UNIVERSIDAD MIGUEL HERNÁNDEZ DE ELCHE

ESCUELA POLITÉCNICA SUPERIOR DE ELCHE

INGENIERÍA TÉCNICA DE TELECOMUNICACIÓN
ESPECIALIDAD EN SISTEMAS DE TELECOMUNICACIÓN

"ESTUDIO Y DISEÑO DE LA CONFIGURACIÓN DE LINK ADAPTATION PARA SERVICIOS MULTIMEDIA EN SISTEMAS AVANZADOS DE COMUNICACIONES MÓVILES"

PROYECTO FIN DE CARRERA

Septiembre – 2003

AUTOR: Miguel López Benítez
DIRECTOR: Dr. Javier Gozálvez Sempere

Índice

Lista de Acrónimos 8

CAPÍTULO 1 – Introducción 15

1.1.- Introducción ... 17
1.2.- Estructura del trabajo ... 18

CAPÍTULO 2 – General Packet Radio Service 21

2.1.- Introducción ... 23
2.2.- ¿Qué es GPRS? ... 26
2.3.- Arquitectura de red ... 27
2.3.1.- Arquitectura de la red GSM ... 27
2.3.2.- Arquitectura de la red GPRS ... 29
2.4.- Arquitectura de protocolos ... 31
2.5.- Interfaz aire ... 35
2.5.1.- Aspectos básicos ... 35
2.5.1.1.- Canales ... 35
2.5.1.1.1.- Canales físicos (PDCH) ... 35
2.5.1.1.2.- Canales lógicos ... 35
2.5.1.1.2.1.- Canales de tráfico (PTCH) ... 36
2.5.1.1.2.2.- Canales de control (PCCH) ... 36
2.5.1.1.2.2.1.- Canales de difusión (PBCH) ... 36
2.5.1.1.2.2.2.- Canales de control común (PCCCH) ... 37
2.5.1.1.2.2.3.- Canales de control dedicado (PDCCH) ... 37
2.5.1.2.- Gestión de los recursos radio ... 38
2.5.1.2.1.- Compartimiento de los recursos radio ... 38
2.5.1.2.1.1.- Concepto de canal maestro/esclavo ... 38
2.5.1.2.1.2.- Asignación de recursos bajo demanda ... 39
2.5.1.2.2.- Modos de operación radio ... 41
2.5.1.2.2.1.- Modo inactivo ... 42
2.5.1.2.2.2.- Modo transferencia ... 42
2.5.1.2.3.- Acceso múltiple ... 42
2.5.1.2.3.1.- Jerarquía temporal ... 44
2.5.1.3.- Flujo de datos ... 46
2.5.2.- La capa RLC/MAC ... 48
2.5.2.1.- Servicios ... 48
2.5.2.2.- Funciones ... 48

2.5.2.3.- Identificadores 49
2.5.2.3.1.- Uplink State Flag (USF) 49
2.5.2.3.2.- Temporary Block Flow (TBF) 50
2.5.2.3.3.- Temporary Flow Identity (TFI) 50
2.5.2.4.- Modos de operación 50
2.5.2.4.1.- Modo de operación con confirmación 50
2.5.2.4.2.- Modo de operación sin confirmación 51
2.5.3.- La capa física 51
2.5.3.1.- La capa PLL 51
2.5.3.1.1.- Servicios 51
2.5.3.1.2.- Funciones 51
2.5.3.1.2.1.- Codificación de canal 52
2.5.3.1.2.1.1.- Esquema de codificación CS-1 53
2.5.3.1.2.1.2.- Esquema de codificación CS-2 53
2.5.3.1.2.1.3.- Esquema de codificación CS-3 54
2.5.3.1.2.1.4.- Esquema de codificación CS-4 55
2.5.3.1.2.1.5.- Comparación de los esquemas CS-1 a CS-4 55
2.5.3.1.2.2.- Entrelazado 56
2.5.3.2.- La capa RFL 57
2.6.- Sumario 58

CAPÍTULO 3 – Gestión de Recursos Radio: Link Adaptation 61

3.1.- Introducción 63
3.2.- Principios de la estructura celular 64
3.3.- Radiopropagación en el canal móvil 67
3.4.- Gestión de recursos radio 69
3.4.1.- Formulación del problema 69
3.4.2.- Principios de gestión 70
3.4.3.- Técnicas de gestión 71
3.4.3.1.- Control de admisión 71
3.4.3.2.- Control de la congestión 71
3.4.3.3.- Operación multislot 72
3.4.3.4.- Asignación de canales 72
3.4.3.5.- Salto en frecuencia 73
3.4.3.6.- Selección de estación base 73
3.4.3.7.- Control de potencia 74
3.4.3.8.- Redundancia incremental 74
3.4.4.- Integración de diversas técnicas 75
3.5.- Link Adaptation 76
3.5.1.- Introducción 76
3.5.2.- ¿Qué es Link Adaptation? 76
3.5.3.- Estimación de la calidad del canal 79
3.5.4.- Elección del modo de transporte 81
3.5.5.- Periodo de actualización 84
3.5.6.- Evaluación de la eficacia 85
3.5.7.- Beneficios de Link Adaptation 87
3.5.8.- Aplicación a otros sistemas 87
3.6.- Sumario 89

CAPÍTULO 4 – Servicios de Vídeo y Calidad de Servicio 91

4.1.- Introducción ... 93
4.2.- Servicios de vídeo ... 94
4.2.1.- Introducción ... 94
4.2.2.- Generalidades ... 94
4.2.3.- Fundamentos de codificación H.263 ... 95
4.2.4.- Modos opcionales H.263 ... 97
4.3.- Calidad de servicio para vídeo H.263 ... 98
4.4.- Sumario ... 99

CAPÍTULO 5 – Modelos y Herramientas de Simulación 101

5.1.- Introducción ... 103
5.2.- Simulación a nivel de enlace ... 104
5.3.- Simulación a nivel de sistema ... 105
5.3.1.- Arquitectura del simulador ... 106
5.3.1.1.- Entorno celular ... 106
5.3.1.2.- Enlace radio ... 107
5.3.1.3.- Estación base ... 108
5.3.1.4.- Estación móvil ... 109
5.3.2.- Modelos de tráfico ... 110
5.3.2.1.- Tráfico World Wide Web (WWW) ... 111
5.3.2.2.- Tráfico de correo electrónico ... 112
5.3.2.3.- Tráfico de vídeo H.263 ... 113
5.3.3.- Utilización de ARQ ... 117
5.3.4.- Configuración del simulador ... 118
5.4.- Interfaces entre simulación a nivel de enlace y nivel de sistema ... 121
5.4.1.- Interfaz a nivel de bloque RLC ... 122
5.4.2.- Interfaz a nivel de ráfaga ... 123
5.5.- Sumario ... 125

CAPÍTULO 6 – Estudio y Diseño de Link Adaptation 127

6.1.- Introducción ... 129
6.2.- Algoritmos Link Adaptation ... 130
6.2.1.- Algoritmo basado en throughput ... 130
6.2.2.- Algoritmo basado en BLER ... 135
6.2.3.- Primer algoritmo propuesto ... 137
6.2.4.- Segundo algoritmo propuesto ... 149
6.3.- Sumario ... 164

CAPÍTULO 7 – Análisis de Resultados 167

7.1.- Introducción ... 169
7.2.- Algoritmos de referencia ... 169
7.2.1.- Causa de la ineficacia del algoritmo basado en BLER ... 179
7.3.- Primer algoritmo propuesto ... 183
7.4.- Segundo algoritmo propuesto ... 193
7.5.- Sumario ... 202

CAPÍTULO 8 – Conclusiones y Trabajo Futuro 205

ANEXO A – Codificación de Vídeo 213

A.1.- Objetivo del anexo 215
A.2.- Codificación de vídeo H.263 215

ANEXO B – Número Necesario de Canales por Usuario 221

B.1.- Objetivo del anexo 223
B.2.- Número de canales por usuario 223

ANEXO C – Demostración 227

C.1.- Objetivo del anexo 229
C.2.- Demostración 229

BIBLIOGRAFÍA 235

Lista de Acrónimos

– A –

AMR	Adaptive Multi Rate
APC	Adaptive Power Control
APN	Access Point Name
ARQ	Automatic Repeat reQuest
ATDMA	Advanced Time Division Multiple Access
ATM	Asynchronous Transfer Mode
AUC	AUthentication Center

– B –

BCS	Block Check Sequence
BEC	Backward Error Correction
BER	Bit Error Rate
BG	Border Gateway
BH	Block Header
BLER	BLock Error Rate
BSC	Base Station Controller
BSN	Block Sequence Number
BSS	Base Station Subsystem
BSSGP	Base Station Subsystem GPRS application Protocol
BTS	Base Transceiver Station

– C –

C/I	Carrier-to-Interference Ratio
CAC	Call Admission Control
CBR	Constant Bit-Rate
CC	Conmutación de Circuitos
CCCH	Common Control CHannel
CDF	Cumulative Distribution Function
CDMA	Code Division Multiple Access
CDR	Call Detail Record
CIF	Common Intermediate Format
CIR	Carrier-to-Interference Ratio
CNCL	Communications Networks Class Library
CP	Conmutación de Paquetes
CS	Coding Scheme
CS	Circuit Switching

– D –

DB	Data Bits
DCA	Dynamic Channel Allocation
DCT	Discrete Cosine Transform
DNS	Domain Name Server
DRA	Dynamic Resource Allocation
DRX	Discontinuous Reception
DTX	Discontinuous Transmission

– E –

EDGE	Enhanced Data-rates for Global Evolution
EGPRS	Enhanced General Packet Radio Service
EIR	Equipment Identity Register
ETS	Errors in Training Sequence
ETSI	European Telecommunications Standards Institute

– F –

FCA	Fixed Channel Allocation
FCFS	First-Come First-Served
FCS	Frame Check Sequence
FDD	Frecuency Division Duplex
FDMA	Frecuency Division Multiple Access
FEC	Forward Error Correction
FER	Frame Error Rate
FH	Frame Header
FR	Full-Rate
FTP	File Transfer Protocol

– G –

GGSN	Gateway GPRS Support Node
GMSC	Gateway Mobile Switching Center
GMSK	Gaussian Minimum Shift Keying
GP	Guard Period
GPRS	General Packet Radio Service
GSM	Global System for Mobile communications
GSN	GPRS Support Node
GTP	GPRS Tunnel Protocol

– H –

HCA	Hybrid Channel Allocation
HDLC	High level Data Link Control
HLR	Home Location Register
HR	Half-Rate
HSCSD	High Speed Circuit Switched Data
HTTP	Hyper Text Transfer Protocol

– I –

IMT	International Mobile Telecommunication
IP	Internet Protocol
IPv4	Internet Protocol version 4
IPv6	Internet Protocol version 6
IR	Incremental Redundancy
ISO	International Standardization Organization
ITU	International Telecommunication Union

– J –

– K –

KBPS	Kilo Bits Por Segundo

– L –

LA	Location Area
LA	Link Adaptation
LAPD	Link Access Protocol D
LLC	Logical Link Control
LUT	Look-Up Table

– M –

MAC	Medium Access Control
MCS	Modulation and Coding Scheme
ME	Mobile Equipment
MM	Mobility Management
MPDCH	Master Packet Data CHannel
MPEG	Motion Picture Experts Group
MS	Mobile Station
MSC	Mobile Switching Center
MT	Mobile Terminal

– N –

NS	Network Service
NSS	Network Switching Subsystem

– O –

OSI	Open System Interconnection

– P –

PACCH	Packet Associated Control CHannel
PAGCH	Packet Access Grant CHannel
PBCCH	Packet Broadcast Control CHannel
PBCH	Packet Broadcast CHannel
PCCCH	Packet Common Control CHannel
PCCH	Packet Control CHannel
PCU	Packet Control Unit
PDCCH	Packet Dedicated Control CHannel
PDCH	Packet Data CHannel
PDCH/F	Packet Data CHannel / Full-rate
PDCH/H	Packet Data CHannel / Half-rate
PDF	Probability Distribution Function
PDN	Packet Data Network
PDP	Packet Data Protocol
PDTCH	Packet Data Traffic CHannel
PDU	Protocol Data Unit
PER	Packet Error Rate
PH	Packet Header
PLL	Physical Link Layer
PLMN	Public Land Mobile Network
PLR	Packet Loss Ratio
PNCH	Packet Notification CHannel
PPCH	Packet Paging CHannel
PRACH	Packet Random Access Channel
PS	Packet Switching
PSK	Phase Shift Keying
PSNR	Peak Signal-to-Noise Ratio
PSTN	Public Switched Telephone Network
PTCCH	Packet Timing-advance Control CHannel
PTCCH/D	Packet Timing-advance Control Channel / Downlink
PTCCH/U	Packet Timing-advance Control Channel / Uplink
PTCH	Packet Traffic CHannel
PTM	Point-To-Multipoint
PTM-G	Point-To-Multipoint Group
PTM-M	Point-To-Multipoint Multicast
PTP	Point-To-Point
PTP-CLNS	Point-To-Point ConectionLess Network Service
PTP-CONS	Point-To-Point Conection Orientated Network Service

– Q –

QoS	Quality of Service

– R –

RA	Rural Area
RACE	Research in Advanced Communication in Europe
RCA	Random Channel Allocation
RDSI	Red Digital de Servicios Integrados
RFL	Radio Frecuency Layer
RGB	Red Green Blue
RLC	Radio Link Control
RRA	Random Resource Allocation
RRM	Radio Resource Management

– S –

SB	Stealing Bits
SDH	Synchronous Digital Hierarchy
segSNR	segmented Signal-to-Noise Ratio
SGSN	Serving GPRS Support Node
SIM	Subscriber Identity Module
SMS	Short Message Service
SNDCP	Sub Network Dependent Convergence Protocol
SNR	Signal-to-Noise Ratio
SOA	Soft Output Averages
SPDCH	Slave Packet Data CHannel

– T –

TB	Tail Bits
TBF	Temporary Block Flow
TCH	Traffic CHannel
TCP	Transmission Control Protocol
TDMA	Time Division Multiple Access
TE	Terminal Equipment
TETRA	TErrestrial Trunked RAdio
TFI	Temporary Flow Identity
TLLI	Temporary Logical Link Identifier
TS	Training Sequence
TU	Typical Urban

– U –

UDP	User Datagram Protocol
UMTS	Universal Mobile Telecommunications System
USF	Uplink State Flag

– V –

VBR	Variable Bit-Rate
VLR	Visitor Location Register

– W –

WAP	Wireless Application Protocol
WCDMA	Wideband Code Division Multiple Access
WLAN	Wireless Local Area Network
WWW	World Wide Web

– X –

– Y –

– Z –

Capítulo 1

Introducción

1.1.- INTRODUCCIÓN

Los sistemas de comunicaciones móviles existentes hoy en día son capaces de adaptarse a un amplio abanico de necesidades, permitiendo de forma extremadamente cómoda y sencilla la comunicación en cualquier lugar, momento y circunstancia, hecho que explica la rápida expansión que ha experimentado de forma reciente así como la importancia que posee para la sociedad actual, cada vez más compleja y dinámica, este tipo de sistemas.

La telefonía móvil, tal y como la conocemos hoy día, es el resultado de un largo proceso evolutivo en el que la adopción de la tecnología digital y el impresionante crecimiento experimentado en el número de abonados fueron, sin lugar a dudas, los dos factores más significativos. Este aumento del número de usuarios vino de la mano del sistema de segunda generación GSM (*Global System for Mobile communications*), con el que la telefonía móvil adquirió gran popularidad y aceptación. La implantación de este estándar, primero europeo y luego mundial, convirtió el mercado de la telefonía móvil en una economía de escala con millones de consumidores en todo el mundo. Si bien actualmente la velocidad de expansión ha decrecido, la demanda de servicios de voz hoy día todavía sigue incrementándose año tras año. Junto a este crecimiento, la tendencia actual del mercado apunta hacia la convergencia de Internet y las redes móviles inalámbricas en los futuros sistemas de tercera generación. La incorporación de estos servicios de datos (entre los que se hallan los atractivos servicios multimedia, caracterizados por requerir un significativo ancho de banda) y el continuo ascenso que se prevé en el número de usuarios crean nuevos desafíos para los operadores móviles, los cuales se ven obligados a implementar métodos que permitan un aprovechamiento, cada vez más flexible y eficiente, de los escasos recursos radioeléctricos. Éste es el objetivo que persiguen las denominadas técnicas de gestión de recursos radio.

Una de estas técnicas, y objeto del presente trabajo, recibe el nombre de *Link Adaptation*. La base de *Link Adaptation* consiste en evaluar, de forma periódica, la calidad del enlace radio y seleccionar aquel esquema de codificación de canal que, bajo dichas condiciones, resulta óptimo de acuerdo a un criterio predefinido. El criterio seguido para la elección del esquema de codificación de canal es un factor clave en el comportamiento de *Link Adaptation*. Uno de los criterios más ampliamente empleados hasta la fecha consiste en elegir el esquema de codificación que proporciona la mayor cantidad de bits correctamente transmitidos por unidad de tiempo, es decir, el mayor throughput. Si bien este criterio puede resultar satisfactorio para servicios con escasos requisitos en términos de calidad de servicio, su aplicación a otro tipo de servicios con importantes requisitos ha sido cuestionada.

Algunos servicios multimedia, como aquellos que involucran la transmisión de vídeo (caso aquí considerado) o audio, se caracterizan, en cuanto a calidad de servicio, por presentar unos requisitos bastante más exigentes que los habituales para otro tipo de servicios. En general, suelen ser más vulnerables ante los errores de transmisión y requieren un retardo de transmisión considerablemente menor, el cual llega a jugar un papel determinante en la calidad de servicio experimentada por el usuario. En este caso, el criterio que busca la maximización del throughput a largo plazo no garantiza la consecución de una calidad de servicio aceptable para este tipo de servicios. Por ello, la aplicación de nuevos criterios que tengan presente otros factores relevantes, en especial el retardo, es una opción a considerar.

El principal objetivo de este trabajo consiste, así pues, en el estudio y análisis de diversos criterios de selección de *Link Adaptation* considerando un tráfico de datos en tiempo real, sensible al retardo. Primero se analiza el criterio tradicional cuyo objetivo es la maximización del throughput a largo plazo; el rendimiento de este algoritmo, inicialmente propuesto para servicios en los que el retardo no era un factor clave, se evalúa aquí al ser aplicado a servicios sensibles al retardo. Se examina también un segundo algoritmo cuyo criterio de selección pretende preservar la tasa de errores por debajo de un determinado valor, el cual se considera como valor máximo tolerable para que la calidad experimentada por el usuario pueda ser considerada aceptable.

Los resultados ofrecidos por estos dos algoritmos (propuestos en la literatura por otros autores) se utilizan como referencia para la comparación de las prestaciones ofrecidas por dos nuevos algoritmos que se proponen y que constituyen la principal aportación del presente trabajo.

El primer algoritmo propuesto tiene, como principal objetivo, alcanzar un compromiso entre los dos algoritmos anteriores, intentando lograr unas prestaciones intermedias a las ofrecidas por ambos. Partiendo con este objetivo, el diseño finaliza obteniendo un algoritmo en el cual el grado de compromiso puede ser modificado en función de cuál de los dos parámetros, throughput o tasa de errores, se considere de mayor relevancia respecto al otro.

El segundo de los algoritmos que aporta este trabajo se diseña teniendo en cuenta, no sólo el throughput, tal y como ha venido siendo habitual, sino también el retardo, parámetro trascendental para el tipo de servicio considerado. Lo que se busca con este planteamiento es que la elección del esquema de codificación no se base únicamente en consideraciones de throughput sino que el retardo sea también un elemento influyente en la decisión final ya que, de hecho, también constituye un factor importante para el tipo de servicio considerado. El resultado final es un algoritmo en el cual el proceso de selección del esquema de codificación se adapta de forma dinámica a las características del flujo de datos transmitido en función de la cantidad de recursos asignados a cada usuario.

1.2.- ESTRUCTURA DEL TRABAJO

Para la evaluación del rendimiento de una técnica de gestión de recursos radio mediante simulación, como es el caso, es necesaria la implementación de la interfaz aire de un sistema de comunicaciones móviles concreto. El simulador empleado para la realización de este trabajo está basado en la interfaz aire del sistema GPRS (*General Packet Radio Service*). Por este motivo, el capítulo segundo de este trabajo ofrece una descripción global del sistema GPRS centrando la atención sobre aquellos aspectos que guardan una mayor relación con los fines aquí perseguidos. Inicialmente se presenta el sistema GPRS enmarcado dentro del proceso evolutivo hacia los sistemas de tercera generación. Tras tratar algunos aspectos básicos de GPRS (servicios, arquitectura y protocolos de transmisión de datos), el capítulo se centra sobre temas relativos a la interfaz aire, entre los que cabe resaltar los diversos esquemas de codificación de canal.

En el tercer capítulo se describe la parte fundamental de este trabajo, es decir, *Link Adaptation*. Puesto que es una técnica destinada a mejorar el aprovechamiento de los escasos recursos radioeléctricos, se presenta enmarcada en el contexto de la gestión de recursos radio. Sin embargo, el estudio de estas técnicas de gestión implica la utilización de algunos conceptos fundamentales relacionados con la estructura celular, utilizada hoy día por todos los sistemas de comunicaciones móviles. Así pues, se inicia el capítulo con un repaso general de la estructura celular para, a continuación, tratar la gestión de recursos radio. Se plantea el problema, se establece una clasificación de las diversas técnicas de gestión de recursos radio y se presenta, de forma breve, algunas de las más importantes. El resto del capítulo se dedica a tratar en profundidad la técnica *Link Adaptation*, entidad entorno a la que gira todo este trabajo.

El estudio de *Link Adaptation* se realiza considerando, como ya se ha dicho, la transmisión tráfico de datos en tiempo real. Se decidió emplear para dichos fines un modelo de vídeo H.263, razón por la cual el cuarto capítulo trata aquellos aspectos relativos a la codificación de vídeo H.263 que resultan de mayor relevancia para este trabajo. En la transmisión de este tipo de datos a través de una tecnología basada en conmutación de paquetes como es el caso de GPRS (y de los futuros sistemas de tercera generación), no se asegura la disponibilidad de los recursos en el momento en que se necesitan, pudiendo producirse retardos adicionales que deterioran la calidad de la comunicación. En este contexto cobra especial relevancia el concepto de calidad de servicio. Este tema también se trata en este capítulo, presentando varios parámetros de calidad asociados a la transmisión de vídeo H.263 a través de la interfaz aire.

El quinto capítulo está dedicado a la descripción de los modelos y herramientas de simulación empleados en la evaluación de los diversos algoritmos *Link Adaptation*. El estudio mediante simulación por computador suele realizarse en dos partes: a nivel de enlace y a nivel de sistema. La atención de los estudios realizados se centra sobre las simulaciones a nivel de sistema, razón por la cual en este capítulo se describen con más detalle los aspectos relativos a este nivel mientras que los relativos al nivel de enlace simplemente se esbozan. El capítulo concluye con la descripción de dos interfaces entre ambos niveles que suelen emplearse en las simulaciones. La segunda, más compleja y precisa, es la que se implementa en el simulador utilizado.

En el sexto capítulo se detalla la forma en que los diversos criterios estudiados se materializan en la definición de un algoritmo *Link Adaptation*. El procedimiento se ilustra empleando el primer tipo de interfaces descrito en el capítulo quinto (ya que así la descripción resulta más fácilmente comprensible) y, a continuación, empleando el segundo tipo de interfaces, que son las que se implementan en el simulador.

Los resultados ofrecidos por las simulaciones de los diferentes algoritmos *Link Adaptation* estudiados se muestran y analizan en el séptimo capítulo. Por último, el presente trabajo concluye con un octavo y último capítulo en el cual se señalan las conclusiones más importantes y se sugieren algunas futuras investigaciones que se pueden derivar del trabajo aquí realizado.

Capítulo 2

General Packet Radio Service

2.1.- INTRODUCCIÓN

En la última década la telefonía móvil experimentó un espectacular crecimiento de la mano de GSM (*Global System for Mobile communications*), sobrepasando las expectativas más optimistas. Este estándar, primero europeo y luego mundial, se basó de forma casi exclusiva en la provisión de servicios de voz y de sus servicios suplementarios y relegó los servicios de datos a un ámbito marginal, constituyendo el servicio de mensajes cortos o SMS la única excepción a esta regla [1, 2]. Este proceso es lo que se conoce como primera fase de GSM, cuyo objetivo principal fue conseguir una cobertura adecuada para el servicio de voz.

Una vez consolidado el uso de la telefonía móvil celular, y ante la incipiente importancia de Internet, los operadores apostaron por la introducción de servicios de datos, cada vez más avanzados, como elemento de distinción ante la competencia. Pero la red existente presentaba una serie de inconvenientes.

Se había dado prioridad a la voz en el proceso de estandarización y la red GSM se basaba en conmutación de circuitos. La asignación fija de recursos que esto supone era inadecuada para comunicaciones de datos, caracterizadas por la transmisión intermitente de ráfagas de información[1]. Además, el estándar contemplaba una tasa binaria máxima de 9.6 kbps[2], claramente insuficiente para la mayor parte de las aplicaciones de datos [1, 2].

Ante esta situación existían dos alternativas: esperar al desarrollo de los sistemas de tercera generación sin realizar nuevas inversiones o evolucionar GSM de forma progresiva. Finalmente se optó por esta última, elección justificada por motivos como la amortización de las inversiones en la red GSM, la posibilidad de crecimiento a corto plazo, la oferta de nuevas prestaciones y funcionalidades de forma paralela a la demanda, la introducción progresiva de las características de los sistemas de tercera generación que facilitaría la adaptación del usuario y la viabilidad técnica [3].

Se sentaron como bases para este proceso las siguientes [3]:

- Necesidad de **generar nuevos servicios portadores de datos** más eficientes, sin perder de vista la escasez de los recursos radioeléctricos y el hecho de que la incorporación de estos servicios requiere una mayor cantidad de recursos por usuario.
- **Mejora de las prestaciones**, en términos de régimen binario, ofrecidas por GSM.
- Estudio de **nuevas técnicas de diseño de red** que mejoren su capacidad y flexibilidad.

[1] Durante la mayor parte del tiempo no se transmite información alguna, pero existen intervalos durante los cuales se transmite gran cantidad de información en poco tiempo. La asignación fija de un recurso implica su ociosidad durante el tiempo en que no se transmite, lo cual supone un desperdicio.

[2] En GSM se dispone de 9.6 kbps/canal o *time-slot* en la interfaz aire. Además, sólo se permite la asignación de una única ranura temporal o *time-slot* por usuario en cada comunicación, por lo que la tasa máxima es de 9.6 kbps.

Bajo estos principios se inició el proceso de incorporación de la infraestructura necesaria para el soporte de servicios de datos con vistas a los sistemas de tercera generación (GSM Fase 2+). El proceso evolutivo se muestra en la figura 2.1.

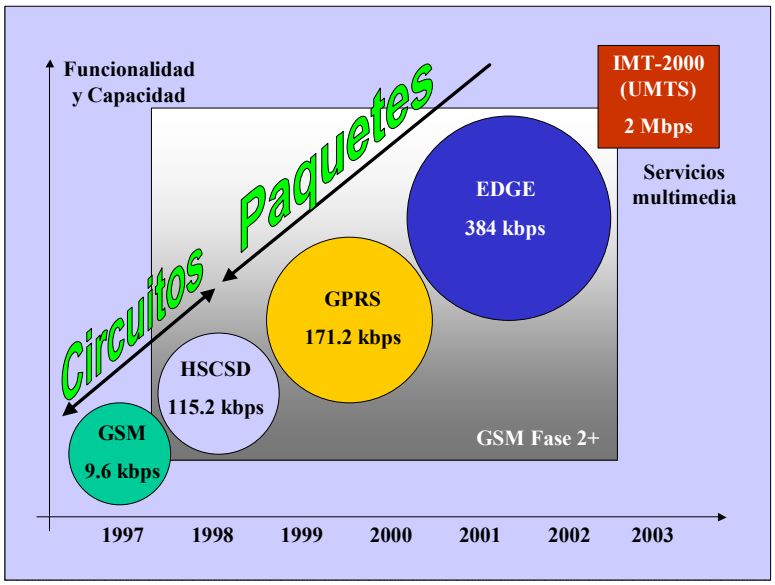


Figura 2.1.- Evolución hacia los sistemas de tercera generación [3]

El primer paso hacia la provisión de servicios de datos lo constituyó el sistema *High Speed Circuit Switched Data* (HSCSD). En HSCSD [3] se modifica la codificación de canal, pasando de los 9.6 a los 14.4 kbps/time-slot. Además, se ofrece la posibilidad de agrupar varias ranuras temporales o time-slots en una misma comunicación, dando lugar a un régimen binario máximo teórico de 14.4 × 8 = 115.2 kbps. El número de canales asignados en *uplink* puede ser diferente del asignado en *downlink*[3] (asignación asimétrica), mejorando la eficiencia espectral. Esta tecnología es idónea para aplicaciones que requieren un régimen binario constante (desafortunadamente, no es el caso para la mayoría de las aplicaciones de datos) y un retardo de transmisión mínimo. La tecnología HSCSD no tuvo éxito debido a lo siguiente [3]:

- Está **basada en conmutación de circuitos** que, como ya se ha dicho, es inadecuada para comunicaciones de datos (infrautilización de recursos).
- **Dificulta la operación de *handover***, ya que al agrupar *k* time-slots en una misma comunicación, se requieren *k* time-slots libres en la celda de destino, lo cual no siempre puede ser posible. Ante esta situación se degrada la calidad del servicio experimentada por el usuario.
- Lo anterior supone la **necesidad de implementar complejos mecanismos adicionales** que eviten la fragmentación en la asignación de recursos radio y que controlen la admisión de llamadas (CAC, apartado 3.4.3.1) en aquellos casos de *handover* con recursos insuficientes.

[3] Se denomina *uplink* o enlace ascendente al enlace en la dirección estación móvil a estación base y *downlink* o enlace descendente al enlace en la dirección contraria, es decir, en dirección estación base a estación móvil.

