación en el espacio de la señal presenta la menor distancia euclídea con el vector de símbolos recibido.

En la práctica la decodificación convolucional se realiza mediante una simplificación del decodificador ML, la cual selecciona, por definición, el valor estimado de la salida del decodificador,  $y'$ , que maximiza la probabilidad de haber transmitido una entrada concreta,  $r$ , condicionada a esa salida  $P(r|y')$ .

Esta simplificación del decodificador ML es el Algoritmo de Decodificación de Viterbi. El proceso consiste en desechar algunos de todos los caminos posibles. Lo que se consigue aplicando este método es reducir el número de cálculos.

Según el algoritmo de Viterbi, para reducir el número de cálculos, cada vez que dos trayectos, también llamados ramas, se junten en un estado en el diagrama de enrejado, el de mayor métrica acumulada, bien sea métrica basada en distancia *Hamming* o en distancia euclídea, se desecha en la búsqueda del trayecto óptimo. Esto se debe a que el estado actual resume la historia de todos los estados anteriores en cuanto a su influencia en los estados posteriores. Esto se hace en los  $2^{k-l}$  estados, se pasa al intervalo  $t_{i+1}$  y se repite el proceso.

Realmente, esta eliminación de posibles caminos no comienza hasta una determinada etapa del enrejado, por ejemplo para un código convolucional de 8 estados, la eliminación de caminos empezaría en la tercera etapa. Esto se debe a que hasta ese instante no han podido converger dos ramas en un estado. ya que por cada  $k$  bits de palabra de datos surgen  $2^k$  estados. Es decir, del primer estado, el 00 por definición, se dispone de dos estados en el siguiente nivel de representación, y por cada uno de estos, otros dos en el posterior nivel. Por tanto será en el tercer nivel cuando se empiece a tomar decisiones porque será entonces cuando convergerán dos ramas en cada estado.

Se puede calcular el valor de las métricas de dos formas distintas. Teniendo en cuenta el número de bits diferentes entre dos secuencias dadas, distancia *Hamming*, o considerando la métrica como la distancia euclídea entre dos secuencias en el espacio de señal. La primera de ellas constituye el núcleo de lo que se conoce como decodificación firme, mientras que la segunda constituye el núcleo de la decodificación indecisa.

Aunque no se va a entrar más en detalle en la descripción de estas técnicas de detección, en esta Tesis Doctoral se han implementado ambos métodos de decodificación de Viterbi de máxima verosimilitud, aunque obviamente la decodificación indecisa obtiene mejores prestaciones que la firme, por lo que será la utilizada para obtener las LUTs características del MNB.

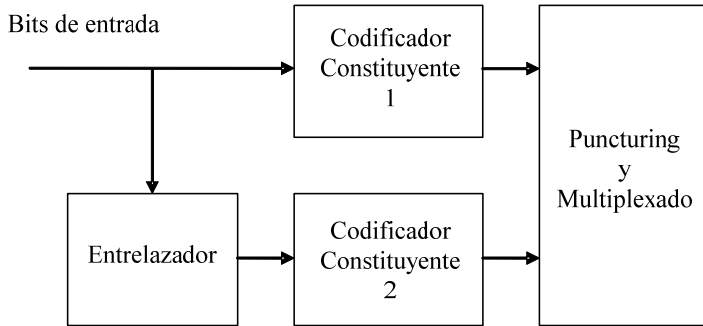
#### **A.1.4.2 Turbo codificación**

La turbo codificación fue propuesta en 1993 por *Berrou* y *Glavieux*, quienes demostraron la existencia de una excelente ganancia en la codificación [2], acercándose a las predicciones de *Shannon*. Basaron el funcionamiento de un turbo código en codificar una secuencia dos veces, con un entrelazador entre las dos codificaciones con el fin de independizar estadísticamente las dos secuencias codificadas. Para la codificación, propusieron el uso de codificadores convolucionales sistemáticos recursivos, o codificadores RSC (*Recursive Systematic Convolutional*).

Para el turbo decodificador propusieron dos decodificadores RSC, los cuales debían emplear algoritmos de decodificación especiales, los cuales aceptasen entradas indecisas y proporcionasen a su salida salidas indecisas. *Berrou* y *Glavieux* establecieron que la salida de los decodificadores debía realizar no solo una estimación de los bits, sino que además debían proporcionar la probabilidad de que los bits hubieran sido correctamente decodificados. Años después, *Berrou* propuso como algoritmo de decodificación una versión modificada del algoritmo MAP (*Maximum a Posteriori*) con una elevada complejidad computacional.

Desde la aparición de los turbo códigos, se han realizado un gran número de investigaciones a fin de reducir la complejidad del turbo decodificador, lo que dio lugar a la aparición de distintos algoritmos de decodificación como el Max-Log-MAP, Log-MAP o SOVA.



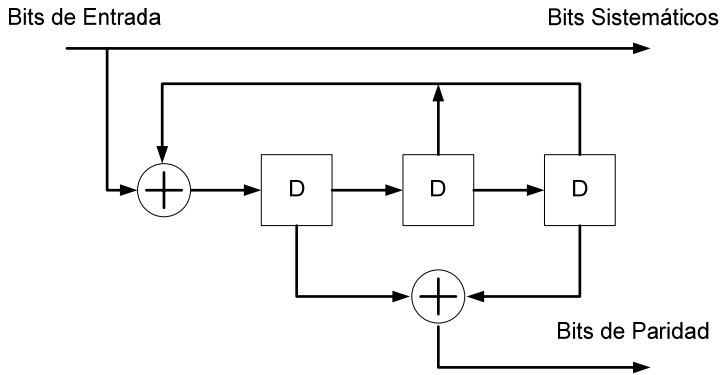


**Figura A.6 Estructura del turbo-codificador**

La estructura general de un turbo codificador puede verse en la Figura A.6. El turbo codificador emplea dos codificadores RSC o convolucional sistemático recursivo, para codificar la misma secuencia de bits de entrada. Sin embargo, en el caso de segundo RSC existe un entrelazador previo. Generalmente en este tipo de codificadores se emplean, como hemos dicho anteriormente varios RSC, aunque es posible obtener buenos resultados empleando códigos bloque. En este apartado se describe únicamente la estructura del turbo codificador estándar, es decir, aquella que emplea como codificador constituyente un RSC y que además es el utilizado en UMTS.

La salida de los RSC es, en algunos casos, perforada (*punctured*) y multiplexada. Normalmente, los dos codificadores convolucionales sistemáticos constitutivos son de tasa  $1/2$ , proporcionando un bit de paridad y un bit sistemático por cada bit de entrada. Si se deseara obtener un turbo codificador de tasa  $1/2$  la mitad de los bits de salida de los dos RSC deberían ser perforados. Normalmente, los bits que son eliminados son los de paridad, ya que, eliminar los bits sistemáticos supondría una disminución excesiva de las prestaciones del código. La Figura A.7 muestra un ejemplo de un codificador RSC con  $k=3$ , empleado en los turbo codificadores.

El rendimiento de un turbo código depende de varios factores, como pueden ser el diseño del bloque entrelazador empleado antes del segundo RSC, los codificadores convolucionales sistemáticos recursivos empleados, y el perforado o *puncturing* aplicado. Todos estos aspectos están especificados en el estándar UMTS.



**Figura A.7 Codificador RSC**

El esquema del turbo codificador empleado en UMTS está descrito en [1]. El turbo codificador está constituido por un codificador convolucional paralelo concatenado, formado por dos codificadores constituyentes de 8 estados y un entrelazador interno. La función de transferencia de los codificadores RSC es

$$G(D) = \left[ 1, \frac{g_1(D)}{g_0(D)} \right] \quad (\text{A.13})$$

La función de transferencia de un codificador, especifica cuantas y cuales son las salidas del mismo. Analizando la función de transferencia anterior, puede verse como la salida del codificador esta formada por el bit de entrada, representado en la función de transferencia con el 1, y un bit de paridad, fruto de realizar una serie de operaciones con los registros de desplazamiento, operaciones especificadas por el término  $g_1(D)/g_0(D)$ . Los polinomios generadores del código tienen como valor

$$\begin{aligned} g_0(D) &= 1 + D^2 + D^3 \\ g_1(D) &= 1 + D + D^3 \end{aligned} \quad (\text{A.14})$$

Como puede verse, cada uno de estos polinomios especifica las conexiones existentes entre los distintos registros de memoria del codificador. Para saber la estructura de las conexiones, basta con estudiar las potencias de los términos en  $D$ . Analizando el

polinomio generador  $g_0$ , éste indica la existencia de conexiones entre la entrada al codificador (término con exponente 0 en  $D$ , es decir el 1), la salida del segundo registro de almacenamiento (término con exponente 2 en  $D$ , el  $D^2$ ) y la salida del tercer registro de almacenamiento (término con exponente 3 en  $D$ , el  $D^3$ ). De forma similar, se analiza el polinomio  $g_1$ , el cual especifica las conexiones entre la entrada al codificador, la salida del primer registro de almacenamiento y la salida del tercero. Destacar también que el polinomio  $g_0$  representa las conexiones de realimentación o *feedback*, mientras que el polinomio  $g_1$  representa las conexiones hacia delante o *forward*.

La estructura del turbo codificador especificado en UMTS puede verse en la Figura A.8. El estado inicial del registro de desplazamiento de los codificadores constituyentes debe ser de todo ceros. La salida del turbo codificador es de la forma  $x_1, z_1, z'_1, x_2, z_2, z'_2, \dots, x_K, z_K, z'_K$ . Donde  $x_1, x_2, x_3, \dots, x_K$  son los bits de entrada del turbo codificador,  $K$  es el número de bits, y  $z_1, z_2, \dots, z_K$  y  $z'_1, z'_2, \dots, z'_K$  son los bits de salida del primer y segundo codificador constituyente respectivamente. Los bits  $x'_1, x'_2, \dots, x'_K$  son los bits de salida del entrelazador, los cuales son pasados al segundo codificador constitutivo como bits de entrada.

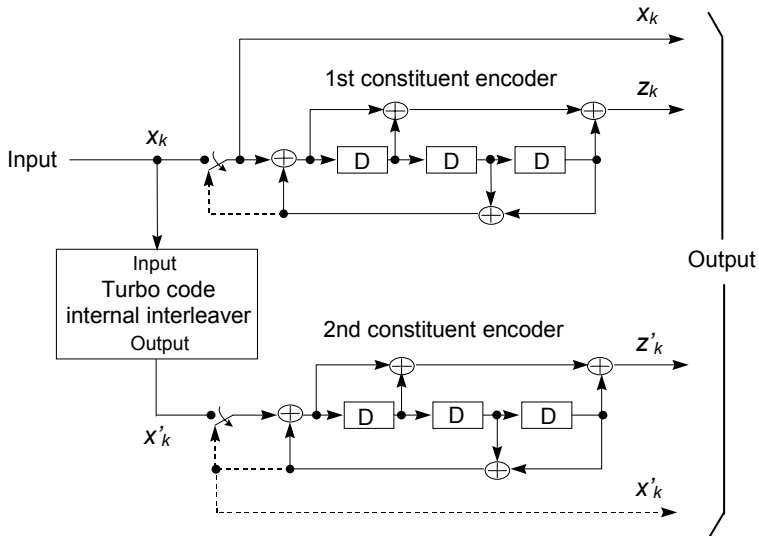
El funcionamiento del entrelazador también viene descrito en [1] y se basa en la creación y manipulación de una matriz, realizando una permutación de filas y columnas en función de una secuencia de número primos y de una serie de vectores fijados por la norma.

El incremento del rendimiento de los Turbo Códigos es debido, por una parte, al entrelazado empleado entre los codificadores y, por otra, a los códigos recursivos empleados.

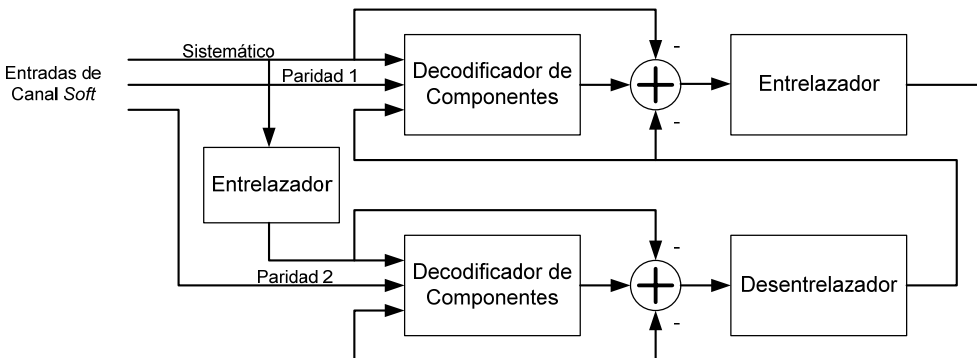
## **Turbo Decodificación**

La estructura del turbo decodificador no está especificada en la norma 3GPP, en la cual únicamente se especifica la estructura del codificador. Por lo tanto, su implementación queda abierta a la consideración de los distintos fabricantes,

permitiendo así el aumento de la competencia entre ellos. En la literatura existen múltiples posibles implementaciones, aunque para esta Tesis Doctoral se ha elegido la estructura propuesta por *Berrou* que se puede observar en la Figura A.9.



**Figura A.8 Estructura del turbo codificador de UMTS [1]**



**Figura A.9 Turbo decodificador de *Berrou***

Como puede verse, al igual que en el turbo codificador, existen dos decodificadores unidos mediante entrelazadores. Cada decodificador tiene tres entradas:

- bits sistemáticos transmitidos
- bits de paridad transmitidos por el codificador asociado en transmisión
- información a-priori

Cada decodificador debe proporcionar a su salida los bits decodificados y, adicionalmente, información conocida como “*soft output*” o “salida indecisa”, la cual representa la probabilidad de que un bit haya sido correctamente decodificado. Estas “salidas indecisas” suelen representarse por medio de los *Log Likelihood Ratios* (LLRs). Los LLRs son una forma sencilla de pasar información de un decodificador a otro. El signo de los mismos determina la naturaleza del bit, mientras que su amplitud cuantifica la probabilidad de haber tomado una decisión correcta.

El decodificador propuesto por *Berrou* funciona de forma iterativa. En la primera iteración, el primer decodificador calcula las estimaciones de los bits transmitidos y las *soft outputs* a partir de la información que le llega del canal. Las *soft outputs* del primer decodificador son usadas por el segundo junto con la información recibida del canal para calcular su estima de los bits de datos. En este punto puede empezar la segunda iteración, en la cual el primer decodificador estima los bits transmitidos a partir de la información transmitida por el canal y de la información adicional que le proporciona el segundo decodificador. Esta información adicional permite al primer decodificador obtener unas *soft outputs* más precisas, las cuales son empleadas posteriormente por el segundo decodificador como información a-priori. Este ciclo se repite, y con cada iteración la probabilidad de error en la decodificación de los bits recibidos mejora. Sin embargo, a medida que aumenta el número de iteraciones la mejora experimentada es menor, por lo que en la práctica se suele considerar un número máximo de 8 iteraciones de decodificación.

Existe la posibilidad de emplear turbo decodificadores no iterativos, pero la mejora en la ganancia que se obtiene con ellos es únicamente de 0.35dB respecto a los iterativos, mientras que su complejidad es mucho mayor.

En concreto, en el simulador de enlace implementado se ha incluido los dos mecanismos de turbo decodificación más extendidos: la turbo decodificación SOVA (*Soft-Output Viterbi Algorithm*) y LOG-MAP (*Logarithmic Maximum A Posteriori*). En el apartado de validación se compararán las prestaciones de ambos decodificadores, comprobando como su comportamiento se corresponde a lo encontrado en la literatura.

Si se desea ampliar la información respecto a la turbo codificación se recomienda la referencia [3].

## A.1.5 Ajuste de tasa de transmisión

Una vez codificados los datos es necesario introducirlos en los correspondientes canales físicos. Para ello, en ocasiones habrá que eliminar o incluir bits adicionales para así completar las capacidades de transmisión del canal físico seleccionado.

Dentro del ajuste de la tasa de transmisión, UMTS contempla dos posibles técnicas: la del relleno de ceros o *bit stuffing* y la de eliminación de información redundante o *puncturing*.

El *bit stuffing* consiste en rellenar el bloque de información ya codificado con 0s de manera que se completen los bits de información física que permite transmitir el factor de ensanchado seleccionado. La Figura A.10 muestra claramente esta técnica.

Por otro lado el ajuste de la tasa de transmisión mediante *puncturing* se basa en la eliminación de determinados bits redundantes en la codificación. Como se ve en la Figura A.11, a los bits sistemáticos no se les ajusta la tasa pero sí al resto. Finalmente hay que incluir indicadores de discontinuidad para que el decodificador, mucho más complejo en cuanto a su implementación, sea capaz de decodificar correctamente la secuencia a pesar de la falta de información.

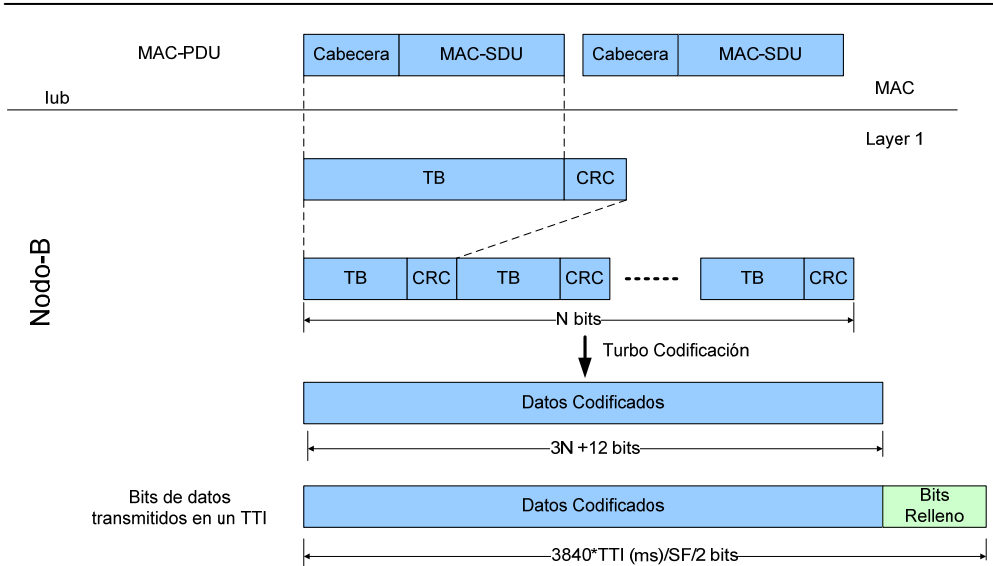


Figura A.10 Conformación de la trama TTI y Bit Stuffing

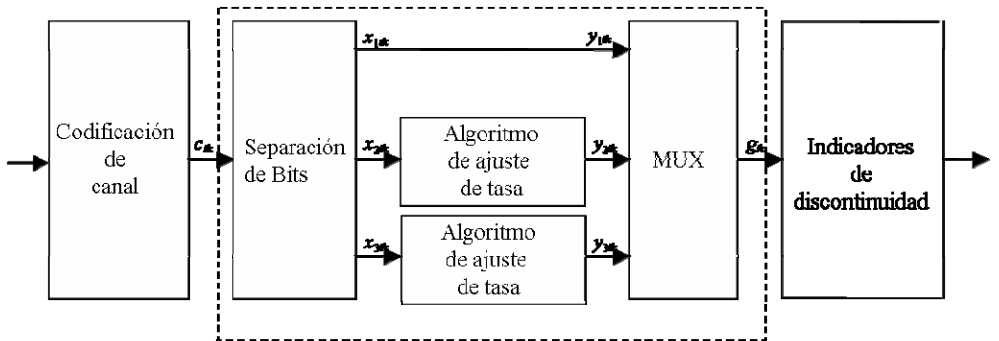


Figura A.11 Técnica de Puncturing

Como se ha comentado anteriormente, es importante tener en cuenta la tasa de transmisión de bits de información de cara a calcular la eficiencia energética, es decir, la relación entre la energía de bit transmitido a nivel físico y la energía de los bits de información. Para evaluar dicha eficiencia energética se necesita calcular el factor  $r$ , definido como la relación entre los bits útiles transmitidos en un TTI respecto al número total de bits físicos transmitidos. El número de bits útiles por TTI se calcula como el número de bits de información por TB ( $N_{PU}$ ) multiplicado por el número de

TBs por TTI, es decir, multiplicado por el TF. El número total de bits transmitidos en *downlink* por TTI es igual al número de *chips* por TTI dividido por el factor de ensanchado, o SF, y multiplicado por dos, ya que en UMTS se utiliza una modulación QPSK. Finalmente  $r$  queda definido como:

$$r = N_{PU} \cdot TF \cdot \frac{SF}{2 \cdot 3840 \cdot TTI(ms)} \quad (A.15)$$

Esta expresión asume que no se aplica la técnica de eliminación de bits redundantes o *puncturing*, sino que se utiliza el relleno de ceros para el ajuste de la tasa de transmisión. Por este motivo, y con el objetivo de simplificar el simulador, sólo se ha incluido el *bit stuffing* como técnica de ajuste de tasa de transmisión.

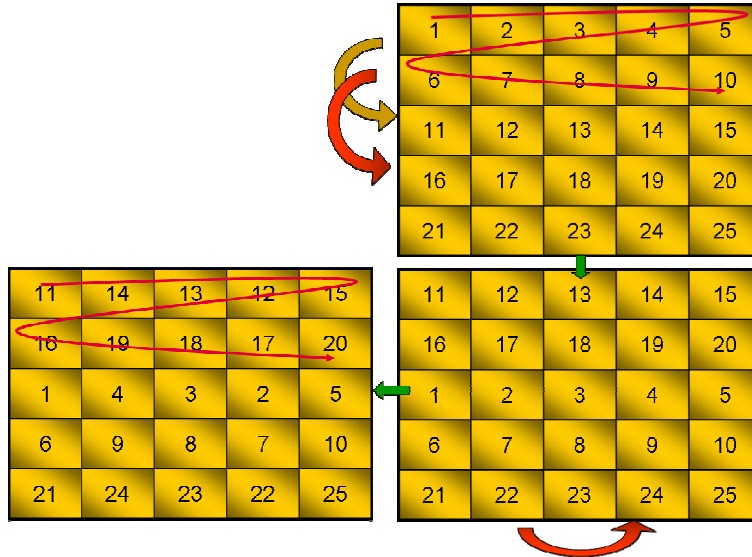
Al emplearse la técnica de relleno de ceros, una vez seleccionado un determinado RAB, el SF se ajusta de manera que se garantice que la capacidad de transmisión por TTI es mayor o igual al tamaño del bloque de información de usuario tras la codificación.

## A.1.6 Entrelazado

El entrelazado se realiza una vez calculado completamente el bloque de bits de información física que se van a transmitir en el siguiente TTI. Básicamente, consiste en una reordenación de los bits a transmitir con el único objetivo de decorrelar el canal, consiguiendo así la eliminación de ráfagas de errores y, por tanto, una mejor corrección de bits en el módulo de decodificación del receptor.

El proceso para realizar el entrelazado es muy sencillo. Tal y como se muestra en la Figura A.12, consiste en rellenar una matriz con la secuencia de bits del bloque TTI, permutar filas y columnas de una manera preestablecida y conocida también por el receptor y, finalmente, definir la secuencia entrelazada leyendo ordenadamente la matriz resultante.





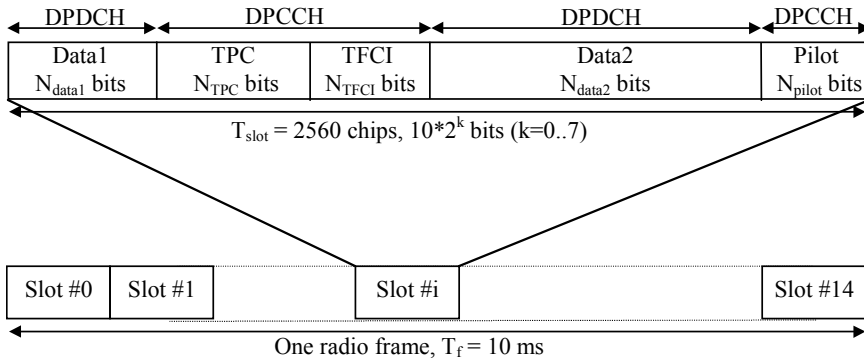
**Figura A.12 Proceso de entrelazado**

A diferencia de otros simuladores, como [4] o [5], donde la profundidad de entrelazado es solamente de 10ms, en el simulador implementado es posible definir la profundidad de entrelazado arbitrariamente, tal y como quede especificado en la RAB correspondiente.

### A.1.7 Segmentación del canal físico

Cualquiera de los canales físicos establecidos por el estándar [6] tiene una estructura de trama predefinida que se deberá de respetar. En concreto, en esta Tesis Doctoral solamente se van a implementar tres canales físicos, el canal dedicado DPCH (*Dedicated Physical Channel*), el canal compartido PDSCH (*Physical Downlink Shared Channel*) y el canal piloto CPICH (*Common Pilot Channel*).

La estructura de trama del canal físico DPCH se muestra resumida en la Figura A.13. El bloque TTI ya entrelazado debe de ser segmentado e insertado consecutivamente en los campos Data 1 y Data 2 reservados para la transmisión de información física.