Gran cantidad de fabricantes ignoró HSCSD y apostó directamente por la implantación de GPRS (*General Packet Radio Service*). GPRS [2, 3] supone una ruptura respecto a las técnicas anteriores de transferencia de datos en redes de telefonía móvil en la medida en que introduce, de forma novedosa, la conmutación de paquetes para implementar servicios portadores para aplicaciones de datos. Este rasgo hace a GPRS idóneo para dichas aplicaciones, ya que los recursos se asignan sólo cuando se necesita transmitir una *ráfaga* de información y se liberan cuando termina la transferencia. De esta forma los recursos se explotan bajo demanda, pudiendo multiplexar varios usuarios sobre un mismo canal físico y mejorando así el aprovechamiento de los recursos radio. La asignación asimétrica introducida con HSCSD sigue siendo posible.

En cuanto a régimen binario, desde el punto de vista teórico se pueden alcanzar velocidades máximas de 171.2 kbps, empleando 8 time-slots (que pueden ser asociados igual que en HSCSD) y una codificación CS-4 (aplicable en condiciones ideales) que proporciona 21.4 kbps/time-slot. En el ámbito práctico esta velocidad es mucho menor, ya que la mayoría de terminales del mercado manejan hasta un máximo de 4 time-slots [2] llegando a los 85.6 kbps. Además, un operador sólo puede permitirse un número pequeño de time-slots dedicados a una misma comunicación, siendo habituales velocidades de pico de 48 kbps [3]. No obstante, GPRS supone una mejora palpable en el aprovechamiento de los recursos radio y un incremento aceptable en la tasa binaria.

Este incremento en la tasa binaria no es suficiente de cara a la tercera generación, y es aquí donde interviene EDGE (*Enhanced Data-rates for Global Evolution*). EDGE [3] tiene como objetivo incrementar los regímenes binarios en la interfaz aire mediante el uso de modulaciones con una mayor eficiencia espectral; además de la modulación GMSK (*Gaussian Minimum Shift Keying*) de GSM, ahora también se emplea una modulación multinivel: 8-PSK (*8-Phase Shift Keying*). El resultado final es un régimen de 48 kbps/time-slot y una tasa máxima teórica de $48 \times 8 = 384$ kbps, aunque pagando el precio de una menor robustez frente a ruido e interferencia cocanal propia de una modulación multinivel y la necesidad de amplificadores altamente lineales dada la mayor susceptibilidad de este tipo de modulación a las no-linealidades [3].

Así pues, mientras HSCSD y GPRS siguen definiendo el tipo de servicio portador de datos (de circuitos el primero y de paquetes el segundo), EDGE se limita a mejorar las velocidades de transmisión en el ámbito radioeléctrico. Esta tecnología es totalmente compatible con GSM, permite reutilizar la infraestructura existente (lo cual no implica la exención de inversiones en hardware, además de nuevo software) y se caracteriza por un tiempo de implementación necesario muy reducido. El escalón final de este proceso será el estándar europeo UMTS (*Universal Mobile Telecommunications System*), perteneciente a la familia mundial de estándares de 3ª generación IMT-2000, para el que se prevén tasas de 144 kbps (vehículos), de 384 kbps (peatones) en entornos exteriores (*outdoor*) y de 2 Mbps teóricos en entornos interiores (*indoor*) [3] que soporten todo tipo de complejos servicios de datos sobre la red de telefonía móvil.

El objetivo de esta introducción ha sido ubicar a GPRS dentro del proceso evolutivo de GSM hacia la tercera generación, dar a entender los motivos que lo justifican y ofrecer una visión global de la problemática que entraña. El resto de este capítulo estará dedicado a proporcionar una descripción de GPRS, incidiendo en aquellos aspectos que resultan de especial interés para los objetivos de este trabajo.

2.2.- ¿QUÉ ES GPRS?

Las siglas GPRS (*General Packet Radio Service*) responden a una extensión de la red GSM tradicional definida por el ETSI (*European Telecommunications Standardization Institute*) en las especificaciones GSM Fase 2+. Es una red con tecnología basada en conmutación de paquetes superpuesta a la red GSM existente (basada en conmutación de circuitos). Su objetivo es proporcionar un nuevo servicio portador inalámbrico que sirva de soporte para la transmisión eficaz de datos. Con el sistema GPRS un usuario de telefonía móvil puede acceder a redes públicas y privadas de datos basadas en protocolos estándar como IPv4, IPv6 o X.25. Los servicios contemplados en las especificaciones se subdividen en dos grupos:

- **Punto a punto (*Point-To-Point*, PTP):** Permite la transmisión de mensajes individuales entre un único usuario de origen y un único usuario de destino. A su vez, se subdivide en dos grupos:

a) ***Point-To-Point ConnectionLess Network Service*, PTP-CLNS*:*** Permite el envío de paquetes sencillos. Los paquetes se envían en modo datagrama, sin ningún tipo de relación entre ellos y, por tanto, sin negociar ningún tipo de conexión de forma previa a la transferencia de información. Pensado para soportar IP.

b) ***Point-To-Point Connection Orientated Network Service*, PTP-CONS:** Permite el intercambio de múltiples paquetes entre usuario y destino. Los paquetes se transfieren en modo circuito virtual, previo establecimiento de la conexión lógica. Pensado para soportar X.25.

- **Punto a multipunto (*Point-To-Multipoint*, PTM):** Permite la transmisión de mensajes entre un usuario del servicio y múltiples receptores. Se subdivide en:

a) ***Point-To-Multipoint Multicast*, PTM-M:** El sentido de los mensajes es unidireccional y va destinado a una determinada área geográfica.

b) ***Point-To-Multipoint Grupo de llamadas*, PTM-G:** El sentido de los mensajes puede ser unidireccional, bidireccional o multidireccional y va destinado a un grupo concreto de usuarios dentro de un área geográfica concreta.

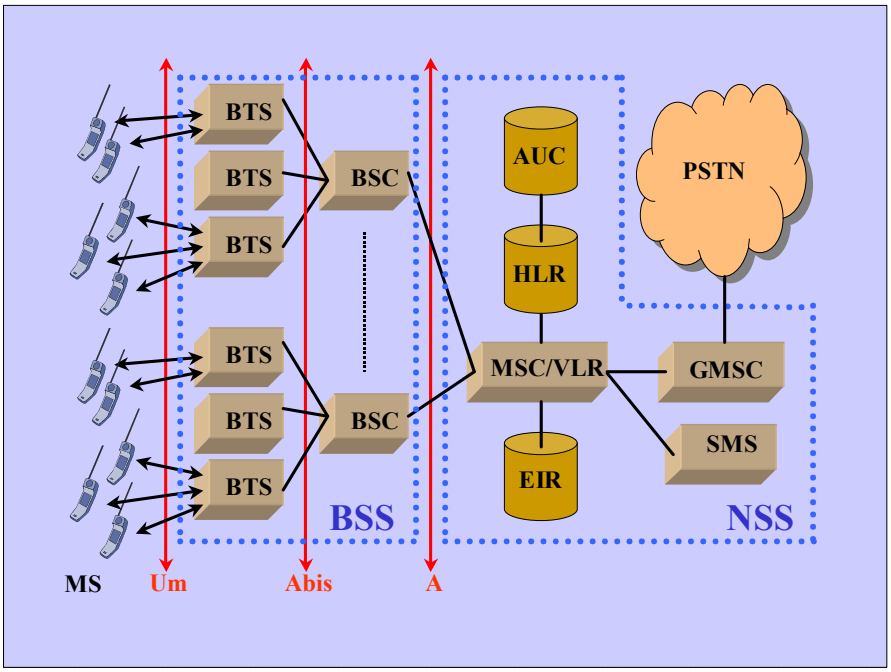
A grandes rasgos [2, 3], la introducción de GPRS permite tasas máximas teóricas de transmisión de hasta 171.2 kbps, la posibilidad de compartir time-slots entre diferentes usuarios conforme a la demanda de recursos así como la asignación de recursos *uplink/downlink* de forma asimétrica, mejorando la eficiencia espectral. Al estar basado en conmutación de paquetes, un aumento de usuarios supone una posible menor calidad del servicio experimentada por cada uno de ellos, frente al bloqueo (incapacidad de aceptar más usuarios por encima de un determinado número) propio de la conmutación de circuitos. Además, la transferencia de pequeñas cantidades de información es más rápida que en conmutación de circuitos, donde en estos casos el establecimiento del camino físico puede necesitar más tiempo y recursos que la propia transferencia de información. Tal y como se muestra en el siguiente apartado, todas estas ventajas se logran con algunas modificaciones sobre la red GSM existente.

2.3.- ARQUITECTURA DE RED

GPRS es básicamente una funcionalidad añadida a la red GSM mediante la introducción de unos pocos elementos en lo que a hardware respecta y la actualización del software en algunos de los nodos GSM ya existentes. Por ello, la comprensión del funcionamiento de la arquitectura GPRS requiere una descripción de la red GSM que, de forma general, se proporciona a continuación.

2.3.1.- ARQUITECTURA DE LA RED GSM

La figura 2.2 muestra la estructura genérica de una red GSM [4 – 6].



AUC: AUthentication Center
BTS: Base Transceiver Subsystem
HLR: Home Location Register
NSS: Network Switching Subsystem
BSC: Base Station Controller
EIR: Equipment Identity Register
MS: Mobile Station
VLR: Visitor Location Register
BSS: Base Station Subsystem
GMSC: Gateway MSC
MSC: Mobile Switching Center

Figura 2.2.- Arquitectura de la red GSM

La estructura se divide en varios subsistemas. Los más importantes son el subsistema de conmutación de red o NSS (*Network Switching Subsystem*) y el subsistema de estación base o BSS (*Base Station Subsystem*), interconectados mediante la interfaz A. Un usuario o subscriptor GSM necesita una estación móvil o MS (*Mobile Station*) para comunicarse con la red a través de la interfaz Um. Dentro del BSS se define la interfaz Abis que, a diferencia de las interfaces A y Um, no está estandarizada en GSM.

El terminal móvil de usuario o **MS** está compuesto por el teléfono propiamente dicho (*Mobile Equipment* o ME) y por una tarjeta inteligente denominada SIM (*Subscriber Identity Module*). La tarjeta SIM contiene toda la información referente al subscriptor de cara a la identificación ante la red y constituye un elemento exclusivo del usuario. El ME es un elemento hardware que se encarga, básicamente, de las funciones de muestreo, cuantificación, codificación, modulación y transmisión de la señal de voz así como del proceso inverso en recepción. El ME, a diferencia de la SIM, no es un elemento exclusivo del usuario en el sentido en que el usuario puede cambiar de ME manteniendo la misma SIM y conservar, de esta forma, su identidad. La MS permite al usuario comunicarse, a través de la interfaz aire (Um), con el BSS.

El subsistema **BSS** desempeña todas las labores relacionadas con la interfaz de radiofrecuencia y está constituido por los nodos *Base Transceiver Station* (BTS) y *Base Station Controller* (BSC).

La **BTS** maneja la interfaz de radiofrecuencia (Um) con la MS. Se encarga principalmente de establecer la conexión radio con el usuario, de las etapas de potencia, modulación y demodulación así como de realizar medidas del nivel de potencia de transmisión de la MS. Incorpora el sistema radiante de la red GSM. La zona geográfica dentro de la cual una BTS es capaz de comunicarse con una MS constituye una celda o célula. Una o más BTSs pueden estar controladas por una única BSC.

La **BSC** provee las funciones lógicas de control de la interfaz radio y los enlaces físicos entre BTS y MSC. Se encarga, entre otras labores, de la asignación y liberación de canales radio, de controlar la potencia de transmisión de la MS así como de decidir y coordinar los procesos de *handover*. Igualmente, se ocupa de transmitir el tráfico de voz hacia la central de conmutación, incorporada dentro del NSS.

El subsistema **NSS** es responsable del control y encaminamiento de llamadas y del control del servicio dentro de la red GSM, además de las funciones de gestión de la movilidad del usuario. Incluye los nodos *Mobile Switching Center* (MSC), *Home Location Register* (HLR), *Visitor Location Register* (VLR), *AUthentication Center* (AUC), *Equipment Identity Register* (EIR) y *Gateway Mobile Switching Center* (GMSC).

La **MSC** se encarga de las funciones de conmutación telefónica propiamente dichas, coordinando el establecimiento de llamadas entre los usuarios de la red GSM así como el de las llamadas establecidas con otras redes externas de telefonía, ya sean fijas o de otros operadores móviles. Una única MSC puede controlar el tráfico asociado a una o más BSCs.

El **HLR** es una base de datos que almacena y administra los datos permanentes relacionados con todos los usuarios de la red GSM de un operador, tales como servicios contratados, perfiles o estado de actividad. El **VLR** es también una base de datos pero que almacena y administra los datos temporales relacionados sólo con los usuarios que, en un momento dado, se encuentran bajo el área de servicio de una MSC. El VLR es utilizado por la MSC para obtener información acerca de los usuarios a los que da servicio en un momento dado. Cuando un usuario cambia de MSC, el VLR de la nueva MSC solicita información acerca del subscriptor al HLR, eliminándose dicha

información en el VLR de la MSC antigua. Así pues, existe un VLR por cada MSC mientras que el HLR es único para toda la red de un operador GSM, si bien se encuentra físicamente distribuido por razones lógicas de seguridad ante fallos del sistema. Es frecuente encontrar el VLR como elemento integrado dentro de la propia MSC, en cuyo caso se utiliza la nomenclatura MSC/VLR.

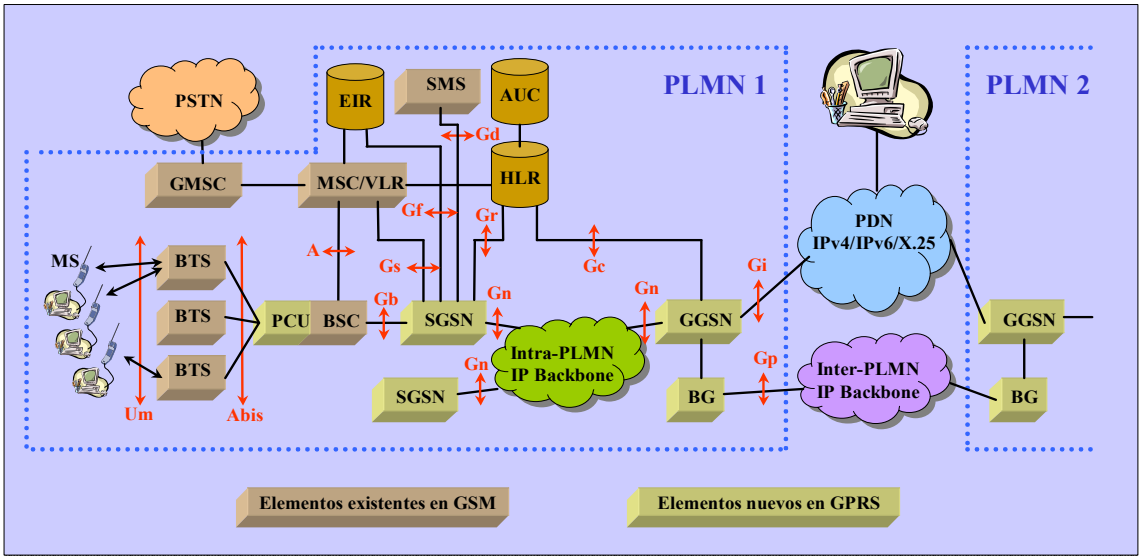
El **AUC** es una base de datos conectada al HLR cuya función principal es proporcionarle parámetros de autentificación y claves de cifrado. Los parámetros de autentificación se utilizan para limitar el acceso a la red a clientes GSM y proteger a los operadores contra el fraude. Las claves de cifrado son necesarias para la encriptación de la información transmitida en la interfaz aire (susceptible de ser interceptada por personal no autorizado) asegurando así la privacidad y seguridad de la comunicación.

El **EIR** es otra base de datos que contiene identidades de los terminales móviles de usuario (ME). Está formado por una lista blanca (de terminales registrados y en correcto funcionamiento), una lista gris (de terminales defectuosos o no homologados, en observación) y una lista negra (de terminales cuyo robo ha sido denunciado). Este nodo posibilita el bloqueo de llamadas desde terminales robados o no autorizados.

Por último, el **GMSC** es una pasarela o *gateway* que actúa como dispositivo de interconexión entre la red GSM y otras redes tales como la RPTC (*Red Pública Telefónica Conmutada*) o la PLMN (*Public Land Mobile Network*) de otro operador de telefonía móvil.

2.3.2.- ARQUITECTURA DE LA RED GPRS

La figura 2.3 muestra la estructura genérica de una red GPRS [4 – 6].



SGSN: Serving GPRS Support Node **GGSN**: Gateway GPRS Support Node
PCU: Packet Control Unit **PLMN:** Public Land Mobile Network **BG:** Border Gateway **PDN:** Packet Data Network

Figura 2.3.- Arquitectura de la red GPRS

Examinando la figura 2.3 es posible apreciar que la estructura de la red GPRS [4 – 6] se obtiene mediante la adición de unos pocos nodos a la red GSM y su interconexión con los elementos correspondientes. Estos nodos son *Packet Control Unit* (PCU), *Border Gateway* (BG), *Serving GPRS Support Node* (SGSN) y *Gateway GPRS Support Node* (GGSN), siendo estos dos últimos los más importantes. El usuario del servicio GPRS necesitará un nuevo terminal, específico para GPRS.

En GPRS, el concepto de estación móvil o **MS** sufre una modificación, distinguiéndose ahora tres conceptos. El equipo terminal de datos o **TE** (*Terminal Equipment*) es un dispositivo (como un portátil o un PDA por ejemplo) que origina los paquetes de datos transmitidos hacia la red y hacia el que van dirigidos los paquetes recibidos de la red. El terminal móvil o **MT** (*Mobile Terminal*) es el dispositivo que comunica al TE con la red, haciendo las veces de un *radio-módem* o módem inalámbrico. La unión del TE y el MT constituye la MS o estación móvil de usuario. Generalmente el TE y el MT están integrados en un único dispositivo, aunque pueden existir de forma aislada. La nueva MS GPRS está preparada para soportar la transmisión en modo multislot de forma asimétrica[4].

En el BSS, la modificación más relevante es la introducción de la unidad de control de paquetes o **PCU**. Esta tarjeta se incorpora en cada BSC con la finalidad de discriminar el tráfico de voz del de datos. El primero se encaminará hacia la MSC mientras que el segundo irá destinado a los nuevos nodos GPRS. La PCU es la responsable de asignar los canales radio dedicados a la transmisión de datos en modo paquete así como de soportar las funciones de las capas RLC/MAC, que se tratarán con detalle en el apartado 2.5.2.

Los elementos más importantes de la red GPRS son los *GPRS Support Nodes* o GSNs. El **SGSN** es el responsable de gestionar el encaminamiento de los paquetes, tanto entrantes como salientes, relacionados con las MSs que se encuentran bajo su área de servicio. Otras funciones son el soporte de las sesiones, la gestión de la movilidad, la autentificación de usuarios, la facturación y el control del enlace lógico con la MS.

El **GGSN** actúa, fundamentalmente, como una interfaz hacia una red de conmutación de paquetes basada en un protocolo estándar como IPv4, IPv6 o X.25. Traduce el paquete recibido del SGSN al formato de la red externa (IP o X.25) y viceversa. Desde el punto de vista de la red de paquetes, el GGSN se comporta como un router. Otras funciones que desempeña son la gestión de las sesiones, la seguridad y protección frente a la red externa, la facturación y el encaminamiento de un paquete hacia el SGSN correcto mediante la técnica de *tunneling*[5].

[4] Recordar que GPRS, a diferencia de GSM, soporta la utilización de más de un time-slot para una misma comunicación. Los terminales de usuario pueden, así pues, clasificarse en función del número de time-slots que son capaces de manejar de forma simultánea. El número de time-slots agrupados en *uplink* y *downlink* puede ser diferente (asignación asimétrica) para mejorar el aprovechamiento de los recursos radio. Estas diferentes combinaciones determinan la clase de multislot de la MS GPRS.

[5] Básicamente consiste en encapsular las PDU del nivel de red (IP o X.25) en el GSN de origen y desencapsularlas en el GSN de destino. La transferencia dentro del backbone se realiza mediante el protocolo IP (el paquete IP o X.25 de la red de datos se encapsula en un paquete IP interno al backbone). De esta forma se logra establecer una conexión lógica individual o túnel SGSN-GGSN especificando la dirección IP del destino dentro del backbone.

Los nodos GSN de la red GPRS están intercomunicados mediante una columna vertebral o **backbone**. Cuando interconecta nodos GSN pertenecientes a la misma PLMN se denomina backbone intra-PLMN, siendo una red IP privada del operador GPRS. Cuando interconecta nodos GSN pertenecientes a diferentes PLMNs recibe el nombre de backbone inter-PLMN, siendo necesario en este caso un acuerdo previo entre operadores y el uso de un puente o **BG** (*Border Gateway*) que acometa las funciones de seguridad, protegiendo los backbones intra-PLMN privados contra ataques o usuarios no autorizados. La única restricción que se impone a este backbone es el uso del protocolo IP a nivel de red, quedando a elección del operador los protocolos de capas inferiores: ATM, Frame Relay, Ethernet, SDH, MPLS, etc.

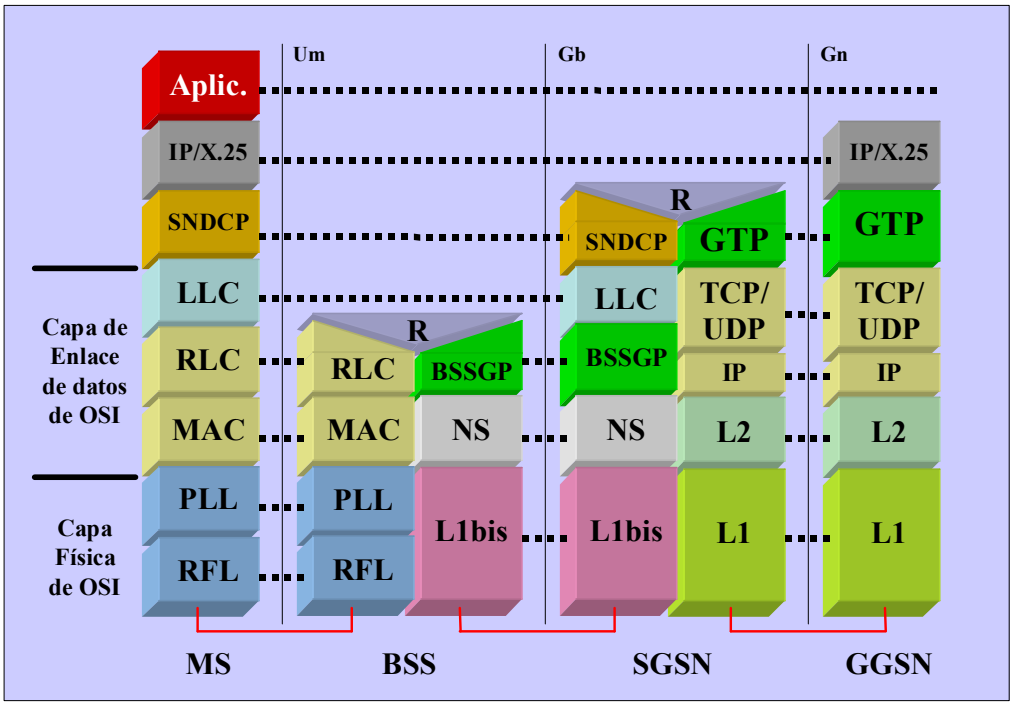
Es importante destacar que los elementos introducidos en la red GSM tienen como objetivo añadir la funcionalidad de la conmutación de paquetes a la red móvil, funcionalidad que antes no poseía. Los servicios basados en conmutación de paquetes se utilizan para soportar aplicaciones de datos, mientras que las llamadas de voz se siguen cursando a través de los elementos tradicionales de GSM, basados en conmutación de circuitos. Así pues, los terminales GSM podrán seguir utilizando el servicio de telefonía sobre la red GSM y tendrán la impresión de estar utilizando la red GSM convencional. Los terminales GPRS cursarán las llamadas de voz igual que los terminales GSM pero además tendrán la posibilidad de acceder a determinados servicios de datos utilizando los nodos introducidos por GPRS.

2.4.- ARQUITECTURA DE PROTOCOLOS

La arquitectura de protocolos de GPRS se basa en el difundido modelo por capas. Las especificaciones cubren hasta la capa de red según el modelo de referencia OSI (*Open System Interconnection*) de la ISO (*International Standardization Organization*). La definición de las capas superiores queda fuera de los objetivos de GPRS, considerándose capas de aplicación y pudiéndose elegir arbitrariamente cualquier protocolo estándar para su implementación.

Las especificaciones GPRS distinguen dos planos de protocolos: el plano de señalización y el plano de transmisión. El plano de señalización está conformado por el conjunto de protocolos que controlan y sustentan toda la señalización asociada a la transmisión de información de usuario (algunas funciones relevantes son la reserva de recursos, el encaminamiento o el control de la conexión). El plano de transmisión cubre el conjunto de protocolos para la transmisión de la información de usuario y procedimientos de control asociados como control de flujo, corrección y detección de errores. Así pues, dado que el objetivo de este trabajo se centra en la gestión de los recursos radio con la finalidad de mejorar las condiciones de transmisión de datos en la interfaz aire, se estudiará el plano de transmisión, más específicamente las capas RLC, MAC, PLL y RFL, careciendo de interés para dicho fin el plano de señalización.

La figura 2.4 muestra la arquitectura de protocolos GPRS en el plano de transmisión [5 – 8].



Aplic.: Capas de aplicación **IP:** Internet Protocol **R:** Relay (Conversión de protocolos)
GTP: GPRS Tunneling Protocol **TCP:** Transmission Control Protocol **UDP:** User Datagram Protocol
L2: Protocolo de capa 2 (no especificado) **L1:** Protocolo de capa 1 (no especificado) **L1bis:** Protocolo de capa 1 (no especificado)
LLC: Logical Link Control **RLC:** Radio Link Control **MAC:** Medium Access Control
PLL: Physical Link Layer **RFL:** Radio Frecuency Layer
BSSGP: Base Station Subsystem GPRS Protocol **NS:** Network Service (Frame Relay)
SNDCP: Sub Network Dependent Convergence Protocol

Figura 2.4.- Arquitectura de protocolos en el plano de transmisión

No hay que olvidar que GPRS es un servicio portador (por ello se define en las capas más bajas) y, como tal, está pensado para transportar la información de usuario de capas superiores independientemente de su contenido o aplicación. El conjunto de aplicaciones de usuario que hacen uso de los servicios de GPRS se representa mediante la capa superior (**aplicación**), que engloba el resto de capas que quedan por encima de la capa de red (IP/X.25).

La información proveniente de las capas de aplicación se empaqueta mediante el protocolo utilizado en la red de datos externa (**IP, X.25** u otro, según corresponda), gracias al cual el paquete de datos es encaminado a través de la red de datos hasta llegar al GGSN correspondiente. Allí es necesario determinar el nodo SGSN que en ese momento está dando servicio a la MS para la cual va dirigida el paquete. Una vez determinado el destino, el GGSN utiliza el protocolo **GTP** (*GPRS Tunneling Protocol*) para encaminar las PDU a través del túnel establecido dentro del backbone hasta llegar al SGSN correcto. Este protocolo se encarga de establecer, utilizar, gestionar y liberar los túneles. Además de un campo donde se especifica el tipo de PDU y parámetros relativos a la calidad del servicio, la cabecera GTP contiene un identificador de punto terminal del túnel para paquetes punto a punto o un identificador de grupo para paquetes punto a multipunto.

Por debajo de GTP se emplea el protocolo **TCP** (*Transmisión Control Protocol*) o el protocolo **UDP** (*User Datagram Protocol*). Generalmente, para el acceso a redes basadas en X.25 (que utilizan frecuentemente transmisión mediante circuito virtual por razones históricas) se emplea el protocolo TCP, orientado a la conexión, mientras que el protocolo UDP (no orientado a la conexión) es empleado para el acceso a redes basadas en la transmisión mediante datagramas, como es el caso de redes IP. Sea cual fuere la elección, se aplica **IPv4** como protocolo de capa de red dentro del backbone, por debajo de TCP/UDP. Los protocolos de la capa de enlace de datos y de la capa física según el modelo de referencia OSI, señalados en la figura 2.4 como **L2** y **L1** respectivamente, quedan a libre elección del operador, pudiéndose aplicar Ethernet, RDSI, ATM, SDH, etc. La elección de IP sobre ATM es una de las alternativas más consideradas de cara a la reutilización del backbone en el proceso de evolución hacia la tercera generación.

La información transportada entre GGSN y SGSN mediante el protocolo GTP es manipulada por un *relay*[6] para proporcionar datos inteligibles por el protocolo **SNDCP** (*Sub Network Dependent Convergence Protocol*), el cual lleva a cabo una nueva encapsulación de dicha información a fin de soportar su transferencia entre SGSN y MS. Este protocolo provee funciones de multiplexación de varios mensajes de capa 3 sobre una única conexión lógica virtual de la capa LLC subyacente. Además, aporta funcionalidades de cifrado, segmentación y compresión-descompresión tanto de datos de usuario como de información de cabecera redundante.

El protocolo **LLC** (*Logical Link Control*), como su propio nombre indica, habilita los mecanismos necesarios para proporcionar una conexión lógica fiable entre el SGSN y la MS capaz de transportar los datos de la capa superior. Está basado en el protocolo LAPD, perteneciente a la familia de protocolos orientados a bit HDLC. Realiza un control de flujo y corrección de errores mediante la técnica de ARQ (*Automatic Repeat reQuest*) y soporta tramas de longitud variable así como la transmisión de datos en modo *acknowledged* (con confirmación) y *unacknowledged* (sin confirmación). La confidencialidad de los datos está garantizada por las funciones de cifrado que se implementan.