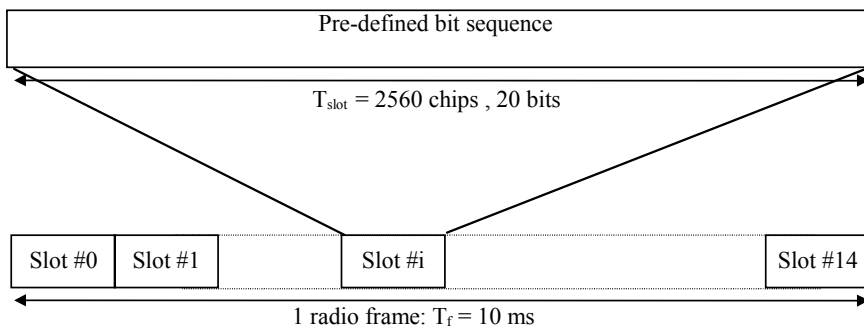


**Figura A.13 Estructura de trama del canal DPCCH [6]**

Al haberse ajustado previamente la tasa, la información contenida en el bloque TTI rellenará por completo todos los *slots* abarcados en el TTI (15 para TTI=10ms, 30 para TTI=20ms, etc.).

La estructura de los otros dos canales físicos considerados es mucho más sencilla ya que sólo incluyen un campo de datos. En el canal de CPICH se ensanchará una secuencia constante por un SF 256, mientras que en canal PDSCH el factor de ensanchado será variable en función de la decisión tomada por el gestor de recursos. La estructura de ambos canales se muestra también en las Figuras A.14 y A.15.

Aunque en el canal PDSCH no existe información de control, ésta está incluida en el canal dedicado asociado a todo usuario que trabaja en modo compartido. Gracias a este canal dedicado es posible realizar el control de potencia rápido en sentido descendente.



**Figura A.14 Estructura de trama del canal CPICH [6]**

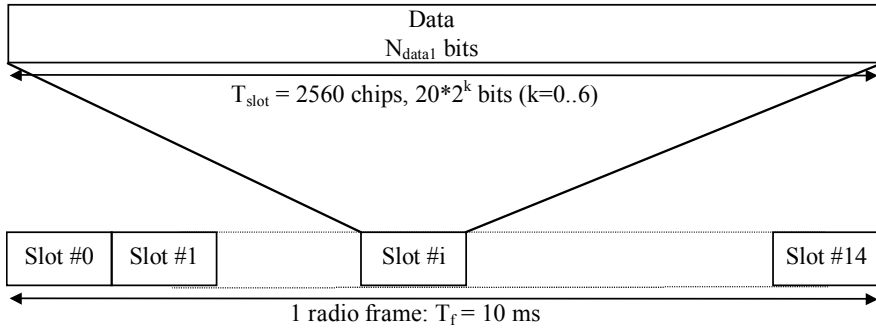


Figura A.15 Estructura de trama del canal PDSCH [6]

### A.1.8 Multiplexación del canal de control

En el enlace dedicado descendente, el canal DPDCH (*Dedicated Physical Data Channel*) y DPCCH (*Dedicated Physical Control Channel*) está multiplexados en el tiempo dentro de la trama, tal y como se puede observar en la Figura A.13. De esta manera, este bloque funcional se encarga de rellenar los campos asociados al canal de control, es decir los bits piloto y los asociados al control de potencia. Los bits reservados al TFCI y los bits del control de potencia se rellenan aleatoriamente puesto que el receptor conoce exactamente el formato de la transmisión y si en el anterior *slot* se ha satisfecho o no la relación  $E_b/N_0$  objetivo, por lo que no es necesario utilizar estos datos. Por el contrario, los bits piloto se introducirán tal y como establece el estándar, de manera que se evalúen los posibles errores cometidos en este canal.

### A.1.9 Ensanchado de la señal y modulación

Como se especifica en el estándar, el proceso de ensanchado por secuencia directa utilizado en WCDMA se caracteriza por la utilización de dos secuencias distintas. La primera de ellas, conocida como secuencia de canalización, utiliza códigos ortogonales

OVSF que en sentido descendente discrimina al receptor. La segunda secuencia, no ortogonal, se conoce como secuencia de aleatorización y sirve para identificar la fuente de la información, que en *downlink* es la estación base.

Los códigos OVSF se calculan de manera sencilla mediante el árbol de *Walsh* tal y como se recoge en las especificaciones. Por otro lado, los códigos de aleatorización se generan de manera pseudoaleatoria de acuerdo a un algoritmo especificado en el apartado 5.2.2 de [7]. Según el estándar, estos códigos se obtienen mediante la implementación mostrada en la Figura A.16 que hace uso de dos registros de desplazamiento a fin de emular los *buffers* circulares.

El proceso de ensanchado y la modulación se realizan simultáneamente. Así, tras separar los bits generados en dos ramas, una en fase y otra en cuadratura, se multiplicarán todos los bits por la misma secuencia de ensanchado. Una vez modulada y ensanchada, la señal se multiplicará por la secuencia de aleatorización (secuencia de números complejos), obteniendo la señal final lista para ser transmitida. La Figura A.17 resume este proceso donde  $w_{i,j}$  es la secuencia de canalización mientras que  $g_{cell}$  representa la secuencia de aleatorización. La modulación se realiza simplemente mediante un conversor serie-paralelo y la multiplicación por un número complejo.

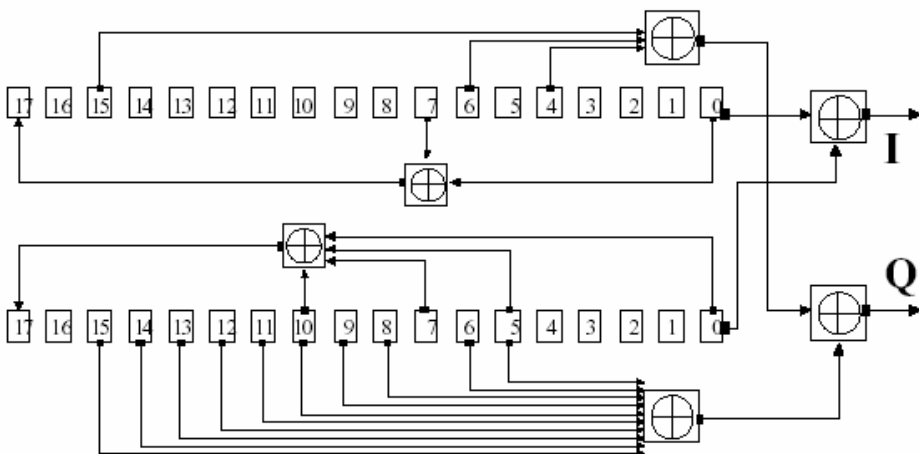


Figura A.16 Generador de códigos de aleatorización [7]

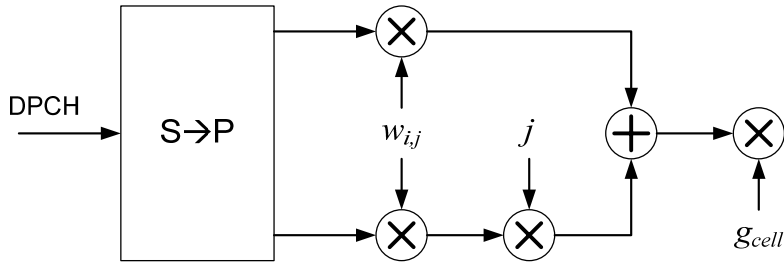


Figura A.17 Proceso de modulación y ensanchado en enlace descendente

### A.1.10 Filtrado Nyquist

Después del ensanchado y modulación banda base de la señal, ésta se pasa por un filtro raíz de coseno alzado con un factor de *roll-off*  $r = 0.22$ . Para realizar este proceso de conformación del pulso es necesario interpolar la señal discreta original con  $M-1$  ceros y filtrarla posteriormente con un pulso de coseno alzado con un factor de sobremuestreo respecto a la frecuencia de *chip* igual a  $M$ .

En esta Tesis Doctoral se ha seleccionado un valor de sobremuestreo  $M=5$ . La Figura A.18 muestra un ejemplo del pulso raíz de coseno alzado con sobremuestreo 5 y también el proceso seguido en el filtrado, desde la secuencia original,  $d[n]$ , hasta la secuencia ya filtrada,  $s[n]$ .

Finalmente, cabe destacar que en las simulaciones llevadas a cabo en esta Tesis Doctoral se ha trabajado siempre con el equivalente paso bajo del sistema, es decir, no se ha realizado la conversión en frecuencia a 2 GHz, por el ahorro computacional que esta práctica conlleva. Además, no se ha considerado el modelado de posibles fallos en la recuperación de la portadora, tarea que queda pendiente para futuras versiones del simulador.

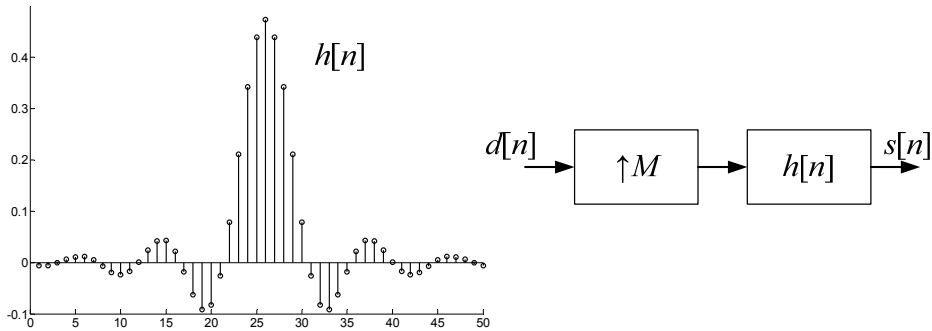
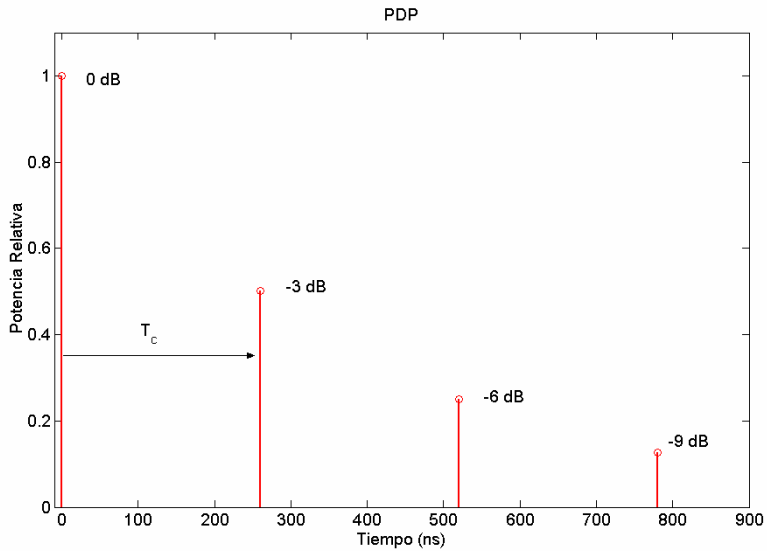


Figura A.18 Filtrado de Nyquist

### A.1.11 Canal de propagación

En cualquier sistema de comunicaciones móviles, entre la estación base y el móvil aparecen numerosos elementos influyentes (difusores o *scatterers*) en la propagación radioeléctrica, tales como montañas, edificios, coches, personas, etc. En la antena del móvil o de la estación base, se recibirán múltiples contribuciones o réplicas de la señal transmitida que habrán sufrido distintos mecanismos de propagación: reflexiones, difracciones, *scattering* (dispersión), etc. La señal recibida corresponderá a la suma coherente, en módulo y fase, de las distintas contribuciones que alcanzan al receptor. Tales contribuciones estarán desfasadas y atenuadas, unas con respecto a otras, como consecuencia de haber recorrido caminos de propagación diferentes. La superposición temporal de las contribuciones en recepción conforma un canal de comunicaciones complejo que puede variar en el tiempo debido al efecto de movilidad. Es evidente entonces que el entorno de propagación limitará la calidad de los sistemas de comunicaciones móviles, donde habrá que adoptar técnicas de transmisión para reducir los efectos que el canal móvil pueda introducir sobre las características de la señal recibida.



**Figura A.19** Ejemplo de perfil de retardo

En esta Tesis Doctoral, se ha implementado un modelo de canal multicamino con espectro *Doppler* clásico definido completamente por su perfil de retardo o PDP (*Power Delay Profile*), es decir, por el número de caminos considerado, su retardo relativo y la potencia media asignada a cada componente. Un ejemplo de PDP se muestra en la Figura A.19.

Para el modelado del canal se ha trabajado bajo la hipótesis de *scattering* incorrelado en la variable de retardo, lo que permite generar cada contribución multicamino por separado con su estadística y espectro *Doppler* particular.

Considerando que cada contribución recibida se compone de la superposición de infinitas contribuciones elementales que alcanzan al receptor con el mismo retardo, y aplicando el Teorema Central del Límite se puede concluir que la respuesta  $h(t, \tau_n)$ , que modela la variación en el tiempo de la contribución  $n$ -ésima, podrá considerarse como una variable aleatoria compleja gaussiana de la forma:

$$h(t, \tau_n) = h_r(t, \tau_n) + jh_i(t, \tau_n) \quad (\text{A.16})$$

donde la parte real,  $h_r(t, \tau_n)$ , y la parte imaginaria,  $h_i(t, \tau_n)$ , siguen una distribución gaussiana, de modo que la fase de  $h(t, \tau_n)$  presenta una distribución uniforme y su módulo una distribución *Rayleigh*. El problema del modelado del canal se reduce entonces a generar dos variables *gaussianas* con el espectro *Doppler* requerido. El procedimiento seguido en esta Tesis Doctoral consiste en generar dos señales gaussianas independientes, una para la parte real o componente en fase y otra para la parte imaginaria o componente en cuadratura. Ambas señales se pasarán por un filtro de respuesta  $\sqrt{S(f)}$  que define la densidad espectral establecida por el espectro *Doppler* de la contribución multicamino. Cuando se considera *scattering* isotrópico y antenas receptoras omnidireccionales, se tiene que

$$S(f) = \frac{1}{\sqrt{1 - \left(\frac{f}{f_d}\right)^2}} \quad (\text{A.17})$$

donde  $f_d$  corresponde a la frecuencia *Doppler* máxima dada por

$$f_d = \frac{v}{\lambda} \quad (\text{A.18})$$

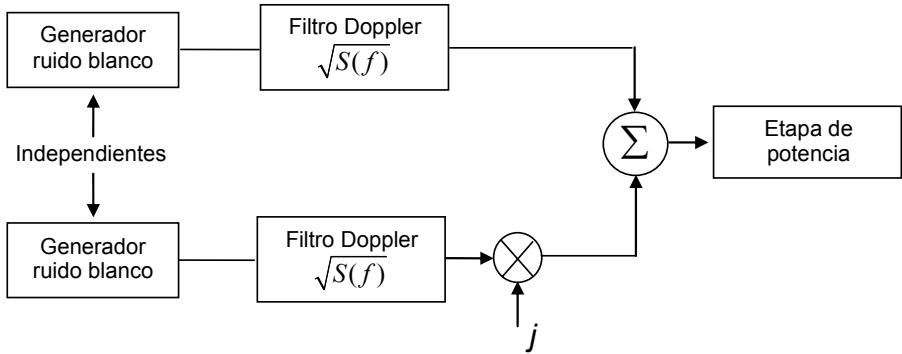
siendo  $v$  la velocidad del terminal en m/s y  $\lambda$  la longitud de onda de la señal electromagnética portadora.

Las componentes en fase y cuadratura filtradas pasarán por una etapa de potencia donde se ajustará la amplitud de la señal recibida manteniendo la forma de su espectro.

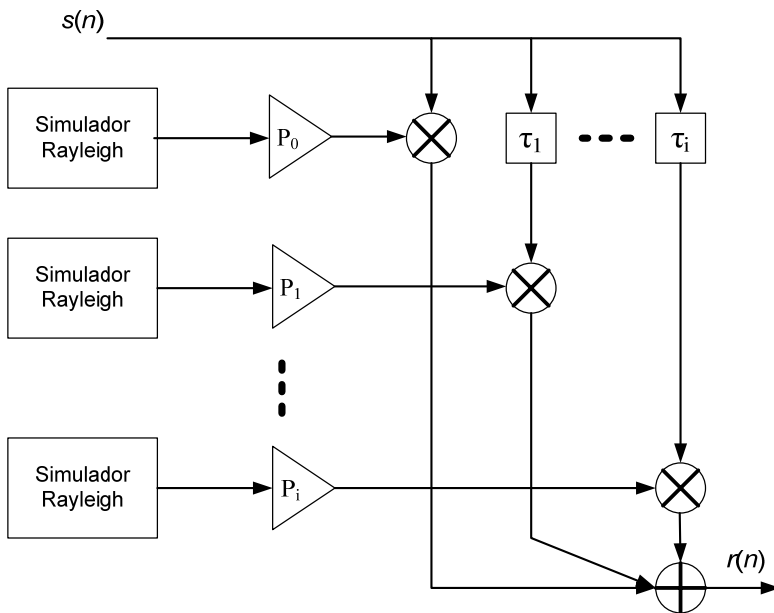
De ese modo se consigue una señal con envolvente *Rayleigh* y potencia la definida por el perfil de retardo. El diagrama de la Figura A.20 resume el procedimiento seguido de generación de las muestras de canal.

Finalmente, y debido a la decorrelación entre las componentes multicamino, la señal recibida se obtiene como la suma de las diversas contribuciones de la señal retardada y multiplicada por la muestra de canal correspondiente a cada uno de los multicaminos. En la Figura A.21 se puede observar claramente este proceso.





**Figura A.20** Generación de señal de envolvente *Rayleigh* y espectro *Doppler* clásico



**Figura A.21** Implementación del canal como filtro FIR

La mayor dificultad del método seguido para generar las muestras de canal reside precisamente en la implementación del filtro *Doppler*. Cuando la respuesta al impulso del filtro se obtiene a partir de la transformada de Fourier inversa, para frecuencias *Doppler* bajas se requiere aumentar el número de muestras del espectro conforme aumenta la frecuencia de muestreo (la inversa del espaciado entre muestras de la

respuesta al impulso del canal), lo que se traduce en un coste computacional elevado con la posible pérdida de resolución en las características estadísticas del *tap*. Este tipo de problemas se puede solucionar fácilmente mediante técnicas de interpolación [8].

Finalmente, es importante destacar que el desvanecimiento asociado al *shadowing* no se ha incluido en la simulación de enlace, ya que se ha asumido que el control de potencia en bucle cerrado es capaz de compensarlo.

### A.1.12 Inserción de ruido

En las simulaciones de enlace, en las que se trabaja con frecuencias de muestreo del orden de 5-10 veces la frecuencia de *chip* ( $f_c = 3.84$  MHz), es imposible plantear un escenario con múltiples estaciones base y usuarios en movilidad. Por el contrario, y acogiéndose al Teorema Central del Límite que postula que la suma de  $N$  variables aleatorias tiene las mismas características estadísticas que una variable gaussiana cuando  $N$  tiende a infinito, se puede suponer un único enlace Estación Base – Usuario. De esta forma, el conjunto de interferencias provenientes de las demás estaciones base y de todos los usuarios del sistema se modela como un ruido AWGN sobre la señal de interés.

La varianza del ruido que es necesario introducir en el canal para obtener un determinado nivel de  $E_b/N_0$  se calcula a partir de la siguiente expresión:

$$\sigma_N^2 = \frac{\sigma_R^2 \cdot G_p}{E_b / N_o} \quad (\text{A.19})$$

donde  $\sigma_N$  es la desviación típica del ruido AWGN,  $\sigma_R$  es la desviación típica de la señal recibida después de atravesar el canal multicamino y  $G_p$  representa la ganancia de procesado definida como la relación entre el ancho de banda efectivo del sistema ( $W=3.84$  MHz) y la tasa binaria del servicio (por ejemplo, para el servicio de voz  $R_{\text{voz}}=12.2$  kbps).

Por otro lado, la potencia de la señal recibida depende del PDP y de las colas asociadas a los pulsos transmitidos, de manera que se puede definir dicha potencia de la siguiente manera:

$$\sigma_R^2 = \sigma_S^2 \cdot \left\{ \sum_{m=-L}^L \left( \sum_{i=1}^N PDP(\tau_i) \cdot p^2(\tau_m - \tau_i) \right) \right\} \quad (\text{A.20})$$

donde  $\sigma_S$  es la desviación típica de la señal transmitida,  $PDP(\tau_i)$  es la potencia asociada al camino de propagación  $i$ -ésimo retrasado  $\tau_i$  con respecto a la primera componente,  $p(t)$  es la respuesta al impulso del filtro raíz de coseno alzado con  $2L-1$  muestras y  $N$  es el número de caminos de propagación considerados.

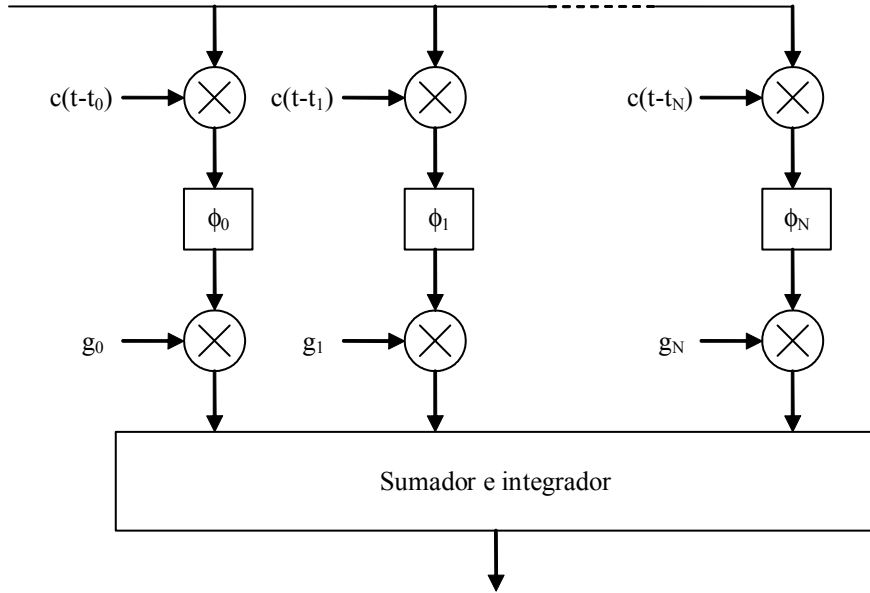
Utilizando la expresión (A.19) y (A.20) es posible calcular exactamente la potencia de ruido que se debe sumar a la señal recibida para emular la relación  $E_b/N_0$  deseada.

### A.1.13 Receptor *Rake*

En sistemas CDMA es posible utilizar un esquema de recepción, conocido como receptor *Rake*, que permite no sólo contrarrestar el efecto de la propagación multicamino, sino además aprovecharlo como mecanismo de diversidad. La idea en la que se basa el receptor *Rake* es intentar recuperar la porción de energía de señal contenida en cada trayecto y sumarlas constructivamente.

El receptor *Rake* consta de un conjunto de  $N$  ramas (*fingers*), en cada una de las cuales se correla la señal recibida con una versión de la secuencia código retardada, donde el desplazamiento temporal es el correspondiente a la componente multicamino que se intenta recuperar. Las salidas de las  $N$  ramas se integran y se combinan para obtener la variable de decisión. La combinación consiste en compensar los desfases relativos y realizar una suma ponderada por la amplitud con que se recibe cada componente, utilizando la estimación de los coeficientes de canal. La Figura A.22 muestra la estructura básica de un receptor *Rake*, donde  $c(t)$  representa el conjunto del código de canalización y aleatorización.

Señal recibida

Figura A.22 Receptor *Rake*

De acuerdo con lo comentado, el receptor puede considerarse un sistema de diversidad de orden  $N$  con combinación óptima MRC, en el que las diferentes réplicas de la señal se separan en virtud de las propiedades de autocorrelación de los códigos utilizados. Este hecho pone un límite a la capacidad del receptor *Rake* para separar las componentes multicamino, puesto que si los ecos se reciben muy próximos, con un retardo inferior a un periodo de *chip*, la función de autocorrelación toma valores elevados y la interferencia entre las contribuciones es importante. Así, la resolución temporal, o capacidad de aislar componentes multi-trayecto retardadas, viene determinada por el periodo de *chip*,  $T_c$ . Otro aspecto que es necesario tener en cuenta es el número de ramas del receptor. En principio al aumentar dicho parámetro mejora la calidad de recepción de la señal, pero también se incrementa la complejidad y el coste del receptor. Por otra parte, según aumenta el número de ramas, la mejora es cada vez más pequeña, puesto que el nivel de las componentes detectadas (*Power Delay Profile*) decrece con el retardo. Por todo ello, es preciso limitar el número de ramas a una cantidad que permita obtener un porcentaje elevado de la energía de la señal, más allá

del cual la adición de nuevos trayectos suponga una mejora limitada. Esto se consigue habitualmente considerando de 4 a 8 trayectos.

Además de contrarrestar el efecto multicamino, el receptor *Rake* juega un papel importante en los procedimientos de trasposos o *handovers*. En esta situación cada terminal móvil recibe la misma señal desde varias estaciones base, cada una de las cuales emplea una secuencia código distinta. El receptor *Rake* permite separar las señales procedentes de las distintas estaciones base utilizando las secuencias de código correspondientes en cada una de sus ramas, para a continuación combinarlas adecuadamente.