El protocolo **BSSGP** (*Base Station Subsystem GPRS Protocol*) deriva del protocolo BSSMAP *(Base Station Subsystem Mobile Application Part*) ya existente en GSM y su función es la de transportar entre BSS y SGSN información bien de encaminamiento o bien relacionada con la calidad del servicio. El protocolo inmediatamente inferior, **NS** (*Network Service*), transporta las PDU generadas por la capa BSSGP y resuelve los casos de congestión, realizando así un control del tráfico. Suele emplearse una técnica *Frame Relay* modificada en la capa NS con la finalidad de minimizar los retardos mientras que la capa física (**L1bis**) se implementa mediante enlaces E1 de la jerarquía digital plesiócrona PDH (*Plesiochronous Digital Hierarchy*).

[6] El *relay* lleva a cabo una función de conversión entre diferentes protocolos de una misma entidad y realiza una labor de almacenamiento guardando las PDUs. El tamaño y tiempo de almacenamiento de cada PDU viene determinado por la calidad del servicio (*QoS*) asociada.

Es conveniente recordar en este punto que la idea básica de un servicio de conmutación de paquetes de datos es lograr el compartimiento de un mismo recurso por varios usuarios (a fin de mejorar la eficiencia espectral), que en la interfaz aire se traduce en la multiplexación de varias MS sobre un mismo canal físico. Dicha función es llevada a cabo por la capa de enlace de datos de la interfaz Um, la cual proporciona servicios con y sin confirmación para la transmisión de información sobre la capa física y define los procedimientos para llevar a cabo dicha multiplexación. Esta capa se divide a su vez en dos subcapas: RLC (*Radio Link Control*) y MAC (*Medium Access Control*).

La capa **RLC** se encarga de la segmentación y reensamblado de las tramas LLC en bloques RLC (*RLC blocks*) y proporciona servicios para garantizar una conexión lógica segura entre MS y BSS así como los procedimientos para la corrección de errores hacia atrás o BEC (*Backward Error Correction*) consistente en la retransmisión selectiva de bloques de información erróneos mediante ARQ.

La capa **MAC** emplea el protocolo ALOHA ranurado y básicamente se encarga de arbitrar el acceso de las diferentes MSs (*uplink*) o de la red (*downlink*) al medio físico radioeléctrico en busca de un aprovechamiento máximo de los escasos recursos radio. Emplea algoritmos para la resolución de colisiones y de la congestión, para la multiplexación de usuarios y para la gestión de la prioridad en función de la calidad de servicio negociada entre la red y cada usuario.

Para finalizar, la capa física [10, 11] de la interfaz Um está divida en las subcapas PLL (*Physical Link Layer*) y RFL (*Radio Frecuency Layer*).

El protocolo de la capa **PLL** ofrece servicios para la transmisión de información sobre el canal físico entre MS y red. Estas funciones incluyen la codificación de canal (apartado 2.5.3.1.2.1), la detección de errores de transmisión del medio físico, la corrección de errores hacia delante o FEC (*Forward Error Correction*), el entrelazado (apartado 2.5.3.1.2.2) de un bloque sobre cuatro ráfagas (*bursts*) consecutivas en tramas TDMA (según se define en [13]) y procedimientos para la detección de la congestión de los enlaces físicos.

A la capa **RFL** compete [8], según las especificaciones de la serie GSM 05, las frecuencias portadoras y la estructura del canal radio de GSM (05.02), la modulación, demodulación y tasas de datos de los canales GSM (05.04) y los requisitos técnicos del transmisor y receptor (05.05).

2.5.- INTERFAZ AIRE

Este trabajo se centra en el estudio de la interfaz aire [10] de forma general y de las capas RLC/MAC y física de forma específica. La interfaz aire (o Um según las especificaciones) es la interfaz entre la estación móvil de usuario (MS) y la red. Aunque en dicha interfaz aparecen más capas (ver figura 2.4), la comunicación entre la MS y la red se establece por medio de las capas RLC, MAC, PLL y RFL, siendo éstas las que ocupan nuestra atención. Antes de entrar en detalle en estas capas, se tratarán aspectos básicos relacionados con la interfaz aire, algunos de los cuales conciernen tanto a la capa RLC/MAC como a la capa física.

2.5.1.- ASPECTOS BÁSICOS

2.5.1.1.- CANALES

2.5.1.1.1.- CANALES FÍSICOS (PDCH)

Se entiende por canal físico [10] de la interfaz radioeléctrica la unidad mínima en que se divide el total de recursos radioeléctricos disponibles. En GPRS dicha división se realiza de forma bidimensional, tanto en frecuencia como en tiempo (ver apartado 2.5.1.2.3). Así pues, un canal físico está limitado dentro de un margen de frecuencias (dimensión espectral) y dentro de un intervalo de tiempo (dimensión temporal). Además, el concepto de canal físico se define independientemente del tipo de información que transporta (a diferencia de los canales lógicos).

Un canal físico dedicado al intercambio de paquetes de datos recibe la denominación de PDCH (*Packet Data CHannel*).

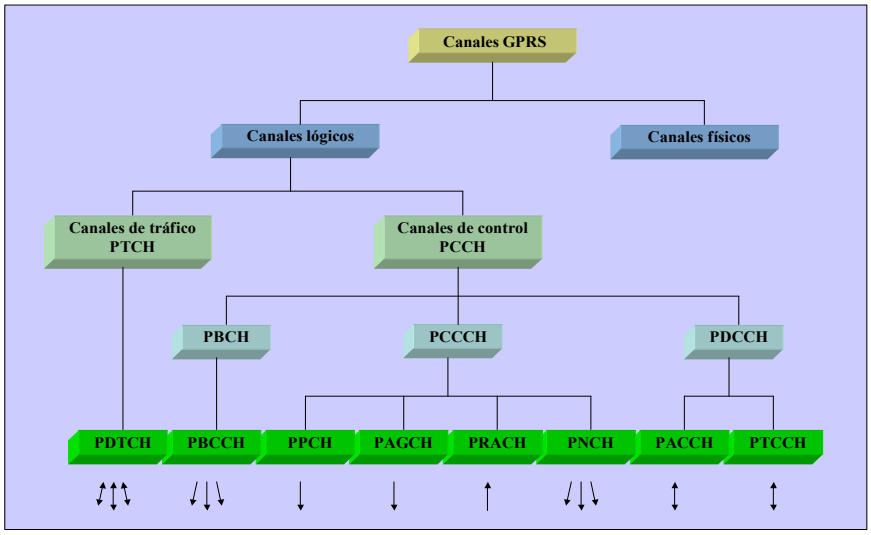
2.5.1.1.2.- CANALES LÓGICOS

El concepto de canal lógico [10] está relacionado con el de canal físico, si bien en este caso la distinción entre canales lógicos se hace atendiendo a la naturaleza o finalidad de la información que transporta. Así pues, un canal físico o PDCH puede ser utilizado por uno o más canales lógicos (no de forma simultánea).

En GPRS se establece la siguiente clasificación de los canales lógicos:

- Canales de tráfico.
- Canales de control.
    - Canales de difusión.
    - Canales de control común.
    - Canales de control dedicado.

La figura 2.5 muestra la clasificación completa de canales en GPRS y el sentido de cada uno (*uplink*, *downlink* o bidireccional) así como su carácter punto a punto (PTP) o punto a multipunto (PTM).



PTCH: Packet Traffic CHannels **PCCH:** Packet Control CHannels
PBCH: Packet Broadcast CHannels **PCCCH:** Packet Common Control CHannels **PDCCH:** Packet Dedicated Control CHannels
PDTCH: Packet Data Traffic CHannel **PBCCH:** Packet Broadcast Control CHannel
PPCH: Packet Paging CHannel **PAGCH:** Packet Access Grant CHannel
PRACH: Packet Random Access CHannel **PNCH:** Packet Notification CHannel
PACCH: Packet Associated Control CHannel **PTCCH:** Packet Timing-advance Control Channel
↑: Canal utilizado en *uplink* ↓: Canal utilizado en *downlink* ↕: Canal utilizado en *uplink* y *downlink*
Una flecha: Carácter Punto-a-Punto (PTP) **Varias flechas:** Carácter Punto-a-Multipunto (PTM)

Figura 2.5.- Clasificación de los canales GPRS

2.5.1.1.2.1.- CANALES DE TRÁFICO (PTCH)

Los canales de tráfico o PTCHs (*Packet Traffic CHannels*) son aquellos dedicados a transportar datos de usuario. En GPRS se define un único tipo de PTCH.

Un canal de tráfico de datos o PDTCH (*Packet Data Traffic CHannel*) se reserva para transferir datos de usuario. Se dedica temporalmente a una MS (PTP) o se comparte secuencialmente por un grupo de MSs (PTM). Un terminal que soporte multislot puede utilizar varios PDTCHs en paralelo para la transferencia individual de información. Este canal se utiliza en ambos sentidos, si bien se realizan dos asignaciones individuales, una en sentido *uplink* y otra en sentido *downlink*.

2.5.1.1.2.2.- CANALES DE CONTROL (PCCH)

Los canales de control o PCCHs (*Packet Control CHannels*) se utilizan para transmitir la información de control necesaria para el establecimiento de la comunicación y la gestión de un PDTCH.

2.5.1.1.2.2.1.- CANALES DE DIFUSIÓN (PBCH)

Los canales de difusión o PBCHs (*Packet Broadcast CHannels*) se utilizan para transmitir información de control a un conjunto específico de MSs. Un canal de control de la difusión o PBCCH (*Packet Broadcast Control CHannel*) se utiliza en *downlink* y con carácter PTM para transmitir información de interés general a todas las MS, como identificador de celda, parámetros de control de potencia, de recepción discontinua (DRX, *Discontinuous Reception*) o de reselección de celda.

2.5.1.1.2.2.2.- CANALES DE CONTROL COMÚN (PCCCH)

Los canales de control común o PCCCHs (*Packet Common Control CHannels*) transportan la información de señalización necesaria para dialogar el establecimiento de la comunicación entre MS y red. Los PCCCHs establecen una relación entre la red y varias MSs.

El canal de búsqueda o PPCH (*Packet Paging CHannel*) es utilizado por la red cuando desea localizar una MS para establecer con ella la conexión y realizar una posterior transferencia de paquetes de información en sentido *downlink*.

El canal de acceso garantizado o PAGCH (*Packet Access Grant CHannel*) es soportado en sentido dowlink y con carácter PTP para asignar a una MS uno o más PDTCHs en los que intercambiar paquetes de datos.

El canal de acceso aleatorio o PRACH (*Packet Random Access CHannel*) es utilizado por la MS cuando desea solicitar uno o más PDTCHs u obtener información acerca del Timing Advance[7].

El canal de notificación o PNCH (*Packet Notification CHannel*) se envía en *downlink* para informar a la MS del inicio de la transferencia de información con carácter punto a multipunto (PTM).

2.5.1.1.2.2.3.- CANALES DE CONTROL DEDICADO (PDCCH)

Los canales de control dedicado o PDCCHs (*Packet Dedicated Control CHannels*) transportan la información de señalización exclusiva entre la red y una única MS. Los PDCCHs establecen, así pues, una relación entre la red y una MS concreta.

El canal de control asociado o PACCH (*Packet Associated Control CHannel)* transporta información de señalización relacionada con una MS específica (como control de potencia, asignación y reasignación de recursos, etc.) y siempre va relacionado con uno o más PDTCHs, con los cuales comparte los recursos. También se utiliza por la red para solicitar el establecimiento de comunicación de conmutación de circuitos cuando la MS se encuentra transmitiendo o recibiendo información en modo paquete.

El canal de control del Timing Advance o PTCCH (*Packet Timing-advance Control CHannel*) se utiliza para mantener la sincronización entre la MS y la red. Se distingue entre PTCCH/U (en *uplink*) y PTCCH/D (en *downlink*). En los PTCCH/U se transmiten ráfagas de acceso aleatorio (*random access bursts*) que permiten a la red evaluar el retardo introducido por el canal radioeléctrico. En los PTCCH/D se envía información relativa al Timing Advance que la MS utiliza para transmitir en el instante adecuado y mantener así la sincronización con la red.

[7] Los retardos de propagación propios del medio radioeléctrico pueden causar la pérdida de sincronismo entre la red y la MS. Por ello, en función de diversas medidas realizadas, la red informa a la MS del momento exacto en el que debe transmitir la información para que ésta llegue en el instante correcto y se mantenga así la sincronización. Esta información se proporciona en el Timing Advance.

2.5.1.2.- GESTIÓN DE LOS RECURSOS RADIO

No hay que olvidar que GPRS supone la adición a la red GSM tradicional de una serie de funcionalidades basadas en la conmutación de paquetes que antes no poseía, lo cual no actúa en detrimento de los servicios iniciales de GSM, que coexisten con los nuevos de GPRS[8]. Sin embargo la introducción de nuevas prestaciones supone un mayor gasto de recursos en toda la red, lo cual ha provocado una mayor repercusión sobre aquellos ámbitos más limitados. Por ello, aunque GPRS sigue utilizando algunas técnicas de gestión empleadas en GSM, es necesaria la introducción de otras nuevas a fin de mejorar el aprovechamiento de los recursos radio.

2.5.1.2.1.- COMPARTIMIENTO DE LOS RECURSOS RADIO

En el caso del medio radio, y dada su gran limitación, la tecnología GPRS se ha visto obligada a compartir los recursos radioeléctricos ya existentes para GSM. En GPRS no se ha ampliado la banda de frecuencia empleada ni se ha reorganizado el espectro disponible: los recursos radio disponibles en una celda siguen siendo los mismos que en GSM, salvo que ahora son compartidos por los servicios de GPRS y de GSM que se ofrecen dentro de ella. Para ello se ha hecho necesaria la introducción de nuevas técnicas de gestión para mejorar el compartimiento de los recursos disponibles y optimizar al máximo su aprovechamiento.

2.5.1.2.1.1.- CONCEPTO DE CANAL MAESTRO/ESCLAVO

A fin de simplificar el concepto de canal lógico, los canales físicos o PDCHs se pueden clasificar como canal maestro (MPDCH, *Master Packet Data CHannel*) o como canal esclavo (SPDCH, *Slave Packet Data CHannel*) en función del tipo de información que transporte, es decir, en función del canal lógico acomodado dentro del propio PDCH [7, 8, 10].

- **Canal maestro, MPDCH:** habitualmente es el CCCH[9] quien transporta la información de señalización y control de tráfico dentro de una celda. En aquellos casos en que el CCCH no se encuentre activo o no se utilice, al menos un PDCH, actuando como maestro, se encarga de acomodar el PCCCH para la transmisión de esta señalización que, en principio, compete al CCCH. De esta forma se comparte siempre un único canal para señalización GSM y GPRS.

[8] GPRS usa la misma red de acceso que GSM tradicional, la misma modulación radio, las mismas bandas de frecuencia, la misma estructura de ráfaga, las mismas leyes de salto en frecuencia (*frecuency hopping*), la misma estructura de trama y, aunque las agrupaciones de jerarquía superior a la trama son diferentes, es posible su reagrupación y amoldamiento a la hipertrama de GSM. En resumen, *GPRS es una característica más de GSM*.

[9] El CCCH o *Common Control CHannel* es un canal lógico definido para GSM que realiza para conmutación de circuitos la misma función que el PCCCH realiza para conmutación de paquetes, es decir, transportar la información de señalización necesaria para establecer la conexión entre móvil y red e iniciar la transferencia de información.

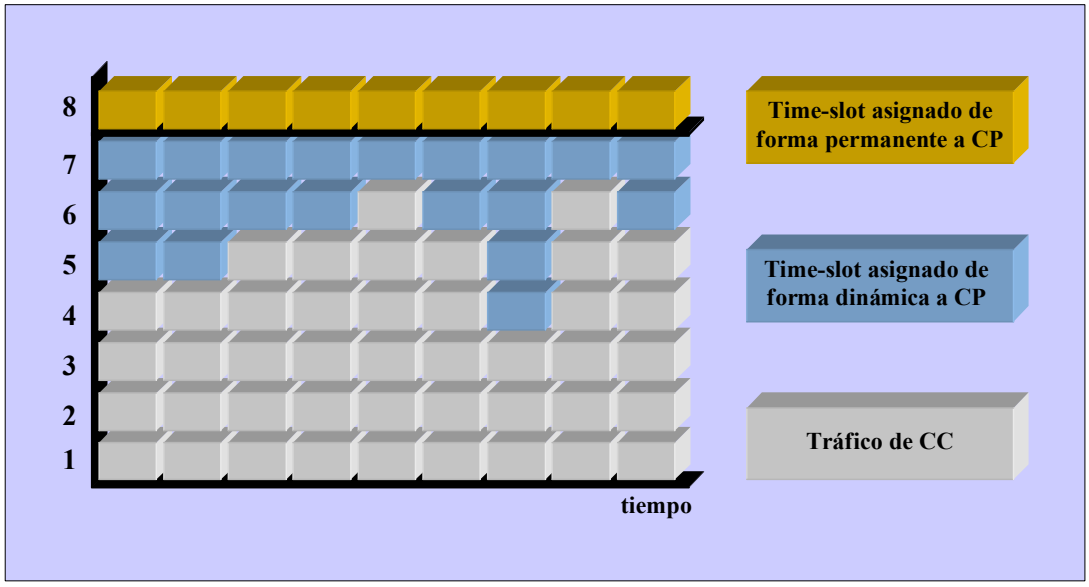
La transmisión de señalización no suele ser tan frecuente como la de información de usuario. Por ello, el MPDCH, cuando se encuentra inactivo, puede ser reutilizado para transportar datos de usuario (PDTCH) o para transportar señalización dedicada (PACCH). No obstante, en un MPDCH, el transporte de PCCCHs es prioritario sobre el transporte de los demás canales lógicos.

- **Canal esclavo, SPDCH:** un canal esclavo es un PDCH que se reserva única y exclusivamente para el soporte de PDTCHs y PACCHs, es decir, se utiliza sólo para transportar datos de usuario y señalización individual punto a punto entre la red y una única MS.

En una celda se define al menos un MPDCH (aunque puede ser más de uno si así lo requieren las circunstancias) y el resto de PDCHs se emplean como SPDCHs para transportar datos de usuario y señalización dedicada. Cuando el MPDCH se encuentra desocupado, se reutiliza en calidad de SPDCH, teniendo siempre prioridad sus funciones como MPDCH.

2.5.1.2.1.2.- ASIGNACIÓN DE RECURSOS BAJO DEMANDA

Como ya se ha dicho anteriormente, los recursos radio de una celda son compartidos por servicios GPRS y GSM. Por lo tanto, cuando un operador decide dar servicios GPRS debe reservar uno o más canales físicos dentro de la celda para utilizarlos como PDCHs, es decir, destinar uno o más canales físicos a la transferencia de información en modo paquete. Esta distribución de recursos entre servicios de conmutación de circuitos y de conmutación de paquetes se realiza de forma dinámica según en el principio de *capacity-on-demand* [6, 8], tal y como se muestra de forma gráfica en la figura 2.6.



CP: Conmutación de paquetes **CC:** Conmutación de circuitos

Figura 2.6.- Asignación de recursos a GPRS según el principio de *capacity-on-demand* [2]

De entre todos los recursos disponibles en la celda, el operador puede decidir reservar uno o más time-slots de forma permanente para la transferencia de datos en modo paquete. El resto de time-slots estarán disponibles para ser utilizados por servicios de conmutación de paquetes (actuando como PDCHs) o de conmutación de circuitos (actuando como TCHs). En caso de que haya insuficiencia de time-slots para atender a toda la demanda de conmutación de circuitos (CC) y de conmutación de paquetes (CP), los servicios de CC tienen prioridad.

Así pues, si todos los recursos asignados para CP (tanto permanente como temporalmente) resultan insuficientes en un momento dado pero sin embargo existen recursos disponibles en la celda, la red puede entonces reservar más time-slots para ser utilizados como PDCHs. Si una vez utilizado el PDCH ya no se necesita para transmitir más información en modo paquete, entonces es inmediatamente liberado y puesto nuevamente a disposición de servicios tanto de CC como de CP. Los time-slots no utilizados por GSM pueden ser utilizados por GPRS para incrementar la calidad del servicio.

Sin embargo, si en cualquier instante se necesitan más recursos para servicios de CC y el resto de time-slots asignados de forma dinámica se encuentran en ese momento utilizados como PDCHs, la red puede proceder a liberar los PDCHs necesarios y utilizar esos time-slots como TCHs, ya que los servicios de CC tienen prioridad sobre los de CP[10]. De esta forma, la distribución de los recursos de una celda entre GSM y GPRS se realiza de forma dinámica en función la capacidad que GSM deja libre y de la demanda realizada por GPRS, y esto es lo que se conoce como principio *capacity-on-demand*.

El mecanismo de prioridad en que se basa este principio requiere la implementación de dos funciones primordiales [10]:

- **Supervisión de carga:** La función de supervisión de carga monitoriza la carga de los PDCHs para determinar, en función de la demanda, si es necesario reservar nuevos time-slots como PDCHs. Esta función puede ser implementada como parte de la capa MAC.

- **Asignación dinámica de PDCHs:** De entre los canales que se asignan de forma dinámica a CC o CP, aquellos no utilizados por GSM pueden ser reservados como PDCHs para mejorar la calidad del servicio de GPRS. Por otra parte, si servicios de CC (de mayor prioridad) requieren más recursos de los que disponen en ese momento, entonces puede tener lugar la liberación de uno o más PDCHs para que los time-slots se utilicen por GSM.

[10] Los motivos por los cuales los servicios de CC de GSM tienen prioridad sobre los servicios de CP de GPRS son los siguientes: en CC, al disminuir los recursos disponibles aumenta la probabilidad de bloqueo (impidiendo nuevas comunicaciones), mientras que en CP al disminuir los recursos disponibles disminuye la calidad del servicio experimentada por cada uno de los usuarios, si bien la comunicación sigue siendo posible.

Por otra parte, el concepto de *asignación de recursos bajo demanda* también adquiere otro matiz en GPRS. El principio de *capacity-on-demand* determina cómo se distribuyen los recursos de una celda entre GSM y GPRS. Posteriormente, la forma en que cada sistema asigna sus recursos a los usuarios de sus servicios (CC en GSM y CP en GPRS) es diferente. Efectivamente, podemos observar que:

- **En GSM** se asigna un único time-slot por MS. Esta asignación es fija, es decir, el recurso permanece reservado para la misma MS mientras dura la comunicación establecida, incluso cuando no se transmite información por él. Dicho recurso sólo es liberado cuando la sesión finaliza.

- **En GPRS** la asignación se realiza de forma dinámica, es decir, sólo cuando se va a transmitir uno o más paquetes de información. Inmediatamente después de terminar la transmisión se libera el recurso, dejándolo disponible para ser utilizado por otra comunicación distinta. Así pues, la sesión puede durar indefinidamente sin ocupar recursos (siempre que no se transmita información). Para aplicaciones de datos, de naturaleza intermitente, este procedimiento proporciona un mejor aprovechamiento de los escasos recursos radio, ya que con este principio varias MSs pueden compartir de forma alternada un mismo PDCH (multiplexación estadística).

  Además, es posible asignar uno o más time-slots a una misma MS dependiendo de la carga de la celda, de la prioridad del servicio y de las capacidades multislot de la MS. Una ventaja adicional reside en el hecho de que la asignación de time-slots en *uplink* y en *downlink* puede ser asimétrica para mejorar la eficiencia espectral, ya que generalmente el flujo de información en sentido red a MS es mayor que en sentido MS a red (por ejemplo, navegación web).

Como consecuencia de lo anterior, se puede afirmar que una disminución de los recursos disponibles para servicios GSM implica la aparición de una mayor probabilidad de bloqueo o incapacidad para atender a más usuarios, mientras que una disminución de los recursos disponibles para GPRS resulta en un empeoramiento de la calidad del servicio experimentada por el usuario, si bien la comunicación sigue siendo posible, razón por la cual los servicios de CC tienen prioridad sobre los de CP.

2.5.1.2.2.- MODOS DE OPERACIÓN RADIO

Los procedimientos de gestión de los recursos radio se caracterizan por dos modos de operación diferentes (*Radio Resource operating modes* [10]), cada uno de los cuales define un cierto grado de funcionalidad y de información. Estos modos reciben el nombre de modo inactivo y modo transferencia.

2.5.1.2.2.1.- MODO INACTIVO

En modo inactivo (*Packet idle mode* [10]) la estación móvil de usuario no transmite paquetes de datos, por lo que no hay asignado ningún canal. Así pues, en modo inactivo no se utiliza ningún recurso radio. Si las capas superiores requieren la transmisión de una trama LLC, entonces se solicitará el establecimiento de uno o más canales y se pasará a modo transferencia.

2.5.1.2.2.2.- MODO TRANSFERENCIA

En modo transferencia (*Packet transfer mode* [10]) la estación móvil de usuario tiene asignado uno o más time-slots y, por tanto, dispone de recursos radio reservados para la transmisión de información de usuario. Se pueden asignar time-slots simultáneos en direcciones opuestas. Ante un cambio de celda la MS pasa primero a modo inactivo, obtiene la información del sistema y vuelve nuevamente a modo transferencia.

2.5.1.2.3.- ACCESO MÚLTIPLE

El acceso múltiple define la forma en que el medio radio es accedido, de forma controlada, por multitud de usuarios. En GPRS [11] el acceso múltiple es idéntico a GSM: consiste en una mezcla de división en frecuencia (FDMA) y división en tiempo (TDMA). Además, el duplexado se establece mediante división en frecuencia (FDD), por lo que se utilizan portadoras diferentes en sentido *uplink* y *downlink*. La figura 2.7 muestra la organización del acceso múltiple.

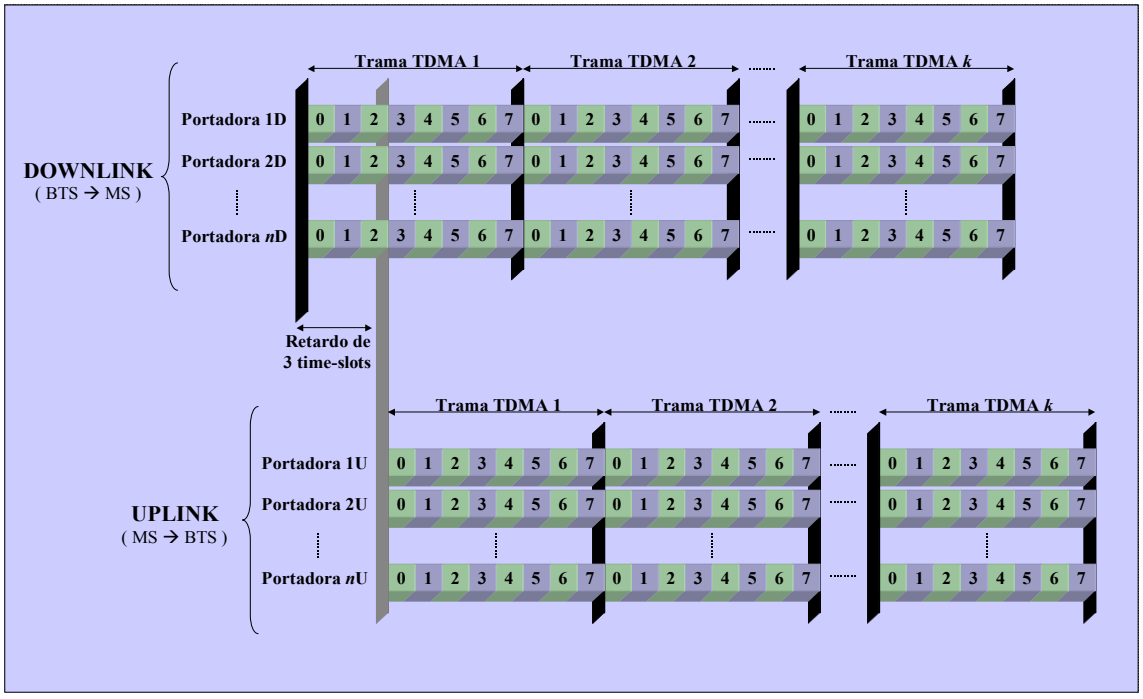


Figura 2.7.- Acceso múltiple [12]

La distribución del espectro de GPRS [6, 9] es la misma que la realizada en GSM900, para el que inicialmente se hicieron reservas en la banda de 900 MHz. Dos bandas separadas de 25 MHz cada una se utilizan para transmisión en *uplink* (890 – 915 MHz) y en *downlink* (935 – 960 MHz)[11]. Otras bandas adicionales en la banda de 1800 MHz se reservaron posteriormente para GSM (también denominado en este caso DCS1800). Una separación entre portadoras de 200 kHz da lugar a un total de

$$\frac{25MHz + 25MHz}{200kHz} = 125 \text{ canales } uplink + 125 \text{ canales } downlink = 125 \text{ comunicaciones}$$

En la práctica se utilizan n = 124 canales (tanto en *uplink* como en *downlink*) para la transmisión. Un canal es subdividido en dos mitades de 100 kHz que se utilizan como bandas de guarda en los límites inferior y superior de ambas bandas de 25 MHz. La función de estas bandas de guarda es dejar un margen de separación espectral para prevenir posibles solapamientos con otras bandas espectrales adyacentes.

En el aspecto temporal del acceso múltiple, cada portadora se divide en ocho intervalos temporales o time-slots de aproximadamente 0.577 ms de duración cada uno, constituyendo la trama TDMA, que se repite cíclicamente de forma indefinida y cuya duración es de 8 × 0.577 ms = 4.615 ms. La dimensión temporal incluye otros tipos de agrupaciones temporales más complejos, constituyendo una jerarquía temporal cuya descripción tiene lugar en el apartado 2.5.1.2.3.1.

Cuando una MS tiene algún tipo de información a transmitir de las capas de aplicación, previamente la subdivide en bloques de bits más pequeños que, tras ser codificados, se fragmentan y distribuyen en ráfagas de bits cuya duración coincide con la de un time-slot (este proceso, mostrado en la figura 2.9, se describe con más detalle en el apartado 2.5.1.3). La MS GPRS tiene asignado uno o más números identificadores de time-slot (del 0 al 7) que utiliza para identificar los intervalos de tiempo (time-slots) en los que puede transmitir o recibir. Así pues, si tiene asignado el time-slot número 5, transmitirá una ráfaga en el time-slot 5 de la trama k, la siguiente ráfaga la transmitirá en el time-slot 5 de la trama k + 1, y así sucesivamente, acomodando sus ráfagas de bits en la misma posición relativa dentro de tramas consecutivas. Además de transmitir en el instante adecuado, debe hacerlo en la frecuencia adecuada. En GSM se utiliza una técnica denominada salto en frecuencia (*frecuency hopping*) [11] que consiste en variar continuamente la portadora por la que se transmiten los time-slots siguiendo una secuencia conocida de antemano tanto por la MS como por la red. De esta forma, los time-slots de una trama se reparten en diversas frecuencias portadoras, conservando siempre la misma posición temporal relativa dentro de la trama[12].