En todo sistema de comunicaciones digitales, el receptor basado en un filtro adaptado debe incluir circuitos de recuperación de portadora y de sincronismo de símbolo. El primero es necesario para conocer la frecuencia y fase de la portadora, de modo que pueda hacerse una demodulación coherente. El segundo determina el instante en el que se muestrea la salida del correlador o filtro adaptado, que debe coincidir con el final del periodo de símbolo. En sistemas CDMA, o de espectro ensanchado en general, debe existir además un subsistema de sincronismo de código que permita alinear la señal código en el receptor con la señal recibida, de modo que pueda llevarse a cabo el desensanchado de ésta. El proceso de sincronización de la señal código, realizado localmente a partir de la señal ensanchada recibida, suele realizarse en dos etapas:

1. Etapa de adquisición, o sincronismo grueso, en la que se consigue una alineación con un error del orden del periodo de *chip*  $T_c$ . El funcionamiento de esta etapa se basa en las propiedades de autocorrelación de las secuencias de código utilizadas.

2. Etapa de seguimiento, o sincronismo fino, que trata de mantener el error de alineación lo más pequeño posible. Habitualmente se realiza mediante un bucle de realimentación, que compara las dos señales y corrige las desviaciones entre ambas. Esta etapa sólo funciona correctamente si el error de entrada está acotado y comprendido en un margen reducido, normalmente  $\pm T_c/2$ . De ahí la necesidad de la etapa de adquisición.

Debe destacarse que el sincronismo de *chip* debe llevarse a cabo antes que la recuperación de portadora (y por tanto de forma no coherente), ya que sin el sincronismo de *chip* no es posible desensanchar para elevar la relación señal a ruido a los niveles adecuados para la recuperación de portadora y demodulación.

El receptor debe incluir también un subsistema de estimación de canal que proporcione los valores estimados de los coeficientes de canal, que permiten llevar a cabo la combinación MRC de las salidas de las ramas. Para ello se utilizan símbolos piloto, conocidos por el receptor, intercalados en la señal. La estimación se basa en una operación de correlación, que proporciona información sobre la amplitud y fase de cada una de las componentes de señal. En los periodos entre símbolos piloto puede utilizarse la estimación correspondiente al grupo de símbolos piloto más reciente o bien un método guiado por decisión, que actualiza la estimación en función de los símbolos de información y las decisiones del circuito de decisión. Además, es habitual la realización de un filtrado de los valores estimados, con objeto de disminuir en lo posible el ruido de estimación, manteniendo una capacidad de seguimiento suficiente para adaptarse a las condiciones cambiantes del entorno de propagación.

En concreto en el simulador desarrollado presenta una resolución temporal de un periodo de *chip* (260 ns). Las ramas o *fingers* del *Rake* se actualizan dinámicamente para ajustarse temporalmente a los puntos de máxima energía de la respuesta al impulso del canal, asumiendo pues un conocimiento perfecto de los retardos asignados a cada componente multicamino. Además el número de ramas consideradas en el *Rake* es un parámetro de la simulación que puede ser modificado.

Por otro lado, es muy importante el proceso de estimación de canal realizado en recepción. Los símbolos piloto utilizado para tal fin se han implementado en el canal físico de control DPCCCH siguiendo el formato de *slot* correspondiente [6]. Estos símbolos piloto se utilizan para estimar el canal mediante técnicas de estimación de error de fase y amplitud. Asimismo se transmite el canal piloto (CPICH), por lo que se puede hacer un uso adicional de esta información para realizar la estimación del canal.

El resto de bloques funcionales no son más que el proceso inverso de los ya mencionados en los apartados anteriores, por lo que se omite una descripción pormenorizada de los mismos.

Para más información sobre simulación de sistemas de comunicaciones, y en particular sobre los fundamentos seguidos en el simulador diseñado en esta Tesis Doctoral, se recomienda la lectura adicional de [9].

## A.2 Validación de los Bloques Funcionales

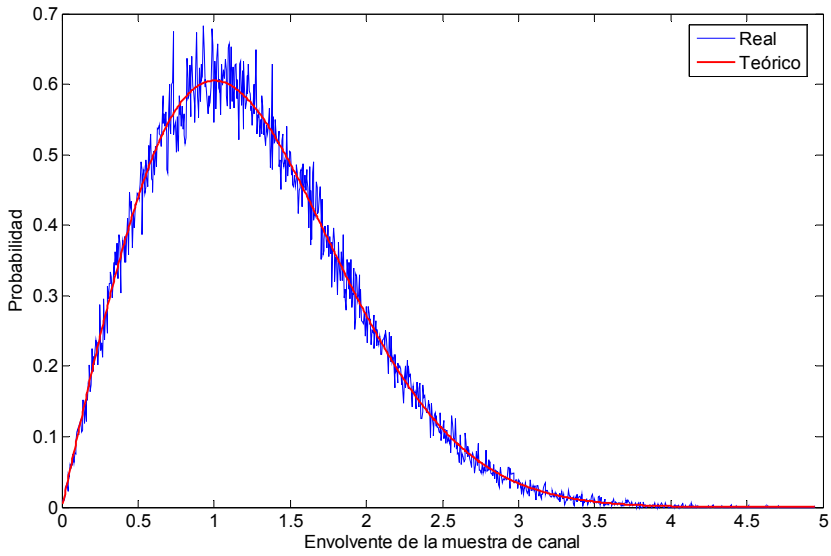
En este apartado se van a analizar los dos modelos clave implementados en el simulador: el modelo de canal y la codificación de canal. El resto no requieren de validación ya que su implementación sigue fielmente el estándar. En este anexo la validación se va a realizar a nivel funcional de los bloques desarrollados, mientras que la validación de los resultados obtenidos mediante el simulador se incluye en el capítulo 3 de la Tesis.

### A.2.1 Validación del modelo de canal

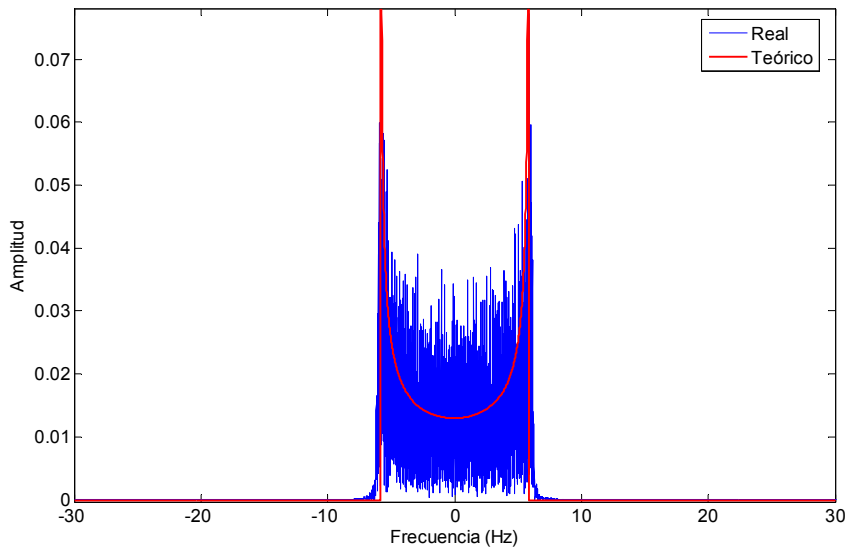
Para validar el modelo de canal se debe comprobar todo el proceso de generación de las muestras de canal. En primer lugar se ha analizado la generación de la envolvente compleja con distribución *Rayleigh*. Para ello se ha analizado, a partir de información extraída del simulador, la envolvente de la componente de canal de un tap concreto. La Figura A.23 muestra la validez del modelo de generación de las muestras de canal *Rayleigh* a partir de variables de la suma compleja de dos distribuciones gaussianas ya que, como se puede observar, la secuencia generada se aproxima claramente a una distribución *Rayleigh* teórica.

Por otro lado, en la Figura A.24 se representa la densidad espectral de potencia de un canal UMTS experimentado por un terminal que se desplaza a una velocidad de 3 km/h ( $f_d = 5.83\text{Hz}$ ). Como se puede observar, el método de generación basado en un filtro FIR de la señal *Rayleigh* provee una aproximación al espectro deseado bastante satisfactoria, incluso para frecuencias *doppler* bajas, situación más conflictiva en la generación de las muestras de canal.





**Figura A.23** Envolvente *Rayleigh*



**Figura A.24** Espectro de un tap tipo *Rayleigh* a 3 km/h

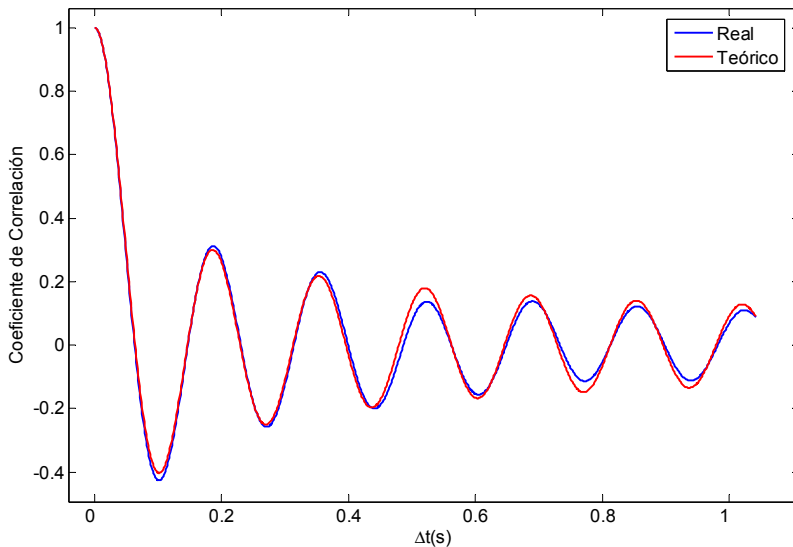
Una vez comprobados estos dos aspectos del modelado del canal, sólo resta asegurarse de que los estadísticos de segundo orden se corresponden también con los esperados. La función de autocorrelación teórica de la componente real o imaginaria del canal es:

$$R_{x_r} = J_0(2\pi f_d \tau) \quad (\text{A.21})$$

En la Figura A.25 se muestra la correlación experimental calculada a partir de una secuencia de canal generada en el simulador, comparándola con su expresión teórica. Como se puede observar, los resultados obtenidos se aproximan mucho a los esperados, lo que corrobora la validación del modelo de canal implementado.

## A.2.2 Validación de la codificación de canal

Para validar la codificación de canal se ha separado el análisis de la codificación convolucional, característica del servicio de voz, de la turbo codificación, empleada para el resto de servicios.



**Figura A.25** Coeficiente de correlación de la componente real del canal

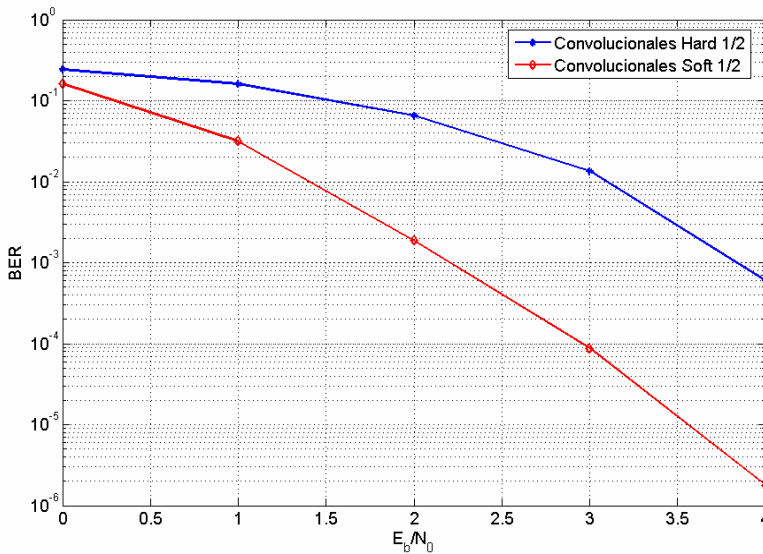
### ***Validación de la codificación convolucional***

Para la obtención de los resultados de la codificación convolucional, se han implementado una serie de programas tanto en Matlab como en C++. Los programas realizados en MATLAB han servido para validar los resultados a nivel físico, mientras que los de C++ han sido incluidos en el simulador de nivel de enlace para obtener las prestaciones de este tipo de codificación en un entorno simulado de UMTS.

Antes de incluir el codificador y decodificador convolucional en el simulador de nivel de enlace, se validaron sus resultados a nivel físico comparándolos con los obtenidos utilizando la herramienta matemática Matlab. Para ello, se implementó en Matlab un sistema que realizaba la codificación convolucional en un entorno con ruido blanco aditivo gaussiano, AWGN y con una modulación BPSK, es decir, los niveles de la señal eran +1 y -1, como ocurre en el simulador de nivel de enlace. Estos datos pasaban a ser codificados y decodificados tanto por las propias herramientas de Matlab como por los objetos de codificación y decodificación implementados en esta Tesis Doctoral.

Se han implementado las dos versiones de la codificación convolucional, la firme y la indecisa, considerando para ambas dos diferentes tasas de codificación, tasa 1/2 y tasa 1/3.

La Figura A.26 compara la decodificación convolucional firme e indecisa para una tasa de 1/2. Como se puede observar, la mejora de la codificación indecisa respecto a la codificación firme es ligeramente superior a 1.5dB. Esto quiere decir, que para un  $E_b/N_0$  dada, se obtiene una menor probabilidad de error en el caso de la decodificación indecisa. Dicho de otro modo, una vez fijada una determinada BER, con la codificación indecisa se puede trabajar con  $E_b/N_0$  menores, es decir, en canales con un mayor nivel de ruido. Estos resultados están en consonancia con las conclusiones alcanzadas en [5], por lo que se refuerza la validez de la implementación realizada. Es importante destacar que los resultados obtenidos coinciden con los obtenidos con los programas proporcionado por Matlab, lo que aumenta aún más su validez. Además, los resultados también están en consonancia con otros estudios encontrados en la bibliografía.



**Figura A.26 Comparativa codificación firme 1/2 frente a codificación indecisa 1/2**

A continuación, se ha realizado un estudio de la evolución de las gráficas de BER al aumentar el número de niveles de cuantificación, previos a la decodificación, tanto para codificadores de tasa 1/2 como para codificadores de tasa 1/3. El número de niveles en la cuantificación influye en el comportamiento del decodificador. En particular, se han simulado niveles de cuantificación de 2 (decodificación firme), 4 y 8.

La ganancia obtenida gracias a la codificación no aumenta linealmente con el número de niveles de cuantificación. La mejora entre utilizar 4 niveles de cuantificación frente a 2 niveles es mayor que la ganancia de pasar de 4 a 8 niveles. Por ello, es necesario encontrar un compromiso entre niveles de cuantificación y complejidad computacional del algoritmo. La gráfica marcada con trazo rojo, corresponde con la situación en la que el número de niveles de cuantificación es infinito, es decir, no existe cuantificación. Como se puede ver, la diferencia entre esta situación y usar 8 niveles de cuantificación es menor de 0.25dB.

Además, una tasa de codificación de 1/3 protege más los bits de información, lo que se traduce en un menor nivel de  $E_b/N_0$  requerido para una BER determinada.

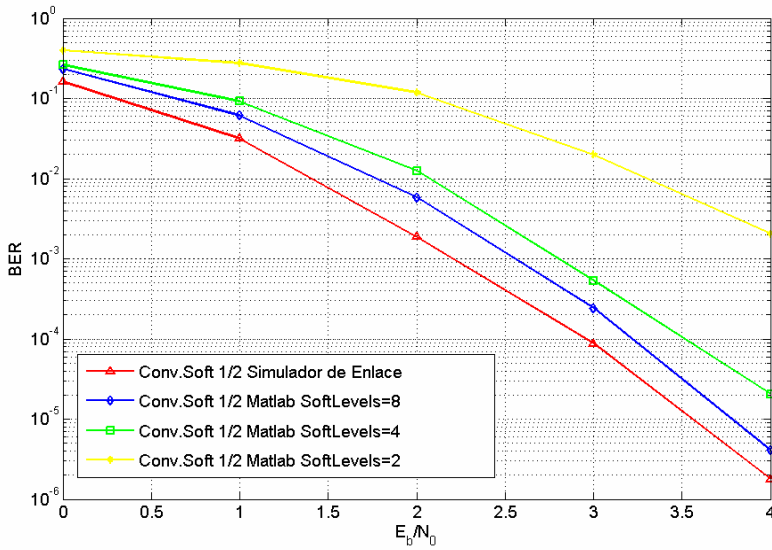


Figura A.27 Evolución de la BER con varios niveles de cuantificación de tasa 1/2

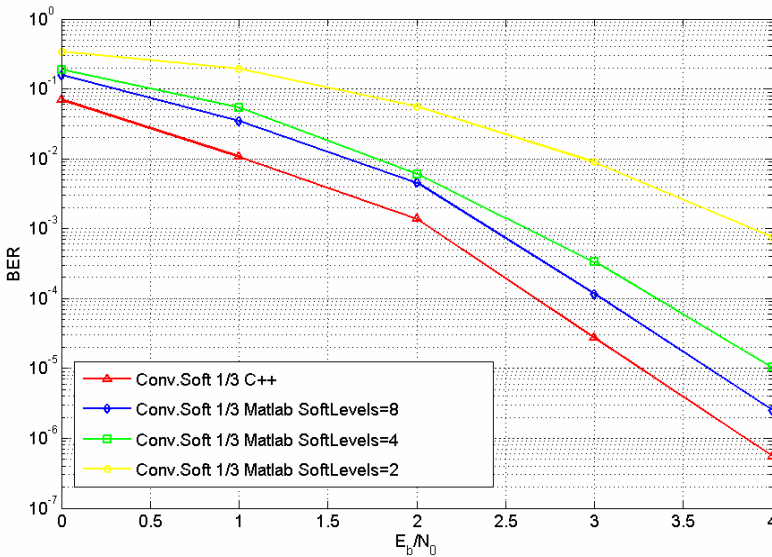
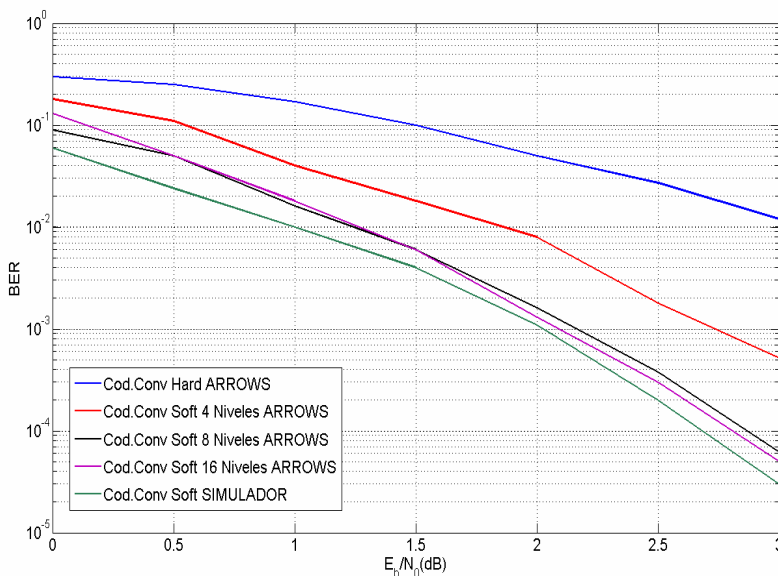


Figura A.28 Evolución de la BER con varios niveles de cuantificación de tasa 1/3

Estos resultados sirven para validar la implementación realizada de los bloques de codificación, ya que se han comparado los resultados obtenidos con las herramientas de codificación y decodificación incluidas en Matlab (curvas amarilla, verde y azul) con los del simulador de nivel de enlace implementado (curva roja). Como se puede ver, los resultados obtenidos con el simulador de enlace son mejores que los obtenidos por Matlab. Esto es debido a que el simulador implementado no realiza una cuantificación de la señal en niveles indecisos, al contrario que Matlab. Observando más detalladamente las curvas, se ve como al aumentar el número de niveles indecisos en Matlab, las curvas tienden a converger hasta la solución del simulador de nivel de enlace. Por todo ello, puede decirse que los resultados del codificador convolucional son correctos.

A continuación se muestran los resultados de una codificación de tasa 1/3 considerando modulación QPSK. En este caso se han comparado los resultados obtenidos mediante el simulador de enlace desarrollado en la Tesis con los resultados del proyecto europeo ARROWS [4]. De nuevo, el simulador desarrollado presenta los mejores resultados ya que no realiza cuantificación.

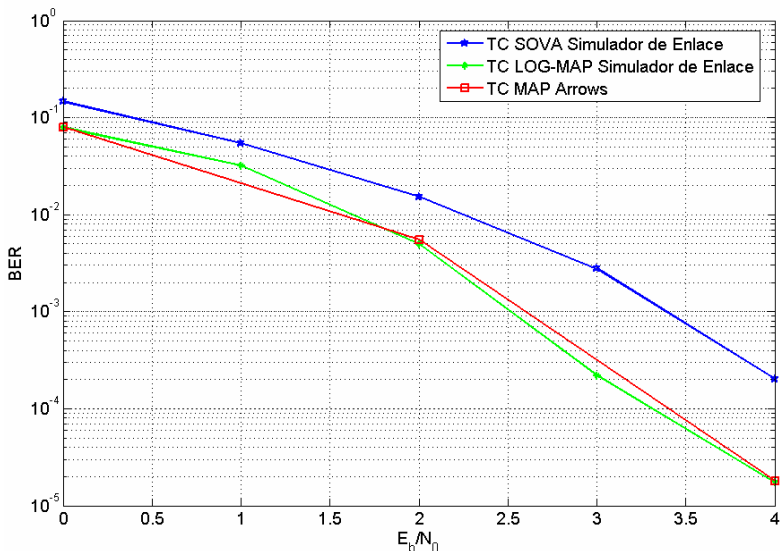


**Figura A.29 Resultados a nivel físico de la codificación convolucional en el proyecto ARROWS y en el simulador desarrollado**

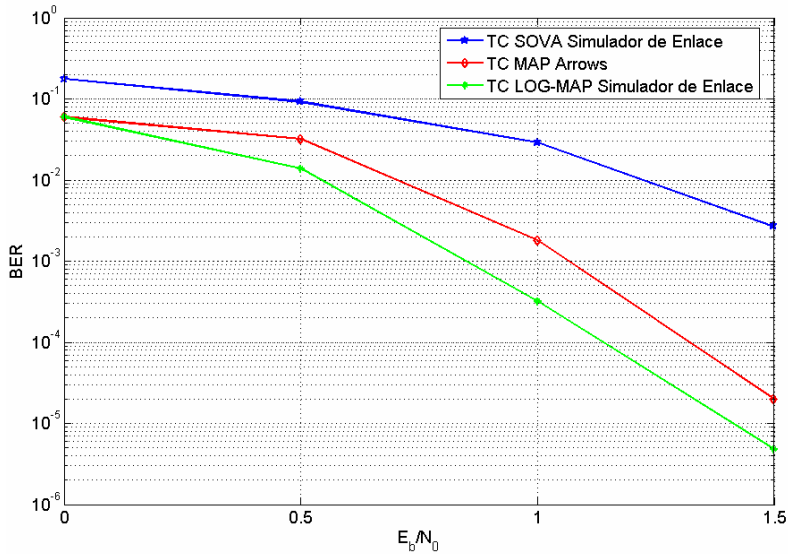
### Validación de la turbo-codificación

Respecto a la turbo codificación, se ha seleccionado el trabajo presentado en [4] para validar las prestaciones de los codificadores y decodificadores. Las Figuras A.30 y A.31 muestran los resultados obtenidos para un SF 32 y una velocidad del terminal de 50 km/h considerando bloques de información de 40 y 512 bits respectivamente. Lo primero que llama la atención es la enorme importancia que tiene el algoritmo de decodificación empleado. En este sentido se llega a la misma conclusión postulada en [3]: la decodificación Log-Map es la más potente. Por este motivo esta decodificación ha sido la implementada en el simulador de enlace, siendo mejor que la decodificación MAP utilizada en [4].

Además, los resultados obtenidos demuestran que la decodificación obtiene mejores prestaciones al aumentar el tamaño del bloque de información a codificar. Esta conclusión también está en consonancia con el trabajo realizado por Hanzo en [3]. Finalmente, es importante destacar que la decodificación implementada en el simulador desarrollado obtiene mejores prestaciones incluso que las obtenidas en [4] ya que la técnica de decodificación es más potente.



**Figura A.30 Comportamiento de los decodificadores para bloques de información de 40 bits**



**Figura A.31 Comportamiento de los decodificadores para bloques de información de 512 bits**

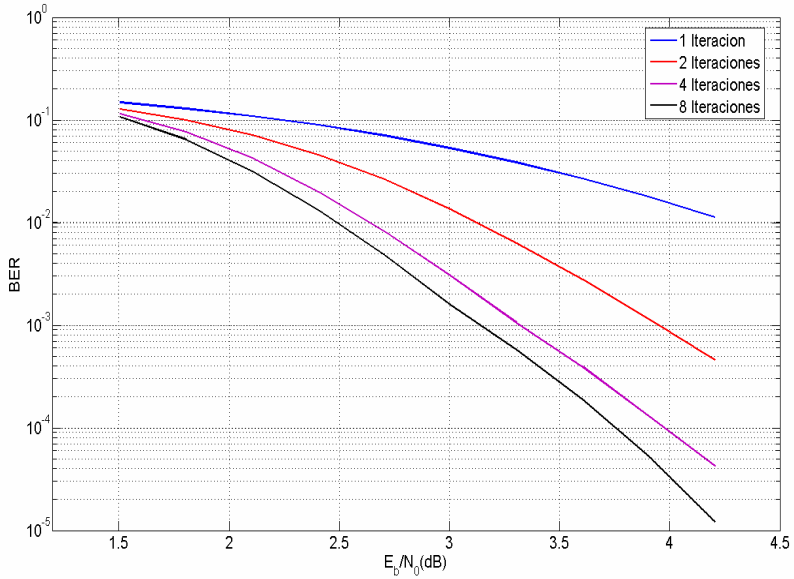
Además, la mejora respecto a [4] se incrementa conforme aumenta el tamaño del bloque de información codificado. Este hecho está también en consonancia con otros estudios [3] que defienden el mejor comportamiento del algoritmo Log-MAP para tamaños crecientes del bloque de datos codificado.

Por último, se ha evaluado el comportamiento del algoritmo de decodificación Log-Map en función del número de iteraciones empleadas en la decodificación. Los resultados pueden verse en la Figura A.32.

Como se puede observar, el número de iteraciones influye considerablemente en las prestaciones del turbo código. Para  $E_b/N_0$  bajas, las prestaciones obtenidas por el turbo codificador son similares independientemente de las iteraciones. Para  $E_b/N_0$  altas, a mayor número de iteraciones, mejor comportamiento de la BER, aunque la mejora de prestaciones deja de ser significativa cuando el número de iteraciones sobrepasa un determinado umbral. En la práctica suelen tomarse entre 5-8 iteraciones, ya que, para estos valores, un aumento en el número de iteraciones conlleva una mejora en las



prestaciones insignificante, que no compensa el aumento de la complejidad y tiempo de respuesta del decodificador.



**Figura A.32 Comportamiento del decodificador Log-Map en función del número de iteraciones de decodificación**

## Referencias

- [1] 3GPP TS 25.212 V5.6.0, “Multiplexing and channel coding (FDD)”, 2003.
- [2] C.Berrou, A. Glavieux, “Near Optimum Error Correcting Coding and Decoding: Turbo Codes”, IEEE Transactions on Communications, vol. 44, pp. 1261-1271, 1996.
- [3] L. Hanzo, “Turbo Coding, Turbo Equalisation and Space-Time Coding for Transmission over Fading Channels”, John Wiley & Sons, 2002.
- [4] IST-2000-25133, “Advanced Radio Resource Management for Wireless Services (ARROWS)”.
- [5] J.J.Olmos, S.Ruiz, COST273: MORANS, “Description of UTRA-FDD Link Level Simulator”, May 2003.
- [6] 3GPP TS 25.211 V3.12.0, “Physical Channels and Mapping of Transport Channels onto Physical Channels (FDD)”, 2002.
- [7] 3GPP TS 25.213 V7.1.0, “Spreading and Modulation (FDD)”, 2007.
- [8] Lorenzo Rubio, “Técnicas de ecualización adaptativa en comunicaciones móviles basadas en estimación y seguimiento de canal”, Tesis Doctoral, 2003.
- [9] M. C. Jeruchim, P. Balaban and K. S. Shanmugan, “Simulation of Communication Systems. Modeling, methodology, and techniques”, Hlumer Academic/ Plenum publishers, 2000.

# Anexo B

## *Descripción del Simulador de Sistema*

Para evaluar los algoritmos de gestión de recursos desarrollados específicamente para el MNB es necesario disponer de un simulador a nivel de sistema que permita conocer su comportamiento global. Dado que el modelado del sistema es un tema muy específico cuya descripción no afecta al entendimiento de los algoritmos propuestos y sus prestaciones, se ha preferido relegar la mayor parte de su descripción a este anexo.

Es importante destacar que algunos de los modelos propuestos en el modelado del sistema UMTS han sido utilizados con posterioridad en distintos foros de trabajo europeos, como COST 273 y NEWCOM, lo que demuestra la importancia que el trabajo de modelado ha supuesto en el desarrollo de la Tesis Doctoral.

Este anexo está compuesto básicamente de dos apartados. El primero describe cada uno de los módulos del simulador de sistema y el segundo muestra cómo se realiza la interacción entre las simulaciones de enlace y de sistema.

### **B.1 Modelo del Sistema**

A la hora de definir el modelo de simulación empleado para evaluar el comportamiento de los algoritmos de gestión adaptativa de recursos en redes MNB, es necesario alcanzar un equilibrio entre la complejidad computacional del simulador y la validez de los resultados alcanzados. Dado que se pretende evaluar el efecto producido por la red distribuida de antenas que conforman el MNB en términos de capacidad

global del sistema, parece razonable seleccionar un modelo de simulación que garantice que las características de macro-diversidad del MNB, resumidas en el Capítulo 3 en la forma de *look-up tables*, queden perfectamente modeladas. El modelo de simulación de sistema es por tanto un elemento clave en el desarrollo de la Tesis Doctoral.

El objetivo de este anexo es describir, con el máximo rigor posible, el simulador utilizado para modelar el enlace descendente de la red de acceso radio de UMTS, UTRAN y su evolución arquitectónica MNB propuesta en esta Tesis Doctoral. La finalidad principal de dicho simulador es estudiar distintos algoritmos de gestión de recursos específicos para la red MNB, evaluando la posible ganancia que el uso de dicha arquitectura puede suponer respecto a la arquitectura de red de acceso radio UMTS convencional.

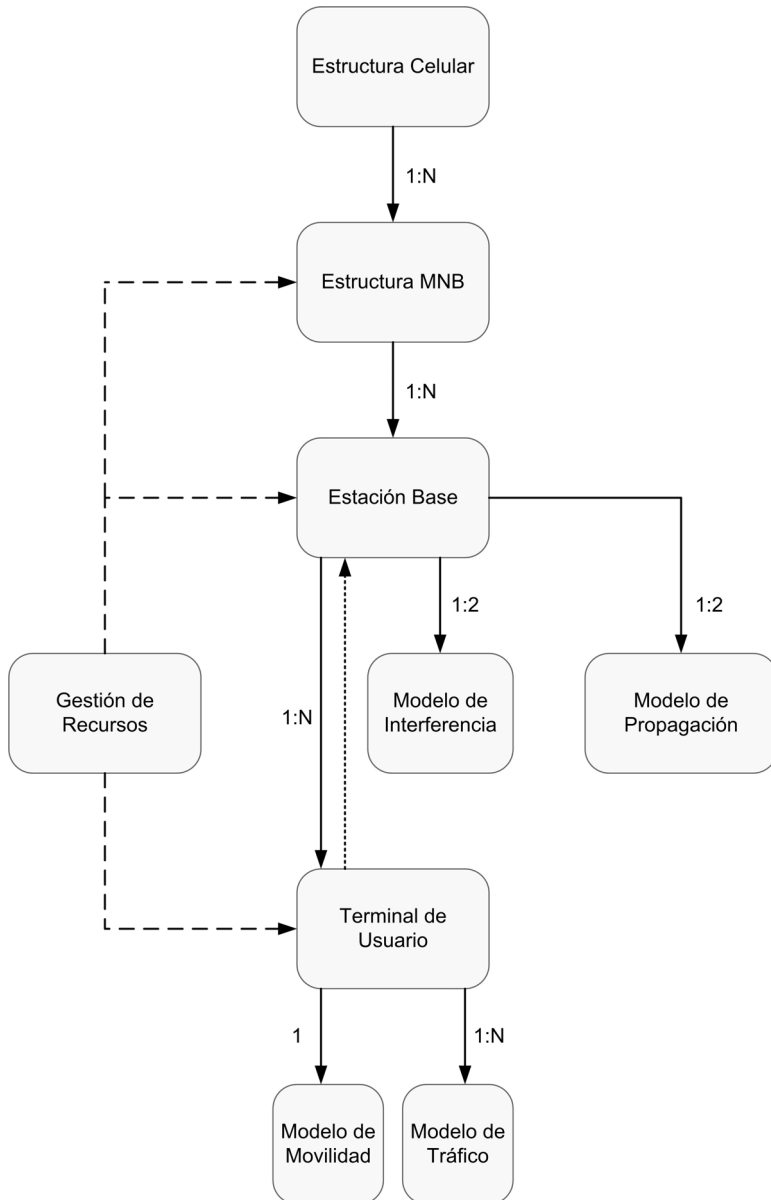
## **B.1.1 Entidades funcionales**

Una fase fundamental en el desarrollo de una herramienta de simulación es identificar las entidades funcionales que integran el sistema que se desea simular y desarrollar modelos o abstracciones que permitan caracterizar el comportamiento de cada una ellas. El grado de detalle con el que se modela cada entidad debe estar en consonancia con su importancia sobre el aspecto que se desea estudiar, con el objetivo de lograr un buen compromiso entre la complejidad del simulador y la precisión de los resultados obtenidos.

La Figura B.1 muestra las principales entidades consideradas en el simulador desarrollado en la presente Tesis Doctoral. Las líneas continuas representan posesión, las líneas discontinuas control y la línea de puntos conocimiento.

La estructura celular define el escenario sobre el cual se desarrollan las simulaciones, estableciendo la disposición geográfica de los Nodos B que conforman la red de acceso del sistema simulado. Cada uno de los Nodos B puede estar estructurado como un Multi-Nodo B así que de manera genérica se puede decir que la estructura celular

dispone de un conjunto de estructuras MNB desplegadas sobre el escenario de simulación. Cada MNB a su vez está formado por un conjunto de transmisores o estaciones base y ejecuta los algoritmos definidos de gestión de recursos dentro del MNB para realizar la distribución de carga entre los transmisores.



**Figura B.1 Entidades modeladas en el simulador de sistema**

Por su parte, la entidad estación base se encarga de administrar el conjunto de recursos asignados por el MNB, según determinen los algoritmos de gestión de recursos. El modelo de propagación caracteriza las condiciones de propagación entre una estación base y un terminal móvil. Cada transmisor, en función de las características de su emplazamiento, se considerará como un emplazamiento macro-celular o micro-celular, existiendo para cada uno de ellos un modelo de propagación y de interferencias particular.

La entidad terminal móvil modela el comportamiento de cada uno de los usuarios del sistema, centrándose en la caracterización de su movilidad y del tráfico generado por ellos. Cada terminal móvil conoce todas las estaciones base de las que está recibiendo potencia, tanto útil como interferente, y así, conociendo la potencia transmitida, el modelo de interferencias y su modelo de canal, puede calcular la calidad de señal recibida.

Por último, el bloque de gestión de recursos engloba el conjunto de algoritmos empleados para administrar los recursos del sistema entre los usuarios que acceden a él.

En los siguientes apartados se describirán los modelos empleados para caracterizar cada uno de los elementos que componen el simulador, salvo la descripción de los algoritmos de gestión de recursos que se realiza en los Capítulos 4 y 5. En la medida de lo posible, se han intentado emplear modelos de uso generalizado, que permitan comparar los resultados obtenidos con los de otros autores, aunque adaptando dichos modelos a las necesidades particulares del simulador desarrollado. En este sentido, se ha tomado en especial consideración los modelos propuestos por el 3GPP en la especificación ETSI TR 101 112 [1], así como los recogidos en el proyecto europeo MORANS [2][3] y en la red de excelencia europea NEWCOM (*Network of Excellence in Wireless Communications*) [4]. La especificación ETSI TR 101 112 fue elaborada en la fase inicial de definición del sistema UMTS, con el objetivo de definir escenarios de prueba comunes sobre los cuales evaluar las distintas propuestas de tecnologías de acceso radio sobre las que trabajaban los diferentes grupos de trabajo. Por su parte, MORANS (*Mobile Radio Acces Network Reference Scenarios*) es el nombre con el que

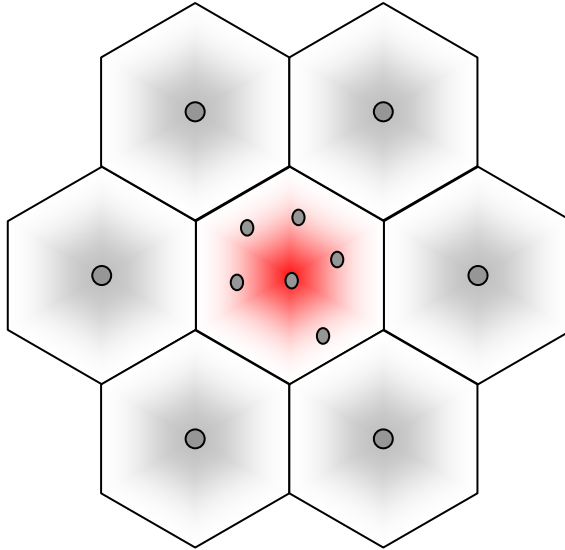
se conoce a una iniciativa incluida dentro del programa europeo COST 273. Este proyecto tiene el objetivo de definir escenarios de referencia que permitan a los distintos grupos de investigación europeos desarrollar y evaluar algoritmos en unas condiciones comunes, de forma que los resultados obtenidos sean comparables entre sí. Este proyecto tuvo su continuación natural dentro de la red de excelencia NEWCOM, donde se amplió la definición de escenarios a otras tecnologías móviles. Cabe destacar de nuevo que algunos de los modelos de simulación desarrollados en la presente Tesis Doctoral han sido incluidos en dichos escenarios, tanto en MORANS como en NEWCOM [5]-[9].

## **B.1.2 Estructura celular**

En general, la definición de la estructura celular depende del entorno de operación en el que se desee evaluar el sistema. Dado que el interés de esta Tesis Doctoral se centra en estudiar distintos algoritmos de gestión de recursos radio en un entorno macro-celular, se ha tomado como referencia inicial una única celda central rodeada por un anillo de celdas interferentes.

En la estructura celular implementada, las estaciones base se sitúan sobre una rejilla hexagonal, siendo la distancia entre los puntos centrales de las celdas un parámetro configurable que permite obtener distintos tamaños de celda. En concreto en las simulaciones llevadas a cabo se ha considerado un radio celular de 500m.

En la estructura celular sólo se han considerado celdas omnidireccionales puesto que la sectorización sólo afectaría a la distribución de interferentes pero no al comportamiento de los algoritmos de gestión de recursos evaluados. Las estaciones base interferentes están situadas en el centro de la celda hexagonal a la que dan servicio, mientras que en la celda bajo estudio las distintas antenas que configuran el MNB se distribuirán de manera que se optimice la capacidad del sistema. Las estaciones base emplean antenas con diagramas de radiación isotrópicos, siendo por lo tanto su ganancia de 0 dBi.



**Figura B.2 Estructura Celular**

Las seis celdas interferentes realmente no están cursando tráfico sino que radian a una potencia constante, definida también como un parámetro de la simulación. De esta manera, toda la actividad estudiada se desarrolla dentro de la celda central bajo estudio.

La Figura B.2 muestra la estructura celular omnidireccional considerada. Las posiciones de las estaciones base se han representado mediante círculos.

### **B.1.3 Estructura del MNB**

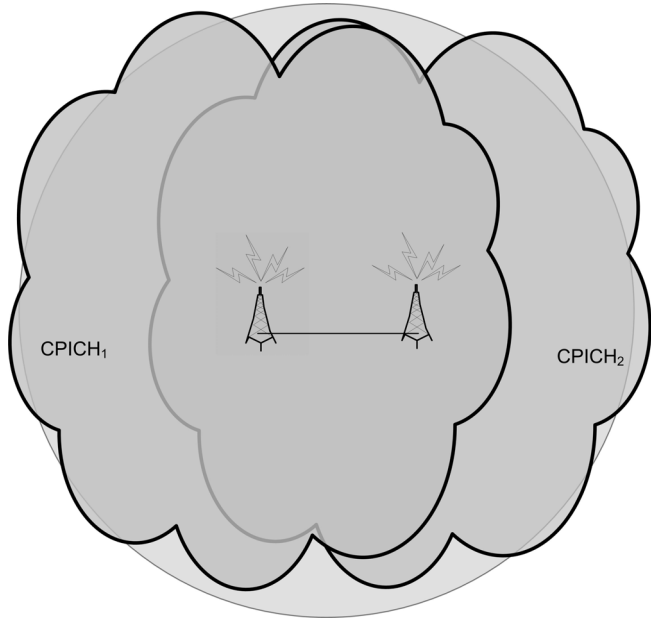
Tal y como se explicó en el Capítulo 2, la nueva arquitectura MNB se plantea como objetivo realizar las mínimas modificaciones necesarias en la arquitectura UTRAN convencional especificada en la *Release 99*. Así, para permitir la transmisión simultánea desde varias antenas, se pretende aprovechar las técnicas de diversidad en transmisión especificadas y de obligada implementación en los terminales UMTS. En [10] se especifica que para soportar la técnica de diversidad CLMTD (*Closed Loop Mode Transmit Diversity*), en la que el Nodo-B transmite a través de dos antenas



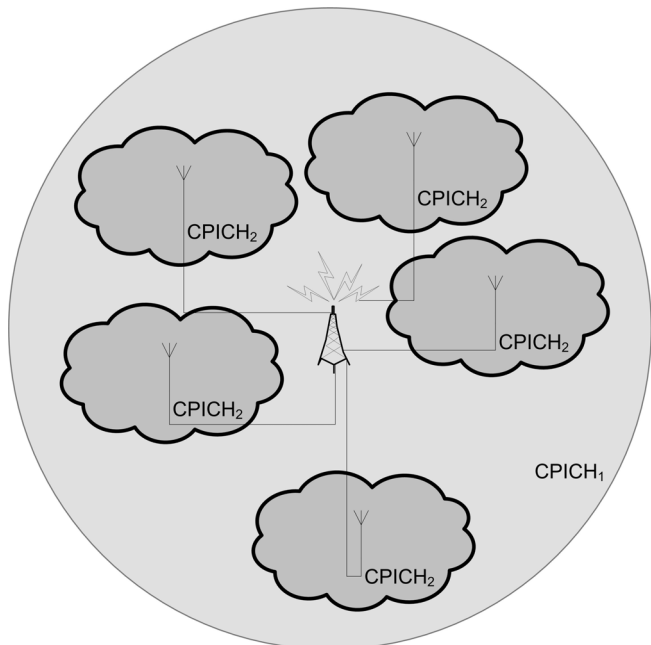
localizadas en el mismo emplazamiento, el terminal móvil debe ser capaz de monitorizar el canal piloto primario y secundario de la celda, permitiendo así la sincronización de la transmisión entre las dos antenas. La arquitectura MNB pretende hacer uso de esta capacidad de los terminales, transmitiendo también de manera simultánea pero esta vez desde distintos emplazamientos. La estructura del MNB por lo tanto presenta dos claras limitaciones. La primera limitación radica en que sólo se va a considerar la transmisión a un terminal desde dos antenas. Aunque el MNB puede estar conformado por más de dos transmisores por celda, a un usuario no se le puede estar transmitiendo simultáneamente desde más de dos ya que, en ese caso, las contribuciones podrían ser destructivas, empeorando gravemente la calidad de servicio experimentada. La segunda limitación estriba en que, para poder sincronizar la transmisión, todas las antenas del MNB deberán de transmitir un canal piloto, o bien el primario o el secundario, mientras que el terminal en cualquier posición sólo debe de recibir información de, como máximo, dos pilotos de la misma celda. Esto obliga a que en la zona de cobertura del MNB solamente se pueda percibir señal piloto de dos nodos simultáneamente, es decir, no puede existir más de dos coberturas solapadas.

Fruto de estas limitaciones se plantean dos posibles escenarios del MNB:

- *Escenario Macro-celular*: en este caso se consideran antenas situadas sobre azoteas, de manera que desde los emplazamientos se dispone de total cobertura sobre la celda. En este escenario, el MNB sólo estará compuesto por dos antenas ya que de otro modo se incumpliría la condición de únicamente dos canales piloto recibidos simultáneamente por un terminal de usuario. Este escenario se ha representado gráficamente en la Figura B.3.
- *Escenario Mixto*: en este escenario se dispone de un emplazamiento macro-celular sobre azotea con cobertura sobre toda la celda y de múltiples subceldas micro-celulares que no presenta solapamiento. El emplazamiento macro-celular será imprescindible para garantizar la continuidad de servicio en todo el área de cobertura del MNB. Este escenario queda reflejado en la Figura B.4.



**Figura B.3 Escenario MNB macro-celular**



**Figura B.4 Escenario MNB mixto**

## **B.1.4 Estación base**

La entidad estación base modela el conjunto de elementos de la red de acceso radio que administra los recursos de cada una de las celdas del escenario. Básicamente, el principal recurso con el que trabaja el simulador es la potencia y el número de códigos disponibles. Así pues, la estación base recibirá la asignación de recursos realizada en el MNB y distribuirá la potencia y los códigos disponibles en la celda entre los usuarios a los que presta servicio. Para ello, la estación base debe disponer de cierta información, como por ejemplo:

- Potencia disponible en la celda.
- Potencia libre en la celda.
- Potencia transmitida en los canales de control común de la celda.
- Potencia máxima por enlace.
- Número de códigos totales.
- Número de códigos utilizados por los canales de control común.
- Conjunto de usuarios servidos por la celda.
- Potencia asignada a cada uno de los usuarios de la celda.
- Código asignado a cada usuario.
- Distribución de la potencia de cada usuario entre las antenas del MNB.
- Distribución de la potencia total de la celda entre las antenas del MNB.

## **B.1.5 Modelo de propagación**

Una de las principales diferencias entre los sistemas de comunicaciones fijas y los móviles radica en el canal de propagación. En los sistemas móviles la propagación radioeléctrica tiene lugar a través de un medio hostil, que experimenta variaciones

aleatorias en sus características físicas, lo cual tiene un gran impacto en las prestaciones del sistema.

La caracterización de un canal móvil puede abordarse desde distintos puntos de vista. El método tradicional consiste en modelar estadísticamente los fenómenos aleatorios que tienen lugar en el canal, centrandó el interés en la atenuación que éstos producen sobre las señales radiadas. Este enfoque es particularmente adecuado para las simulaciones a nivel de sistema, por lo que ha sido el adoptado en la caracterización de canal incluida en la herramienta de simulación.

Desde este punto de vista, la propagación radio a través de un canal móvil puede caracterizarse por medio de tres fenómenos: las pérdidas básicas de propagación o *path loss*, un término de pérdidas a largo plazo, conocido como *shadowing*, y un término de pérdidas a corto plazo que recibe el nombre de *fast fading*.

Las pérdidas básicas de propagación proporcionan una estimación de las pérdidas medias que experimenta la señal. Este término se obtiene mediante algún modelo de propagación que relacione de forma determinista las pérdidas medias con parámetros del entorno tales como la distancia entre el transmisor y el receptor, la frecuencia de la señal, altura de las antenas y el perfil del terreno.

El *shadowing*, también conocido como *large scale fading* o como *long term fading*, es un término de pérdidas aleatorio que se añade a las pérdidas medias, y representa el efecto de los desvanecimientos producidos por la presencia de obstáculos en el trayecto entre el transmisor y el receptor que no han sido tenidos en cuenta en el modelo de propagación determinista.

El *fast fading* o *multipath fading* se debe a lo que se conoce como efecto multicamino. El efecto multicamino consiste en que, debido a los mecanismos de propagación que tienen lugar en el canal, en el receptor se recibirán múltiples contribuciones o réplicas de la señal transmitida que habrán sufrido distintos procesos de *scattering* y difracción. Cada una de estas contribuciones o componentes multicamino presentarán ángulos de incidencia, atenuación, desfases y retardos de propagación diferentes, por lo que su combinación en el receptor originará desvanecimientos rápidos y profundos en la señal recibida. Las fluctuaciones causadas

por la propagación multicamino pueden ser combatidas por medio de distintas técnicas, como por ejemplo, el control de potencia rápido, la combinación MRC efectuada en el receptor *Rake*, las técnicas de diversidad, la codificación con entrelazado, etc.

El modelo de canal implementado en el simulador de sistema sólo tiene en cuenta las pérdidas básicas de propagación y los desvanecimientos causados por el *shadowing*, ya que los desvanecimientos rápidos se tuvieron en cuenta en las simulaciones a nivel de enlace llevadas a cabo en el Capítulo 3. A continuación, se describen los modelos utilizados en la herramienta de simulación para caracterizar las pérdidas de propagación. Dada la importancia que el modelado del *shadowing* ha tenido en el desarrollo de la presente Tesis Doctoral, su descripción se ha incluido en un capítulo específico, el Capítulo 6, plenamente dedicado al estudio del *shadowing* y su efecto sobre la evaluación de las prestaciones del sistema UMTS convencional.

### B.1.5.1 Pérdidas de propagación

Dado que se van a considerar dos escenarios distintos de propagación, tanto un escenario micro-celular como uno macro-celular, se ha de establecer un modelo de propagación para cada uno de ellos.