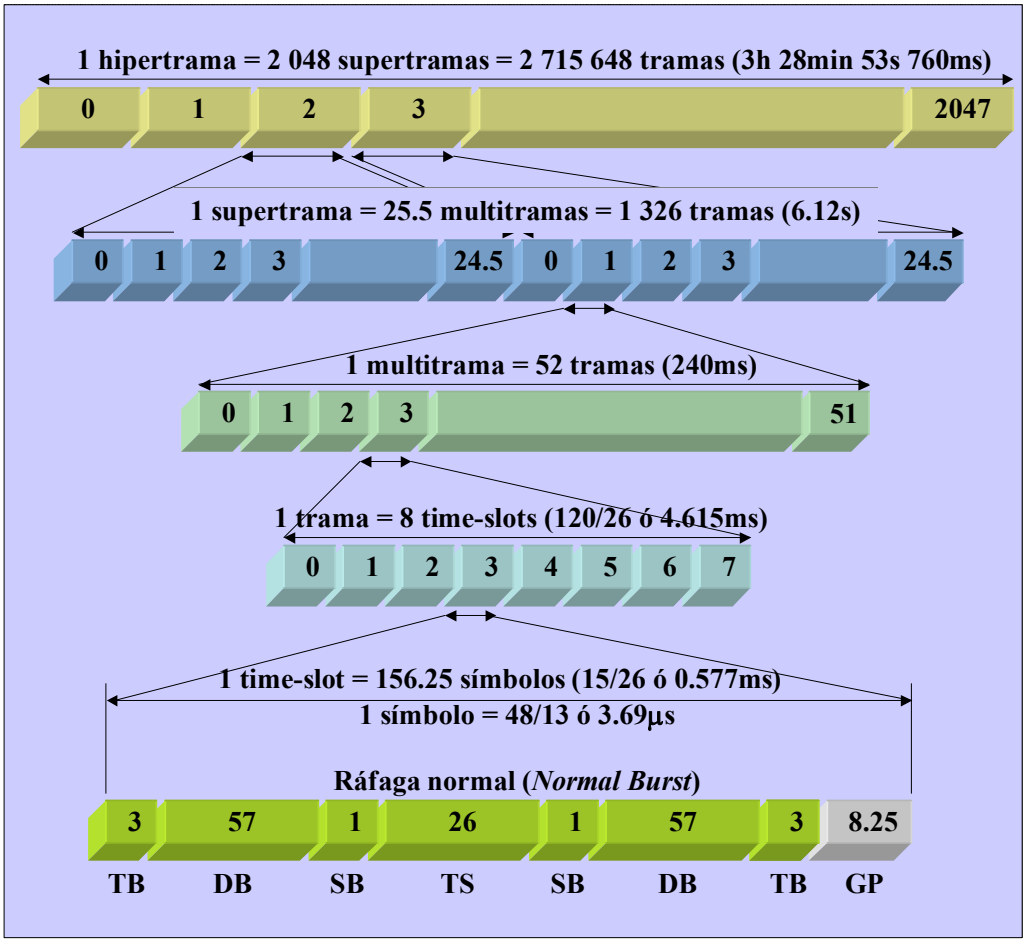
[11] Observar que los canales *uplink* (frecuencias bajas) y *downlink* (frecuencias altas) de una misma comunicación se ubican en bandas separadas. El motivo es que las frecuencias bajas sufren una menor atenuación por el canal, por lo que es preferible reservarlas para el *uplink*, donde la potencia de transmisión es mucho menor que en el *downlink*.

[12] Esta técnica pretende disminuir el efecto de las interferencias. Supongamos que existe una señal que interfiere con una frecuencia portadora *f*. Todos los time-slots transmitidos por esa portadora sufrirán la interferencia. Si todos los time-slots de un usuario se transmiten siempre por la misma portadora, entonces la información de aquel que lo haga en la portadora *f* quedará seriamente dañada. Si realizamos el salto en frecuencia, las comunicaciones que utilicen la portadora *f* se verán afectadas por la interferencia pero sólo en los time-slot transmitidos por esa portadora. En este caso la cantidad de información distorsionada dentro de una misma comunicación es menor.

Observar en la figura 2.7 que todos los canales de radiofrecuencia en sentido *downlink* están temporalmente alineados entre sí. Lo mismo sucede en el *uplink*. Además, desde el punto de vista de la BTS el comienzo de la trama TDMA en *uplink* sufre un retardo constante de 3 time-slots (3 × 0.577 ms = 1.73 ms) respecto al comienzo de la trama TDMA en sentido *downlink*, ya que la MS transmite en el instante adecuado para que esto suceda siempre así. En cambio, desde la MS este retardo será variable según la información de *Timing-Advance*, para hacer así que la BTS reciba la información con un retardo constante de 3 time-slots y mantener de esta forma la sincronización con la red. La razón de este desfase entre las tramas en *uplink* y en *downlink* es hacer posible que una MS utilice el mismo número de time-slot en ambos sentidos de la comunicación sin necesidad de transmitir y recibir de forma simultánea, lo cual supondría una mayor complejidad tanto en software como en hardware [11].

2.5.1.2.3.1.- JERARQUÍA TEMPORAL

Las subdivisiones y agrupaciones temporales vinculadas a GPRS, que se contemplan en las especificaciones [11, 12], se muestran en la figura 2.8.



TB: *Tail Bits* (bits de cola) **DB:** *Data Bits* (bits de datos) **SB:** *Stealing Bits* (bits robados)
TS: *Training Sequence* (secuencia de entrenamiento) **GP:** *Guard Period* (periodo de guarda)

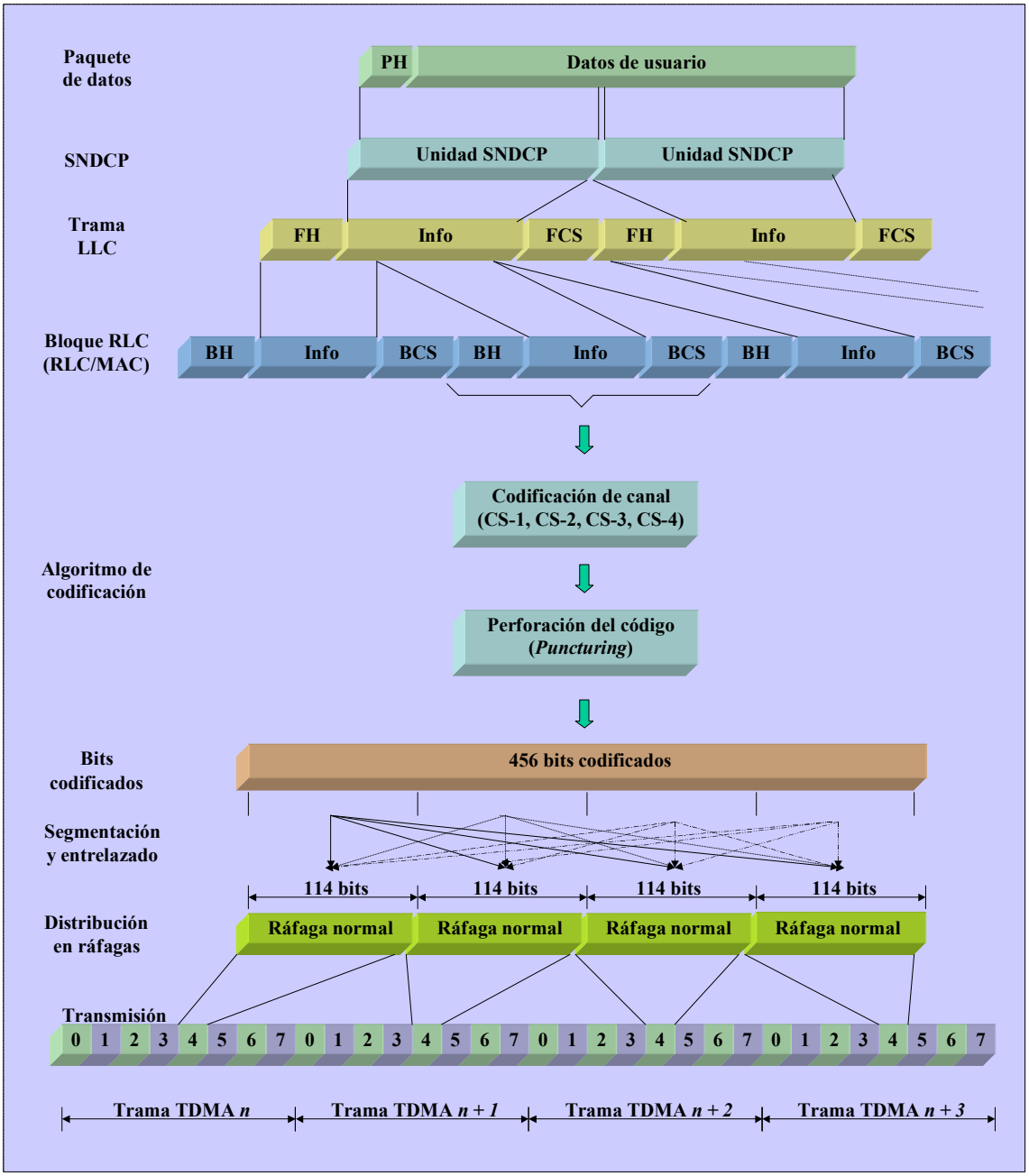
Figura 2.8.- Jerarquía temporal [11]

Los bits transmitidos a través del medio radio se agrupan en ráfagas que tienen la duración de un time-slot. En GSM se especifican cinco tipos de ráfagas[13]: ráfagas normales (*normal bursts*), ráfagas de acceso aleatorio (*random access bursts*), ráfagas de sincronización (*synchronisation bursts*), ráfagas de corrección de frecuencia (*frecuency correction bursts)* y ráfagas de relleno (*dummy bursts*). La única utilizada para transmitir datos de usuario es la ráfaga normal (figura 2.8). Ésta será la única de interés para la labor que nos ocupa[14]. La ráfaga normal [9] se utiliza para transportar información de usuario o de señalización asociada. Dicha información se transmite en los dos campos DB (*Data Bits,* bits de datos) de 57 bits cada uno. Los dos bits SB (*Stealing Bits*, bits robados) se utilizan en los canales PDTCH de GPRS para indicar el esquema de codificación empleado. El campo TS (*Training Sequence*, secuencia de entrenamiento) contiene una secuencia de 26 bits que es conocida de antemano tanto por emisor como por receptor. Esta secuencia se utiliza tanto para sincronización como para ecualización del canal. Para esta segunda finalidad, puesto que se conoce de antemano cuál ha de ser el contenido del campo TS, comparando la secuencia recibida con la esperada es posible realizar una estimación de cómo afecta el canal a los símbolos transmitidos (respuesta al impulso del canal) y realizar así una ecualización adecuada. También puede utilizarse como parámetro de la calidad del canal en *Link Adaptation* (apartado 3.5.3). Los bits del campo TB (*Tail Bits*, bits de cola) son introducidos por el proceso de codificación. Finalmente, el campo GP (*Guard Period*, periodo de guarda) tiene una duración temporal equivalente a 8.25 símbolos de canal ($8.25 \times 3.69\ \mu s = 30.44\ \mu s$) y es necesario para acomodar la región de transición en la conexión y desconexión de la etapa de potencia (antes y después de transmitir en un time-slot) y evitar así el solapamiento de dos ráfagas transmitidas en time-slots adyacentes. Un símbolo de canal (equivalente a un bit codificado ya que la modulación GMSK utilizada es mononivel) tiene una duración de $48/13 \approx 3.69\ \mu s$. Sumando la longitud de todos los campos de una ráfaga normal se obtiene una duración equivalente a 156.25 símbolos, los cuales se acomodan dentro de un time-slot, por lo que un time-slot tiene una duración total de $156.25 \times 48/13\ \mu s = 15/26\ ms \approx 0.577\ ms$. Los time-slots se agrupan en grupos de ocho constituyendo una trama, la cual dura $8 \times 15/26\ ms = 120/26\ ms \approx 4.615\ ms$. La concatenación de 52 tramas da lugar a una multitrama de $52 \times 120/26\ ms = 240\ ms$ de duración. Observar que la supertrama agrupa a 25.5 multitramas, un número no entero. La razón es que en las especificaciones GSM [11] iniciales se contemplaba la posibilidad de agrupar 26 multitramas de 51 tramas cada una (la multitrama duraba $3060/13 \approx 235.35\ ms$) o 51 multitramas de 26 tramas cada una (la multitrama duraba 120 ms). Cualquiera de las dos agrupaciones daba lugar a una supertrama de $26 \times 51 = 1326$ tramas que duraba $1326 \times 120/26\ ms = 6.12\ s$. En GPRS la multitrama se ha definido de 52 tramas y con una duración de 240 ms. A fin de hacer compatible esta nueva multitrama con la hipertrama existente es necesario mantener la duración de la supertrama en 6.12 s, por lo que se ha definido la supertrama como una agrupación de 6.12 s / 240 ms = 25.5 multitramas. No obstante la multitrama de 52 tramas puede verse como la concatenación de 2 multitramas de 26 tramas cada una que se agrupan en una supertrama de $25.5 \times 2 = 51$ multitramas (con 26 tramas cada multitrama). Finalmente, la agrupación de 2048 supertramas conforma una hipertrama, que es la agrupación de más alta jerarquía. La hipertrama tiene una duración total de $2048 \times 6.12\ s = 12533.76\ s = 3\ h\ 28\ min\ 53\ s\ 760\ ms$.

[13] No confundir el término ráfaga (*burst*) con el de ráfaga de datos relativa al tráfico con carácter intermitente.
[14] Una descripción detallada de los diferentes tipos de ráfaga se halla en la especificación técnica GSM 05.02 [10].

2.5.1.3.- FLUJO DE DATOS

Desde que un paquete de datos es originado por una aplicación de usuario hasta que es transmitido por el medio radioeléctrico sufre una serie de diferentes subdivisiones y procesos. La figura 2.9 muestra el esquema global del flujo de datos que sufre la información hasta llegar a las capas más bajas que, a grandes rasgos, se describe a continuación.



FCS: Frame Check Sequence (*Secuencia de comprobación de trama*) **BCS:** Block Check Sequence (*Secuencia de comprobación de bloque*)
PH: Packet Header (*Cabecera de paquete*) **FH:** Frame Header (*Cabecera de trama*) **BH:** Block Header (*Cabecera de bloque*)

Figura 2.9.- Flujo de datos

El paquete de datos contiene la información de usuario procedente de las capas de aplicación. Este paquete es segmentado en unidades más pequeñas sobre las que el protocolo SNDCP realiza funciones de compresión adicional de cabecera y de datos así como una segmentación y encriptación. Las unidades procesadas en la capa SNDCP se encapsulan dentro de tramas LLC constituyendo el campo de información de dichas tramas. En la capa LLC, además de la cabecera correspondiente (FH, *Frame Header*) se añade un campo (FCS, *Frame Check Sequence*) para la comprobación de trama.

Las PDUs de la capa LLC se dividen en varios fragmentos antes de ser encapsulados en PDUs de la capa RLC/MAC subyacente. En esta capa se añade una cabecera (BH, *Block Header*) y un campo (BCS, *Block Check Sequence*) para comprobación de errores[15]. Una explicación más detallada del funcionamiento de estas capas será el tema objeto del apartado el 2.5.2.

La secuencia de bits que la capa MAC entrega a la capa física se denomina bloque RLC y constituye un bloque de entrada al codificador de canal. Tal y como se verá en el apartado 2.5.3.1.2.1, en GPRS existen cuatro esquemas de codificación, cada uno de los cuales admite como entrada secuencias de bits de diferente longitud. Sin embargo, un bloque de datos de la capa RLC/MAC no es fragmentado antes de ser codificado según el tamaño de datos requerido por el codificador de canal, sino que es introducido entero al codificador. La forma en que se consigue ajustar el tamaño de la secuencia de entrada del codificador consiste en variar la longitud de cada una de las partes en que se divide una trama LLC.

Así pues, las tramas LLC se dividen en fragmentos con el tamaño adecuado de tal forma que, tras ser procesados por la capa RLC/MAC, dan lugar a bloques con el tamaño correspondiente a la longitud de entrada del esquema de codificación empleado. Un bloque RLC es entonces introducido íntegramente al codificador de canal dando lugar a la secuencia de bits codificados.

Para los cuatro posibles esquemas de codificación, independientemente del tamaño de la secuencia de entrada de cada uno, se obtienen siempre 456 bits a la salida. Estos 456 bits, tras un proceso de segmentación y entrelazado (descrito en el apartado 2.5.3.1.2.2) se reparten en cuatro grupos de 456/4=114 bits, cada uno de los cuales puede acomodarse en los dos campos de 114/2=57 bits de datos de una ráfaga normal. Las cuatro ráfagas obtenidas (transportando en total un bloque RLC codificado) se transmiten en el mismo time-slot relativo dentro de cuatro tramas TDMA consecutivas.

[15] Según las especificaciones GPRS [8], un fragmento LLC que se antecede por una cabecera RLC se denomina *bloque RLC* (de datos). Cuando a un bloque RLC se le añade una cabecera MAC al principio y un campo BCS al final, entonces la PDU se denomina *bloque radio*. No obstante, algunos autores utilizan indistintamente los términos bloque RLC y bloque radio.

2.5.2.- LA CAPA RLC/MAC

2.5.2.1.- SERVICIOS

La capa de enlace de datos de la interfaz aire está dividida en varias subcapas, siendo las subcapas RLC y MAC [10] las de nivel inferior dentro de dicha capa.

En la capa MAC (*Medium Access Control*) se definen los procedimientos que permiten a varias MSs compartir un medio común de transmisión, el cual consiste en varios canales físicos. La capa MAC provee el arbitraje entre varias MSs que intentan transmitir de forma simultánea, aportando los procedimientos necesarios para evitar y detectar colisiones así como para lograr su recuperación. Las funciones de la capa MAC también permiten a una estación móvil de usuario utilizar uno o más PDCHs en paralelo. La capa RLC (*Radio Link Control*) se encarga de segmentar (y reensamblar) las tramas LLC en PDUs que a este nivel se denominan bloques RLC (*RLC blocks*), añadiendo la información necesaria para posibilitar la retransmisión selectiva de bloques erróneos (corrección de errores hacia atrás o BEC, *Backward Error Correction*).

La capa RLC/MAC proporciona dos modos de operación diferentes: uno con confirmación (*acknowledged*) y otro sin confirmación (*unacknowledged*). Estos modos se tratan, respectivamente, en los apartados 2.5.2.4.1 y 2.5.2.4.2.

2.5.2.2.- FUNCIONES

La capa MAC tiene la responsabilidad de ejercer las siguientes funciones [10]:

- Proveer una multiplexación eficiente de los datos y la señalización de control relacionada con diversas fuentes, tanto en sentido *uplink* como en *downlink*. Dicha función de control reside en el lado de la red. En sentido *downlink* la multiplexación es controlada por un mecanismo programado. En el sentido *uplink* la multiplexación se controla en función de la reserva de recursos (en respuesta a la demanda del servicio).

- En sentido *uplink*, solventar la contienda en el acceso al medio por varias MSs que intentan acceder de forma simultánea a un mismo recurso, detectando las posibles colisiones y habilitando los mecanismos necesarios para su recuperación.

- En sentido *downlink*, proveer los mecanismos para acceder de forma coordinada a los recursos existentes, incluyendo la puesta en cola de los accesos fallidos.

- Manejar la prioridad de los servicios, haciendo que la información sea transmitida por el medio en función de la carga existente y de la prioridad asociada a la calidad de servicio contratada.

Por otra parte, la capa RLC tiene como obligaciones [10]:

- Proporcionar las primitivas que permiten entregar las PDUs de la capa LLC (LLC-PDUs) a las funciones de la capa MAC.

- Segmentar las LLC-PDUs en bloques RLC en el extremo de origen, así como reensamblar los bloques RLC en LLC-PDUs en el extremo de destino.

- Supervisar la corrección de errores hacia atrás (BEC), mecanismo que permite la retransmisión de los bloques radio dañados por el medio físico[16].

2.5.2.3.- IDENTIFICADORES

2.5.2.3.1.- UPLINK STATE FLAG (USF)

El campo USF (*Uplink State Flag* [10]) de un bloque RLC (3 bits de longitud) es utilizado para permitir la multiplexación en *uplink* de la información procedente de varias MSs sobre un mismo PDCH. Cada estación móvil que desea transmitir cualquier información en *uplink* tiene asignado un valor de entre los $2^3 = 8$ posibles valores que se pueden codificar en el USF. Así pues, la estación móvil permanece de forma constante en escucha del USF (que sólo se transmite en los bloques RLC en sentido *downlink*) y cuando detecta que en dicho campo aparece el valor que tiene asignado, entonces reconoce que la red ha reservado ese recurso para ella y realiza la transmisión.

Como puede verse, el USF se transmite sólo en sentido *downlink* con la finalidad de coordinar la transmisión en sentido uplink. Este mecanismo no sólo permite la multiplexación de varias MSs sobre un mismo PDCH, sino que también puede ser utilizado para gestionar la prioridad, ya que si una MS se encuentra transmitiendo y detecta que en el USF el valor cambia a otro distinto del que tiene asignado, entonces cesa su actividad y no sigue transmitiendo por ese PDCH.

Cuando un PDCH transporta un PCCCH se reserva un valor de USF para indicar un PRACH (ver apartado 2.5.1.1.2.2.2) y quedan disponibles siete valores para ser asignados a diferentes MSs. Cuando el PDCH no transporta un PCCCH, entonces los ocho posibles valores del USF se asignan a diferentes MSs.

[16] La secuencia de comprobación de bloque (BCS) no es generada por la capa RLC sino que es una función desempeñada por la capa física, ya que dicha secuencia se genera a partir de aplicar diversos algoritmos de codificación sobre el campo de datos (ver apartado 2.5.3.1.2.1).

2.5.2.3.2.- TEMPORARY BLOCK FLOW (TBF)

Un TBF (*Temporary Block Flow* [10]) es una conexión física establecida de forma unidireccional entre la MS y la red que soporta la transferencia de LLC-PDUs sobre PDCHs. Durante la transmisión de datos, la MS envía y/o recibe uno o más paquetes de datos empleando un conjunto de PDCHs que la red ha reservado para ella. Esta asignación temporal recibe el nombre de TBF. El TBF, por tanto, es una asignación transitoria, ya que cuando la transmisión de datos cesa, los recursos asignados son inmediatamente liberados. Además, puesto que el TBF se define unidireccionalmente, es necesario establecer un TBF para el *downlink* y otro TBF diferente para el *uplink*, lo cual posibilita una asignación asimétrica y una eficiente explotación de los recursos.

2.5.2.3.3.- TEMPORARY FLOW IDENTITY (TFI)

Un TFI (*Temporary Flow Identity* [10]) es simplemente un identificador de TBF que es asignado por la red. Cada TFI es único respecto a los TFI que identifican a TBFs en la misma dirección, si bien un mismo TFI puede ser utilizado para identificar TBFs diferentes en direcciones opuestas (*uplink* y *downlink*). Cuando a una estación móvil se le envía un mensaje de asignación de recursos, recibe un TFI. Dicho TFI se incluirá tanto en la cabecera de los bloques RLC destinados al TBF apuntado por dicho TFI como en los mensajes de control asociados a la transferencia de tramas LLC. La MS utilizará este TFI para identificar los TBFs que la red le ha reservado y reconocer así los mensajes destinados a ella, mientras que en la red se utilizará para identificar a la MS.

2.5.2.4.- MODOS DE OPERACIÓN

2.5.2.4.1.- MODO DE OPERACIÓN CON CONFIRMACIÓN

La transferencia de información en modo confirmado [10] está controlada por un mecanismo de retransmisión ARQ asociado a la numeración de los bloques RLC dentro de un TBF. Los bloques transmitidos por la parte emisora son confirmados por la parte receptora mediante la devolución de mensajes *Packet Uplink/Downlink Ack/Nack* según corresponda. Cada mensaje de confirmación positiva (*Ack*) confirma todos los bloques RLC recibidos correctamente hasta un cierto número máximo de secuencia (BSN, *Block Sequence Number*), lo que provoca el desplazamiento del comienzo de la ventana de transmisión (el receptor también desplaza el comienzo de su ventana una posición por cada bloque RLC recibido correctamente). Cada mensaje de confirmación negativa (*Nack*) especifica el número de secuencia que se espera recibir que, al ser comparado por el emisor con el último número de secuencia enviado, permite determinar qué bloque RLC o conjunto de bloques RLC es necesario retransmitir. Tras la retransmisión y confirmación positiva de los bloques, ambas ventanas (de emisor y de receptor) se desplazan. Los mensajes de confirmación (tanto positivo como negativo) no incluyen ningún cambio en la asignación existente. Además, la pérdida de estos mensajes no representa ningún problema crítico ya que son enviados con cierta frecuencia y los bloques no confirmados en el mensaje perdido pueden ser confirmados (junto con otros bloques enviados posteriormente) en el mensaje de confirmación siguiente.

2.5.2.4.2.- MODO DE OPERACIÓN SIN CONFIRMACIÓN

La transferencia de información en modo de operación sin confirmación [10] está controlada por la numeración de los bloques RLC dentro de un TBF y no contempla ningún tipo de retransmisión. La parte receptora extrae la información de usuario de los bloques RLC recibidos, sustituyendo aquellos recibidos con error por secuencias de bits que se utilizan a modo de relleno con la finalidad de preservar la longitud de la información recibida. El que no exista retransmisión en este modo no implica que no se envíen mensajes de confirmación por parte del receptor. De hecho, la parte emisora, cada cierto número de bloques RLC transmitidos, solicita a la parte receptora una confirmación. Estos mensajes de confirmación, que utilizan el mismo mecanismo y formato que los enviados en modo confirmado, contienen campos mediante los que se especifica el número de bloques RLC recibidos erróneamente. Esta información puede ser utilizada por la parte emisora para evaluar la calidad del canal (por ejemplo, como parámetro para *Link Adaptation*) y emprender las acciones correctivas oportunas que eviten el acontecimiento de nuevos errores en bloques RLC posteriores.

2.5.3.- LA CAPA FÍSICA

La capa física en la interfaz aire de GPRS [10, 11] se subdivide en dos capas: la capa PLL (*Physical Link Layer*) y la capa RFL (*Radio Frecuency Layer*).

2.5.3.1.- LA CAPA PLL

2.5.3.1.1.- SERVICIOS

El propósito de la capa PLL es el de transportar información a través de la interfaz radio de GPRS (incluyendo información proveniente de la capa RLC/MAC), proporcionando los procedimientos necesarios para soportar la comunicación sobre el canal radioeléctrico. En esta capa se habilitan los mecanismos necesarios para que varias MSs puedan compartir un mismo canal físico.

2.5.3.1.2.- FUNCIONES

Las funciones principales que se encomiendan a la capa PLL son [10]:

- La realización de una codificación de canal eficaz que posibilite la corrección de errores hacia delante (FEC, *Forward Error Correction*), permitiendo así la detección y corrección de palabras código alteradas o, por lo menos, la detección en aquellos casos en que no sea posible su corrección. Este tema se trata en el apartado 2.5.3.1.2.1.

- El entrelazado rectangular de los bits resultantes de la codificación sobre cuatro ráfagas normales (apartado 2.5.1.2.3.1) y la distribución de estas ráfagas a lo largo de cuatro tramas TDMA consecutivas. Este aspecto se examinará en el apartado 2.5.3.1.2.2.

Otras funciones que también son desempeñadas por la capa PLL son [11]:

- Soportar aquellos procedimientos necesarios para detectar la congestión en el enlace físico.

- Controlar el nivel de potencia de transmisión de la MS y, opcionalmente, también el de la BTS a fin de conseguir la calidad requerida con la menor potencia (e interferencia) posible.

- Asegurar la sincronización entre MS y BTS, realizando para ello medidas del retardo de propagación del canal y calculando el Timing Advance.

- Realizar algunas medidas que puedan ser usadas como parámetro para evaluar la calidad de la señal en el enlace radio.

- Realizar medidas utilizadas para decidir la selección y reselección de celda.

- Proporcionar procedimientos para el ahorro de la batería de usuario (por ejemplo, recepción discontinua o DRX, *Discontinuous Reception*).

2.5.3.1.2.1.- CODIFICACIÓN DE CANAL

En las especificaciones relativas a GPRS se contemplan diversos algoritmos de codificación para cada canal lógico. Dada la finalidad de este trabajo, resulta de especial interés aquel destinado a transportar la información de usuario, es decir, el canal PDTCH.

En GPRS se definen cuatro esquemas de codificación de canal diferentes para los canales PDTCH, denominados CS-1, CS-2, CS-3 y CS-4[17] [9, 10, 13]. Cada uno de ellos se ha diseñado con unas características diferentes para ofrecer, en función de la calidad concreta del canal de transmisión en cada momento, diferentes grados de compromiso entre robustez frente a errores y tasa de información[18]. Así pues, CS-1 es el más robusto frente a errores pero también el que proporciona menor velocidad de transmisión de información. En cambio, CS-4 es el que posee una mayor tasa de información pero también el más vulnerable frente a las hostilidades del canal físico. A continuación se explica el funcionamiento general de cada uno de estos esquemas de codificación de canal.

[17] Los cuatro esquemas de codificación de canal son de obligada implementación para MSs. Para la red, sólo CS–1.

[18] Si las condiciones de transmisión son buenas es necesaria menos información redundante destinada a la corrección de errores y se pueden transmitir más bits de información. Si las características son peores, parte de esta información es eliminada para poder introducir en su lugar más bits de redundancia, disminuyendo la tasa de información para aumentar así la fiabilidad frente a errores.