Respecto a la propagación macro-celular, el modelo utilizado para calcular las pérdidas básicas de propagación es el descrito en [11], siendo la ecuación de pérdidas básica del usuario  $i$ -ésimo

$$L_{p,i}(\text{dB}) = 137.4 + 35.2 \log_{10}(d_i) \quad (\text{B.1})$$

donde  $d_i$  es la distancia en km del usuario  $i$ -ésimo al Nodo-B de la celda bajo estudio.

Para escenarios micro-celulares, característicos del escenario MNB mixto, se ha considerado el modelo propuesto en [1]. En este caso, la ecuación que describe las pérdidas de propagación es la siguiente:

$$L_{p,i}(\text{dB}) = 142 + 40 \log_{10}(d_i) \quad (\text{B.2})$$

## **B.1.6 Modelo de interferencias**

Como ya se ha comentado, el sistema UMTS, por el hecho de emplear la técnica de acceso WCDMA, es un sistema limitado por interferencias. El nivel de interferencias percibido por cada usuario afecta a la calidad de sus comunicaciones y, en último término, limita la capacidad del sistema. Como consecuencia, el modelado de las interferencias es un proceso fundamental en el desarrollo de un simulador diseñado para evaluar las prestaciones de un sistema UMTS.

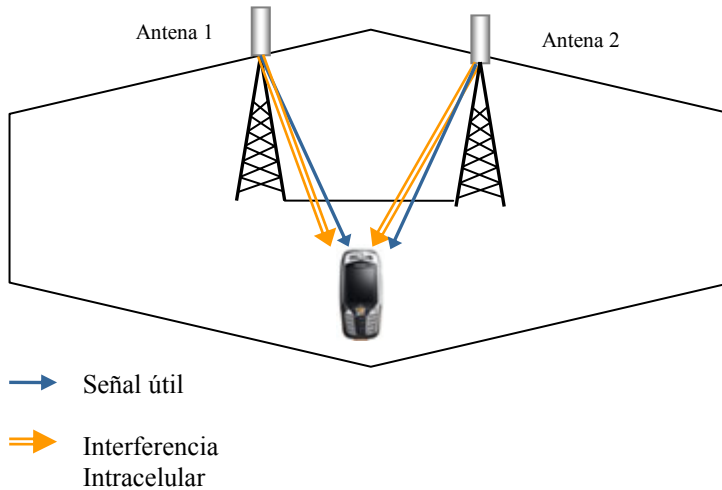
El modelo de interferencias implementado permite obtener la interferencia que degrada la calidad en cualquier enlace establecido en el sentido descendente. En un escenario clásico, un enlace radio está definido entre la estación base origen del enlace, y el terminal móvil de destino. No obstante, para que el cálculo de interferencias sea consistente con la arquitectura Multinodo B, hay que tener en cuenta que cada enlace radio está definido entre una de las antenas de una celda y un terminal móvil.

A continuación, se analiza los modelos de interferencias correspondientes a cada uno de los escenarios contemplados en esta Tesis Doctoral.

### **B.1.6.1 Escenario MNB macro-celular**

Considerando esta estructura, la interferencia total experimentada en un enlace puede descomponerse en dos términos: la interferencia intracelular y la interferencia intercelular.

La interferencia intracelular es toda aquella potencia transmitida por el MNB que no está destinada al envío de información útil al usuario. En el caso del escenario MNB macro-celular, se pueden presentar dos situaciones: o que se transmita simultáneamente información útil a un terminal desde ambas antenas o que solamente transmita una de ellas. En cualquier caso, las fuentes de interferencia intracelular serán ambas antenas. La Figura B.5 muestra un esquema de las interferencias intracelulares.



**Figura B.5 Fuentes de interferencia intracelular – MNB macro-celular**

La interferencia intercelular percibida en un determinado enlace está provocada por la potencia transmitida en las celdas cercanas al terminal, exceptuando aquella a la que pertenece el enlace considerado. Sin pérdida de generalidad se ha asumido que solamente existe interferencia intercelular proveniente del primer anillo de celdas cocanal, tal y como se estableció en el apartado B.1.2.

Dado que la transmisión de datos está claramente desbalanceada, siendo el sentido descendente el más cargado y por tanto en más limitante en cuanto a capacidad, sólo se ha realizado este estudio para enlace descendente. Además, la gestión de recursos radio en enlace ascendente se corresponde con una simple priorización de colas FIFO por lo que no necesita de optimización.

En el escenario macro-celular, en sentido descendente, se dispone de dos estaciones base que transmiten con potencia  $P_{T1}$  y  $P_{T2}$  respectivamente. La suma de ambas potencias,  $P_T$ , no podrá superar la potencia máxima disponible en el sistema. La relación de energía de bit de información respecto a la densidad espectral de potencia de ruido más interferencias, denominada  $(E_b/N_0)$ , tendrá, para el usuario  $i$ -ésimo, la siguiente expresión:

$$\left(\frac{E_b}{N_0}\right)_i = \frac{\left(\frac{P_{1i}}{L_{1,i}} + \frac{P_{2i}}{L_{2,i}}\right)W}{P_N + \chi_i + \rho \left(\frac{P_{T1} - P_{1,i}}{L_{1,i}} + \frac{P_{T2} - P_{2,i}}{L_{2,i}}\right)} \quad (\text{B.3})$$

donde  $P_{1,i}$  y  $P_{2,i}$  son las potencias transmitidas por la antena 1 y 2 respectivamente al usuario  $i$ ,  $L_{1,i}$  y  $L_{2,i}$  son las pérdidas de propagación en ambos enlaces,  $W$  el ancho de banda de transmisión de UMTS (3.84MHz),  $R_i$  la tasa binaria de servicio asignada al usuario,  $P_N$  el ruido térmico,  $\chi_i$  la potencia interferente intercelular a la salida de la antena del usuario  $i$  y  $\rho$  el factor de ortogonalidad, siendo  $\rho=0$  para códigos perfectamente ortogonales. Como se puede ver en el numerador, se ha considerado que la potencia recibida por el usuario es exactamente la suma de las potencias recibidas desde cada una de las antenas del MNB. Realmente, y tal y como se demostró en el Capítulo 3, gracias al preprocesado realizado en el MNB, las señales recibidas por el usuario están relativamente ajustadas en fase, por lo que la potencia de la señal no se corresponde exactamente con la suma incoherente de las potencias recibidas. Sin embargo, esta expresión es válida ya que, en las simulaciones de enlace realizadas, el cálculo de la potencia de ruido interferente se realizó utilizando esta misma expresión, lo que garantiza que, utilizando la expresión (B.3) junto con los resultados obtenidos en el Capítulo 3, la caracterización del comportamiento del sistema es correcta.

A partir de (B.3) y considerando  $P_i$  como la potencia total asignada por el MNB al usuario  $i$  y  $q_i$  como la proporción de dicha potencia transmitida por la estación base 1 resulta que

$$\begin{aligned} P_{1i} &= q_i \cdot P_i \\ P_{2i} &= (1 - q_i) \cdot P_i \end{aligned} \quad (\text{B.4})$$

Si, además, se define  $Q$  como la proporción de la potencia total  $P_T$  transmitida por la estación base 1, finalmente se puede calcular la relación de energía de bit respecto a densidad espectral de ruido más interferencia como:



$$\left(\frac{E_b}{N_0}\right)_i = \frac{P_i \left( \frac{q_i}{L_{1,i}} + \frac{(1-q_i)}{L_{2,i}} \right) \frac{W}{R_i}}{P_N + \chi_i + \rho \left( \frac{Q \cdot P_T - P_{1,i}}{L_{1,i}} + \frac{(1-Q) \cdot P_T - P_{2,i}}{L_{1,i}} \right)} \quad (\text{B.5})$$

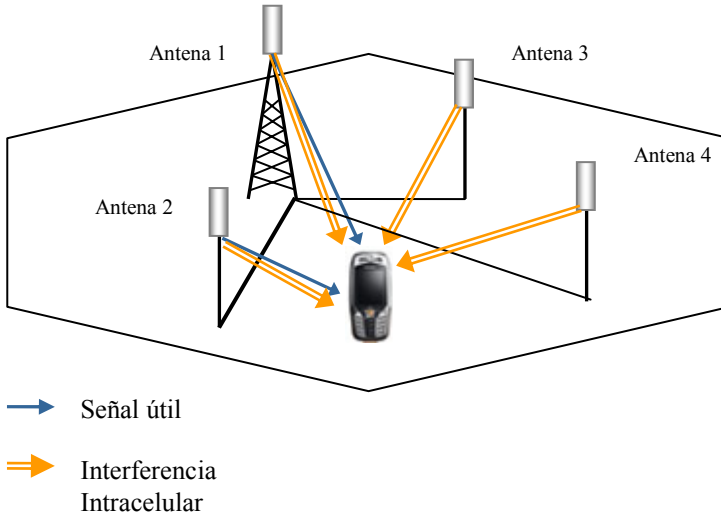
El caso mixto no es más que una generalización de este escenario. A continuación, se analizará la expresión que caracteriza las  $E_b/N_0$  en el escenario MNB mixto para pasar finalmente a estudiar el significado y la utilidad de las expresiones obtenidas.

### B.1.6.2 Escenario MNB mixto

La Figura B.6 muestra un esquema de las interferencias intracelulares experimentadas por un terminal servido simultáneamente por dos antenas del MNB dentro del escenario mixto. Aunque en el apartado 5.4.2 se mostró cómo diseñar la potencia de CPICH para no tener coberturas solapadas, realmente la potencia transmitida por una antena micro-celular del MNB puede recibirse en cualquier parte del sistema, aunque dicha señal estará atenuada en mayor o menor medida.

Se define  $P_{pi}$  como la potencia que la antena  $p$ -ésima del MNB está transmitiendo al usuario  $i$ -ésimo y  $q_{pi}$  como la proporción que  $P_{pi}$  representa respecto a la potencia total transmitida por la celda al usuario  $i$ ,  $P_i$ , de manera que se cumple que:

$$\begin{aligned} P_{pi} &= q_{pi} \cdot P_i \\ \sum_{p=1}^P P_{pi} &= \sum_{p=1}^P q_{pi} \cdot P_i = P_i \cdot \sum_{p=1}^P q_{pi} = P_i \end{aligned} \quad (\text{B.6})$$



**Figura B.6 Fuentes de interferencia intracelular – MNB mixto**

Además, se define  $P_{Tp}$  como la potencia total transmitida por la antena  $p$ ,  $P_T$  como la potencia total transmitida por la celda y  $Q_p$  como la proporción de dicha potencia transmitida por la antena  $p$ -ésima. Es importante destacar que, tal y como se justifica en [12], el factor de ortogonalidad medio,  $\rho$ , es distinto para las celdas macro y micro-celulares. De hecho, en las celdas micro-celulares existe por lo general visión directa entre transmisor y receptor y la distancia de cobertura es mucho menor, por lo que las contribuciones apenas se dispersan temporalmente lo que reduce el factor de ortogonalidad. Teniendo en cuenta esto y para caracterizar correctamente el escenario mixto, será necesario considerar un factor de ortogonalidad diferente para cada tipo de celda. Así pues, se define  $\rho_p$  como el factor de ortogonalidad de la antena  $p$ -ésima. Finalmente, la relación de energía de bit respecto a densidad espectral de ruido más interferencia que experimenta el usuario  $i$  se puede calcular como:

$$\left(\frac{E_b}{N_0}\right)_i = \frac{\sum_{p=1}^P \left(\frac{P_{pi}}{L_{p,i}}\right) \frac{W}{R_i}}{P_N + \chi_i + \sum_{p=1}^P \left(\frac{(P_{Tp} - P_{pi})\rho_p}{L_{p,i}}\right)} = \frac{\sum_{p=1}^P \left(\frac{q_{pi}P_i}{L_{p,i}}\right) \frac{W}{R_i}}{P_N + \chi_i + \sum_{p=1}^P \left(\frac{(Q_p P_T - q_{pi}P_i)\rho_p}{L_{p,i}}\right)} \quad (B.7)$$

siendo  $P$  el número total de antenas que conforman el MNB y  $L_{p,i}$  las pérdidas de propagación experimentadas por la señal transmitida desde la antena  $p$  al usuario  $i$ .

A partir de (B.7) se puede obtener de manera sencilla la expresión correspondiente a la potencia transmitida al usuario  $i$  como:

$$P_i = \frac{P_N + \chi_i + P_T \sum_{p=1}^P \frac{\rho_p Q_p}{L_{p,i}}}{\sum_{p=1}^P \left( \frac{q_{pi}}{L_{p,i}} \cdot \left( \rho_p + \frac{W}{\left( \frac{E_b}{N_0} \right)_i R_i} \right) \right)} \quad (\text{B.8})$$

Es esta ecuación la que define la potencia que el MNB debe reservar para la transmisión de datos al usuario  $i$  con una cierta tasa binaria  $R_i$ . La relación  $(E_b/N_0)_i$  necesaria para garantizar la calidad de servicio objetivo en términos de BLER dependerá de la relación entre las potencias recibidas desde cada antena por lo que no puede definirse sin conocer previamente los valores de  $q_{pi}$  y las pérdidas de propagación experimentada en todos los enlaces involucrados.

También resultaría muy útil contar con una expresión cerrada de la potencia total transmitida por el MNB para controlar que una determinada asignación de recursos no excede la potencia máxima disponible en el sistema. La potencia total transmitida por el MNB será la suma de todas las potencias transmitidas a los  $N$  usuarios conectados a la celda más la suma de las potencias consumidas por los canales de control de todas las antenas del MNB. Si  $P_{C,p}$  es la potencia de los canales comunes de control transmitida por la antena  $p$ -ésima, la potencia total resulta:

$$P_T = \sum_{i=1}^N P_i = \sum_{i=1}^N \frac{P_N + \chi_i + P_T \sum_{p=1}^P \frac{\rho_p Q_p}{L_{p,i}}}{\sum_{p=1}^P \left( \frac{q_{pi}}{L_{p,i}} \cdot \left( \rho_p + \frac{W}{\left( \frac{E_b}{N_0} \right)_i R_i} \right) \right)} + \sum_{p=1}^P P_{C,p} \quad (\text{B.9})$$

o bien

$$P_T = \sum_{i=1}^N \frac{P_N + \chi_i}{\sum_{p=1}^P \left( \frac{q_{pi}}{L_{p,i}} \cdot \left( \rho_p + \frac{W}{\left( \frac{E_b}{N_0} \right)_i R_i} \right) \right)} + \sum_{i=1}^N \frac{P_T \sum_{p=1}^P \frac{\rho_p Q_p}{L_{p,i}}}{\sum_{p=1}^P \left( \frac{q_{pi}}{L_{p,i}} \cdot \left( \rho_p + \frac{W}{\left( \frac{E_b}{N_0} \right)_i R_i} \right) \right)} + \sum_{p=1}^P P_{C,p} \quad (\text{B.10})$$

Con lo que finalmente la potencia total del sistema se puede calcular como

$$P_T = \frac{\sum_{i=1}^N \frac{P_N + \chi_i}{\sum_{p=1}^P \left( \frac{q_{pi}}{L_{p,i}} \cdot \left( \rho_p + \frac{W}{\left( \frac{E_b}{N_0} \right)_i R_i} \right) \right)} + \sum_{p=1}^P P_{C,p}}{1 - \sum_{i=1}^N \frac{\sum_{p=1}^P \frac{\rho_p Q_p}{L_{p,i}}}{\sum_{p=1}^P \left( \frac{q_{pi}}{L_{p,i}} \cdot \left( \rho_p + \frac{W}{\left( \frac{E_b}{N_0} \right)_i R_i} \right) \right)}} \quad (\text{B.11})$$

Es importante destacar también que el sumatorio

$$\sum_{p=1}^P \left( \frac{q_{pi}}{L_{p,i}} \cdot \left( \rho_p + \frac{W}{\left( \frac{E_b}{N_0} \right)_i R_i} \right) \right) \quad (\text{B.12})$$

sólo contará, como máximo, con dos términos no nulos, puesto que en la arquitectura MNB solamente se permite la transmisión simultánea desde dos antenas.

### B.1.6.3 Análisis del modelo de interferencias

Analizando las ecuaciones (B.8) y (B.11) se puede apreciar como tanto la potencia asignada a un usuario como la potencia total del sistema se minimizan al maximizar la función:

$$D = \sum_{p=1}^P \left( \frac{q_{pi}}{L_{p,i}} \cdot \left( \rho_p + \frac{W}{\left( \frac{E_b}{N_0} \right)_i R_i} \right) \right) \quad (\text{B.13})$$

De manera que, para una determinada localización de los transmisores y del terminal móvil, existe una combinación de potencias de transmisión que hace que el sistema esté consumiendo mínimos recursos.

La Figura B.7 muestra el valor de la potencia transmitida por el MNB macro-celular a un usuario, calculándola a partir de (B.8), para distintos valores de  $q$ , es decir, para distintas distribuciones de potencia entre las dos antenas involucradas en la transmisión. En abscisas se representa la diferencia en dB entre las pérdidas de propagación de las dos antenas. Por ejemplo,  $x=0$  representa el caso en que los dos caminos de propagación presentan exactamente las mismas pérdidas, mientras que  $x=3$  representa el caso en el que las pérdidas desde la antena secundaria son el doble que las experimentadas desde la antena principal. Con línea gruesa discontinua se muestra la potencia que se tendría que transmitir si no se considerase el MNB. Como se puede observar existe un valor de  $q$  óptimo que permite minimizar la potencia transmitida, siempre que la antena secundaria presente unas condiciones de propagación o mejores que la de la antena principal o al menos no peores que 2.3 dB.

Este valor  $q$  óptimo de distribución de potencia entre las dos antenas se ha representado en la Figura B.8. Si la antena que tiene mejores condiciones de propagación es la principal (valores de abscisas positivos y altos) entonces la encargada de la transmisión será ella misma, mientras que cuando la diferencia de pérdidas de propagación esté entre -2.3 y 2.3 entonces las dos antenas transmitirán simultáneamente. Finalmente, para valores inferiores a -2.3 será la antena secundaria la que asumirá toda la transmisión. Es muy significativo señalar que, cuando se transmite simultáneamente desde dos antenas, la potencia transmitida es, en la mayor parte de las ocasiones, muy similar, estando acotado el valor de  $q$  entre 0.4 y 0.6.

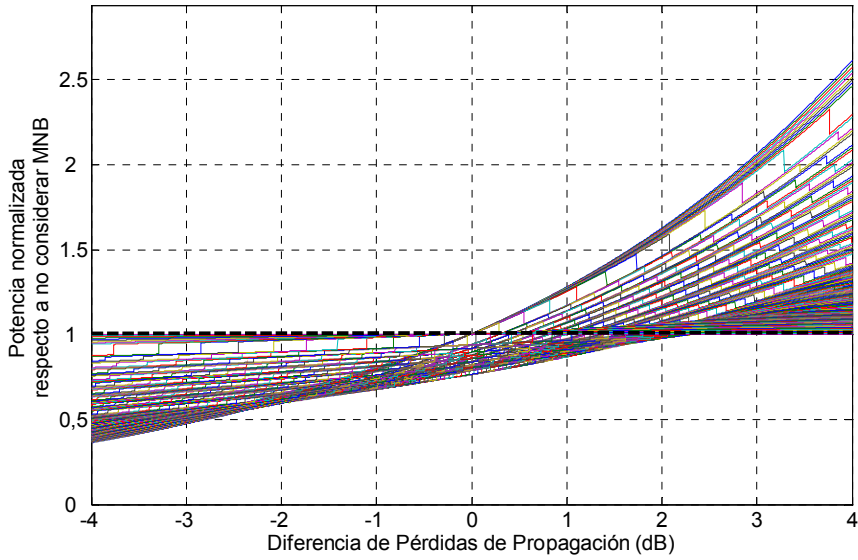


Figura B.7 Potencia transmitida en función de las pérdidas de propagación relativas

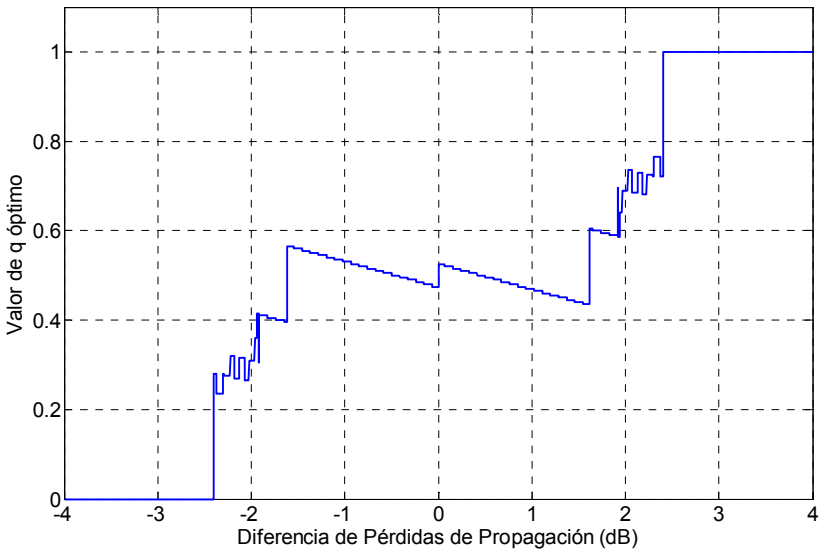
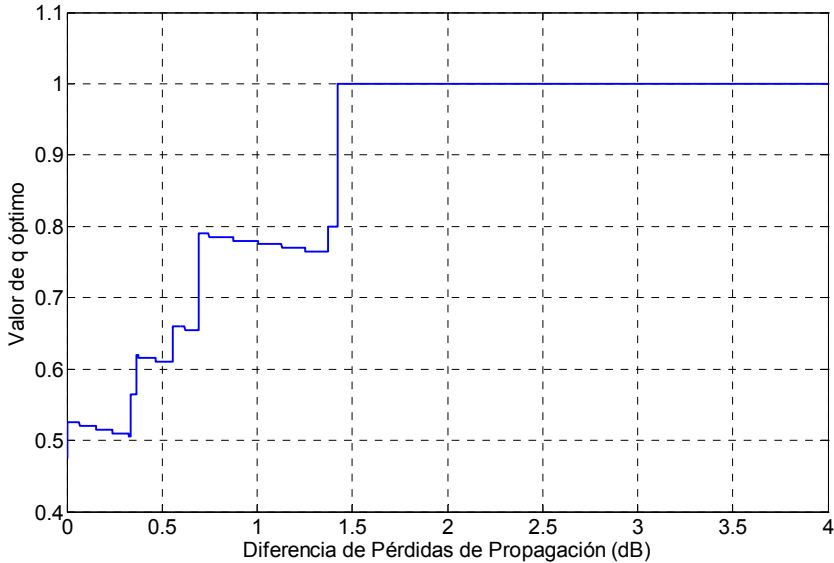
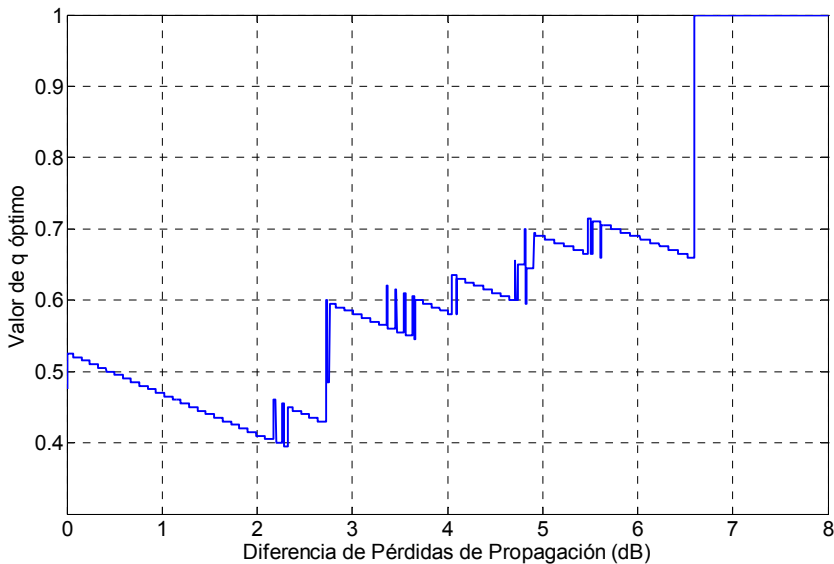


Figura B.8 Distribución de potencia óptima en el MNB macro-celular

En el escenario MNB mixto la distribución óptima de potencia dependerá de si se recibe más potencia por el emplazamiento macro-celular o por el micro-celular. Ambos casos se recogen en la Figura B.9 y Figura B.10 respectivamente.



**Figura B.9 Distribución óptima en el MNB mixto con dominio macro-celular**



**Figura B.10 Distribución óptima en el MNB mixto con dominio micro-celular**

En estas figuras se puede apreciar que, en el caso de que el emplazamiento macro-celular sea el dominante, la contribución del emplazamiento micro-celular es muy poco significativa por lo que con una diferencia de menos de 1.5 dB ya se prefiere transmitir únicamente desde el emplazamiento macro-celular. Por el contrario, complementar la transmisión desde el emplazamiento micro-celular con la señal proveniente de la antena macro-celular supone una muy importante ganancia, tal y como se constató en el Capítulo 3, aspecto que se ve reflejado en que con una diferencia de hasta 6.6 dB sea aun interesante transmitir simultáneamente desde ambas antenas.