2.5.3.1.2.1.1.- ESQUEMA DE CODIFICACIÓN CS-1

El procedimiento empleado en el esquema de codificación CS-1 se representa de forma gráfica en la figura 2.10. El mensaje entregado al codificador es un bloque conformado por 181 bits de datos y los 3 bits que constituyen el campo USF. Sobre estos 184 bits se aplica una código Fire que proporciona como resultado 40 bits de información redundante, utilizada en el extremo receptor para la detección y corrección de errores (FEC). Estos 40 bits son acomodados dentro del campo BCS (*Block Check Sequence*). La adición de 4 bits de cola da como resultado una secuencia de 3 + 181 + 40 + 4 = 228 bits sobre los que se aplica un código convolucional de factor 1/2 . El resultado final es una secuencia de 228 × 2 = 456 bits codificados.

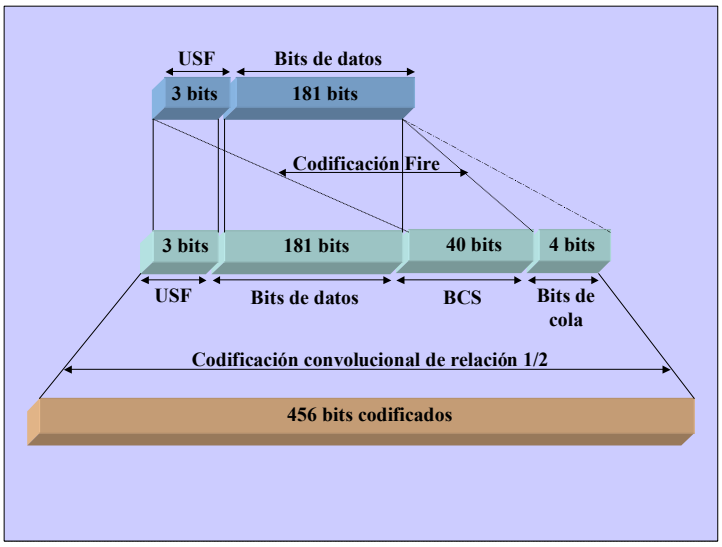


Figura 2.10.- Esquema de codificación CS-1 [9]

2.5.3.1.2.1.2.- ESQUEMA DE CODIFICACIÓN CS-2

El procedimiento empleado en el esquema de codificación CS-2 se representa de forma gráfica en la figura 2.11. A la entrada del codificador llegan los 3 bits del USF y una secuencia de 268 bits de datos. Sobre estos 271 bits se aplica un código cíclico que proporciona 16 bits (utilizados para detección y corrección de errores hacia delante) los cuales se insertan en el campo BCS. Además, los tres bits del USF son precodificados[19] convirtiéndose en secuencias de 6 bits. La concatenación de los 6 bits del USF precodificado, los 268 bits de datos, los 16 bits del BCS y la adición de 4 bits de cola proporciona una secuencia de 6 + 268 + 16 + 4 = 294 bits sobre los que se aplica el mismo código convolucional de factor 1/2 que en CS-1. El resultado es una secuencia de 294 × 2 = 588 bits, 12 bits procedentes del USF precodificado más 576 bits del resto. Sobre estos 576 bits se aplica un código perforado (para ajustar el tamaño final a los 456 bits de un bloque radio) que elimina un total de 132 bits, dejando 576 – 132 = 444 bits (la perforación no se aplica a los 12 bits del USF precodificado). Nuevamente se vuelven a obtener a la salida del codificador 12 + 444 = 456 bits codificados.

[19] La precodificación del USF se lleva a cabo por motivos de seguridad ante fallos de transmisión. La correcta recepción de este campo resulta crucial en el acceso *uplink* por parte de los terminales móviles, ya que si el campo USF recibido es erróneo, un móvil intentará transmitir cuando debe hacerlo otro diferente, provocando una colisión.

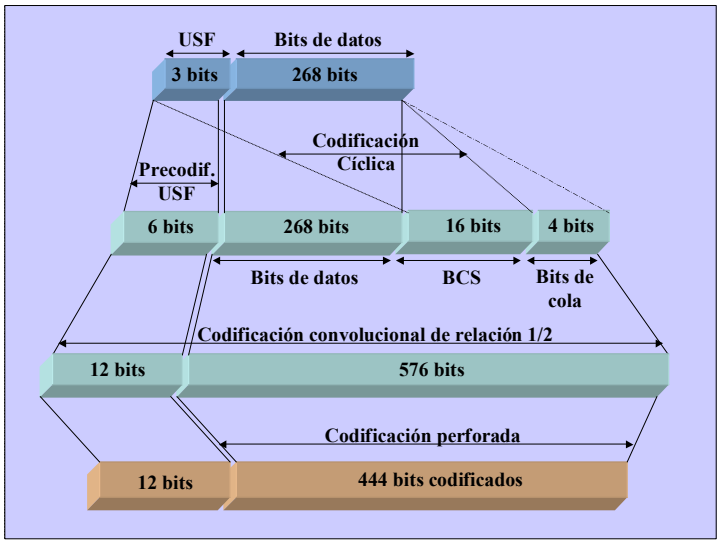


Figura 2.11.- Esquema de codificación CS-2 [9]

2.5.3.1.2.1.3.- ESQUEMA DE CODIFICACIÓN CS-3

El procedimiento empleado en el esquema de codificación CS-3 se representa de forma gráfica en la figura 2.12.

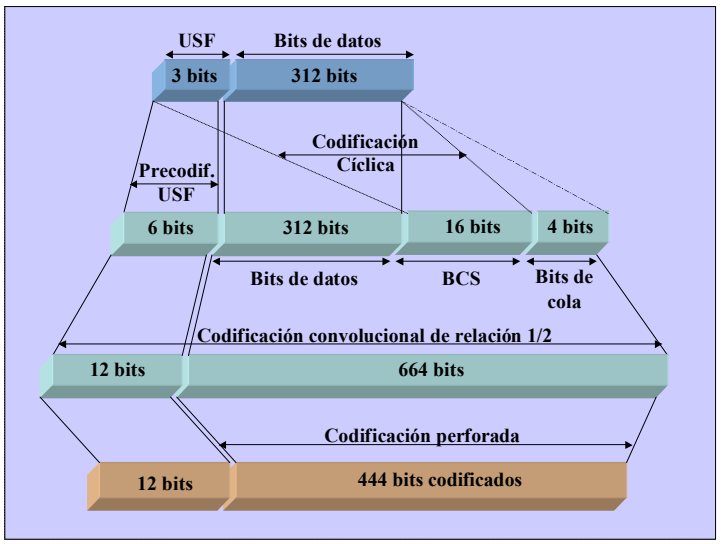


Figura 2.12.- Esquema de codificación CS-3 [9]

El procedimiento seguido en el esquema de codificación CS-3 es idéntico al seguido para CS-2, con dos salvedades. La primera es que a la entrada del codificador de canal el número de bits de datos no es 268 sino que ahora es de 312 bits. La segunda es que el código perforado no elimina 132 bits sino 220 bits para proporcionar, nuevamente, una secuencia de 456 bits codificados a la salida.

2.5.3.1.2.1.4.- ESQUEMA DE CODIFICACIÓN CS-4

El procedimiento empleado en el esquema de codificación CS-4 se representa de forma gráfica en la figura 2.13.

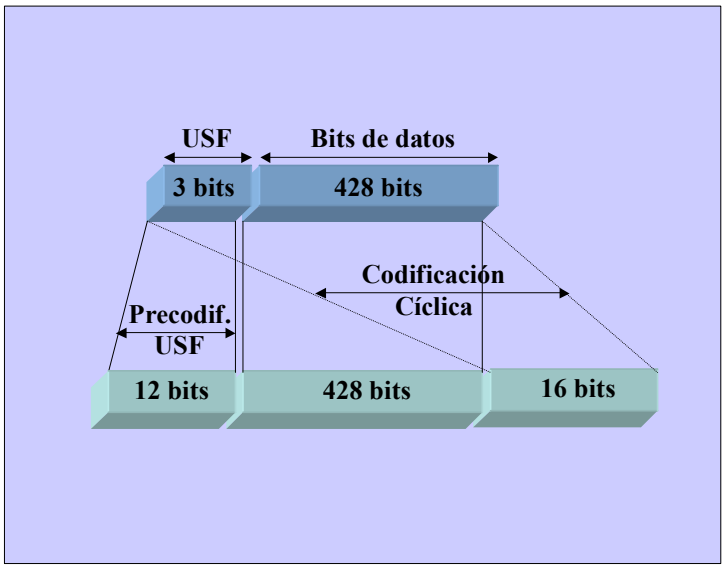


Figura 2.13.- Esquema de codificación CS-4 [9]

El codificador CS-4 recibe un bloque de entrada de 428 bits de datos más los 3 bits del USF, el cual es precodificado en una secuencia de 12 bits. Al bloque de entrada de 431 bits se le aplica un código cíclico que proporciona los 16 bits incluidos en el campo BCS, utilizados para detección y corrección de errores. La concatenación de los 12 bits del USF precodificado, los 428 bits de datos y los 16 bits del BCS proporcionan los 12 + 428 + 16 = 456 bits que se obtienen a la salida del codificador CS-4, el cual no aplica ningún tipo de código convolucional, haciéndolo muy vulnerable ante los errores de transmisión.

2.5.3.1.2.1.5.- COMPARACIÓN DE LOS ESQUEMAS CS-1 A CS-4

La tabla 2.1 muestra los parámetros que caracterizan el proceso de codificación de canal para cada uno de los diferentes esquemas de codificación.

Esquema	Factor	Bits de USF	Bits de USF precod.	Bits de datos	Bits para FEC	Bits de cola	Bits codific.	Bits perf.	Tasa (kbps)
CS-1	1/2	3	3	181	40	4	456	0	9.05
CS-2	≈ 2/3	3	6	268	16	4	588	132	13.4
CS-3	≈ 3/4	3	6	312	16	4	676	220	15.6
CS-4	1	3	12	428	16	0	456	0	21.4

Tabla 2.1.- Parámetros de los diversos esquemas de codificación GPRS para PDTCH/F

Las observaciones a realizar respecto a los diferentes esquemas son varias:

- El esquema CS-1 emplea un código Fire para obtener los bits del BCS mientras que CS-2 a CS-4 emplean un código cíclico CRC-16 [10].

- Los esquemas CS-2 y CS-3 son una versión perforada del mismo código convolucional empleado en CS-1 para FEC. El esquema CS-4 no aplica codificación para FEC [10].

- El factor del código se obtiene según la expresión [6]:

$$\frac{\text{Bits USF} + \text{Bits Datos} + \text{Bits FEC} + \text{Bits Cola}}{456}$$

- El número de bits perforados tras la codificación es tal que a la salida del codificador siempre se obtienen 456 bits, es decir,

Bits codificados – Bits perforados = 456

- La tasa binaria se obtiene como

Bits codificados / Duración de 1 bloque radio (20 ms)

- La salida de todos los codificadores siempre es de 456 bits, mientras que la cantidad de bits de datos tomados para realizar la codificación es diferente en cada caso. El resultado es que, al tomar menos bits de datos, disminuye la tasa de información a la salida del codificador, ya que de los 456 bits son menos los que contienen información, mientras que el resto de bits son de información redundante para obtener una mayor robustez frente a errores. Por tanto, se observa que existe un compromiso entre ambos factores.

2.5.3.1.2.2.- ENTRELAZADO

Tal y como se ha podido observar más arriba, la salida de todos los codificadores es una secuencia de 456 bits. Esta secuencia se puede dividir en cuatro fragmentos de 456 / 4 = 114 bits, cada uno de los cuales puede acomodarse en los dos campos DB de 57 bits de una ráfaga normal (ver apartado 2.5.1.2.3.1 y figura 2.8), por lo que cada secuencia de salida del codificador (debida a un bloque RLC) requiere 4 ráfagas normales para ser transmitida. Estas ráfagas son repartidas en 4 tramas TDMA consecutivas (ver figura 2.9).

Sin embargo, en el procedimiento real no se dividen los 456 bits en cuatro bloques continuos y se introducen directamente en cuatro ráfagas, sino que existe una operación de dispersión de estos bits de forma alternada entre ellas. Este proceso se denomina entrelazado [9], y la forma en la cual es llevado a cabo dicho reparto entre las cuatro ráfagas se muestra en la tabla 2.2.

Número de bit dentro de la secuencia de 456 bits	Posición dentro de la estructura de trama
0, 8, 16 ... 448	Bits pares de la ráfaga N
1, 9, 17 ... 449	Bits pares de la ráfaga N + 1
2, 10, 18 ... 450	Bits pares de la ráfaga N + 2
3, 11, 19 ... 451	Bits pares de la ráfaga N + 3
4, 12, 20 ... 452	Bits impares de la ráfaga N
5, 13, 21 ... 453	Bits impares de la ráfaga N + 1
6, 14, 22 ... 454	Bits impares de la ráfaga N + 2
7, 15, 23 ... 455	Bits impares de la ráfaga N + 3

Tabla 2.2.- Proceso de entrelazado [9]

La razón por la cual se realiza esta operación es para mejorar la corrección de errores. Las técnicas de corrección soportan un número bastante limitado de bits erróneos dentro de una determinada secuencia y pierden eficiencia conforme los errores están más próximos entre sí. Efectivamente, la capacidad de un código corrector para detectar y corregir errores es directamente proporcional a la distancia entre bits erróneos. Desafortunadamente, en el ámbito práctico, los errores no se producen en bits aislados, sino que se producen en agrupaciones de bits consecutivos, haciendo inútiles estas técnicas de corrección. El entrelazado representa la solución a este problema. Los bits, antes de ser transmitidos, son distribuidos en diferentes ráfagas de forma salteada. Puede suceder que una de estas ráfagas que contiene bits alternados se vea afectada por errores durante la transmisión, pero esto no supone un problema crítico debido a que, posteriormente, al reordenar los bits en el receptor, esta secuencia continua de errores se dispersa al conformar la secuencia original. El resultado es que, en la secuencia obtenida, los bits erróneos quedan distribuidos con cierta separación entre sí, es decir, los bits erróneos están relativamente alejados. Sobre esta secuencia es posible aplicar de forma más eficaz las técnicas de corrección, lo cual hubiera resultado inútil si no se hubiera empleado el entrelazado, ya que los bits erróneos serían consecutivos.

2.5.3.2.- LA CAPA RFL

La capa de radiofrecuencia o RFL (*Radio Frecuency Layer*) tiene como función primordial la modulación de una portadora según la secuencia de bits recibidos desde la capa PLL, su amplificación y su transmisión al canal de comunicación en forma de onda electromagnética. Esta capa proporciona la conexión física entre las estaciones móviles de usuario y las BTSs. El esquema de modulación empleado es GMSK (*Gaussian Minimum Shift Keying*) con un ancho de banda por periodo de bit BT = 0.3 [14]. Este tipo de modulación es mononivel, por lo que un símbolo de canal codifica un único bit de información. Una modulación mononivel proporciona una velocidad binaria menor que una multinivel pero también una mayor robustez e inmunidad frente a ruido e interferencias. La capa RFL queda determinada por la serie GSM 05 en la cual se especifican, entre otras cosas, las frecuencias portadoras y la estructura del canal radio de GSM (GSM 05.02 [12]); la modulación, demodulación y tasas de datos de los canales GSM (GSM 05.04 [14]) y los requisitos técnicos del transmisor y receptor (GSM 05.05 [15]).

2.6.- SUMARIO

En la introducción de este capítulo se ha explicado la necesidad de evolucionar GSM hacia la tercera generación, se han señalado los pasos establecidos para ello (HSCSD, GPRS y EDGE) y sus principales características así como el papel de GPRS dentro de este proceso. Tras explicar qué es exactamente GPRS se ha pasado a describir su arquitectura de red, basada sobre la arquitectura existente de GSM. Después de una descripción general de la torre de protocolos en el plano de transmisión, se ha centrado la atención sobre el ámbito de mayor interés para los fines aquí perseguidos: la interfaz aire, más concretamente, las capas RLC/MAC y física. Una rápida descripción de los canales de GPRS ha sido sucedida por la introducción de los aspectos básicos de la interfaz aire de GPRS: los nuevos conceptos en la gestión de recursos radio, sus modos de operación y la técnica de acceso múltiple.

El resto del capítulo se ha dedicado al estudio específico de las capas RLC/MAC y física, ya que éstas son las capas utilizadas y sobre las que se basa el presente trabajo. Para las capas RLC/MAC se han citado sus servicios y sus funciones, además de los modos de operación (con y sin confirmación). El capítulo concluye con la descripción de los servicios y funciones que conciernen a la capa más baja de la interfaz aire: la capa física. La explicación de este punto se ha centrado sobre la codificación de canal y el entrelazado.

En definitiva, este capítulo se ha dedicado a proporcionar una descripción de GPRS centrando la atención en aquellos aspectos que guardan una mayor relación con los objetivos de este trabajo y tratando aquellos puntos que resultan de mayor necesidad.

Capítulo 3

Gestión de Recursos Radio: Link Adaptation

3.1.- INTRODUCCIÓN

En el capítulo anterior se citó que, ante la incipiente influencia de Internet en todos los ámbitos, GPRS nace para posibilitar la prestación de nuevos servicios de datos en la red GSM de telefonía móvil tradicional. Debido al impacto de la navegación web, el correo electrónico o la multiconferencia, la provisión de servicios de Internet se ha convertido en el principal paradigma a la hora de concebir y diseñar los sistemas móviles de acceso inalámbrico de tercera generación y posteriores [16]. La demanda de este tipo de servicios junto al continuo incremento de los usuarios de telefonía móvil y la escasez de espectro radioeléctrico ha resaltado la necesidad, a medio y largo plazo, de mecanismos de gestión de recursos cada vez más y más eficaces [16]. El problema se ve agravado por los exigentes requisitos de los servicios multimedia de banda ancha. Para los sistemas 3G se prevén tasas de 144 kbps (vehículos), 384 kbps (peatones) en entornos exteriores y 2 Mbps en interiores [17] (mientras que velocidades de 10 Mbps son habituales en redes de área local); se han llegado a plantear velocidades de 100 Mbps o más para los futuros sistemas de cuarta generación [18] e incluso de hasta 155 Mbps para acceso a redes ATM [19]. Además de necesitar un elevado ancho de banda, los servicios multimedia se caracterizan por diferentes requisitos de calidad de servicio en términos de retardo así como por un tráfico a ráfagas, mixto[20] y asimétrico[21] [16].

Tan comprometida es la situación que se ha llegado al punto de cuestionar la viabilidad de que una red móvil pueda afrontar este reto dados los recursos disponibles [20]. Efectivamente, los recursos radio se hallan claramente limitados mientras que no es tan obvio determinar el número máximo de usuarios que puede soportar un sistema móvil ni existe una delimitación clara de la capacidad máxima sostenible, así como tampoco existe unanimidad acerca de qué parámetros pueden resultar útiles y fiables para que pueda ser evaluada con aceptable precisión [21].

Por todo ello, resulta indiscutible la necesidad de una eficaz gestión de recursos radio (RRM, *Radio Resource Management*). No obstante, si bien es cierto que las técnicas RRM han adquirido recientemente gran relevancia, igualmente cierto es el hecho de que han jugado un papel importante desde la propia existencia de las comunicaciones móviles debido a la problemática inherente a dichos sistemas. El rasgo más distintivo es, sin lugar a dudas, el carácter variante del canal móvil. El canal móvil es en esencia un medio cambiante, sujeto a numerosos procesos aleatorios que le otorgan unas características peculiares y que causan sobre la señal electromagnética varios tipos de alteraciones que no pueden ser ignoradas en el estudio de la gestión de recursos radio y que, por tanto, serán tratadas en el presente capítulo.

Así pues, y mostrada la relevancia de las técnicas RRM, en este capítulo se presenta una explicación de la gestión de recursos radio, haciendo especial hincapié en *Link Adaptation*, tema principal de este trabajo. La descripción de técnicas RRM supone el manejo de los conceptos básicos de la estructura celular en que se basan los sistemas de comunicaciones móviles. Por ello, este capítulo se inicia con un repaso de los fundamentos celulares.

---

[20] Tráfico generado por aplicaciones de diversa naturaleza: voz, web, correo electrónico, vídeo, música, etc.

[21] Tráfico en el que los flujos de información en sentido red a móvil y móvil a red son de diferente magnitud. Habitualmente el flujo de información en sentido red a móvil es considerablemente mayor que en sentido opuesto.

3.2.- PRINCIPIOS DE LA ESTRUCTURA CELULAR

En los sistemas de telefonía móvil celular [22], la zona de cobertura de un operador se divide en zonas más pequeñas, denominadas *celdas* o *células*, que constituyen la unidad básica de cobertura. En cada una de estas células se coloca una estación base de potencia limitada capaz de comunicarse con los terminales móviles que hay dentro de ella. Si la antena de cada estación base es omnidireccional, entonces la forma geométrica de la célula será aproximadamente circular. No obstante, una célula con esta forma supone la existencia de zonas sin cobertura o de zonas donde dos o más células se solapan. En consecuencia, desde el punto de vista teórico, los sistemas celulares se estudian considerando células hexagonales, ya que permiten cubrir por completo todo el plano de cobertura sin solapamientos y, además, de entre los polígonos que cumplen esta condición (junto con el cuadrado y el triángulo), el hexágono es el que posee una mayor relación perímetro-área, lo cual significa que es capaz de cubrir una misma región con menos células. La figura 3.1 muestra el principio de división celular y su representación mediante células hexagonales.

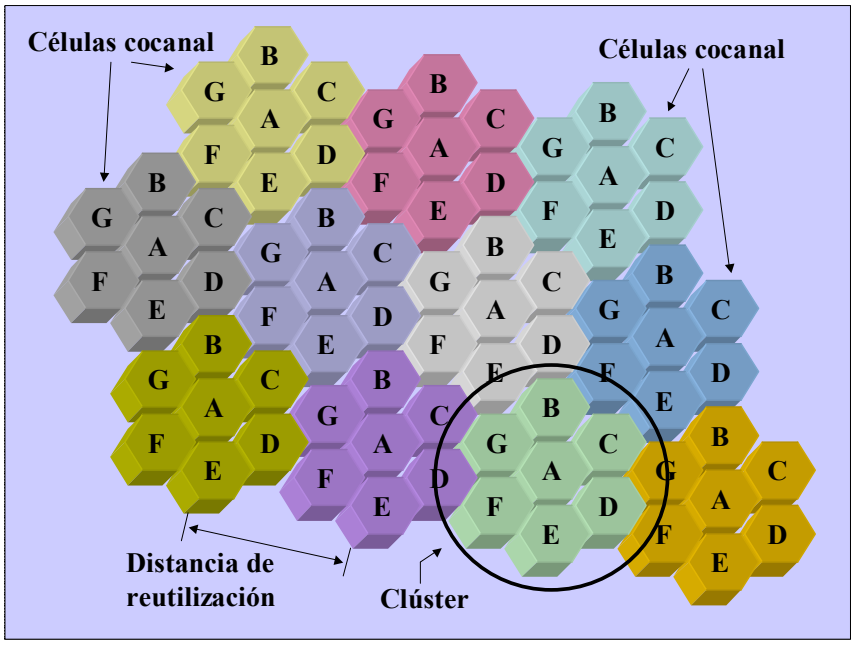


Figura 3.1.- Estructura celular con un tamaño de clúster $K = 7$

El operador asigna una pequeña porción del espectro a una célula y una porción diferente a las células vecinas, de forma que todo el espectro disponible se reparte entre un número muy pequeño de K células. Este conjunto de células se denomina *clúster* o *racimo* de tamaño K y se repite de forma periódica a lo largo de toda la zona de cobertura. El tamaño K del clúster cumple la expresión:

$$K = i^2 + ij + j^2 \tag{3.1}$$

donde i y j son enteros, lo cual da lugar a que sólo ciertos tamaños de clúster sean válidos (K = 1, 3, 4, 7, 9, 12, ...). En la figura 3.1 el tamaño del clúster es $K = 7$.

Esta estructura repetitiva es necesaria (porque resulta imposible asignar una frecuencia diferente a cada célula) e introduce el concepto de *reutilización de frecuencias*, ya que una misma frecuencia se emplea en varias células separadas entre sí por una distancia denominada *distancia de reutilización* (ver figura 3.1), que viene dada por la expresión

$$D = R\sqrt{3K} \tag{3.2}$$

donde R es el radio de la célula y la relación D/R se denomina *factor de reutilización*. De esta forma la estructura celular permite cubrir amplias zonas con un espectro limitado, ya que los recursos radioeléctricos se reutilizan varias veces en zonas alejadas, resultando en un aprovechamiento más eficiente de los recursos y en una mayor capacidad del sistema.

Pero en el esquema celular no todo son ventajas ya que, si bien es cierto que dentro de un mismo clúster las células no interfieren entre sí porque utilizan frecuencias diferentes, no es menos cierto el hecho de que dos células en clústeres adyacentes que utilizan la misma frecuencia solapan sus espectros, experimentando fenómenos de interferencia. Estas células se denominan *células cocanal* (ver figura 3.1) y el fenómeno de interferencia recibe la denominación de *interferencia cocanal*. A diferencia del ruido térmico, cuyo efecto puede paliarse incrementando la potencia de transmisión para mejorar la relación CIR (*Carrier-to-Interference Ratio*, relación portadora a interferencia), la interferencia cocanal no desaparece de esta forma sino que, por el contrario, se acentúa al aumentar la potencia. La solución radica en una buena planificación celular donde la distancia de reutilización cocanal D sea lo suficientemente elevada como para que este efecto sea despreciable.

Además, dos células que utilizan frecuencias consecutivas en el espectro pueden interferir entre sí debido a la selectividad limitada de los filtros reales, dando lugar a fenómenos de *interferencia por canal adyacente*. Nuevamente, la solución pasa por una buena planificación celular en la que los canales contiguos no sean asignados a células adyacentes sino, preferiblemente, a células lo más separadas posible dentro del mismo clúster.

Por definición, los sistemas celulares son sistemas limitados por la interferencia y, aunque esta es una situación normal y controlada, impone restricciones al sistema. Un factor de reutilización pequeño implica un tamaño de célula elevado y permite aumentar la potencia de transmisión, consiguiendo una mejor relación CIR y, consecuentemente, una mejora en la calidad de transmisión. Sin embargo, un tamaño elevado de célula supone que las frecuencias sólo pueden reutilizarse a distancias mayores, con lo que la capacidad total del sistema disminuye. Un aumento de capacidad, por tanto, se consigue diminuyendo el tamaño de las células y aumentando así el factor de reutilización, pero en este caso la potencia ha de reducirse para controlar la interferencia cocanal (mientras que el ruido total permanece constante) con lo que se empeora la relación CIR. Así pues, puede observarse que en el sistema celular existe un compromiso entre la capacidad de usuarios y la calidad que se puede ofrecer a cada uno de ellos.

El principio celular permite una adaptación flexible al incremento en el número de usuarios gracias a la *fragmentación celular* (figura 3.2) al mismo tiempo que la *sectorización* (figura 3.3) reduce la interferencia cocanal del sistema.

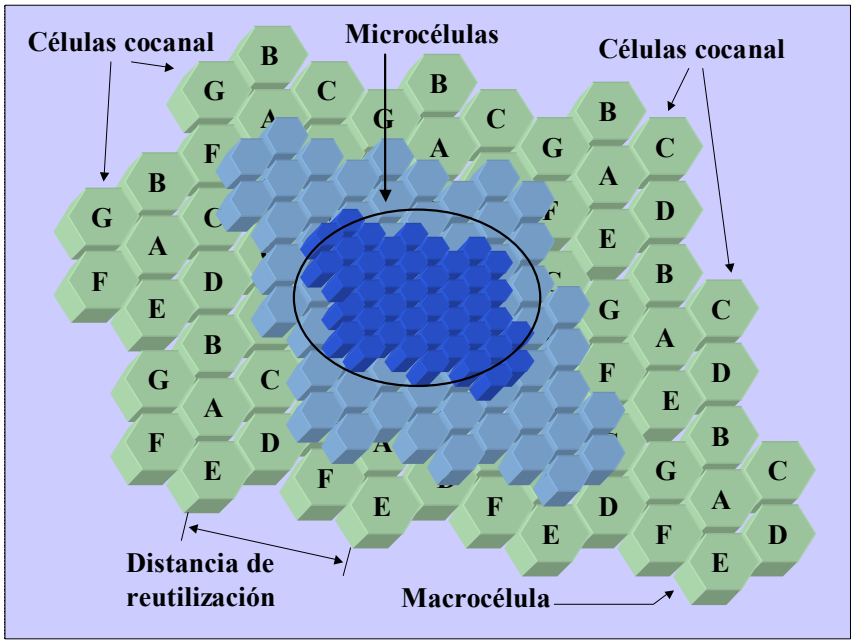


Figura 3.2.- Principio de fragmentación celular

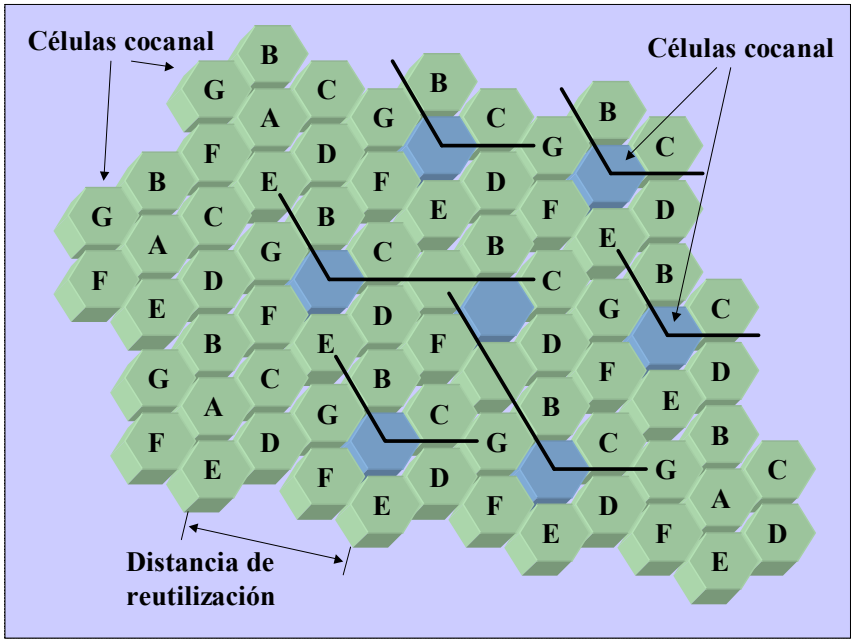


Figura 3.3.- Principio de sectorización

La fragmentación celular (figura 3.2) consiste en dividir una célula en otras de tamaño más reducido. Con ello se consigue que una misma región esté cubierta por más células y, por tanto, aumente la reutilización de frecuencias dentro de la zona consiguiendo una mayor capacidad de usuarios. La aplicación de la fragmentación celular requiere el uso de más estaciones base, lo cual supone un incremento en el coste. Por ello, el proceso de fragmentación se lleva a cabo de forma progresiva allí donde se requiere, dando lugar a *microcélulas* o células de tamaño reducido en zonas urbanas (con alta densidad de usuarios) y a *macrocélulas* o células extensas en zonas rurales (con baja densidad de usuarios).

El principio de sectorización (figura 3.3) consiste en sustituir una antena omnidireccional colocada en el centro de la célula por un conjunto de antenas directivas colocadas en el centro o en los vértices del hexágono, cada una de las cuales cubre sólo un sector de la célula. En cada sector se emplea una única antena (el coste se multiplica) y una única frecuencia (la capacidad se multiplica también). La sectorización suele realizarse empleando tres antenas con un ancho de haz de 120^o (sectores rómbicos) o seis antenas con un ancho de haz de 60^o (sectores triangulares). La figura 3.3 muestra un ejemplo de sectorización 120^o, donde aparecen señaladas siete células cocanal.

Una ventaja de la sectorización se puede apreciar considerando la célula central. Si las antenas fueran omnidireccionales, entonces las seis células restantes producirían interferencia cocanal sobre la central, mientras que con la sectorización se puede apreciar que ahora son sólo dos las que interfieren. Vemos, pues, que la sectorización ha reducido la interferencia cocanal: la sectorización en N partes reduce la interferencia en un factor $1/N$ y aumenta la capacidad N veces, pero multiplicando el coste por N.

3.3.- RADIOPROPAGACIÓN EN EL CANAL MÓVIL

Una de las principales singularidades que caracteriza a las comunicaciones móviles es el carácter cambiante del canal de transmisión con el tiempo y con el espacio. Durante la propagación de la señal por el medio de transmisión, ésta está sujeta a procesos de reflexión, difracción y dispersión (*scattering*) [23]. Los diferentes modos en que estos procesos interactúan entre sí y con agentes externos al medio radioeléctrico dan lugar a diversos fenómenos de perturbación sobre la señal: *atenuación*, *efecto sombra* y *efecto multicamino*.

Las señales electromagnéticas que se propagan a través del medio radioeléctrico sufren un proceso de atenuación debido a que el medio físico está caracterizado por un factor de pérdidas que origina una disminución en la potencia de la onda a medida que ésta se propaga. Así pues, la atenuación o *path loss* se refiere a las pérdidas entre transmisor y receptor introducidas únicamente por el medio físico de transmisión propiamente dicho. El modelo de propagación en espacio libre, según el cual la atenuación es directamente proporcional al cuadrado de la distancia, no resulta válido en este caso ya que las pérdidas responden a una curva logarítmica. Un modelo matemático ampliamente aceptado en su lugar es el modelo empírico Okumura-Hata [23] (se verá en el apartado 5.3.1.2).

En ocasiones, entre transmisor y receptor no existe visibilidad directa sino que suele ser frecuente la presencia de obstáculos interpuestos entre ambos, los cuales ejercen un efecto de disminución en la potencia de la señal electromagnética. Por tanto, además de la atenuación hay que considerar las pérdidas introducidas por estos obstáculos, fenómeno conocido como efecto sombra o *shadowing*. En ocasiones también se conoce como desvanecimiento lento o *slow fading*. En este caso las pérdidas L_S responden a una distribución de probabilidad normal o gaussiana de media L' igual a cero y desviación estándar σ_S, de manera que su función de densidad de probabilidad viene dada por:

$$f_S(L_S) = \frac{1}{\sqrt{2\pi\sigma_S^2}} e^{-\frac{1}{2\sigma_S^2}(L_S - L')^2} \tag{3.3}$$

Finalmente, el efecto multicamino o *multipath fading* se produce cuando al receptor llegan múltiples muestras de una misma señal, cada una de las cuales ha recorrido un camino eléctrico diferente. Ello es debido a que la directividad del transmisor es finita, emitiendo la misma señal en varias direcciones. Puede suceder que varios de estos rayos, que sufren una serie de procesos de reflexión, difracción y dispersión en diferente medida, coincidan en el receptor, donde se suman de forma constructiva o destructiva dependiendo de la fase relativa de cada uno. Dado el carácter móvil del receptor y la cantidad de eventos impredecibles que pueden acaecer, el valor instantáneo de la señal recibida es la suma vectorial de un gran número de componentes no deterministas, por lo que se comporta como una variable aleatoria que experimenta rápidas fluctuaciones. Por ello, este efecto es también conocido como desvanecimiento rápido o *fast fading*.

El desvanecimiento debido al efecto multicamino puede ser *plano* (afecta por igual a todas las componentes de frecuencia de la señal) o *selectivo* (afecta a un grupo concreto de componentes de frecuencia). Tendrá lugar uno u otro tipo dependiendo del ancho de banda de la señal a transmitir y del *ancho de banda de coherencia* del canal. Si el primero es menor que el segundo entonces el desvanecimiento será plano y todas las frecuencias experimentarán la misma perturbación, mientras que en caso contrario (el ancho de banda de la señal es mayor que el ancho de banda de coherencia del canal) el desvanecimiento será selectivo y diferentes frecuencias sufrirán alteraciones diferentes.

En este último caso, para trayectos de varios centenares de longitudes de onda (*long-term fading*) la intensidad del campo eléctrico recibido sigue una distribución gaussiana (3.3) con media L' distinta de cero. Para trayectos de algunas decenas de longitudes de onda (*short-term fading*) son habituales los modelos Rayleigh (cuando predomina una componente sobre las demás, el rayo directo en este caso) y Rice (cuando todas las componentes tienen amplitud instantánea similar) [23].

De forma análoga al dominio de la frecuencia, en el dominio temporal se define el *tiempo de coherencia*, que hace referencia al tiempo durante el cual los parámetros que caracterizan el canal de transmisión permanecen constantes. Los pequeños intervalos de tiempo donde el canal varía de forma abrupta, produciendo un desvanecimiento en la señal, dan lugar a un ensanchamiento en el dominio de la frecuencia conocido como *ensanchamiento Doppler*.

3.4.- GESTIÓN DE RECURSOS RADIO

3.4.1.- FORMULACIÓN DEL PROBLEMA

En la introducción de este capítulo quedó clara la principal problemática de los sistemas móviles del futuro y se puso de relieve la necesidad de un conjunto de técnicas para la eficiente gestión de los recursos radio (en adelante RRM, *Radio Resource Management*). Aunque las ideas clave que definen el objetivo perseguido en la RRM se citaron en dicha introducción, el problema requiere de una formulación precisa.

Por una parte, la popularidad de la telefonía móvil aumenta de forma continua y, con ello, el número de usuarios de este servicio, lo cual impone una necesidad de recursos cada vez mayor. Además, y de forma paralela, aumenta la demanda de servicios multimedia, que requieren en sí una mayor cantidad de recursos por usuario y exigen la oferta de diversas calidades de servicio en función de las necesidades del consumidor. Por el contrario, la reserva actual de espectro destinada a los sistemas móviles se presenta como insuficiente ante las previsiones de futuro. Por tanto,

> *"El objetivo fundamental que las técnicas de gestión de recursos radio persiguen es dar servicio al mayor número posible de usuarios, ofreciendo al mayor porcentaje posible la calidad de servicio exigida (idealmente al 100%) y empleando para ello la cantidad de recursos que resulte imprescindible"*

Para lograr este objetivo es necesario tomar decisiones acerca de qué acciones sobre el sistema son las oportunas en cada momento. En [16] y [21] el problema de la RRM consiste básicamente en decidir:

a) ¿Qué canal y forma de onda asignar para cada usuario[22]?
b) ¿Qué estación base servirá como punto de acceso a la red?
c) ¿Con qué nivel de potencia debe transmitir cada usuario?

No obstante, puede resultar más conveniente realizar una consideración más desglosada, ya que alguno de los puntos anteriores engloba otros aspectos más específicos. Concretamente, dentro de a) se puede incluir el mecanismo de control de admisión (apartado 3.4.3.1), el cual resulta necesario ya que en el ámbito práctico un sistema puede alcanzar su límite de capacidad y se han de habilitar mecanismos que rechacen la conexión en lugar de realizar una asignación de canal; además, una vez aceptado el usuario, hay que determinar con qué prioridad se le asignan los canales (apartado 3.4.3.2), en qué cantidad (apartado 3.4.3.3) y cómo (apartado 3.4.3.4 y 3.4.3.5).

Siguiendo esta formulación, en el apartado 3.4.3 se presentan diversas técnicas RRM. La técnica de *Link Adaptation*, dada la relevancia que tiene para este trabajo, se ha pospuesto para una sección aparte (apartado 3.5) donde se trata con más detalle. No hay que olvidar que todas ellas persiguen un objetivo común, el enunciado más arriba, si bien la forma en que se intenta alcanzar en cada caso es diferente.

[22] Se refiere a la selección para cada usuario de una señal lo más incorrelada posible con el resto de señales (señales ortogonales) para minimizar los efectos de interferencia.

3.4.2.- PRINCIPIOS DE GESTIÓN

Los primeros sistemas celulares FDMA basaban su RRM en principios que, si bien no pueden ser tildados de estáticos, sí es cierto que eran muy poco dinámicos y adaptables [21], ya que los recursos eran asignados de forma más o menos permanente. Aunque estos métodos ofrecen buenos resultados en sistemas de carácter macrocelular con tráfico elevado y homogéneo, su funcionamiento, en cambio, deja mucho que desear en sistemas microcelulares, donde cada estación base soporta un tráfico menor y éste tiende a variar de forma mucho más abrupta, especialmente si se trata de tráfico multimedia [21].

Una mejor solución a las técnicas RRM estáticas lo constituyen las técnicas adaptables, las cuales se pueden clasificar en técnicas de asignación dinámica de recursos o DRA (*Dynamic Resource Allocation*) y asignación aleatoria de recursos o RRA (*Random Resource Allocation*) [16].

Las técnicas DRA son aquellas que actúan en respuesta a las diversas solicitudes o acciones realizadas por los terminales móviles y en función de cuáles sean las condiciones particulares de cada situación concreta. Un ejemplo de DRA se encuentra en el capítulo 2 (apartado 2.5.1.2.1.2), donde se explicaba el reparto de recursos bajo demanda entre GSM y GPRS en respuesta a las diversas solicitudes y según los recursos disponibles en cada instante. Otros ejemplos se hallan en los apartados 3.4.3.1, 3.4.3.2, 3.4.3.3, 3.4.3.4, 3.4.3.6, 3.4.3.7 y 3.4.3.8. Las técnicas RRA, en cambio, varían la asignación de recursos siguiendo una secuencia pseudo-aleatoria preestablecida. Un ejemplo de RRA se explica en el apartado 3.4.3.5. Este principio, por ejemplo, encuentra gran aplicación en los sistemas CDMA, donde cada grupo de bits se transmite utilizando formas de onda pseudoaleatorias o *chips* diferentes.

La comparación de los beneficios relativos de cada principio es algo que no resulta inmediato. El inconveniente de las técnicas DRA es que requieren una elevada señalización, un notable consumo de potencia así como un muestreo periódico de las condiciones del canal. Su eficacia está bastante condicionada por dicha periodicidad y la calidad de la estimación. Por otra parte, muestran buen comportamiento cuando los parámetros del tráfico son bien conocidos. Cuando dicha información es limitada resulta difícil obtener conclusiones. Las técnicas RRA, en cambio, se caracterizan por una gran simplicidad de implementación pero requieren un mayor ancho de banda. En general, bajo pobres condiciones de CIR (macrocélulas) las técnicas RRA alcanzan un throughput mayor, mientras que en condiciones más favorables (microcélulas) la tasa binaria de información es superior con técnicas DRA [16].

A continuación se explican diversas técnicas de gestión de recursos radio. Salvo la explicada en el apartado 3.4.3.5 (salto en frecuencia), el resto de técnicas del apartado 3.4.3 se cataloga dentro de las basadas en el principio de asignación dinámica de recursos o DRA. *Link Adaptation* también pertenece a este conjunto.

3.4.3.- TÉCNICAS DE GESTIÓN

Son muchas las técnicas de gestión de recursos radio existentes. Sin embargo, en esta sección se presentarán sólo las más importantes, que han demostrado ser eficientes y que actualmente se encuentran implementadas en sistemas móviles reales.

3.4.3.1.- CONTROL DE ADMISIÓN

El procedimiento de control de admisión o CAC (*Call Admission Control*) se encarga de determinar si una solicitud para el establecimiento de una nueva sesión de comunicación es aceptada o, en cambio, es rechazada [21]. Esta solicitud puede deberse tanto a una solicitud de nueva conexión como a un *handover*[23] y varios son los motivos por los cuales puede ser rechazada. El primero y más obvio es la carencia de recursos suficientes como para poder atender la petición con la calidad de servicio solicitada. No obstante, aun existiendo recursos disponibles, el CAC puede decidir no aceptar una nueva solicitud ya que en ocasiones se prefiere mantener una reserva de recursos para posibles *handover*s en lugar de aceptar nuevas conexiones. Esto es así porque que el usuario tiene en peor consideración la caída de una llamada establecida que la indisponibilidad temporal para establecer una nueva. Estas reservas suelen hacerse mediante funciones de seguimiento (*shadow cluster*) de los terminales móviles a lo largo de diversas células y basándose en diversas predicciones sobre movilidad.

La inclusión de una nueva conexión en el sistema también puede decidirse en función de si dicha incorporación supone un excesivo nivel de interferencia al resto de móviles, causando una disminución generalizada de la relación CIR y, por tanto, perjudicando al resto de usuarios, los cuales experimentan un empeoramiento en la calidad. En este caso, el algoritmo CAC puede clasificarse como interactivo o no interactivo [24]. En el primer caso, el CAC reajusta la potencia del terminal (interacción con el control de potencia, apartado 3.4.3.7) hasta que cumpla los requisitos de CIR para ser aceptado, mientras que en el segundo caso el nivel de potencia de transmisión es comparado con un umbral y, si es sobrepasado, la conexión simplemente se rechaza.

3.4.3.2.- CONTROL DE LA CONGESTIÓN

En el ámbito práctico se dan situaciones en las que el total de recursos disponibles en un instante dado resulta insuficiente para atender de forma simultánea a todas las peticiones de transmisión realizadas. Por ello, es necesaria la existencia de algoritmos que controlen el orden en el cual tiene lugar la transferencia de los datos procedentes de diversos usuarios. Dicha función es llevada a cabo por los métodos de control de la congestión o *scheduling*, los cuales determinan la prioridad de cada transmisión en función de sus requisitos de calidad de servicio [25].

En [26] se realiza un estudio comparativo de diferentes técnicas de *scheduling* utilizadas en GPRS mientras que en [27] se plantean algunas técnicas ya evaluadas sobre redes fijas y se examina su posible aplicación a GPRS.

[23] Paso de una célula a otra con una comunicación en curso.

3.4.3.3.- OPERACIÓN MULTISLOT

Como se vio en el capítulo segundo, en GPRS es posible la asignación de más de un time-slot a una misma estación móvil[24]. Ello permite variar el ancho de banda asignado de forma dinámica, incrementando el throughput neto experimentado por el usuario y reduciendo así el retardo de transmisión. Sin embargo, tal y como se discute en [54], la utilización de dos o más time-slots para una misma transmisión también supone el incremento de la ocupación media de time-slots, lo cual tiene dos efectos: un aumento en el tiempo requerido para obtener time-slots libres para un nuevo usuario y un aumento de la interferencia causada al resto de usuarios. Por ello, la asignación multislot se decide de forma dinámica en función de la disponibilidad de recursos y de la carga del sistema, así como de las capacidades de operación multislot del terminal móvil de usuario [25].

3.4.3.4.- ASIGNACIÓN DE CANALES

En el apartado 3.2 se explicó que la totalidad de recursos disponibles por un operador se asigna a las células que integran un clúster. Dependiendo de cómo se realice la asignación de canales[25], se distinguen diferentes esquemas: asignación fija de canales o FCA (*Fixed Channel Allocation*), asignación dinámica de canales o DCA (*Dynamic Channel Allocation*) y asignación híbrida de canales o HCA (*Hybrid Channel Allocation*) [25].

En FCA, la asignación de canales a cada célula es fija, es decir, la totalidad de canales disponibles por el operador se divide en tantos grupos (no necesariamente del mismo tamaño) como células existen en un clúster y cada grupo es asignado de forma permanente a una célula. Este método es sencillo y proporciona buenos resultados en ambientes de carácter macrocelular con tráfico homogéneamente distribuido. No obstante, resulta insuficiente cuando las condiciones del tráfico sufren variaciones abruptas, como es el caso de zonas urbanas.

En DCA, no existe una reserva permanente de canales para cada célula como en FCA, sino que cada canal está disponible para cualquiera de las células del clúster y la asignación se realiza de forma dinámica, a medida que se solicitan nuevas conexiones. Así pues, si una célula experimenta una elevada carga en un momento dado dispondrá de más canales conforme la demanda aumenta y estos canales no estarán disponibles para el resto de células. Este rasgo permite una gran capacidad de adaptación a variaciones rápidas de tráfico, algo imposible con FCA. Sin embargo, resulta más complejo de implementar, ya que para realizar una asignación es necesaria una elevada señalización, un gran número de medidas de la calidad del canal así como considerar los efectos de dicha asignación sobre la interferencia entre células vecinas.

---

[24] En condiciones óptimas (uso del esquema de codificación CS-4) se puede alcanzar un máximo teórico de hasta 171.2 kbps agrupando los ocho time-slots de una trama para un único usuario.

[25] Recordar que en GSM (y en GPRS) un canal está determinado por una frecuencia portadora (FDMA) y por un intervalo temporal (TDMA). Ver capítulo 2, apartado 2.5.1.2.3.

Efectivamente, puede suceder que dos células muy cargadas, perteneciendo a clústeres diferentes, sean adyacentes. En estos casos hay que controlar que no se asignen canales de la misma frecuencia (cocanales) o de frecuencias cercanas (canales adyacentes) a estas células. Generalmente, de entre los canales disponibles en cada célula, se elige aquel que proporciona una mejor relación CIR para el usuario.

Finalmente, HCA consiste en una mezcla de FCA y DCA, donde unos canales son asignados de forma permanente a determinadas células (carácter FCA) y el resto se asigna de forma dinámica a las células del clúster (carácter DCA).

3.4.3.5.- SALTO EN FRECUENCIA

El objetivo que persigue el salto en frecuencia o *Frecuency Hopping* es la minimización del efecto de las interferencias sobre una comunicación concreta. El principio de funcionamiento consiste en variar, de forma pseudoaleatoria o cíclica, la frecuencia por la cual cada usuario transmite, de forma que siempre se realiza la transmisión en time-slots con la misma posición relativa dentro de la trama pero saltando de una frecuencia a otra y de ahí su nombre. La ventaja de este método radica en que el efecto de una frecuencia interferente puede reducirse hasta el punto de ser despreciable. En efecto, si existe una frecuencia interferente que coincide con una portadora concreta y cada usuario transmite de forma permanente por una misma frecuencia portadora, entonces aquel usuario al que le corresponda ésta experimentará una peor calidad en la comunicación. En cambio, gracias al salto en frecuencia, la interferencia afectará a más usuarios, pero sólo durante un pequeño intervalo de tiempo, aquél en el que transmitan por la portadora que sufre la interferencia. En tráfico de voz este hecho puede suponer una distorsión leve o incluso inapreciable por el oído, si bien en tráfico de datos puede suponer un problema más relevante. Una forma de ver el salto en frecuencia [28] es considerando una asignación estática de canales donde los usuarios de una misma estación base los comparten entre sí.

3.4.3.6.- SELECCIÓN DE ESTACIÓN BASE

Los mecanismos de selección de estación base, como su propio nombre indica, determinan el criterio a seguir para establecer a qué estación base debe conectarse cada estación móvil.

En los primeros sistemas, el proceso de *handover* era iniciado al detectar una señal muy débil por parte de la estación móvil. Cuando el nivel de potencia recibida disminuía por debajo de un cierto umbral, entonces el móvil se conectaba a una estación base vecina donde la señal se recibía con niveles más altos. En este caso, el parámetro fundamental que determinaba la selección de estación base era el umbral de potencia mínima necesaria para mantener una comunicación [21]. Lo que se buscaba era conseguir que la mayoría de los *handovers* se realizaran con éxito (no cayera ninguna de las comunicaciones establecidas) y, por otra parte, que la señalización que este proceso conlleva fuera mínima. Tal y como se explica en [29, 30], existe un compromiso entre ambos factores y un aumento o disminución del valor umbral beneficia a uno pero perjudica al otro. La aplicación de histéresis es una alternativa a considerar en estos casos.

En los sistemas actuales, la recepción de una señal débil no constituye un problema ya que el área de cobertura de las estaciones base se solapa en los bordes y una misma estación móvil suele recibir una señal aceptable por parte de más de una estación base. En estos casos el *handover* suele decidirse, no en función del nivel de señal recibido, sino en función de la relación CIR [21]. En aquellas situaciones en las que la relación CIR sea aproximadamente similar en ambas estaciones base, entonces la decisión se toma en función de la carga de cada una de ellas [21]. En ocasiones puede incluso asignarse una estación base con peores condiciones de CIR si la reducción de la carga de usuarios en una celda determinada supone en ese momento una necesidad de fuerza mayor (interacción con CAC, apartado 3.4.3.1).

3.4.3.7.- CONTROL DE POTENCIA

La función de control de potencia o APC (*Adaptive Power Control*) de una estación base se encarga de controlar, en todo momento, cuál ha de ser el nivel de potencia de emisión de cada una de las estaciones móviles que se encuentran bajo su área de cobertura. Lo que se pretende es realizar la transmisión empleando la cantidad de potencia imprescindible que permita lograr una relación CIR aceptable, con la finalidad de alargar la vida de las baterías de las estaciones móviles y, especialmente, reducir al máximo la interferencia cocanal. Hay que señalar en este punto que el aumento de potencia de transmisión por parte de una estación móvil para mejorar la relación CIR no resulta una buena estrategia, ya que el resto de estaciones móviles incrementarán de forma paralela su potencia de transmisión para lograr el mismo objetivo. El resultado final es que todos los móviles transmiten con la máxima potencia, aumentando el consumo de la batería y también la interferencia en el sistema.

Algunos aspectos dignos de ser mencionados son, entre otros, que se puede alcanzar una mayor capacidad en un sistema utilizando (de forma inteligente) APC, tal y como se demuestra en [31], pero en ocasiones puede ser preferible no implementarlo. Un ejemplo de este último caso [21] lo constituye la transmisión de paquetes de pequeño tamaño, donde el uso de APC impone un retardo temporal superior al tiempo necesario para la propia transferencia del paquete[26].

3.4.3.8.- REDUNDANCIA INCREMENTAL

La idea en que se basa el principio de redundancia incremental o IR (*Incremental Redundancy*) es en la adición de más o menos información redundante (para corrección de errores) dentro de cada bloque transmitido basándose en el método de *ensayo y error*.

[26] La técnica APC, como cualquier otra técnica DRA, necesita una medición periódica del canal, en este caso de la potencia de señal recibida, lo cual requiere una determinada cantidad de tiempo.

Inicialmente se transmite un bloque sin ningún tipo de protección, esto es, sin redundancia alguna. Si el receptor indica mediante un mensaje de confirmación (ver apartado 2.5.2.4) que el bloque ha sido recibido con error, entonces el emisor añade más información redundante para que la transmisión sea más fiable y robusta. De esta forma se consigue una adaptación a las condiciones del canal. Una de las ventajas de IR sobre *Link Adaptation*, tal y como se podrá comprobar más adelante (una vez explicado el funcionamiento de *Link Adaptation* en el apartado 3.5), es que no requiere ningún tipo de medición directa sobre el canal, si bien supone un mayor coste derivado de la necesidad de más memoria para almacenar los datos transmitidos hasta que el paquete ha sido confirmado correctamente [25].

3.4.4.- INTEGRACIÓN DE DIVERSAS TÉCNICAS

Las diversas técnicas de gestión de recursos radio anteriormente explicadas no son entidades autónomas que funcionan de forma aislada e independiente entre sí, sino que en múltiples ocasiones actúan de forma coordinada entre sí para lograr un mejor rendimiento general del sistema. Algunos ejemplos, señalados más arriba, son la interacción entre CAC y APC cuando se reajusta la potencia de transmisión del terminal móvil para permitir su acceso a la red (apartado 3.4.3.1) o la interacción entre CAC y la selección de estación base de forma conjunta con APC (apartado 3.4.3.7) o sin APC (apartado 3.4.3.6) con la finalidad de homogeneizar la distribución del tráfico.

Sin embargo, en ocasiones la interacción de varias técnicas RRM puede resultar en la degradación del comportamiento de una de ellas (o ambas). Un ejemplo [43] sucede cuando se combina el salto en frecuencia con *Link Adaptation*. La reducción de la eficacia se debe a que la calidad del enlace radio adquiere una gran aleatoriedad y varianza haciendo que las predicciones de *Link Adaptation* se vuelvan inexactas [25].

Un interesante ejemplo de integración de técnicas RRM aparece en [32], donde se propone un algoritmo genérico que combina FCA/DCA, APC y selección de estación base para la gestión eficiente en áreas metropolitanas, algoritmo a partir del cual resulta inmediata la derivación de múltiples variantes considerando varias combinaciones que pueden darse entre estas técnicas. Con ello se consigue una mayor flexibilidad a la hora de la aplicación práctica. Otros estudios de combinaciones particulares englobadas en el caso anterior, como DCA con APC [33, 34], o APC con selección de estación base [35, 36], también muestran resultados interesantes.

Otro ejemplo lo constituye el estándar EDGE, que contempla el uso conjunto de las técnicas *Incremental Redundancy* y *Link Adaptation* para la obtención de mejores resultados. Sin embargo, en GPRS esta combinación no resulta posible ya que los esquemas de codificación de GPRS (apartado 2.5.3.1.2.1) no fueron pensados para ello. La coordinación de *Link Adaptation* con el control de potencia [37] puede ser de especial interés ya que un incremento controlado en la potencia supone una mejor calidad de transmisión que puede ser utilizada por *Link Adaptation* para lograr tasas mayores.

En definitiva, lo que debe quedar claro es la posibilidad de conseguir mejores resultados cuando las decisiones de diferentes técnicas RRM se coordinan entre sí de forma controlada y apropiada, lo cual resulta posible ya que no son entidades individuales y aisladas dentro de un sistema de comunicaciones móviles.

3.5.- LINK ADAPTATION

3.5.1.- INTRODUCCIÓN

Cuando en Europa se decidió evolucionar hacia los sistemas de tercera generación, se crearon diversos grupos de trabajo, cada uno encargado de estudiar la definición e implementación de un sistema avanzado de telefonía móvil, de entre los cuales se elegiría uno para el futuro estándar UMTS. Uno de los sistemas candidatos se elaboró en el proyecto RACE/ATDMA (*Research in Advanced Communications in Europe / Advanced Time Division Multiple Access*).

En el sistema ATDMA se consideraron diversos niveles de adaptabilidad. El primer nivel consistía en la existencia de un conjunto de ráfagas, estructuras de trama, tasas binarias y modulaciones, cada uno de los cuales se optimizó para un determinado tipo de célula. Un segundo nivel de adaptación, denominado *Link Adaptation*, consistía en la selección dinámica de los parámetros de la interfaz aire de entre un conjunto de posibles combinaciones predefinidas denominadas *modos de transporte*. Cada modo de transporte se caracterizaba por una tasa de información a la salida del vocoder o codificador de voz y un esquema de codificación de canal determinado. El resto de parámetros relativos a la interfaz aire, especialmente el tipo de modulación empleada, se decidió mantenerlos fijos. Un tercer y último nivel de adaptabilidad se implementaba mediante una serie de procesos para la detección de la actividad de la voz y la asignación de recursos para la transmisión sólo durante dichos periodos de actividad[27]. Para ello se utilizaba PRMA++, una versión modificada de PRMA (*Packet Reservation Multiple Access*).

Al presente trabajo tan sólo concierne el segundo nivel de adaptabilidad, *Link Adaptation* (LA de aquí en adelante). Aunque LA fue inicialmente propuesto para el desarrollo del sistema ATDMA, en la actualidad representa un concepto bastante amplio y aplicable a gran cantidad de sistemas de comunicaciones móviles (como se verá en el apartado 3.5.8). Aquí se considerará su aplicación al sistema GPRS.

3.5.2.- ¿QUÉ ES LINK ADAPTATION?

Link Adaptation es una técnica de gestión de recursos radio DRA que selecciona de forma dinámica el modo de operación del enlace radio para ajustarlo a las condiciones particulares del canal de transmisión en cada momento y obtener así un rendimiento óptimo en diversas circunstancias de ruido e interferencia. El principio de funcionamiento de LA consiste en realizar, con carácter periódico, una estimación de la calidad del canal de transmisión y, en consecuencia, seleccionar el modo de transporte (de un conjunto ya preestablecido) que resulta más apropiado para dichas condiciones.

[27] En una conversación normal, el usuario habla sólo durante un 40% o 50% del tiempo que dura la conversación. Durante el resto de la conversación el canal de transmisión permanece ocioso, ya que sólo se transmite silencio.