Tras analizar estas gráficas, parece obvio que se puede obtener una gran ganancia en capacidad si el sistema conoce en todo momento la diferencia de pérdidas de propagación que experimenta el usuario. Para ello, es necesario que en la arquitectura MNB el móvil envíe periódicamente información acerca de la potencia recibida en los dos canales piloto monitorizados. Este procedimiento está estandarizado y se puede ordenar al móvil que envíe estas medidas periódicamente desde cada 200ms hasta decenas de segundos. Con esta información recibida, y ya que el MNB conoce la potencia del canal piloto transmitida, se puede estimar fácilmente la diferencia de pérdidas de propagación entre ambas antenas. Conociendo este valor se puede ajustar el valor de  $q$  que minimice la potencia consumida por usuario. Otra cosa que hay que tener en cuenta es que no siempre la distribución de potencias que supone una menor potencia transmitida para un usuario concreto es la mejor decisión para el sistema global, ya que quizá cargar demasiado una determinada antena puede afectar negativamente a los niveles de interferencia intracelular experimentados por otros usuarios y reducir la capacidad global del sistema. Este efecto global producido por cada decisión solamente puede tenerse en cuenta cuando se apliquen estrategias de gestión de recursos que consideren paralelamente la evolución de todos los usuarios y que no vayan asignando recursos usuario por usuario. Este factor es el que hace que el sistema que se propone en esta Tesis Doctoral en el Capítulo 5 sea mucho más potente que otros algoritmos de gestión de recursos convencionales.

Además, respecto a la decisión óptima de  $q$ , se puede optar por minimizar la potencia por usuario o por aumentar el balanceo de carga entre las antenas del MNB. Para ello

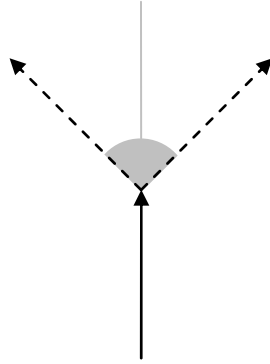


se puede establecer un margen de diversidad, similar al margen de *soft handover*, de manera que, si la diferencia de potencia entre las contribuciones está dentro de ese margen, se transmite simultáneamente por ambas antenas. Este margen de diversidad deberá de adaptarse al escenario concreto de aplicación puesto que, como se ha visto en las figuras, los umbrales dependen mucho del mismo. Por ejemplo, en el escenario macro-celular, si el umbral establecido es igual que 2.3 dB coincidirá con el margen mostrado en la Figura B.8 y por tanto garantizará la minimización de la potencia. Si el margen está por encima de este valor, la potencia total consumida por el usuario no será la mínima posible pero, por el contrario, se conseguirá un mayor balanceo de carga, lo que garantiza un mejor funcionamiento de la red. Obviamente, existirá un valor de este margen a partir del cual los beneficios del balanceo serán menores que el incremento de potencia que supone distribuir el esfuerzo de transmisión entre ambas antenas. En la sección de resultados del Capítulo 5 se estudiará en profundidad todas las posibles alternativas en la asignación de recursos dentro del MNB.

## **B.1.7 Modelo de movilidad**

El modelo de movilidad implementado en el simulador está basado en el modelo propuesto en [1] para el entorno de simulación *vehicular*.

Se trata de un modelo pseudoaleatorio que describe el movimiento del terminal mediante una serie de saltos, entre los cuales se supone que el móvil se desplaza en línea recta. La movilidad de cada terminal está caracterizada por una velocidad constante durante todo su movimiento y por la dirección de desplazamiento que se va actualizando cada paso temporal. La dirección inicial del móvil se determina aleatoriamente cuando el usuario accede al sistema. Cada vez que se actualiza la posición del móvil, su dirección podrá cambiar con una determinada probabilidad. Para ello se suma a la dirección previa un ángulo aleatorio distribuido uniformemente en un rango determinado. Con este mecanismo se simula una trayectoria semi-dirigida. La Figura B.11 ilustra este concepto.



**Figura B.11 Actualización de la dirección**

Dado que se desea obtener una distribución homogénea de usuarios en el sistema, la posición inicial de los móviles cuando acceden al sistema se determina aleatoriamente, siguiendo una distribución uniforme.

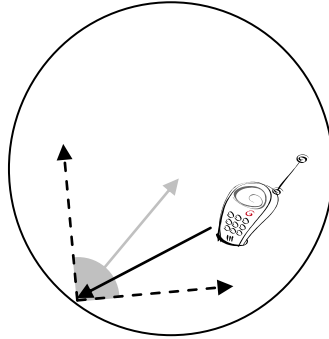
La Tabla B.1 recoge los valores empleados para cada uno de los parámetros definidos por el modelo de movilidad.

En el supuesto de que el terminal llegue al límite de la celda, es decir, alcance el radio celular, la dirección de movimiento se dirige hacia el centro de la celda más un radio distribuido uniformemente entre  $\pm 45^\circ$ . La Figura B.12 muestra este efecto de rebote.

Parámetro	Valor
Velocidad	50 km/h
Probabilidad de cambiar de dirección	0.2
Ángulo máximo de cambio de dirección	$45^\circ$
Distancia de decorrelación	20 m
Distribución inicial de usuarios	Uniforme

---

**Tabla B.1 Parámetros del modelo de movilidad**



**Figura B.12 Efecto de rebote en los límites de la celda**

## **B.1.8 Modelos de tráfico**

Uno de los requisitos fundamentales a la hora de elaborar un modelo de un sistema de comunicaciones es caracterizar el tráfico generado por los usuarios que acceden a dicho sistema. Para conseguir este objetivo, suelen emplearse modelos que describan estadísticamente ciertos parámetros claves, obtenidos a partir del tráfico registrado en las redes reales.

Obviamente, los flujos de tráfico correspondientes a tipos de servicio distintos presentan características diferentes, por lo que será necesario emplear un modelo de tráfico por cada tipo de servicio considerado. En esta Tesis Doctoral se han considerado 4 tipos de servicio para caracterizar los cuatro tipos de tráfico que se desea estudiar. En primer lugar, se ha implementado tráfico de voz para estudiar el comportamiento de los algoritmos de gestión de recursos en la transmisión por conmutación de circuitos. El servicio de navegación web permite estudiar el comportamiento de los servicios interactivos con baja restricción temporal. También se ha implementado el servicio de vídeo conferencia, ya que este servicio se transmitirá también mediante conmutación de paquetes, de manera que se pueda estudiar el comportamiento de los algoritmos de gestión de recursos para servicios con alta restricción temporal. Por último, se ha considerado también la descarga de ficheros

mediante FTP para evaluar el comportamiento de los servicios de *background*. En los siguientes apartados se analizarán uno a uno todos los modelos de tráfico utilizados.

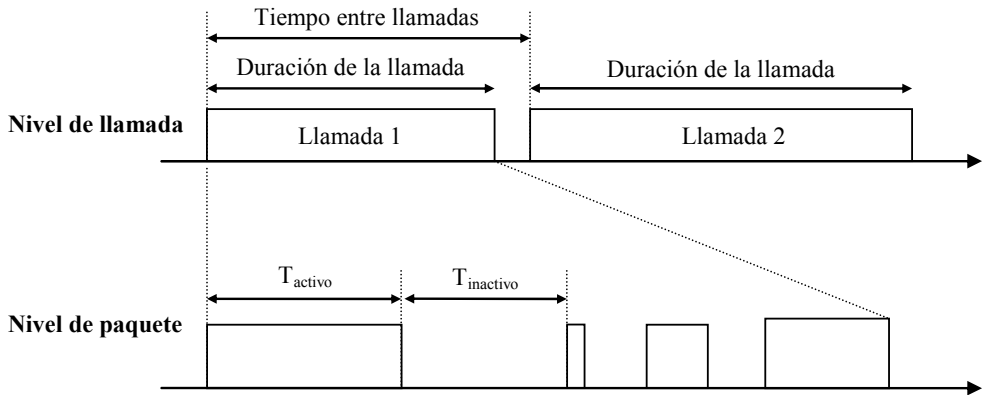
Es importante tener en cuenta que, dado que el objetivo del simulador desarrollado es caracterizar únicamente el enlace descendente, los modelos de tráfico implementados describen el tráfico sólo en este sentido de la comunicación.

En el simulador se ha considerado un número fijo de usuario de cada uno de los servicios considerados. De esta manera, no se han modelado procesos de nacimiento y muerte del usuario sino que el único interés es describir los estados de actividad y silencio de los usuarios presentes en el sistema. De alguna manera se está considerando un número fijo de Erlangs de tráfico en el sistema, de manera que se evitan las clásicas oscilaciones de carga que únicamente provocan la obtención de resultados ambiguos.

### **B.1.8.1 Modelo de tráfico de voz**

Normalmente el tráfico de voz se caracteriza empleando dos niveles de granularidad diferentes. El primero de ellos, conocido como nivel de llamada, modela el proceso de comienzo y finalización de las llamadas de voz establecidas por el conjunto de usuarios del sistema. El segundo de los niveles, que suele conocerse como nivel de paquete, describe el tráfico generado en cada una de las llamadas. Este segundo nivel se modela como una secuencia de periodos alternos de actividad y silencio. Ambos niveles de caracterización se muestran en la Figura B.13.

Dado que en esta Tesis Doctoral se ha considerado un número fijo de usuarios de cada servicio, no es necesario describir el tráfico de voz a nivel de llamada. Así pues, los usuarios de voz incluidos en el sistema estarán continuamente alternando periodos de actividad y silencio hasta el fin de la simulación.



**Figura B.13 Modelo de tráfico de voz**

Para describir el tráfico a nivel de paquete es necesario caracterizar la duración de los procesos de actividad e inactividad (o silencio). El estudio realizado en [13] reveló que estos parámetros siguen una distribución exponencial negativa. En el modelo de tráfico de voz implementado en ese trabajo se definió, a partir de medidas experimentales, una duración media de los periodos de actividad,  $\bar{T}_{activo}$ , de 1.35 segundos, mientras que la duración media de los periodos de silencio considerada,  $\bar{T}_{inactivo}$ , fue de 1.70 segundos. Con estas consideraciones, el factor de actividad de la fuente, que se corresponde con la proporción de tiempo en el que la fuente está activa, vendría dada por [14]:

$$\eta = \frac{\bar{T}_{activo}}{\bar{T}_{activo} + \bar{T}_{inactivo}} = 0.4426 \quad (\text{B.14})$$

Sin embargo, en [13] sólo se considera el tráfico generado por la fuente, siendo esta aproximación no válida para el caso de UMTS ya que, multiplexada con los datos, se transmite información de control adicional. Según [11] el tiempo de actividad de fuente teniendo en cuenta el tráfico de control es de 0.67. Si se considera el mismo periodo de actividad de 1.35 s entonces el periodo de silencio resulta ser de 0.665 s.

La Tabla B.2 resume los parámetros del modelo de tráfico de voz considerado.

Parámetro	Distribución	Valor medio
Generación de llamadas		Número fijo de usuarios
Duración de las llamadas		Tiempo de Simulación
Periodo de actividad	Exponencial negativa	1.35 s
Periodo de silencio	Exponencial negativa	0.665 s
Actividad de fuente		0.67

**Tabla B.2 Parámetros del modelo de tráfico de voz**

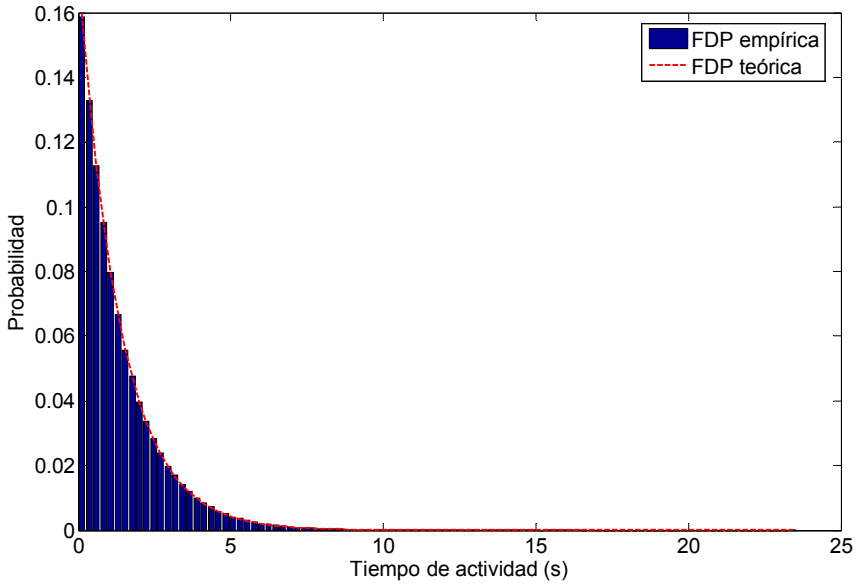
### Validación del modelo

Para validar la implementación del modelo de tráfico de voz se comprobará que los parámetros definidos en el modelo se ajustan a las distribuciones propuestas. Estos parámetros son la duración de los periodos de actividad y la duración de los periodos de silencio.

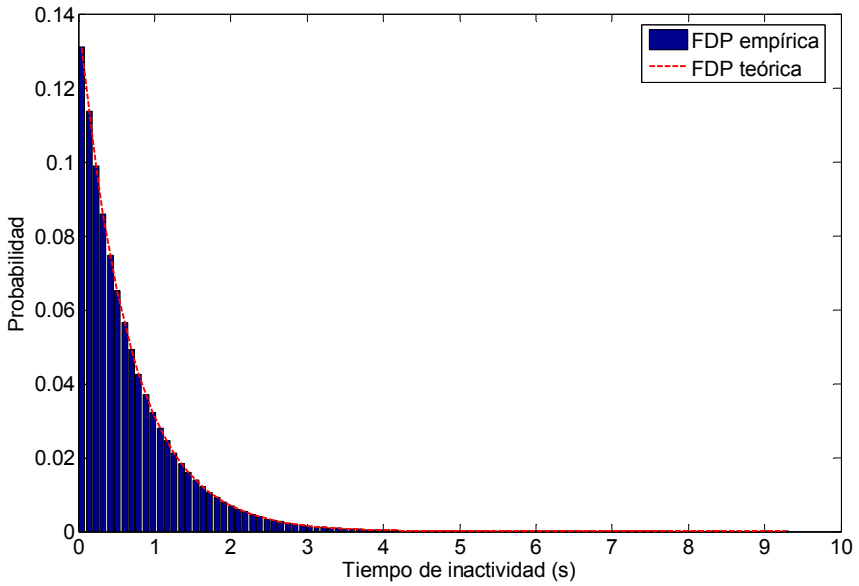
Para ello se ha simulado una llamada de voz en la cual se han registrado 100.000 periodos de actividad y otros tantos de silencio. La Figura B.14 y la Figura B.15 muestran respectivamente las funciones densidad de probabilidad (FDP) de las distribuciones empíricas obtenidas a partir de dichas muestras de los periodos de actividad y silencio, junto a las distribuciones teóricas correspondientes. Se puede comprobar como las muestras se generan siguiendo la distribución esperada.

En este estudio el factor de actividad ha resultado ser de  $\eta = 0.6719$ , frente al 0.67 que establece el modelo teórico. La aproximación es suficientemente buena como para validar el funcionamiento del modelo.

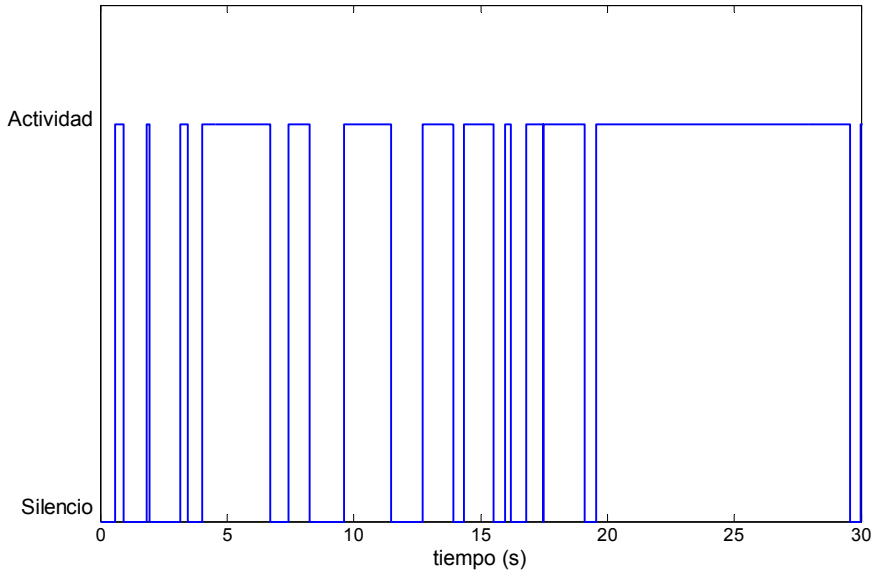
Una vez validado el algoritmo, en la Figura B.16 se muestra un ejemplo de tráfico generado por el modelo. En este ejemplo se puede apreciar claramente la alternancia entre los periodos de actividad y silencio.



**Figura B.14 FDP de la duración de los periodos de actividad**



**Figura B.15 FDP de la duración de los periodos de inactividad**



**Figura B.16** Ejemplo de realización de tráfico de voz

### **B.1.8.2** Modelo de tráfico web

Diversos estudios de tráfico sobre redes de área local [15] y redes de área ancha [16] han evidenciado que el tráfico de datos muestra una propiedad que no está presente en el tráfico de voz, conocida como autosemejanza (*self-similarity*). El tráfico autosemejante se caracteriza porque tiene propiedades estadísticas similares en un amplio rango de escalas de tiempo y presenta dependencia a largo plazo. Como consecuencia de la propiedad de autosemejanza, el tráfico de datos no muestra un tamaño típico de ráfaga, sino que éstas aparecen en un amplio rango de escalas de tiempo. Otra importante propiedad del tráfico autosemejante es que su comportamiento “a ráfagas” se intensifica al aumentar el nivel de tráfico.

La existencia de dependencia a largo-plazo en el tráfico de paquetes tiene importantes implicaciones en el dimensionamiento y evaluación de las redes de datos, por lo que debe ser reflejada en el modelo de tráfico utilizado. Este objetivo se



consigue modelando el tráfico de datos a través un tipo de procesos conocidos como procesos autosemejantes.

A continuación, se realizará una breve descripción teórica de los procesos autosemejantes. En ella se introducen los principales conceptos y definiciones, se analizan las propiedades y consecuencias mas importantes, y a partir de ellas se derivan metodos para analizar el grado de autosemejanza de un proceso. Posteriormente, se describirá con detalle el modelo de tráfico implementado en el simulador y se finalizará con la validación del modelo.

### Definición y propiedades de autosemejanza

Sea  $X = \{X_t, t = 0,1,2,\dots\}$  un proceso estacionario en sentido amplio, con media  $\mu$ , varianza  $\sigma^2$  y función de autocorrelación  $r(k)$ . Para cada valor de  $m=1,2,3,\dots$ , se define el denominado proceso agregado  $X^{(m)} = \{X_k^{(m)}, k = 1,2,3,\dots\}$  promediando  $X$  sobre bloques no solapados de tamaño  $m$ . Es decir:

$$X_k^{(m)} = \frac{1}{m}(X_{km-m+1} + \dots + X_{km}) \quad , k \geq 1 \quad (\text{B.15})$$

Se dice que el proceso  $X$  es H-autosemejante (*H-self-similar*) si  $X^{(m)}$  tiene la misma distribución que  $X$  reescalado por  $m^H$ , es decir:

$$X_t \stackrel{d}{=} m^{-H} X_t^{(m)} \quad , m \in N \quad (\text{B.16})$$

donde  $\stackrel{d}{=}$  indica igualdad en distribución, y  $H$  es el denominado parámetro de Hurst, que mide el grado de autosemejanza del proceso.

Si  $X$  es un proceso autosemejante, se verificará ademas:

$$r^{(m)}(k) = r(k) \quad , k \geq 0, m \in N \quad (\text{B.17})$$

siendo  $r^{(m)}(k)$  la función de autocorrelación del proceso agregado  $X^{(m)}$ .

La presencia de autosemejanza en un proceso tiene una serie de implicaciones equivalentes entre sí. En primer lugar, la autocorrelación del proceso decae hiperbólicamente en lugar de exponencialmente, lo cual implica que la función de autocorrelación es no sumable:

$$\sum_k r(k) = \infty \quad (\text{B.18})$$

Esta propiedad indica que los procesos autosemejantes muestran dependencia a largo plazo (*long-range dependence*), es decir, el valor del proceso en un instante determinado está correlado con todos los valores futuros.

En segundo lugar, la varianza del proceso agregado  $X^{(m)}$  decrece más lentamente que la inversa del nivel de agregación  $m$ , es decir:

$$\text{var}(X^{(m)}) \sim m^{-\beta}, m \rightarrow \infty, 0 < \beta < 1 \quad (\text{B.19})$$

El parámetro  $\beta$  que aparece en la ecuación anterior, está relacionado con el parámetro de Hurst de la siguiente forma:

$$H = 1 - \frac{\beta}{2} \quad (\text{B.20})$$

Por lo tanto, para procesos autosemejantes se verifica  $\frac{1}{2} < H < 1$ , de forma que el grado de autosemejanza del proceso es mayor cuanto más cerca se encuentre  $H$  de 1. Así, la detección de autosemejanza en un proceso se reduce a comprobar si  $H$  es significativamente diferente de  $\frac{1}{2}$ .

Por último, la densidad espectral de potencia obedece a una ley de potencia cerca del origen, como muestra la siguiente ecuación:

$$f(\lambda) \sim \lambda^{-\gamma}, \lambda \rightarrow 0, 0 < \gamma < 1 \quad (\text{B.21})$$

donde el parámetro  $\gamma$  viene dado por:

$$\gamma = 1 - \beta \quad (\text{B.22})$$

A partir de las propiedades exhibidas por los procesos autosemejantes, se han deducido varios métodos para determinar si un proceso es autosemejante, y en caso de que lo sea, estimar su grado de autosemejanza. En [17] se recogen cuatro de estos métodos. En esta Tesis Doctoral se ha empleado uno de ellos, conocido como *variance-time plot*, que consiste en representar la varianza del proceso agregado  $X^{(m)}$  en función del nivel de agregación  $m$  ( $m \geq 1$ ) en ejes logarítmicos, y ajustar los puntos resultantes a una línea mediante regresión por mínimos cuadrados, ignorando los valores pequeños de  $m$ . Una pendiente ( $-\beta$ ) comprendida entre  $-1$  y  $0$  es, de acuerdo con la ecuación (B.19), indicativa de autosemejanza, pudiendo obtener una estima del grado de autosemejanza mediante  $H = 1 - \beta/2$ .

### **Generación de tráfico autosemejante**

En la literatura pueden encontrarse diversos métodos para generar tráfico autosemejante. El método mas extendido consiste en multiplexar un número elevado de fuentes ON/OFF, donde la duración de los periodos ON y OFF sigue distribuciones del tipo *heavy tailed*, y donde las fuentes generan datos a velocidad constante durante el periodo ON. En [15]-[17] puede encontrarse información más detallada.

En [18] se emplea este método para explicar la autosemejanza observada en el tráfico WWW de las redes de datos. Para ello, se supone que la actividad de un usuario se puede describir mediante una secuencia de periodos ON y periodos OFF, donde el estado ON corresponde al periodo de transmisión en el que se transfiere un documento Web, y el estado OFF al periodo de silencio entre transmisiones. Dentro del periodo de silencio habrá que distinguir dos tipos diferentes. En primer lugar se encuentra el periodo de procesado del objeto principal en el que el navegador determina el número de objetos secundarios y la localización de los mismos. El segundo tipo de periodo de silencio es el tiempo de lectura de la página web, es decir, el periodo de tiempo transcurrido desde que se solicita la transferencia de una página web hasta que se solicita otra página distinta. Analizando la distribución de los periodos de transmisión y silencio a partir de datos registrados en redes reales, se propone que la autosemejanza

del tráfico WWW se debe principalmente a la distribución del tamaño de los archivos presentes en los servidores Web y al modo en el que el ser humano procesa la información, que en último termino determina la distribución de los periodos de silencio.

Debido a su simplicidad y facilidad de implementación, el modelo descrito en [18] será el empleado en esta Tesis Doctoral para generar tráfico autosemejante. Las pruebas realizadas mostrarán la validez del mismo.

### **Modelo de tráfico WWW implementado**

Como ya se ha explicado anteriormente, cada usuario web transmite en una sesión de duración infinita, es decir, está continuamente generando peticiones web a lo largo de toda la simulación. La actividad de un usuario durante su sesión se puede representar mediante un proceso que alterna peticiones de descarga de páginas web con periodos de silencio. Una página web está formada por un archivo que contiene el código HTML y, posiblemente, referencias a otros archivos, típicamente gráficos o imágenes, que pueden ser necesarios para visualizar correctamente los contenidos de la página. Por lo tanto, una petición de descarga resultará en la transferencia de varios ficheros desde uno o varios servidores web. La forma en que se transfieren los diferentes ficheros de una página web depende del navegador utilizado y de la versión del protocolo HTTP. HTTP 0.9 y 1.0 utilizan una conexión TCP diferente para cada archivo. Cuando se termina de transmitir un archivo se cierra la conexión, y debe establecerse una nueva para el siguiente archivo. Por el contrario, HTTP 1.1 tras el envío del objeto principal permite transferir varios archivos en una única conexión TCP, por lo que no se pierde tiempo abriendo y cerrando conexiones TCP. En el modelo utilizado en esta Tesis se ha considerado la versión HTTP 1.1, por ser la más actual, por lo que se utiliza una conexión TCP inicial para descargar el archivo principal y posteriormente una única conexión TCP para la transmisión de todos los archivos descargados de una página web.