Un modo de transporte queda definido por un conjunto de parámetros que determinan la transmisión en la interfaz aire. El cambio en el modo de transporte supondrá, por tanto, la modificación de estos parámetros. En este sentido, el concepto de modo de transporte es muy genérico y puede englobar diferentes parámetros tales como la velocidad binaria del codificador de voz, el tipo de codificación de canal, el nivel de protección contra errores, el tipo de entrelazado, el número total de time-slots por trama para cada usuario, el tipo de modulación empleada o incluso el nivel de potencia de transmisión.

En el caso del proyecto ATDMA, donde se propuso por primera vez la utilización de esta técnica, LA fue concebido para trabajar de forma conjunta con el control de potencia. Mientras que este último se encargaba de determinar el nivel de potencia de transmisión con el objeto de lograr la adaptación ante variaciones a corto plazo en la calidad de la señal, LA pretendía dar respuesta a variaciones a largo plazo [38]. Cuatro eran los posibles modos de transporte disponibles, cada uno de los cuales se caracterizaba por un determinado régimen binario neto a la salida del vocoder (codificador de voz) y por el uso de un determinado esquema de codificación de canal. Estos dos parámetros eran, por lo tanto, los que definían los cuatro posibles modos de transporte en ATDMA, fijándose una tasa neta y una tasa bruta para cada uno de ellos.

Modo	Tasa Neta	Tasa Bruta	Time-slots/trama
1	6.4 kbps	26 kbps	2
2	6.4 kbps	13 kbps	1
3	8.0 kbps	13 kbps	1
4	9.6 kbps	13 kbps	1

Tabla 3.1.- Modos de transporte en ATDMA para servicios de voz

La tasa bruta (*gross-rate*) hace referencia a la cantidad de bits por unidad de tiempo que son enviados por el transmisor mientras que la tasa neta (*net-rate*) considera, en este total de bits transmitidos, sólo aquellos que son bits de información, despreciando la cabecera e información redundante. Así mismo, se diferencia entre algoritmos de tipo *net-rate* y algoritmos de tipo *gross-rate*. En los primeros la tasa bruta permanece constante mientras que lo que se varía es la tasa neta (se combinan los modos 2, 3 y 4). En los segundos sucede lo contrario: la tasa bruta se varía asignando diferente cantidad de time-slots por trama para cada usuario mientras que la tasa neta permanece fija (se combinan los modos 1 y 2).

En el sistema GPRS, el principio de LA también encuentra un ámbito de aplicación. La posibilidad de aplicar LA en GPRS se justifica por la existencia de cuatro esquemas de codificación de canal diferentes, cada uno de los cuales introduce información redundante en diferente medida, proporcionando así diversos grados de compromiso entre protección frente a errores de transmisión y tasa binaria neta.

En el capítulo 2 (apartado 2.5.3.1.2.1) se vio que el aumento del nivel de protección ante errores de transmisión se consigue introduciendo más información redundante a costa de reducir la cantidad de bits de datos en cada bloque radio, lo cual repercute en una reducción del throughput. Por tanto, cada esquema de codificación ofrece una tasa neta de datos inversamente proporcional a su grado de robustez ante errores.

Esquema	Factor	Bits de datos	Tasa neta	Tasa bruta
CS–1	1/2	181	9.05 kbps	22.8 kbps
CS–2	≈ 2/3	268	13.4 kbps	22.8 kbps
CS–3	≈ 3/4	312	15.6 kbps	22.8 kbps
CS–4	1	428	21.4 kbps	22.8 kbps

Tabla 3.2.- Modos de transporte en GPRS[28]

Cuando las condiciones del canal de transmisión son favorables (por ejemplo cuando hay visibilidad directa entre estación base y estación móvil) es posible emplear un esquema de codificación con poca protección para lograr velocidades de transmisión mayores utilizando la misma cantidad de recursos. Con ello se puede aumentar la cantidad de datos transmitidos en un intervalo de tiempo dado (throughput) o, de forma equivalente, reducir el tiempo requerido para enviar la misma cantidad de datos (retardo). Si la calidad del canal empeora ocasionando un mayor número de errores en la transmisión, entonces se escoge otro esquema de codificación con más información redundante, lo cual tendrá dos efectos sobre la transmisión: por una parte se conseguirá un aumento en la robustez y fiabilidad ante errores (efecto positivo) pero por otra se reducirá el throughput y aumentará también el retardo de transmisión (efecto negativo).

Esta capacidad de adaptación en la codificación de canal supone un mejor rendimiento respecto a la codificación de canal fija [30, 39, 42], ya que si las condiciones de transmisión son demasiado malas, la información redundante puede resultar insuficiente y la elevada cantidad de errores hace que sea necesario un gran número de retransmisiones, aumentando el retardo y disminuyendo así la velocidad neta de transmisión. En cambio, en condiciones favorables en las que no existan errores, la introducción de información redundante supone un desperdicio de recursos. Por lo tanto, el uso de LA representa un uso más eficiente de los recursos radio, permitiendo que la velocidad de transmisión sea lo más elevada posible de forma acorde a la calidad del canal de transmisión en cada instante.

Debe quedar claro que la definición del concepto de LA admite la adaptación de diversos parámetros de la interfaz aire, tal y como se explicó más arriba. Sin embargo, en lo sucesivo, cuando se hable de LA aplicado a GPRS, se estará haciendo referencia únicamente al mecanismo por el cual se decide cuál de los cuatro posibles esquemas de codificación (CS–1, CS–2, CS–3 o CS–4) se ha de emplear en función de la calidad de transmisión experimentada en el canal radio. El resto de parámetros de la interfaz aire serán indiferentes para LA en este trabajo y permanecerán fijos.

[28] Todos los esquemas tienen la misma tasa bruta (456 bits / 20 ms = 22.8 kbps), ya que todos entregan 456 bits (datos, cabecera e información redundante) cada 20 ms (duración de un bloque radio). Sin embargo, la cantidad de bits de datos (carga útil) dentro de los 456 bits totales es diferente, dando lugar a diferentes tasas netas.

3.5.3.- ESTIMACIÓN DE LA CALIDAD DEL CANAL

Una de las claves que determinan el éxito o el fracaso de la técnica LA reside en la estimación de la calidad del canal. Las medidas realizadas han de proporcionar una información lo más fiel y exacta posible acerca de las condiciones reales del canal de transmisión, ya que si no se conoce con suficiente precisión su estado real, difícilmente podrá tomarse una decisión acertada acerca del esquema de codificación idóneo y, en consecuencia, la eficacia de LA se verá perjudicada. Además de la fidelidad de la métrica utilizada para estimar la calidad del canal, otro factor clave es la necesidad de que la estimación se realice en un intervalo de tiempo prudencialmente reducido o, en caso contrario, tal y como se verá en el apartado 3.5.5, la estimación carecerá de utilidad por muy precisa que ésta sea. Ambos factores se hallan, pues, comprometidos.

Uno de los métodos más sencillos para estimar la calidad del canal se basa en el BLER (*BLock Error Rate*), el cual se define como el cociente entre el número de bloques RLC confirmados por el receptor como erróneos y el número total de bloques RLC confirmados (con o sin error). Por tanto, el BLER proporciona una estimación de la calidad del canal a nivel de bloque RLC. Un bloque RLC es la secuencia de bits mínima sobre la que es posible determinar el BLER. El valor del BLER es fácilmente extraíble a partir de los mensajes de confirmación que el receptor devuelve al transmisor (ver apartado 2.5.2.4). El transmisor tan sólo debe llevar la cuenta de los bloques RLC que transmite y compararla con las confirmaciones recibidas. Sin embargo, la obtención de una estimación precisa a partir del BLER requiere que la evaluación se realice sobre un número considerable de bloques RLC[29]. Esto puede llevar un tiempo demasiado elevado y hacer que el comportamiento de LA empeore (como se verá en el apartado 3.5.5). Una mejora en el comportamiento de LA podría obtenerse, pues, reduciendo el número de bloques RLC sobre los que se evalúa el BLER. No obstante, si la estimación se realiza sobre un número reducido de bloques RLC, entonces la exactitud de la medida disminuye, con lo que el comportamiento de LA empeora igualmente. Por tanto, se concluye que el BLER es una métrica que resulta muy fácil de obtener pero que, sin embargo, tiene un grado de exactitud y fiabilidad bastante limitado.

Otro parámetro también representativo de la calidad del canal es la relación CIR (*Carrier-to-Interference Ratio*), que se define como el cociente entre la potencia de señal recibida y la suma de la potencia total de ruido más la potencia total de señal recibida de otros terminales móviles cocanales (interferencia). En el receptor, la relación CIR se evalúa sobre cada ráfaga recibida, por lo que esta métrica proporciona una estimación de la calidad del canal a nivel de ráfaga. Una ráfaga es la secuencia de bits mínima sobre la que se determina la relación CIR. La principal ventaja de la relación CIR frente al BLER es que la obtención de una estimación aceptable se puede conseguir en un tiempo más reducido: puede ser estimada con un error medio de 1 dB en el tiempo que dura una ráfaga [42]. Sin embargo, cuenta con la desventaja de que su obtención resulta bastante compleja [43] y requiere elevados costes computacionales.

[29] Si evaluamos el BLER sobre un único bloque RLC podemos obtener BLER = 0 (el bloque se transmite sin error) o BLER = 1 (el bloque se transmite con error). Si evaluamos el BLER sobre dos bloques RLC, entonces los posibles valores de BLER son: 0 (los dos se transmiten sin error), 1 (los dos se transmiten con error) o 0.5 (uno se transmite con error y el otro sin error). Queda claro, pues, que al evaluar el BLER sobre un mayor número de bloques RLC la medida posee una mayor resolución y resulta más precisa, ofreciendo una mejor estimación de la calidad del canal.

La relación CIR no se puede obtener de forma directa, tal y como sucede con el BLER, sino que ha de ser *estimada* a partir de otros parámetros que sí son directamente extraíbles del canal de transmisión. Su cálculo puede realizarse utilizando técnicas basadas en la mínima distancia euclídea derivada de un decodificador de Viterbi [45, 46] o, alternativamente, basadas en proyecciones sobre subespacios de señal [47 – 50]. En [47], por ejemplo, se utiliza la secuencia de entrenamiento (*Training Sequence*) de una ráfaga normal (apartado 2.5.1.2.3.1) para realizar una estimación de la relación CIR[30]. En [48] se propone la estimación de la relación CIR a partir de los valores propios de la matriz de covarianza de la señal recibida. Otra técnica basada en la proyección de la señal recibida sobre un subespacio de señal que representa la señal deseada aparece en [49]. En [50] se utiliza la secuencia de autocorrelación de las muestras de señal recibidas. En definitiva, lo que se pretende mostrar es que la obtención del valor de la relación CIR en el ámbito práctico se caracteriza por una considerable complejidad de cálculo, considerablemente mayor que la necesaria para obtener el BLER. La utilización de algoritmos LA basados en la relación CIR se justifica por ofrecer mejores resultados que aquellos basados en el BLER [42].

Si bien el BLER y la relación CIR son los parámetros más ampliamente usados en LA para representar la calidad del canal de transmisión [42], no son los únicos[31]. El BER (*Bit Error Rate*) es otro parámetro de indudable utilidad a la hora de proporcionar una estimación de la calidad del canal. Se define como el cociente entre el número de bits erróneos y el número total de bits recibidos (con o sin error) en el receptor y se obtiene de forma sencilla contando cuántos bits de datos de una ráfaga normal son erróneos a la salida del decodificador de canal. Por lo tanto, el BER proporciona una estimación de la calidad del canal a nivel de ráfaga.

Se ha propuesto [51, 52] el uso de dos métricas alternativas que siguen de forma fiel el comportamiento del BER. La primera de ellas se denomina *Soft Output Averages* (SOA) y se basa en el valor obtenido a la salida de un decodificador de "decisiones blandas" (*soft decisions*), el cual puede ser considerado como un indicador del grado de confianza que se puede depositar en la decodificación realizada. La segunda recibe el nombre de *Errors in Training Sequence* (ETS) y se basa en la secuencia de entrenamiento de una ráfaga normal. La principal ventaja de estas dos métricas es que proporcionan un grado de precisión en la estimación equiparable al proporcionado por la relación CIR pero con una dificultad de cálculo mucho más reducida.

Antes de concluir el presente apartado conviene dejar claro que los parámetros más ampliamente utilizados en la literatura para la estimación de la calidad del canal en estudios de LA son el BLER y la relación CIR. No obstante, ello no quiere decir que no sea posible emplear otros parámetros. A modo de ejemplo se han citado algunas de las alternativas propuestas por algunos autores.

[30] Recordar que la secuencia de entrenamiento es una secuencia de 26 bits dentro de una ráfaga normal cuyo contenido es fijo y conocido de antemano por el receptor. Comparando la secuencia de entrenamiento recibida con la secuencia esperada es posible realizar una estimación de la relación CIR del canal de transmisión.

[31] A título de curiosidad, en [53] se propone el uso del valor del Timing Advance como parámetro que determina el esquema de codificación de canal a emplear.

3.5.4.- ELECCIÓN DEL MODO DE TRANSPORTE

El comportamiento de cada esquema de codificación depende fuertemente de las condiciones del canal [7, 43]. Cada uno de ellos ofrece mejores prestaciones que el resto dentro de una determinada región [41]. Las fronteras que delimitan estas regiones se denominan *umbrales de decisión*. La ubicación exacta de los umbrales de decisión dependerá del criterio que se siga para determinar qué esquema de codificación es *mejor* que los demás. Para el caso de servicios de datos en tiempo no-real, un criterio ampliamente aceptado consiste en escoger aquel esquema de codificación de canal que proporciona el mayor throughput [42, 43, 44, 53]. En este caso, los umbrales de decisión definen las fronteras entre las regiones donde cada esquema de codificación maximiza el throughput. La figura 3.4, obtenida por simulación, muestra el throughput en función de la relación CIR para los cuatro esquemas de codificación.

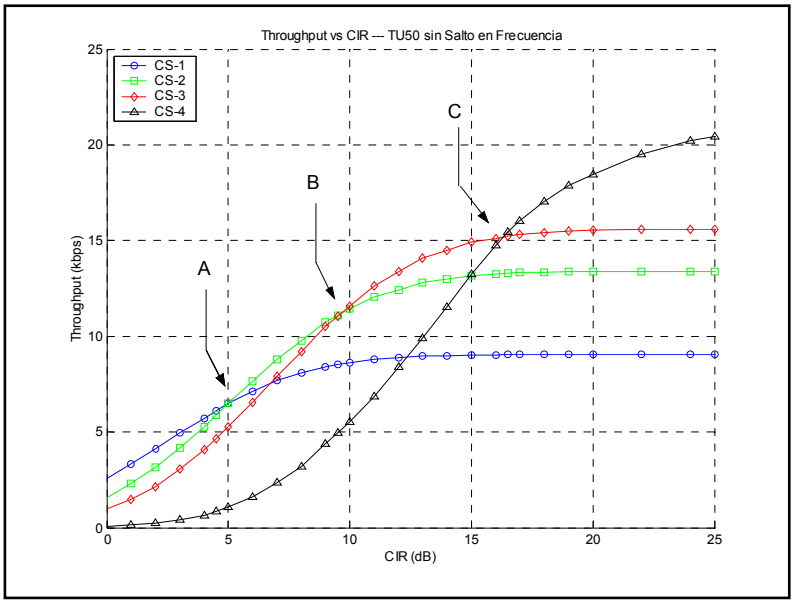


Figura 3.4.- Throughput vs CIR [101]

Puede apreciarse claramente que para una relación CIR pobre el canal introduce muchos errores en la transmisión y el esquema de codificación que proporciona una mayor tasa binaria (por debajo del umbral *A*) es CS–1, ya que es el que incorpora mayor protección contra errores y, a largo plazo, esto supone un menor número de retransmisiones, logrando así un mayor throughput. Conforme mejora la relación CIR, el throughput más elevado se alcanza haciendo uso de esquemas de codificación con menos información redundante, ya que ésta no es tan necesaria y puede ser sustituida por datos. Entre *A* y *B* el mejor throughput lo proporciona CS–2, y entre *B* y *C* lo proporciona CS–3. En casos de visibilidad directa u otras situaciones donde la relación CIR sea muy buena (por encima del umbral *C*) la introducción de información redundante para corrección de errores supone un desperdicio ya que no es necesaria, y el throughput puede maximizarse eliminando todo tipo de protección contra errores, es decir, empleando el esquema de codificación CS–4.

Una expresión matemática para las curvas de la figura 3.4 puede derivarse fácilmente como sigue. Si la transmisión en la interfaz aire ocurriera sin ningún error, entonces el valor del throughput coincidiría con la tasa binaria neta ($R_{CS\text{-}i}$) del esquema de codificación empleado (tabla 3.2). De hecho, si en la figura 3.4 se representara un rango de valores de la relación CIR suficientemente elevado, se podría apreciar claramente que conforme el valor de la relación CIR tiende idealmente hacia infinito (dejan de producirse errores), el valor del throughput para cada esquema de codificación tiende hacia los valores de velocidad binaria neta dados en la tabla 3.2. Sin embargo, en la realidad suceden errores de transmisión que hacen que algunos bloques RLC hayan de ser desechados, por lo que el throughput real es menor que la velocidad binaria del esquema de codificación empleado. Una medida de estos errores se puede obtener a partir del BLER. Teniendo en cuenta que el valor del BLER (definido en el apartado 3.5.3) varía entre 0 y 1, el throughput para el esquema de codificación *CS-i* viene dado por la expresión

$$Throughput_{CS-i} = R_{CS-i} \cdot \left[1 - BLER_{CS-i}(CIR)\right] \tag{3.4}$$

El BLER es una medida de los errores que introduce el canal de transmisión, mientras que la relación CIR nos da una idea acerca de su calidad. Por lo tanto, parece lógico que exista una relación de proporcionalidad inversa entre la calidad del canal de transmisión (CIR) y la cantidad de errores que introduce en la transmisión (BLER). La figura 3.5 muestra esta relación para cada uno de los esquemas de codificación (obtenida por simulación bajo las mismas condiciones que la figura 3.4) en la que se muestra la dependencia entre BLER y CIR. Así pues, a partir de la figura 3.5 y de la expresión (3.4) es posible obtener las curvas mostradas en la figura 3.4, de donde resulta inmediata la obtención de los umbrales de decisión (puntos de intersección entre curvas) que determinan la selección del esquema de codificación a emplear en cada caso en función de la calidad del canal (expresada en términos de relación CIR).

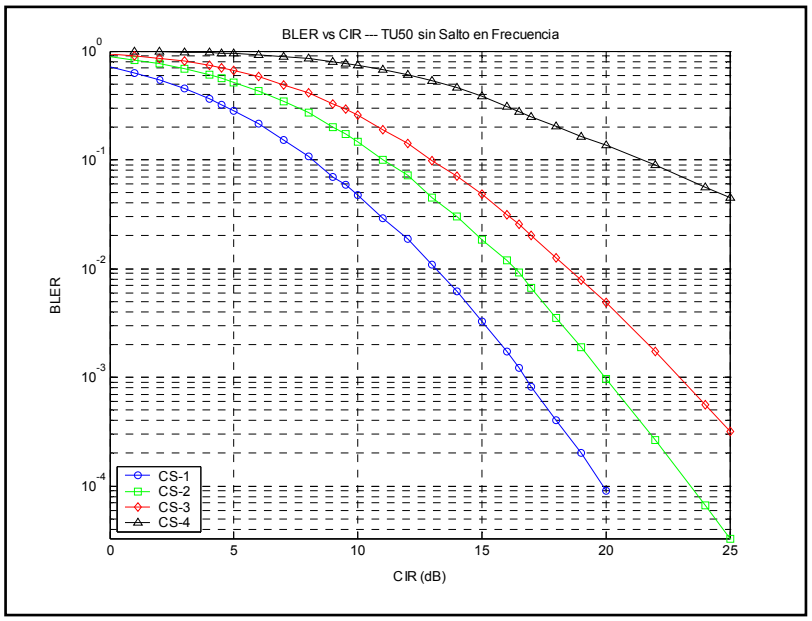


Figura 3.5.- BLER vs CIR [105]

El método descrito para la determinación de los umbrales de decisión conduce a una implementación de LA donde el esquema de codificación se escoge siguiendo el criterio de maximización del throughput, criterio considerado tradicionalmente y que resulta adecuado para servicios de datos en tiempo no real como, por ejemplo, navegación web (WAP), correo electrónico, mensajes SMS o FTP. Sin embargo, en otro tipo de servicios puede ser conveniente la utilización de criterios diferentes. En servicios de voz, por ejemplo, la elección del esquema de codificación se realiza en función de la relación segSNR[32], útil para medir la calidad de la señal de voz que recibe el usuario. Para el caso que ocupa la atención de este trabajo, el retardo es un factor fundamental en la determinación de la calidad de servicio y la elección del esquema de codificación basada en criterios de retardo es una opción a considerar. Si bien es cierto que un throughput mayor supone un menor retardo de transmisión, ello no implica necesariamente que el criterio que maximiza el throughput medio también minimice el retardo de transmisión, tal y como se señala en diversos estudios citados más adelante. La cuestión del retardo, de especial relevancia para los propósitos de este trabajo, será retomada y discutida en un capítulo posterior.

Una situación peculiar relacionada con la elección del modo de transporte y que resulta de interés citar aquí la constituye la elección inicial. Al comienzo de cada periodo de actividad, LA desconoce por completo cuál es la calidad del canal ya que carece de información y, sin embargo, en ese momento debe elegir el esquema de codificación a emplear. En [42] se propone que, si la carga del sistema es baja (aproximadamente por debajo del 50%), entonces una posibilidad interesante puede ser emplear inicialmente el esquema de codificación de canal con menor protección frente a errores, es decir, CS–4. En caso de carga moderada o alta, una posibilidad es adoptar la solución que se propone en [39], según la cual la mejor opción consiste en emplear el mismo esquema de codificación con el que se concluyó el último periodo de actividad. Para iniciar una nueva sesión de comunicación, el esquema CS–2 es el que proporciona un inicio más seguro según [42].

Por último, y para concluir este apartado, conviene comentar algunas peculiaridades del estándar GPRS, el cual impone, en el ámbito práctico, algunas restricciones a LA a la hora de seleccionar el esquema de codificación de canal a emplear. El estándar GPRS no contempla el cambio en el esquema de codificación para efectuar la retransmisión de bloques erróneos, es decir, la retransmisión de un bloque debe realizarse empleando el mismo esquema de codificación que se empleó para su transmisión inicial. Ello es debido a que en GPRS el tamaño del campo de datos de un esquema de codificación de canal es incompatible con el tamaño del campo de datos de cualquier otro esquema, a diferencia de, por ejemplo, EGPRS [60]. En EGPRS (EDGE) estos cambios son posibles entre esquemas de codificación de una misma familia, donde los campos de datos guardan una relación de proporción de múltiplo entero. Además, en GPRS, el esquema de codificación sólo puede ser cambiado entre bloques RLC asociados a dos LLC-PDUs consecutivas [44].

[32] La relación segSNR se define como la media aritmética de la relación señal a ruido o SNR (*Signal-to-Noise Ratio*). La relación SNR es la relación entre la varianza de la señal vocal de entrada y la varianza del error de reconstrucción de dicha señal.

3.5.5.- PERIODO DE ACTUALIZACIÓN

La eficiencia de LA depende, principalmente, de dos factores clave: la exactitud en la estimación realizada de la calidad del canal de transmisión y la capacidad del sistema para adaptarse a sus variaciones. El primero se trató en el apartado 3.5.3; el segundo, condicionado por el periodo de actualización, se trata a continuación.

El periodo de actualización es el intervalo de tiempo que define con qué regularidad se toman las decisiones que determinan qué esquema de codificación de canal resulta óptimo. Este periodo debe estar ajustado a la dinámica del canal móvil.

Un periodo de actualización corto consigue una mejor adaptación ya que permite una respuesta rápida ante cambios bruscos, pero también implica que exista una mayor probabilidad de cambio en el esquema de codificación empleado, con el consecuente aumento de señalización que ello conlleva para el sistema. Este hecho puede dar lugar al efecto *ping-pong*, es decir, a un cambio constante entre dos esquemas de codificación de canal alrededor de un umbral de decisión. En cambio, un periodo de actualización largo disminuye esta cantidad de señalización pero, por el contrario, también supone que la adaptación no sea tan fiel en aquellos casos en los que la calidad del canal varía con rapidez [42].

Por tanto, el periodo de actualización representa un compromiso entre el número de cambios de esquema de codificación de canal (carga de señalización) y la capacidad del sistema para adaptarse a las variaciones rápidas experimentadas por el canal móvil. La implementación de LA será óptima si maximiza la capacidad de adaptación a las variaciones del canal a la vez que minimiza el número de cambios de esquema de codificación (carga de señalización). En [55] se presenta un estudio acerca de esta problemática, aunque considerando un conjunto de condiciones bastante limitado. Este estudio se amplía en [56], donde se consideran otros parámetros que afectan al periodo de actualización como la carga del sistema, la diversidad del tráfico o la velocidad del terminal móvil.

La situación ideal se caracterizaría por la aplicación de un periodo de actualización corto sin que ello supusiera una elevada señalización. En [39] se demuestra que la introducción de histéresis en los umbrales de decisión permite mantener un periodo de actualización relativamente corto, disminuyendo de forma considerable el efecto *ping-pong* sin que ello suponga una reducción apreciable en la calidad experimentada por el usuario.

Por último, comentar que el periodo de actualización mínimo para GPRS es de 20 ms, que es el plazo de tiempo al cabo del cual se transmite un nuevo un bloque radio. Cualquier otro periodo de actualización ha de ser, necesariamente, un múltiplo entero de 20 ms, ya que no es posible cambiar el esquema de codificación de canal durante la transmisión de un bloque radio; cada bloque se codifica íntegramente empleando el mismo esquema de codificación.

3.5.6.- EVALUACIÓN DE LA EFICACIA

El uso de LA proporciona resultados claramente superiores respecto a aquellas situaciones en las que se opta por utilizar una codificación de canal fija [30, 39, 42]. Además, resulta particularmente eficiente en combinación con esquemas de acceso en modo paquete [30] (como GPRS), permitiendo una asignación de recursos realmente flexible y un aprovechamiento máximo. La figura 3.6 ilustra de forma clara el objetivo perseguido mediante la utilización de LA en GPRS.

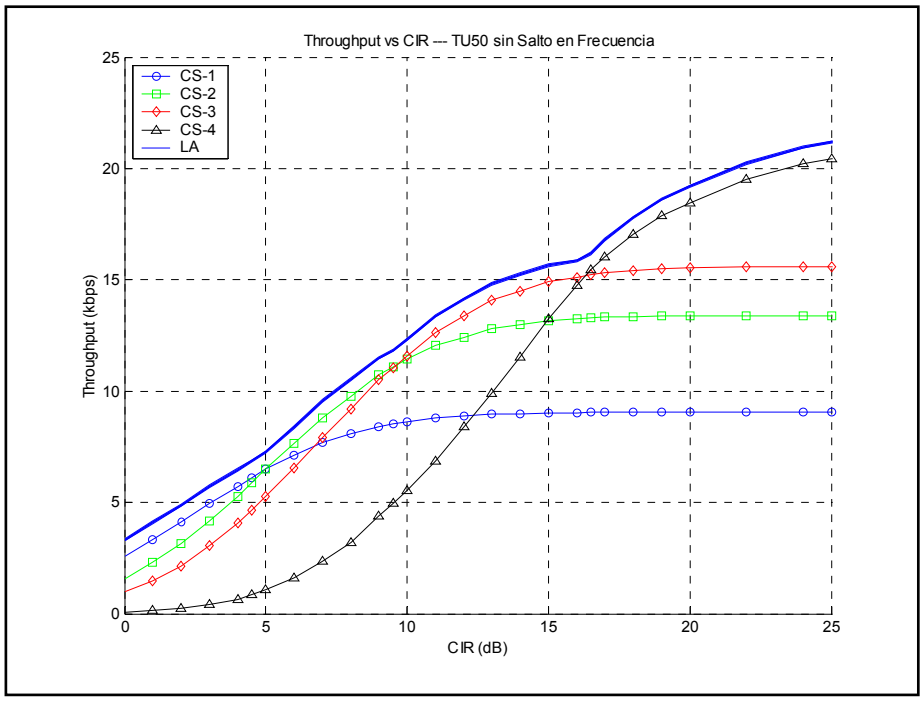


Figura 3.6.- Throughput con *Link Adaptation* [101]

Según [42], el throughput medio resulta importante a la hora de evaluar el comportamiento de un algoritmo, pero en ciertos casos no resulta suficiente para poder realizar una estimación realista acerca de su utilidad. Otro factor importante es su grado de imparcialidad: ¿pueden todos los usuarios, o un porcentaje aceptablemente elevado, esperar aproximadamente el mismo throughput o, sin embargo, existen grandes desigualdades entre ellos? Un algoritmo que permita alcanzar velocidades muy elevadas será poco útil si sólo un porcentaje bajo de usuarios tiene garantizado el disfrute de estas prestaciones. Es preferible proporcionar un throughput menor pero que esté al alcance de un porcentaje aceptable de usuarios (los criterios más habituales establecen entre el 95% y 99% de usuarios). Otra medida importante y que toma en cuenta la consideración anterior es la función de distribución acumulativa o CDF (*Cumulative Distribution Function*) del throughput. En la figura 3.7, obtenida mediante simulación, se representa la CDF del throughput para *Link Adaptation*.