Con las consideraciones anteriores, el tráfico generado por un usuario durante una sesión WWW se puede representar mediante un proceso ON/OFF. El estado ON representa el intervalo de tiempo durante el cual se están transfiriendo los archivos, mientras que el estado OFF representa un periodo en el que el terminal no está recibiendo datos. Un terminal estará en estado OFF o bien porque esté procesando el objeto principal o bien porque el usuario esté examinando la última página transferida.

La composición y el tamaño de la descarga web dependerá del tamaño del objeto principal y del número y tamaño de los objetos secundarios, es decir, los ficheros incrustados en el hipertexto del objeto principal.

Por otro lado, el tiempo de lectura entre dos peticiones de descarga deberá de caracterizarse también experimentalmente.

La Tabla B.3 recoge las distribuciones y parámetros utilizados para caracterizar el tráfico WWW a nivel de conexión. Por su parte, la Figura B.17 muestra el diagrama de flujo del modelo de tráfico web implementado.

	Distribución	Parámetros	Estadísticos
Tiempo entre peticiones de descarga	Exponencial	$\lambda = 0.165$	$\mu = 30 \text{ s}, \sigma = 30 \text{ s}$
Tamaño de objeto principal	Lognormal	$\alpha = 1.37, \beta = 8.35$	$\mu = 10 \text{ kB}, \sigma = 25 \text{ kB}$
Número de objetos secundarios	Pareto Truncada	$\alpha = 1.1, k = 2$	$\mu = 5.55, \text{max} = 55$
Tamaño de los objetos secundarios	Lognormal	$\alpha = 2.36, \beta = 6.17$	$\mu = 7.7 \text{ kB}, \sigma = 126 \text{ kB}$
Duración 'OFF' procesado del OP ( $T_p$ )	Exponencial	$\lambda = 7.69$	$\mu = 0.13 \text{ s}, \sigma = 0.13 \text{ s}$

**Tabla B.3 Parámetros del modelo de tráfico WWW**

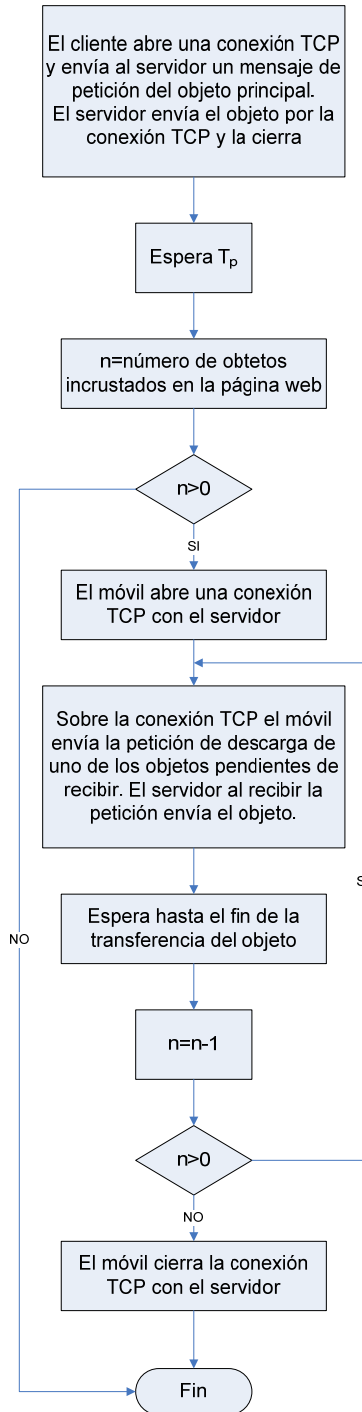


Figura B.17 Proceso de transmisión de página web [18]

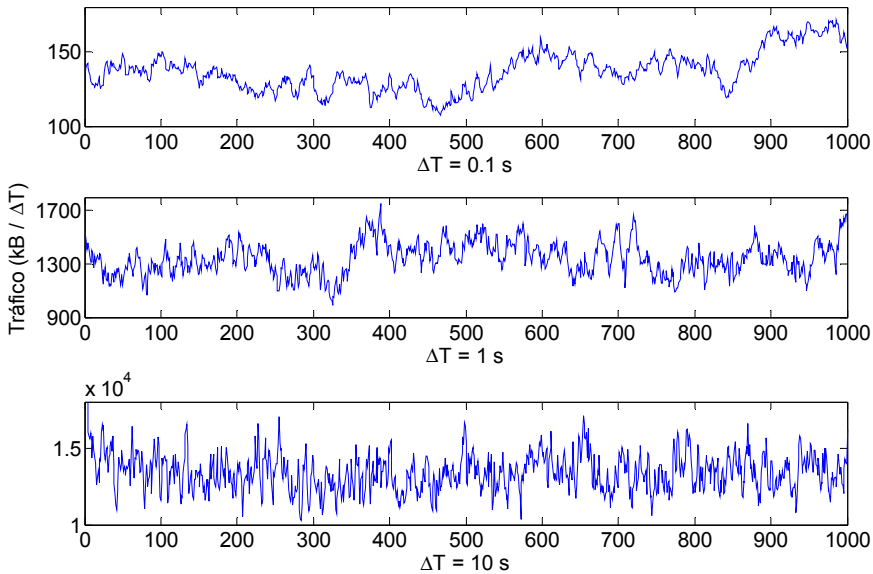
## Validación del modelo de tráfico WWW

En este apartado se describen los resultados de las pruebas realizadas para evaluar el modelo de tráfico WWW. Para que un modelo de tráfico WWW sea válido debe reflejar la propiedad de autosemejanza que exhibe el tráfico de datos de las redes reales.

La primera consecuencia de la autosemejanza es que el tráfico WWW mantiene sus propiedades estadísticas frente a cambios de escala temporales, es decir, la distribución del tráfico se parece a sí misma cuando se observa en diferentes escalas de tiempo. Así pues, la forma más sencilla de valorar la presencia de autosemejanza, consiste en observar el patrón de tráfico generado en varias escalas de tiempo distintas. La Figura B.18 recoge tres gráficas obtenidas a partir del modelo de tráfico WWW implementado. Cada una de las gráficas muestra la intensidad de tráfico, en kBytes por unidad de tiempo, generada por un conjunto de usuarios, donde la unidad de tiempo varía de una gráfica a otra. En la primera de ellas la unidad de tiempo es de 0.1 segundos, en la segunda es de 1 s, mientras que en la tercera es de 10 s.

Se puede observar que las tres gráficas son muy similares. En particular, se puede ver que en las tres escalas de tiempo el tráfico alterna periodos de gran actividad con periodos de menor actividad, constatando el hecho de que no existe un tamaño típico de ráfaga. Tal y como se observó al describir el tráfico WWW, ésta es una propiedad característica del tráfico autosemejante.

Así pues, las gráficas mostradas en la Figura B.18 permiten intuir la presencia de autosemejanza en el tráfico, pero, para poder comprobarla con mayor rigurosidad e incluso valorar su grado de intensidad, se requieren métodos que hagan uso de las propiedades matemáticas de los procesos autosemejantes. Tal y como se comentó antes, en esta Tesis Doctoral se ha empleado un método conocido como *variance-time plot*, cuyo procedimiento y fundamento teórico fue presentado en el apartado dedicado a introducir las propiedades de los procesos autosemejantes.



**Figura B.18** Tráfico generado por el modelo en tres escalas de tiempo

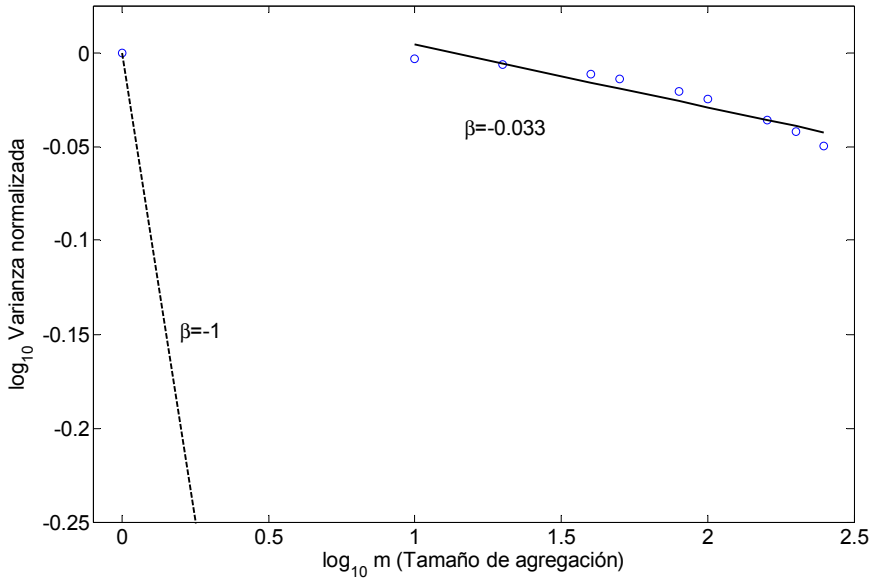
La Figura B.19 muestra el resultado del método *variance-time plot* aplicado al tráfico generado por el modelo de tráfico WWW para distintos números de usuarios.

Como se puede observar, la pendiente de la trazas es mayor que -1, lo que sugiere la presencia de autosemejanza. Se ve además que la pendiente de la traza está muy próxima a cero con lo que el parámetro de Hurst es prácticamente 1, lo que corrobora el alto grado de autosemejanza conseguido por el modelo utilizado.

Las observaciones anteriores demuestran que el modelo de tráfico implementado es capaz de reproducir el comportamiento del tráfico registrado en las redes de datos en cuanto a autosemejanza se refiere.

Por último, es importante destacar que, con los parámetros seleccionados, el tamaño de la descarga web media resulta ser de 51 kB mientras que la tasa binaria por usuario media es de 13.12 kbps. Estos valores están en consonancia con los valores publicados en [19][20]. De igual manera, los tiempos de lectura también se corresponden con los presentados en [19].





**Figura B.19** *Variance-time plot*

### B.1.8.3 Modelo de tráfico de vídeo conferencia

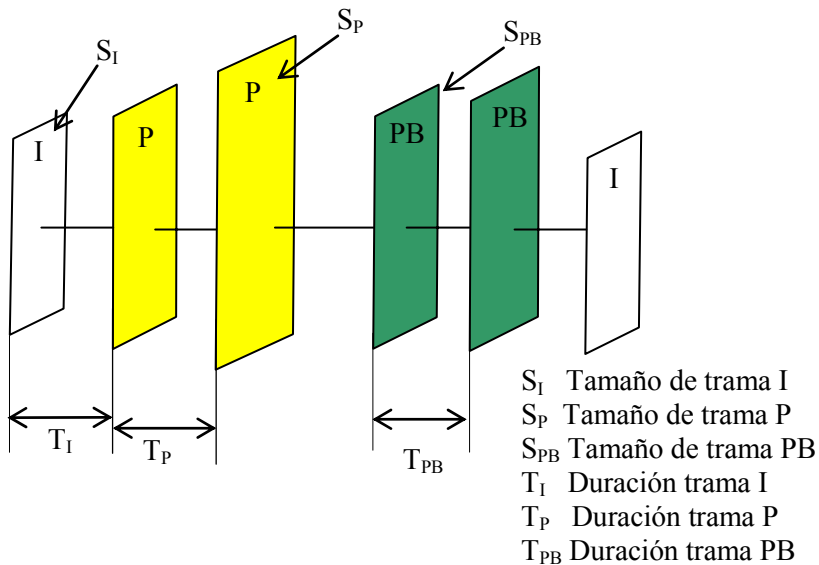
Resulta fácil demostrar que para la transmisión de información de vídeo en una red de comunicaciones móviles es imprescindible la aplicación de algún método de codificación que permita conseguir elevados niveles de compresión ya que, de lo contrario, se requerirían tasas de transmisión varios órdenes de magnitud por encima de las que cualquier sistema de comunicaciones móviles puede proporcionar.

El estándar de codificación H.263 [21], que surgió con el objetivo de mejorar las prestaciones ofrecidas por su predecesor, el estándar H.261, permite relaciones de compresión del orden de 100:1 o incluso 200:1, con una calidad de imagen aceptable [22], reduciendo la velocidad binaria requerida a valores en el rango de 16 a 300 kbps aproximadamente [23]. En el estándar H.263, la codificación de vídeo se realiza imagen a imagen, es decir, cada uno de los fotogramas que componen la secuencia de vídeo es codificado dando lugar a una trama de vídeo que se transmite a través del

interfaz aire hacia el receptor. Dependiendo del modo en que se codifica la imagen de vídeo, la trama resultante puede ser de tipo I (se codifica la imagen completa sin referencia a ninguna otra), tipo P (se codifica sólo la diferencia entre la imagen a transmitir y la imagen anterior, es decir, se codifica utilizando la imagen anterior como referencia) o tipo PB (la imagen se codifica haciendo referencia tanto a la imagen anterior como a la posterior, también incluida en la trama). Además del tipo de trama (I, P o PB), el tamaño de cada trama y el tiempo que transcurre hasta que se genera la siguiente son variables relevantes en el modelado de vídeo H.263 (Figura B.20).

El modelo de tráfico H.263 empleado se basa en el uso de una secuencia de vídeo real utilizada para la extracción de sus parámetros característicos. Los datos que el modelo requiere de dicha secuencia son el instante en el que se genera cada trama de vídeo codificada, su tipo (I, P o PB) y su tamaño. Para la obtención de estos datos es necesario codificar previamente la información de vídeo original sin comprimir mediante un codificador H.263. Sin embargo, dicha tarea no ha sido necesaria ya que los resultados han sido obtenidos de [24]. En esta URL se encuentran disponibles de forma pública los resultados obtenidos al codificar diversas secuencias de vídeo originales, de una hora de duración aproximadamente, con H.263 a 16 kbps, 32kbps, 64 kbps y 256 kbps así como con tasa binaria variable VBR (*Variable Bit-Rate*). Para esta Tesis Doctoral se ha optado por una secuencia de una hora de grabación de vídeo conferencia codificada a 64 kbps ya que esta tasa binaria es la que más se adecua a las tasas ofrecidas por los esquemas de codificación de canal disponibles en UMTS. Los detalles acerca de la configuración del codificador H.263 empleado se encuentran en [25].

A partir de unas tablas disponibles en [24] en las que se especifica cada uno de los tipos de tramas generadas en la codificación H.263 y su tamaño y duración, el modelo de vídeo extrae ciertos valores de probabilidades y CDFs (*Cumulative Distribution Functions*) de los parámetros más característicos en los que se basa posteriormente a la hora de generar las tramas de vídeo en tiempo de simulación. A continuación se explicará el procedimiento mediante el cual el modelo genera las tramas de vídeo y cómo para ello hace uso de los datos relativos a la secuencia de vídeo codificada.



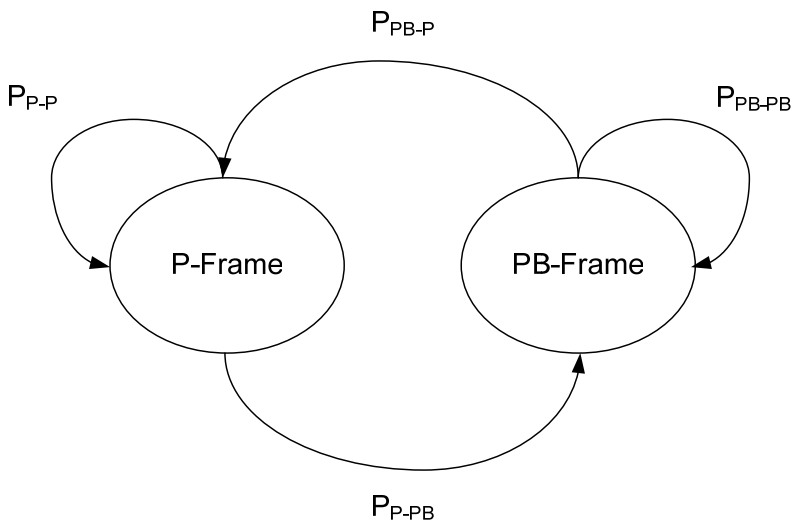
**Figura B.20 Parámetros relevantes del tráfico H.263**

Cuando el modelo ha de generar una trama de vídeo debe seguir varios pasos. El primero de ellos consiste en decidir de qué tipo será la trama generada (I, P o PB). Dicha decisión se toma en base a una cadena de Markov que modela las transiciones entre los diversos tipos de tramas durante la transmisión de la secuencia de vídeo codificada.

En principio, el modelo de Markov que rige la decisión del tipo de trama a generar debería contemplar tres estados, uno por cada tipo de trama. Sin embargo, hay que tener en cuenta varios factores. En primer lugar, las tramas I se generan en instantes de tiempo concretos, es decir, instantes de tiempo que no son aleatorios sino que son perfectamente conocidos de antemano (la frecuencia de refresco de tramas I es un parámetro de configuración del codificador H.263). Así pues, la generación de una trama de tipo I no vendrá determinada por la cadena de Markov sino por la frecuencia de refresco de tramas I utilizada para codificar la secuencia de vídeo. Además, por diseño del estándar H.263 [21], la primera trama generada siempre es de tipo I, seguida inmediatamente por una trama de tipo P, por lo que esta transición no ha de ser

modelada. Por estos dos motivos la cadena de Markov necesaria para representar la decisión del tipo de trama a generar consta de dos estados, tal y como se muestra en la Figura B.21. Un estado representa la generación de una trama de tipo P y el otro representa la generación de una trama de tipo PB.

Cuando el modelo comienza a funcionar genera, en primer lugar, una trama de tipo I e inmediatamente a continuación una trama de tipo P. Posteriormente, el tipo de las tramas siguientes (P o PB) y las transiciones entre ellos vendrán ya determinados por las probabilidades indicadas en el modelo de Markov. Cuando se haya cumplido el periodo de refresco de tramas I, se volverá a generar una nueva trama de tipo I, luego una trama de tipo P y se repetirá el proceso. Las probabilidades  $P_{P-P}$ ,  $P_{P-PB}$ ,  $P_{PB-P}$  y  $P_{PB-PB}$  contempladas en la cadena de Markov de la Figura B.21 pueden ser obtenidas a partir de la secuencia de vídeo codificada, contando el número de veces que ocurre una transición entre la trama  $X$  y la trama  $Y$  y dividiéndolo entre el número total de transiciones ocurridas a lo largo de la secuencia de vídeo. En concreto, estas probabilidades se han obtenido para el caso de una vídeo conferencia real codificada en H.263 disponible en [24], resumiéndolas en la Tabla B.4.



**Figura B.21** Probabilidad de transición entre tramas

	$P_{P-P}$	$P_{P-PB}$	$P_{PB-P}$	$P_{PB-PB}$
Probabilidad	0.6909	0.3091	0.0376	0.9624

**Tabla B.4 Probabilidades de transición para una vídeo conferencia real a 64kbps**

Tras decidir el tipo de trama que se generará, el último paso consiste en obtener su tamaño y duración. El proceso seguido hace uso de varias CDFs obtenidas a partir de los datos de la secuencia de vídeo real codificada (ver [14] para una descripción detallada) y respeta la correlación existente entre el tamaño de una trama y su duración. Así pues, la secuencia de vídeo real se utiliza para la obtención no sólo de estas CDFs sino también del valor de las probabilidades  $P_{X,Y}$  de la cadena de Markov de la Figura B.21. La Tabla B.5 resume los parámetros representativos para la secuencia de vídeo empleada.

Parámetro	Distribución matemática	Parámetros
Tamaño de las tramas tipo I (Bytes)	Normal	$\mu=2301$ $\sigma=648$
Tamaño de las tramas tipo P (Bytes)	Normal	$\mu=837$ $\sigma=125$
Tamaño de las tramas tipo PB (Bytes)	Normal	$\mu=1821$ $\sigma=107$
Duración de las tramas tipo I (ms)	Normal	$\mu=280$ $\sigma=80$
Duración de las tramas tipo P (ms)	Normal	$\mu=138$ $\sigma=61$
Duración de las tramas tipo PB (ms)	Normal	$\mu=224$ $\sigma=87$

**Tabla B.5 Parámetros representativos del modelo de tráfico de vídeo H.263 a 64kbps**

El procedimiento de generación de tráfico será el siguiente:

1. Se genera una trama tipo I
2. Posteriormente se genera una trama tipo P

3. A partir de aquí la sucesión de tramas se establece a partir del proceso de Markov definido en la Figura B.21.
4. Se calcula el tamaño y la duración de cada trama según la Tabla B.5.

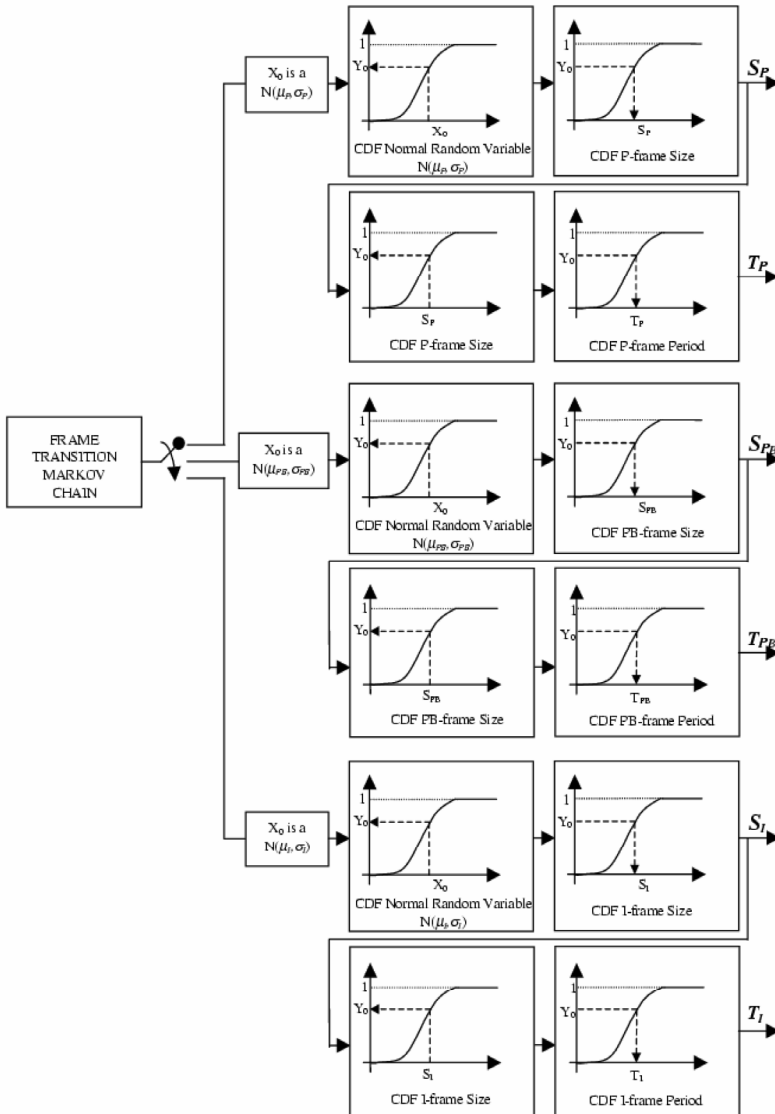


Figura B.22 Modelo de tráfico H.263 [14]

Para implementar el último paso de este algoritmo, es necesario calcular, a partir de los estadísticos obtenidos, la función distribución de probabilidad,  $F$ , tanto del tamaño como de la duración de cada uno de los tipos de trama. Para obtener una realización concreta deberá de hacerse uso de la función inversa,  $F^{-1}$ . Esta idea queda resumida en la Figura B.22.

### Validación del modelo de tráfico de vídeo conferencia

Inicialmente, el modelo de tráfico se ha analizado a través de sus estadísticos de primer orden, es decir, comprobando que la media y la desviación típica de los resultados obtenidos son las esperadas.

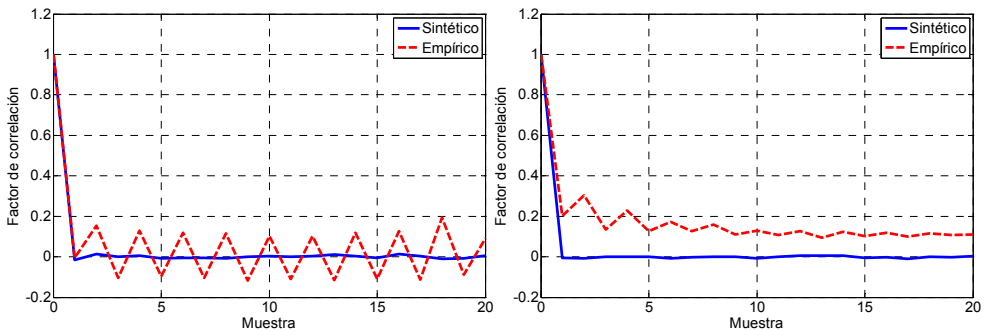
Se ha hecho uso de diez realizaciones diferentes de tráfico para obtener los estadísticos con un intervalo de confianza del 95%. Estos resultados obtenidos, resumidos en la en la Tabla B.6, muestran que tanto los valores empíricos como los sintéticos coinciden, demostrando así que el tamaño y la duración de las tramas están bien representadas con el modelo propuesto.

Parámetro	Empírico	Sintético
Media del tamaño de las tramas P	$\mu=837$	$\mu=838 \pm 2.432$
Desviación del tamaño de las tramas P	$\sigma=125$	$\sigma=125.39 \pm 1.198$
Media del tamaño de las tramas PB	$\mu=1821$	$\mu=1821 \pm 0.582$
Desviación del tamaño de las tramas PB	$\sigma=107$	$\sigma=108.59 \pm 0.3539$
Media de la duración de las tramas tipo P	$\mu=138$	$\mu=138.41 \pm 1.564$
Desviación de la duración de las tramas tipo P	$\sigma=61$	$\sigma=61.75 \pm 0.731$
Media de la duración de las tramas tipo PB	$\mu=224$	$\mu=224.12 \pm 0.72$
Desviación de la duración de las tramas tipo PB	$\sigma=87$	$\sigma=86.59 \pm 0.352$

**Tabla B.6 Estadísticos de primer orden del modelo de tráfico de vídeo H.263**

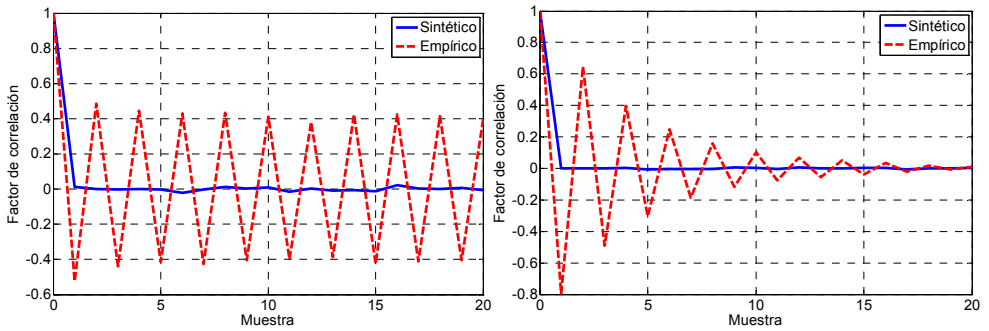
A continuación, se ha investigado la capacidad del modelo para capturar los estadísticos de segundo orden de las secuencias de vídeo, analizando las funciones de autocorrelación. La Figura B.23 y la Figura B.24 muestran respectivamente las funciones de autocorrelación para el tamaño de trama y la duración de las tramas P y PB. Se puede observar que el tráfico sintético presenta una total ausencia de correlación entre las muestras, lo cual es lógico puesto que se trata de una realización aleatoria sin ningún grado de correlación. Sin embargo las muestras empíricas muestran un grado de correlación periódico importante. Este hecho no se veía reflejado en [14] ya que en ese caso se consideró un vídeo correspondiente a una película que carecía de la correlación que presenta una vídeo conferencia, donde la imagen es más bien estática. A pesar de esta limitación del modelo, se puede considerar que el tráfico sintético presenta una similitud suficiente con el tráfico real al promediar las muestras obtenidas de correlación.

Finalmente para realizar un análisis de las distribuciones se ha representado un diagrama Q-Q (*Quantile-Quantile plot*) para las distribuciones empíricas (las de las secuencias de vídeo real) y sintéticas de los tamaños y la duración de las tramas P y PB.



**Figura B.23** Función de autocorrelación del tamaño de las tramas P(izq) y-PB(dcha)

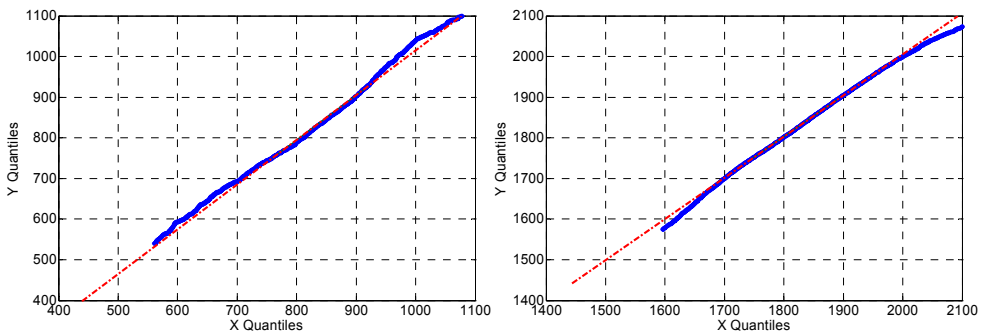




**Figura B.24 Función de autocorrelación de la duración de trama P(izq) y-PB(dcha)**

El diagrama Q-Q se basa en el concepto de cuantil: el cuantil  $q$  de  $F(x)$  es un número  $x_q$  tal que  $F(x_q) = q$ , donde  $F$  es la función de distribución de probabilidad de una variable aleatoria  $X$ . El diagrama Q-Q representa los cuantiles de dos distribuciones una en cada eje de manera que si ambas secuencias siguen la misma distribución estos puntos deben estar cerca de la primera bisectriz.

La Figura B.25 y la Figura B.26 representan los diagramas Q-Q para el tamaño y duración de las tramas P y PB. Se puede apreciar como las distribuciones siguen correctamente la línea recta, reflejando que hay un buen ajuste entre ambas distribuciones la empírica y la sintética. Por lo tanto, el modelo propuesto es capaz de capturar perfectamente las propiedades estadísticas de primer orden de la secuencia de vídeo real y se aproxima a sus estadísticos de segundo orden.



**Figura B.25 Diagrama Q-Q del tamaño de las tramas P(izq) y-PB(dcha)**

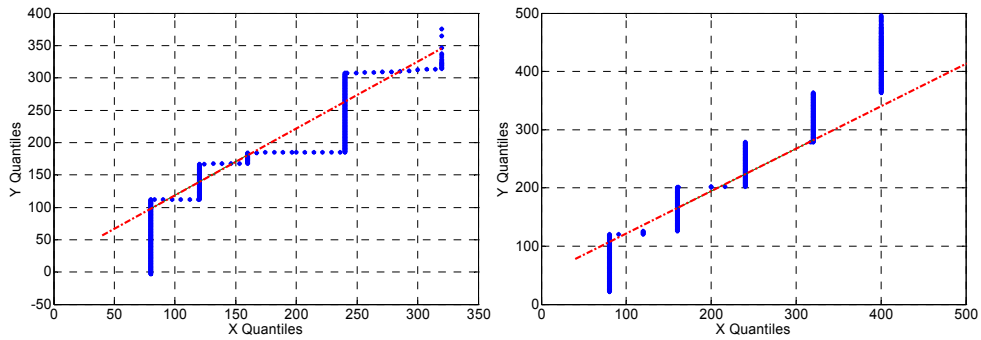


Figura B.26 Diagrama Q-Q de la duración de las tramas P(izq) y-PB(dcha)

### B.1.8.4 Modelo de tráfico FTP

En las aplicaciones FTP, una sesión consiste en una secuencia de transferencia de ficheros, separados por un tiempo entre descargas o tiempo de lectura. Por lo tanto, los dos principales parámetros de la sesión FTP son:

- $S$ : el tamaño del fichero descargado
- $D_{pc}$ : tiempo de lectura, es decir, intervalo de tiempo transcurrido entre dos peticiones consecutivas de descarga de fichero.

Los parámetros utilizados para el modelo de la aplicación FTP se han extraído de [18] y se resumen en la siguiente tabla:

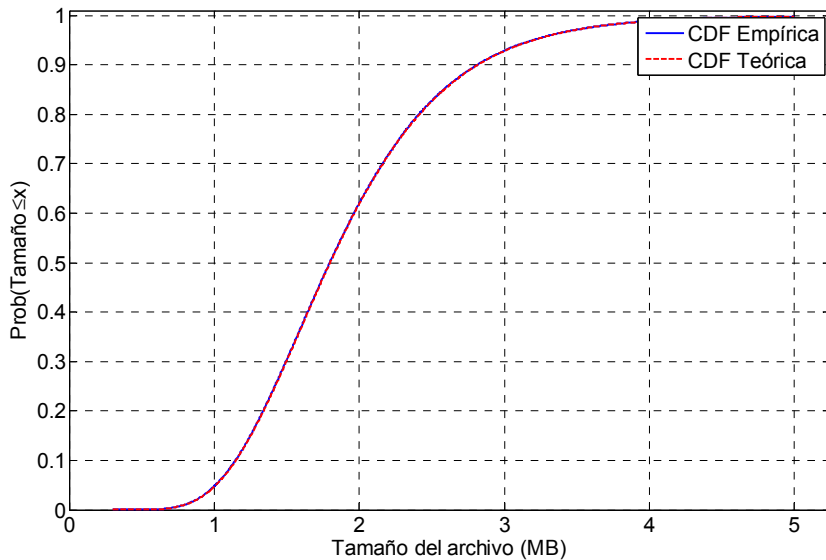
	Distribución	Parámetros	Estadísticos
Tamaño del fichero	Lognormal Truncada	$\alpha = 0.35, \beta = 14.45$ Máximo= 5MB	$\mu=2$ MB, $\sigma=0.722$ MB
Tiempo de lectura	Exponencial	$\lambda = 0.0055$	$\mu=180$ s, $\sigma=180$ s

Tabla B.7 Parámetros del modelo de tráfico FTP

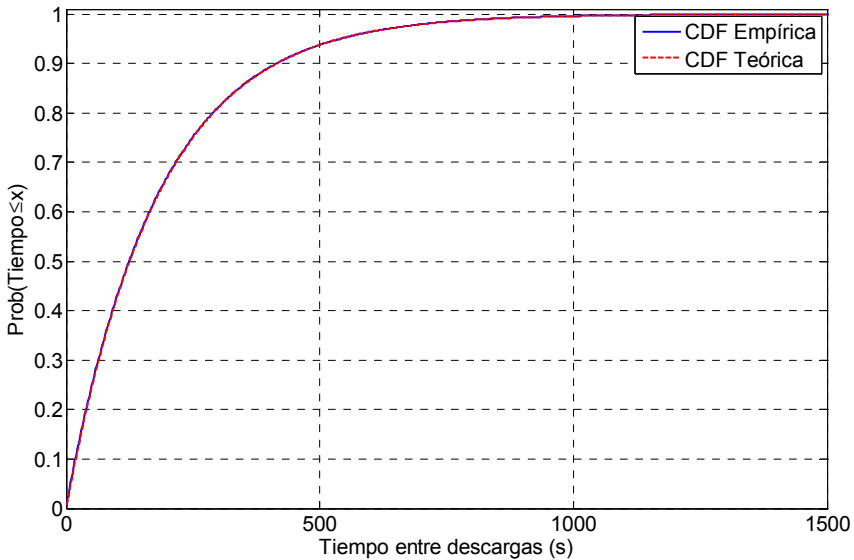
## Validación del modelo de tráfico FTP

Para validar el modelo de tráfico FTP se ha analizado una secuencia de un millón de descargas, comprobando los valores medios y máximos de cada parámetro y comparando también la función de distribución de probabilidad con la teórica esperada.

La Figura B.27 muestra la CDF del tamaño del archivo descargado mientras que la Figura B.28 refleja la distribución del tiempo de lectura de los archivos FTP. En ambas figuras se puede comprobar perfectamente como la distribución generada sigue el mismo comportamiento que la teórica.



**Figura B.27 Distribución de probabilidad del tamaño del archivo descargado**



**Figura B.28 Distribución de probabilidad del tiempo de lectura del archivo**

La Tabla 2.2 resume el valor medio y máximo de los principales parámetros característicos del modelo de tráfico FTP, calculados a partir de las secuencias generadas. Todos los valores obtenidos se encuentran dentro de los márgenes normalmente considerados.

Parámetro	Valor Obtenido
Tamaño medio del archivo descargado	1.92 MB
Tamaño máximo del archivo descargado	5 MB
Tiempo medio de lectura del archivo descargado	180.15 s
Tiempo máximo de lectura del archivo descargado	2416.92 s

**Tabla B.8 Valor medio y máximo de los parámetros considerados**

---

## B.2 Interfaz entre Simulación de Enlace y Simulación de Sistema

Para reducir la complejidad de la simulación se suele descomponer ésta en dos etapas o niveles de abstracción, conocidos como nivel de enlace y nivel de sistema o de red.

La principal razón para proceder de esta manera reside en la diferente resolución temporal requerida en ambos niveles. En las simulaciones a nivel de sistema es necesario extender las simulaciones durante al menos 10-20 minutos [11][26] para obtener unos estadísticos aceptables de los modelos aleatorios de generación de tráfico y de movilidad empleados. Por otra parte, en el nivel de enlace, para obtener una caracterización adecuada del enlace radio, se requiere una resolución a nivel de *chip* o de símbolo. Si se considerara esta resolución temporal en las simulaciones a nivel de sistema el coste computacional del simulador resultante sería inmanejable.

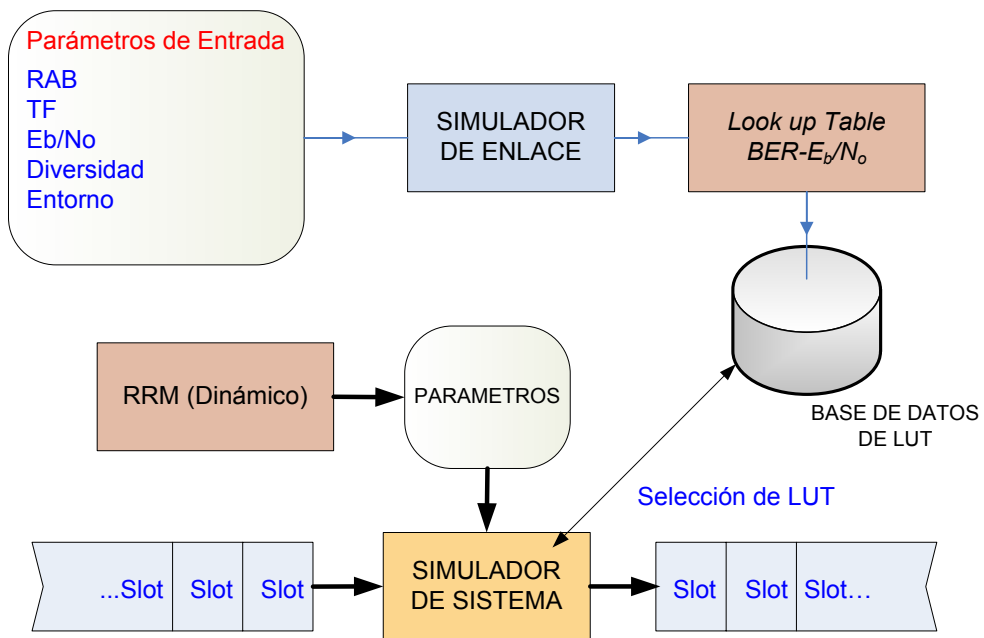
La interacción entre ambos niveles de simulación se lleva a cabo mediante el empleo de interfaces, utilizando los resultados de la simulación a nivel de enlace como información de entrada para la simulación a nivel de sistema. El comportamiento de una transmisión a nivel de enlace se representa de forma simplificada mediante un conjunto de tablas que reciben el nombre de *Look-Up Tables* (LUTs). Estas LUTs no son más que tablas numéricas en las cuales se establece una relación entre la calidad experimentada en el enlace radio (expresada, por ejemplo, en términos de relación *CIR* media o  $E_b/N_0$  media) y su correspondiente tasa de error (expresada, por ejemplo, en términos de BER).

La Figura B.29 muestra claramente los dos pasos necesarios en la simulación de un sistema. Primero se obtienen los estadísticos de primer orden del *BER* para un valor medio de  $E_b/N_0$ , definiendo el conjunto de LUTs necesarias, para posteriormente utilizar dichas tablas en el simulador de sistema.

Para modelar la nueva arquitectura MNB es necesario conocer, en primer lugar, si se está transmitiendo desde una o dos antenas. Si sólo se transmite desde una antena el funcionamiento es el convencional: se calculan las pérdidas de propagación y las

interferencias recibidas y se obtiene la  $E_b/N_0$  media. Finalmente, a través de la LUT que caracteriza esta situación, se calcula la calidad del enlace. Si, por el contrario, se transmite desde dos antenas, se deberá calcular, en primer lugar, la relación  $E_b/N_0$  media como suma de las  $E_b/N_0$  parciales de cada enlace. Conociendo además la relación de potencia entre ambas contribuciones y el tipo de celdas involucradas, se selecciona la LUT que mejor caracteriza esta situación para, finalmente, obtener la calidad del enlace en términos de *BER*.

El procedimiento seguido para la obtención de las LUTs que caracterizan la arquitectura MNB se explicó con detalle en Capítulo 3. El procedimiento seguido para la obtención de la LUT que mejor caracteriza al enlace se resume en la Figura B.30.



**Figura B.29 Interoperabilidad entre simulaciones de Enlace y de Sistema [27]**

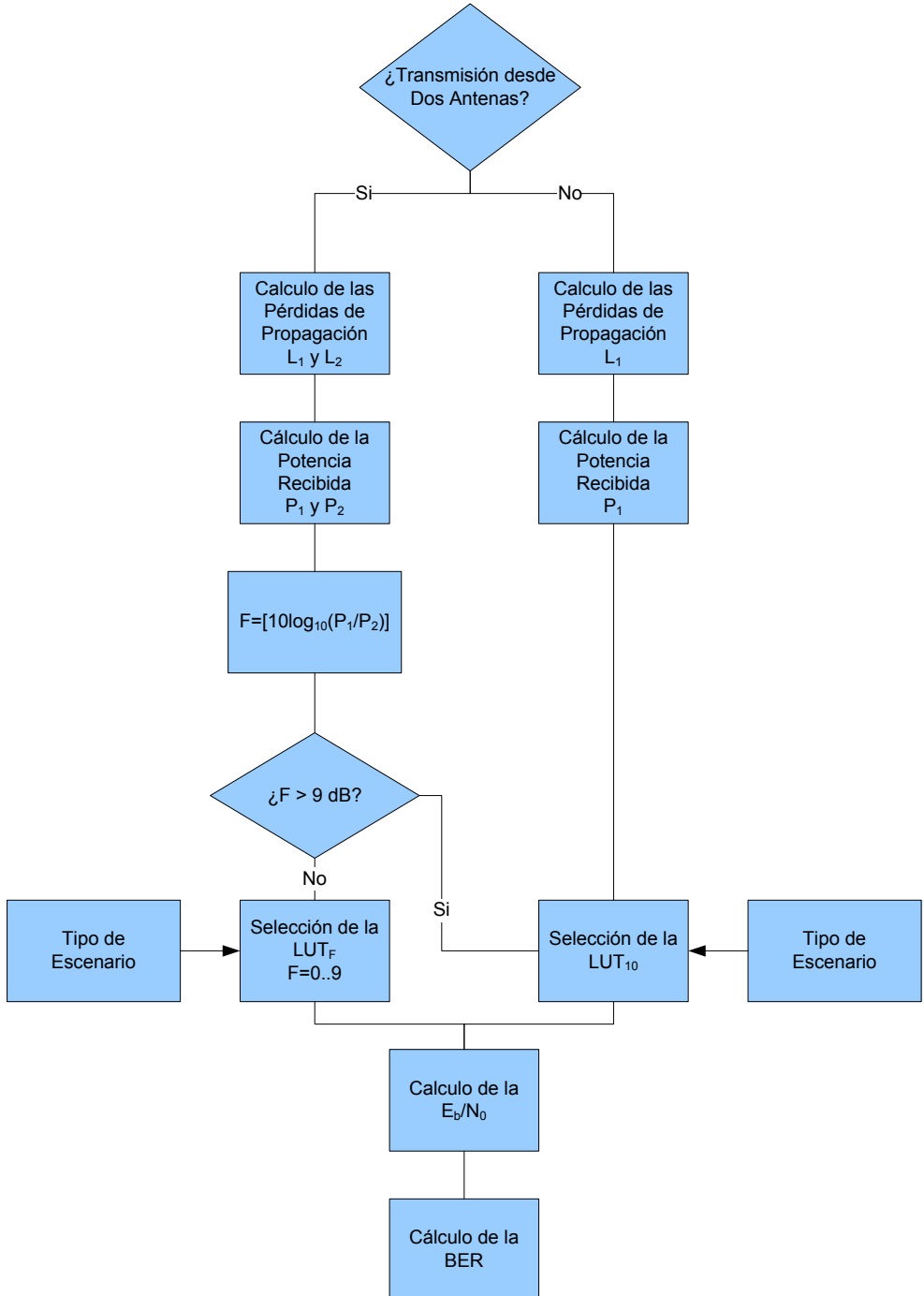


Figura B.30 Selección de la LUT más adecuada

## Referencias

- [1] 3GPP TR 101 112, “Universal Mobile Telecommunications System (UMTS); Selection procedures for the choice of radio transmission technologies of the UMTS”, UMTS 30.03 v 3.2.0, 2004.
- [2] MORANS Task Force. “MORANS White Paper”, COST273 TD (03) 057, 2003.
- [3] MORANS Task Force. “MORANS White Paper – Update”, COST273 TD (04) 062, 2004.
- [4] NEWCOM Project FP6-IST-507325-2002. Available at <http://newcom.ismb.it>.
- [5] J. F. Monserrat et al., “MORANS Layer 2: Traffic Information Synthetic Scenario”, COST273 TD (04) 071, 2004.
- [6] J. F. Monserrat et al., “MORANS Layer 5: Site Locations Synthetic Scenario”. COST273 TD (04) 072, 2004.
- [7] R. Fraile and J.F. Monserrat, “MORANS Layer 6: Propagation Data Synthetic Scenario”, COST273 TD (04) 073, 2004.
- [8] R. Fraile and J.F. Monserrat, “MORANS Layer 7: Node B Parameters Synthetic Scenario”, COST273 TD (04) 074, 2004.
- [9] L. Ferreira, J. F. Monserrat et al., “Definition of Reference Scenarios for the evaluation of radio resource allocation algorithms”, IST-TUL\_WPR7\_068\_04, 2006.
- [10] 3GPP TS 25.214 V5.9.0, “Physical layer procedures (FDD)”, 2004.
- [11] H. Holma and A. Toskala, “WCDMA for UMTS”, John Wiley & Sons Ltd. 2000.
- [12] Jaana Laiho, Achim Wacker and Tomás Novosad, “Radio Network Planning and Optimisation for UMTS”, Ed. Wiley, 2005.
- [13] P.T. Brady. “A Model of On-Off Speech Patterns in Two-Way Conversation”, CBell System Technical Journal, vol. 48, pp. 2445-2472, 1969.
- [14] O. Lázaro. “Dynamic Radio Resource Management Algorithms and Traffic Models for Emerging Mobile Communication Systems”, Phd Thesis, University of Strathclyde, 2001.
- [15] W. Leland, M. Taqqu, W. Willinger, D. Wilson. “On the Self-Similar Nature of Ethernet Traffic”, IEEE/ACM Transactions on Networking, vol. 2, no. 1, pp. 1-15, 1994.



- 
- [16] V. Paxson, S. Floyd. “Wide-Area Traffic: The failure of Poisson Modeling”, IEEE/ACM Transactions on Networking, vol.3, no. 3, pp. 226-244, 1995.
- [17] M. Crovella, A. Bestavros. “Self-Similarity in World Wide Web Traffic Evidence and Possible Causes”, in Proc. of the ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems, pp. 160-169, May 1996.
- [18] 3GPP2-TSGC5. HTTP and FTP Traffic Model for 1xEV-DV Simulations.
- [19] A. Reyes-Lecuona, E. González-Parada, E. Casilari and A. Díaz-Estrella, “A page-oriented WWW traffic model for wireless system simulations”, in Proc. of the 16th International Teletraffic Congress (ITC'16), pp. 1271-1280, 1999.
- [20] N. Vicari, “Measurement and Modeling of WWW-Sessions”, Institute of Computer Science, University of Würzburg, Technical Report No. 184.
- [21] ITU-T, “Video Coding for Low Bit Rate Communication”, Recommendation H.263, 1998.
- [22] B. Girod, N. Farber, E. Steinbach, “Performance of the H.263 Video Compression Standard”, Journal on VLSI Signal Processing: System Signal, Image and Video Technology, vol. 17, pp. 101–111, 1997.
- [23] G. Cote, B. Erol, M. Gallant, F. Kossentini, “H.263+: Video Coding at Low Bit Rates”, IEEE Transactions on Circuits and Systems for Video Technology, vol. 8, no 7, pp. 849–866, 1998.
- [24] “MPEG-4 and H.263 Video Traces for Network Performance Evaluation”, <http://www-tkn.ee.tu-berlin.de/research/trace/trace.html>.
- [25] K. Rijkse, “H.263: Video Coding for Low-Bit-Rate Communication”, IEEE Communications Magazine, vol. 34, no 12, pp. 42–45, 1996.
- [26] S. Hämäläinen, H. Holma, K. Sipilä, “Advanced WCDMA Radio Network Simulator”, in Proc. of the IEEE International Symposium on Personal, Indoor and Mobile Radio Conference (PIMRC'99), pp. 951-955, 1999.
- [27] J.J.Olmos, S.Ruiz, COST273: MORANS, “Description of UTRA-FDD link level simulator”, May 2003.