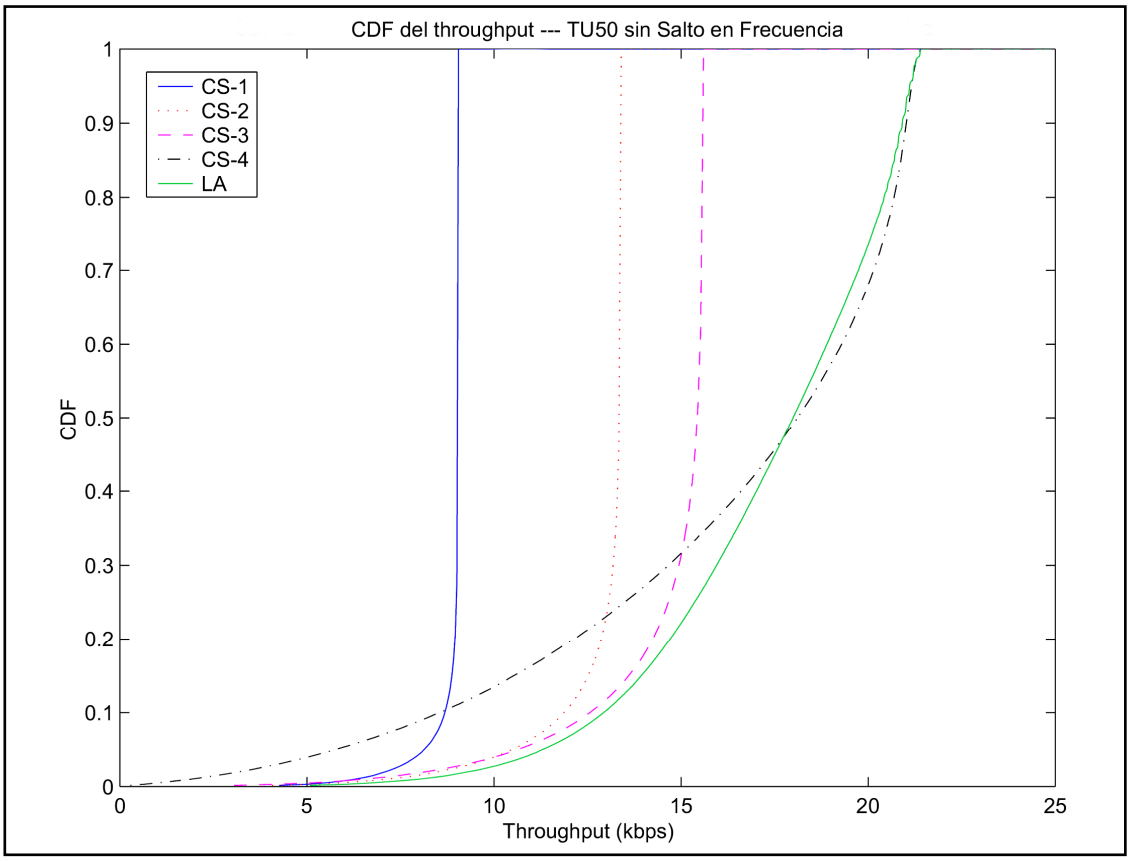


Figura 3.7.- Evaluación de la CDF del throughput con LA [101]

Puede verse que el uso de LA proporciona mejores resultados respecto al uso de cualquier esquema de codificación de canal de forma individual, es decir, el uso de una codificación de canal fija, excepto en una pequeña región de la gráfica en la que LA ofrece unas prestaciones ligeramente peores que CS–4. Ello es debido a que, cuando las condiciones del canal son muy buenas, LA escoge a CS–4 como esquema de codificación, el cual es muy susceptible a errores de transmisión.

En principio, si el canal no introduce ningún error, la trasmisión en estas condiciones será óptima y se alcanzará el throughput máximo. Sin embargo, es suficiente que ocurra un error en un único bit de un bloque RLC para que LA cambie a un esquema más robusto (CS–3), el cual se seguirá utilizando hasta que se cumpla el periodo de actualización. Si no ocurre ningún error durante la transmisión de los bloques posteriores, para los que se utiliza CS–3, entonces no se estará utilizando el esquema de codificación óptimo, es decir, CS–4. Ésta es la razón por la que LA muestra un comportamiento ligeramente inferior en esa región.

No obstante, observando en detalle la figura 3.7, se puede observar que la diferencia de rendimiento de LA respecto a CS–4 en esta zona es relativamente pequeña en comparación con la mejora que muestra en el resto de regiones (respecto a CS-4 y respecto al resto de esquemas de codificación), motivo por el cual sigue resultando provechosa la utilización de LA.

3.5.7.- BENEFICIOS DE LINK ADAPTATION

Las ventajas de utilizar LA son palpables no sólo en términos de throughput, tal y como se ha visto en el apartado anterior, sino que otros aspectos importantes en un sistema de comunicaciones móviles pueden verse mejorados al utilizar LA.

En [54] se realiza el estudio de un algoritmo LA *gross-rate* donde se concluye que con LA es posible lograr un aumento en la capacidad del sistema (salvo cuando el sistema está sobrecargado). Este incremento se logra mejorando la calidad que experimentan aquellos usuarios más desfavorecidos a costa de reducir la calidad de los más afortunados, lo cual permite la incorporación de nuevos usuarios. En este caso, el objetivo de LA no es la mejora de la calidad de todos los bloques RLC transmitidos, sino la reducción del número de aquellos cuya calidad no se considera aceptable.

En [30, 54, 61] se demuestra que en sistemas limitados en interferencia LA puede mejorar la cobertura de una estación base significativamente o, de forma equivalente, reducir la potencia de transmisión, lo cual supone una disminución de la interferencia cocanal. Un ejemplo de aplicación práctica se da en [64].

3.5.8.- APLICACIÓN A OTROS SISTEMAS

Aunque LA nació durante el desarrollo del sistema de tercera generación ATDMA, hoy día representa el principio fundamental sobre el que se apoya una gran cantidad técnicas RRM de otros estándares, como lo demuestra su aplicación en AMR, EDGE y otros más. El concepto de LA, particularizado a GPRS, consiste en la elección del esquema de codificación de canal en función de la calidad experimentada sobre el canal de transmisión. Sin embargo, la definición de LA es mucho más amplia y va más allá, abarcando otros parámetros de la interfaz aire, tal y como se explicó en el apartado 3.5.2. Los parámetros que en GPRS eran indiferentes para LA, en otros sistemas, como se verá a continuación, son competencia de LA.

En el estándar AMR (*Adaptive Multi-Rate*) se usan técnicas de adaptación dinámicas que ofrecen un compromiso entre calidad de voz y capacidad del sistema [3, 57]. Se definen dos niveles de adaptación: nivel de canal y nivel de códec. En el nivel de canal se permite el cambio entre canales de voz *Full-Rate* (FR) y *Half-Rate* (HR). Cuando la carga del sistema es baja, se utilizan canales FR para proporcionar una mejor calidad de voz; si un usuario experimenta una calidad muy buena, entonces se le asignará un canal HR (peor calidad de voz) a fin de conseguir un aumento en la capacidad del sistema. De forma análoga, cuando la carga del sistema sea elevada, se utilizarán canales HR a fin de poder satisfacer toda la demanda de usuarios; si un usuario experimenta una calidad demasiado mala, entonces se le asignará un canal FR. Para cada canal (FR o HR) se define un segundo nivel de adaptación en el que se consideran tres códecs diferentes (I, II y III), cada uno de los cuales introduce una cantidad de información redundante diferente. De esta forma, ajustando el códec empleado, AMR puede reaccionar ante los errores de transmisión del canal, lo cual es equivalente a LA tal y como se entiende en GPRS.

Otro sistema donde LA encuentra un campo de aplicación es EDGE (*Enhanced Data-rates for Global Evolution*). El estándar EDGE contempla nueve esquemas de codificación de canal y modulación o MCSs (*Modulation and Coding Schemes*), los cuales ofrecen diferentes grados de compromiso entre velocidad de transmisión y robustez ante errores. En EDGE, estos diferentes grados de compromiso se logran modificando no sólo la cantidad de información redundante introducida durante el proceso de codificación de canal (como es el caso de GPRS), sino modificando también el tipo de modulación empleada y de ahí su denominación MCS. Tal y como se comentó en la introducción del capítulo segundo, además de la modulación mononivel GMSK que se utiliza en GSM y GPRS, en EDGE también se emplea una modulación multinivel 8-PSK que permite aumentar la tasa binaria, aunque haciendo la transmisión más vulnerable a errores de transmisión, característica inherente a cualquier modulación multinivel.

Una mejora de LA investigada sobre EDGE consiste en establecer, además de los umbrales de decisión convencionales, un umbral de decisión adicional denominado *umbral de corte*, por debajo del cual se impide la transmisión al usuario. Esta variante de LA se denomina *modo-0* [58, 59, 60]. Transmitir por debajo del umbral de corte causa la inestabilidad del sistema a causa a las excesivas retransmisiones. Impidiendo la transmisión a aquellos usuarios cuya relación CIR sea inferior a la del umbral de corte se reduce la interferencia en el sistema. Como consecuencia, el resto de usuarios tienen la posibilidad de emplear otros MCSs que permitan conseguir un throughput mayor, reduciendo de esta manera el tiempo medio de ocupación de los recursos y mejorando así el comportamiento global del sistema.

El modo-0 también puede ser visto como un algoritmo de control de admisión (CAC) que controla la carga del sistema en beneficio de los usuarios activos y, gracias al cual, el comportamiento del sistema mejora de forma notable, especialmente en aquellos entornos limitados en interferencia. El modo-0 es más conservativo que LA, ya que para valores bajos de la relación CIR se impide la transmisión y, además, para relaciones CIR altas, puede intervenir reduciendo el tiempo de ocupación de los recursos a fin de maximizar la eficiencia espectral.

En [61] se presenta la aplicación de los principios de LA al sistema GSM. El principio del algoritmo LA consiste en la utilización de un canal de tráfico FR para acomodar un canal HR sobre el que se aplican procesos adicionales de codificación destinados a introducir mayor información redundante con objeto de incrementar la robustez frente a errores. Los resultados no sólo muestran una resistencia mayor frente a ruido e interferencias, sino que además permite ampliar el radio de cobertura de las estaciones base.

Los principios de LA también han sido propuestos para su aplicación en la interfaz aire de los sistemas síncronos CDMA2000 [62] y en los sistemas asíncronos WCDMA [63]. En los sistemas TETRA (*TErrestrial Trunked RAdio*), LA se puede utilizar para lograr un aumento del radio de cobertura de las células [64]. La versatilidad de LA llega hasta el extremo de ser empleado más allá de las redes móviles. Sirva [65] como un ejemplo de su utilización en redes de área local inalámbricas.

3.6.- SUMARIO

En el presente capítulo se ha descrito el pilar de este trabajo: *Link Adaptation*. *Link Adaptation* es una técnica destinada a mejorar el aprovechamiento de los escasos recursos radioeléctricos y, como tal, se ha explicado enmarcada en el contexto de la gestión de recursos radio. Sin embargo, el estudio de estas técnicas de gestión implica la utilización de algunos conceptos fundamentales relacionados con la estructura celular, utilizada hoy día por todos los sistemas de comunicaciones móviles. Se supone al lector conocedor de estos principios; no obstante, se ha considerado oportuno realizar un breve recordatorio en el apartado 3.2. La explicación de los fenómenos que determinan la propagación en el canal móvil de transmisión (atenuación, efecto sombra y efecto multicamino) cierran las premisas al tema que ocupa nuestro interés: la gestión de recursos radio y, más específicamente, la técnica *Link Adaptation*.

En el apartado 3.4.1 se introduce el tema mediante una formulación clara de los objetivos perseguidos en la gestión de recursos radio, cuya necesidad quedó patente en la introducción de este capítulo. En el apartado 3.4.2 se establece una clasificación genérica de las diversas técnicas de gestión de recursos radio, algunas de las cuales se presentan en el apartado 3.4.3: control de admisión, control de la congestión o *scheduling*, operación multislot, asignación de canales, salto en frecuencia, selección de estación base, control de potencia y redundancia incremental. La descripción de diversos modos de interacción entre las diversas técnicas de gestión cierra el apartado 3.4 para dar paso a un estudio más detallado de *Link Adaptation*.

Tras una breve reseña histórica introductoria, se explica el principio de funcionamiento de *Link Adaptation*, en primer lugar tal y como fue concebido inicialmente (en el sistema ATDMA) y después aplicado a GPRS. *Link Adaptation* se define básicamente como la adaptación dinámica del modo de transporte de forma periódica en función de la calidad experimentada en el canal de transmisión. Los factores clave que determinan la eficacia de *Link Adaptation* son dos: la calidad de las estimaciones de la calidad del canal y la frecuencia con la que se decide el modo de transporte a emplear. Estos dos aspectos clave, la forma en que se utilizan para decidir el modo de transporte y su repercusión sobre el rendimiento de LA se tratan en los apartados 3.5.3 al 3.5.5. Tras discutir la eficacia de *Link Adaptation* y los beneficios derivados de su uso, el presente capítulo se cierra con un último apartado donde se cita la aplicación de LA a otros sistemas como GSM, AMR, EDGE, WCDMA, CDMA2000, TETRA o redes WLAN, mostrando así la versatilidad de aplicación de sus principios.

Capítulo 4

Servicios de Vídeo y Calidad de Servicio

4.1.- INTRODUCCIÓN

En el capítulo de introducción del presente trabajo se señaló que el crecimiento en la demanda de servicios de voz que hoy día siguen experimentando los sistemas de comunicaciones móviles se ve acompañado por una tendencia hacia la convergencia de Internet y las redes móviles inalámbricas, hecho que fomentará en buena medida la prestación de una amplia gama de servicios multimedia sobre los futuros sistemas de comunicaciones móviles de tercera generación.

	Tiempo Real	**Tiempo No Real**
Interactivo	Videoconferencia Telefonía	Navegación web Control remoto
No Interactivo	Radio difusión Vídeo difusión	Correo electrónico Transferencia de archivos

Tabla 4.1.- Clasificación de algunos servicios multimedia [66]

Dentro de este tipo de servicios destacan, por su especial atractivo, aquellos que implican la transmisión de datos de vídeo o audio (tabla 4.1). Esta clase de servicios, habitualmente de tiempo real, se caracterizan por presentar unos requisitos en cuanto a calidad de servicio notablemente más exigentes que los habituales para otro tipo de servicios. En general, suelen ser bastante vulnerables ante los errores de transmisión y, en especial, requieren un retardo de transmisión considerablemente reducido entre transmisor y receptor, además de un flujo de datos entre ambos en el cual es necesario preservar la relación temporal existente entre las diversas entidades de información. Cada bloque de información lleva asociado un plazo de tiempo dentro del cual debe completarse su transmisión. Una vez superado este plazo de tiempo, la información deja de ser útil para el receptor y se considera perdida, lo cual es equivalente a haberla recibida con error y tener que desecharla. Es por este motivo que el retardo llega a jugar un papel clave para este tipo de servicios en la determinación de la calidad de servicio final experimentada por el usuario.

El interés de este trabajo se centra, precisamente, en el estudio de la aplicación de *Link Adaptation* a este tipo de servicios, para el cual se consideró oportuno emplear el estándar de codificación de vídeo H.263 de la ITU-T (*International Telecommunication Union*). Por esta razón se proporcionará en este capítulo una breve descripción de los fundamentos básicos de vídeo que son necesarios y empleados. El anexo A ofrece una breve y sencilla descripción de la codificación de vídeo H.263.

En conmutación de circuitos la asignación de un canal es fija, por lo que los recursos están disponibles siempre que se necesitan, garantizando la eliminación de retardos adicionales durante la transmisión debidos a la ausencia de recursos en el momento en que son requeridos. En conmutación de paquetes la situación es bien diferente, tal y como se vio en el capítulo segundo. Ya que los servicios en tiempo real tendrán que sufrir las características del tráfico intermitente propio de la conmutación de paquetes, es necesario que tengan garantizada la calidad de servicio [67]. Es por ello que en este contexto cobra especial relevancia el concepto de calidad de servicio. En este capítulo se presentará el concepto general así como varios parámetros de calidad asociados a la transmisión de vídeo a través de la interfaz radio.

4.2.- SERVICIOS DE VÍDEO

4.2.1.- INTRODUCCIÓN

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 [68]. En la actualidad existen varios estándares de compresión de vídeo aunque sus características y modo de funcionamiento son bastante similares [69]. Los métodos de codificación estandarizados están dominados por dos familias [70]: la familia H.26x de la ITU (vídeo) y la familia MPEG de la ISO (vídeo y sonido). El método de codificación H.263 [71], que surgió con el objetivo de mejorar las prestaciones ofrecidas por su predecesor [72] (el estándar H.261), permite relaciones de compresión del orden de 100:1 o incluso 200:1 con una calidad de imagen aceptable [73], reduciendo la velocidad binaria requerida a valores en el rango de 16 a 300 kbps aproximadamente [74]. La posibilidad de transmitir con estas velocidades tan reducidas hace que la codificación H.263 sea sumamente atractiva para la transmisión de vídeo en un sistema de comunicaciones móviles [75], motivo por el que se ha considerado aquí.

4.2.2.- GENERALIDADES

Antes de describir el proceso de codificación H.263 conviene comentar de forma breve dos criterios en función de los cuales se puede clasificar de forma genérica una técnica de codificación [76]. El primero distingue entre codificación sin pérdidas o con pérdidas. En el primer tipo de codificación (también denominado reversible) la imagen de vídeo a la salida del decodificador es idéntica, píxel a píxel, a la imagen de vídeo a la entrada del codificador. En el segundo tipo de codificación (también denominado irreversible), la imagen decodificada no es idéntica, píxel a píxel, a la imagen original.

El segundo criterio distingue entre técnicas de codificación con tasa binaria constante (CBR, *Constant Bit-Rate*) o variable (VBR, *Variable Bit-Rate*). En la codificación CBR se mantiene constante la velocidad binaria a la salida del codificador de vídeo a costa de reducir la calidad en aquellas imágenes que requieren una tasa binaria superior a la que se desea mantener. Por el contrario, en la codificación VBR la tasa binaria varía a lo largo del tiempo ya que cada imagen de vídeo se codifica con tantos bits como sean necesarios para mantener constante la calidad de imagen a lo largo de la secuencia de vídeo. Este último tipo de codificación ofrece una mayor calidad que el primero pero requiere mecanismos de gestión de recursos más complejos [76].

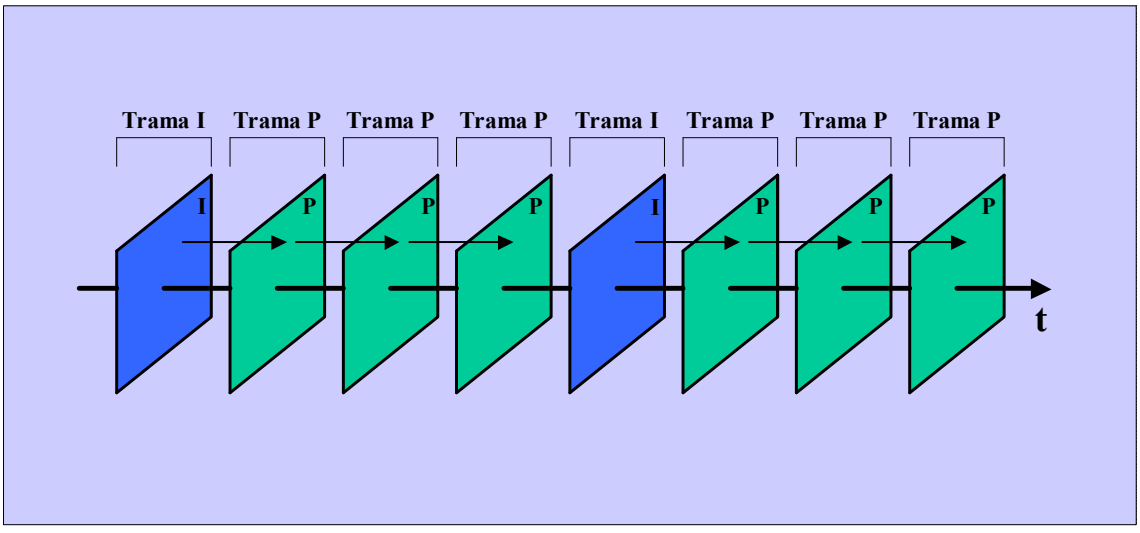
El método descrito en el estándar ITU-T H.263 es un método de codificación con pérdidas o irreversible, ya que la compresión requerida para la transmisión de vídeo se consigue eliminando parte de la información de la secuencia de vídeo original. En el ámbito práctico, existen implementaciones de codificadores H.263 tanto de tasa binaria constante (CBR) como variable (VBR).

4.2.3.- FUNDAMENTOS DE CODIFICACIÓN H.263

Una película de vídeo no es más que una secuencia de imágenes proyectadas a una velocidad lo suficientemente elevada como para que el ojo humano no detecte el cambio de una imagen a otra y perciba la sensación de movimiento. La misión de un codificador de vídeo es, por tanto, la de codificar cada una de las imágenes que componen dicha secuencia. Sin embargo, todas las imágenes no son procesadas de la misma forma por el codificador.

Un codificador de vídeo H.263 tiene dos modos básicos de funcionamiento [71]: modo *Intra* y modo *Inter*. Las imágenes codificadas en el primer modo se denominan imágenes I, mientras que aquellas codificadas en el segundo modo reciben el nombre de imágenes P [71]. Las imágenes de tipo I son aquellas en las que se codifica toda la información de vídeo contenida en ellas, de forma que pueden ser reconstruidas por completo de forma independiente; se dice en este caso que la imagen original se codifica sin referencia a ninguna otra. Por el contrario, las imágenes de tipo P son aquellas que se codifican haciendo referencia a la imagen predecesora. Cuando el codificador funciona en modo *Inter* tan sólo se codifica la diferencia entre la imagen actual y la imagen previa. Las zonas de la imagen que no varían respecto a la imagen anterior no son codificadas, ya que el receptor posee esa información.

La información de vídeo codificado se agrupa, junto con otra información adicional necesaria para reconstruir la imagen en el receptor, en entidades que reciben el nombre de tramas[33]. Dependiendo del tipo de imagen que contengan, se denominan tramas I o tramas P.



I: Imagen de tipo I (se codifica sin referencia a ninguna otra imagen) **P:** Imagen de tipo P (se codifica con referencia a la imagen anterior)

Figura 4.1.- Relación de dependencia entre imágenes de tipo I y P

[33] No confundir el concepto de *trama de vídeo* (entidad que contiene una imagen de vídeo codificada) con el de *trama TDMA* o PDU (*Protocol Data Unit*) a nivel de enlace de datos.

Este procedimiento de codificación tiene ventajas pero también inconvenientes. Por una parte, resulta obvio que las tramas P tendrán un tamaño significativamente menor que las tramas I debido a que se codifican de forma diferencial, sin considerar toda la información que contiene la imagen original. Esto significa que tras transmitir la primera trama de vídeo, forzosamente de tipo I, si el resto de tramas transmitidas son exclusivamente de tipo P entonces es posible lograr un importante ahorro en ancho de banda respecto al caso en el que se transmiten más tramas de tipo I además de la primera. De esta forma se podría transmitir una secuencia de vídeo sin necesidad de enviar toda la información que la compone, incrementando así el aprovechamiento de los recursos radio.

En una transmisión de vídeo los errores provocados por el canal de transmisión pueden llegar a tener consecuencias bastante serias [77, 78], siendo la más evidente la propagación de errores entre imágenes consecutivas. Efectivamente, una imagen P puede ser decodificada de forma incorrecta, incluso habiéndose recibido la trama P correspondiente sin error, en caso de que la trama anterior hubiera sufrido un error de transmisión, ya que una imagen P se codifica (y decodifica) haciendo referencia a la imagen predecesora. Así pues, el error no sólo afecta a la imagen corrupta sino también a la imagen siguiente, la cual servirá, a su vez, de referencia para la que se decodifique a continuación, con lo que el error se propagará nuevamente. Un error de transmisión en una imagen de vídeo puede propagarse durante tiempo suficiente como para que el usuario perciba un efecto desagradable [77]. La acumulación de errores a lo largo de la secuencia de vídeo da como resultado una calidad final bastante pobre para el usuario, aun cuando los errores individuales experimentados sean poco significativos.

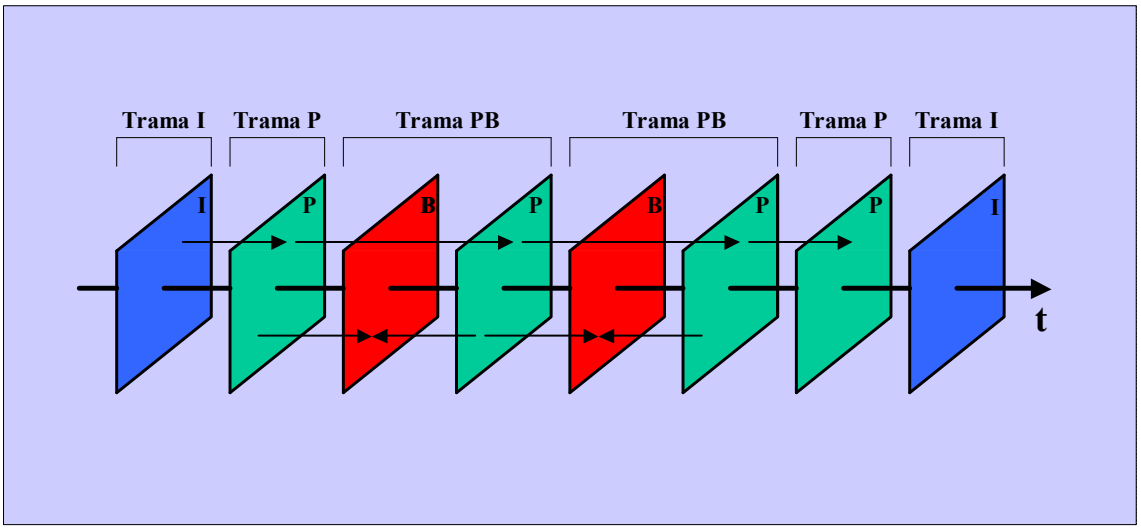
La solución, como fácilmente podrá imaginarse, consiste en insertar tramas I a intervalos regulares de tiempo a lo largo de la secuencia de vídeo. De esta forma, si ocurre un error en alguna de las tramas transmitidas y éste afecta a varias tramas P consecutivas, la distorsión que el error produce en la imagen se detendrá y dejará de propagarse en el momento en que se reciba una trama I, ya que una imagen I se codifica (y decodifica) de forma independiente, es decir, sin hacer referencia a ninguna otra imagen. El error de imágenes predecesoras que pueda acumularse en una imagen concreta no influirá sobre el proceso de decodificación de la imagen I que se reciba a continuación.

Resulta obvio que una trama I posee un tamaño mayor que una trama P, ya que corresponde a una imagen de vídeo que se codifica íntegramente, es decir, incluyendo toda la información de la imagen original, lo cual tiene una implicación bastante importante: al aumentar la *frecuencia de refresco* de las tramas I, si bien es cierto que se mejora la calidad de imagen que el usuario de vídeo experimenta porque se detiene la propagación de errores, también es cierto que se incrementa el ancho de banda requerido para la transmisión de la secuencia de vídeo debido al mayor tamaño de estas tramas. Así pues, queda claro que la variación de la frecuencia de refresco de tramas I de una secuencia de vídeo modifica el grado de compromiso existente entre la calidad de imagen experimentada por el usuario y el ancho de banda requerido para la transmisión de dicha secuencia.

4.2.4.- MODOS OPCIONALES H.263

La recomendación H.263 contempla la posibilidad de aplicar cuatro modos opcionales a la codificación de vídeo H.263 básica [71]: modo vector de movimiento sin restricción, modo de codificación aritmética basada en sintaxis, modo de predicción avanzada y modo de tramas PB. Se omite aquí una descripción de todos estos modos opcionales por resultar excesivo e innecesario. En [73] se halla una descripción bastante detallada. Sí se tratará aquí el último modo opcional citado, el modo de tramas PB, al ser utilizado por el modelo de vídeo empleado en las simulaciones.

El modo de tramas PB permite incrementar la tasa de imágenes por unidad de tiempo sin un aumento significativo de la tasa binaria requerida [72], ya que en la misma trama se transmite la información relativa a dos imágenes de vídeo codificadas. Una trama PB consiste en una imagen P y una imagen B. Las imágenes B son aquellas que, al igual que las imágenes P, se codifican haciendo referencia a otras imágenes. La diferencia está en que la imagen B se codifica haciendo referencia, no sólo a la imagen anterior en la secuencia de vídeo, sino también a la imagen posterior, según se ilustra en la figura 4.2.



I: Imagen de tipo I (se codifica sin referencia a ninguna otra imagen)
P: Imagen de tipo P (se codifica con referencia a la imagen anterior)
B: Imagen de tipo B (se codifica con referencia a la imagen anterior y posterior)

Figura 4.2.- Relación de dependencia entre imágenes de tipo I, P y B

Para generar una trama PB es necesario que el receptor posea las dos imágenes que se van a codificar en dicha trama. La información codificada de la imagen B y la de la imagen P es entrelazada a nivel de macrobloque (ver anexo A) sobre la trama PB, de tal manera que a cada macrobloque de la imagen P le sucede inmediatamente el macrobloque de la imagen B correspondiente a la misma posición dentro de la imagen. En el decodificador, primero se reconstruye el macrobloque de la imagen P y luego el de la imagen B, tomando como referencia para ello el macrobloque obtenido de la imagen P y el correspondiente a la imagen anterior. Cuando concluye la reconstrucción de la trama PB entonces se muestra primero la imagen B y, a continuación, la imagen P.

4.3.- CALIDAD DE SERVICIO PARA VÍDEO H.263

Existen diferentes nociones del concepto de calidad de servicio (QoS, *Quality of Service*) dependiendo del nivel considerado [80]. En la capa física, la calidad de servicio es sinónimo de una relación CIR o un BER en el receptor aceptables. En la capa de enlace de datos la calidad de servicio se suele expresar a menudo en términos de BLER (*Frame Error Rate*) o FER (*Frame Error Rate*) y, en la capa de red, en función del PER (*Packet Error Rate*). Para capas más altas, la calidad de servicio se percibe habitualmente como un requisito de throughput mínimo o retardo máximo. Según las especificaciones técnicas de GPRS [81], el concepto de calidad de servicio se define como el "*efecto colectivo de los rendimientos de un servicio que determina el grado de satisfacción de un usuario del servicio*" o "*el conjunto de parámetros de rendimiento que se pueden observar y medir directamente en el punto en el que el servicio es accedido por el usuario*". Esta definición lleva implícita la consideración de que es el usuario quien determina si la calidad de servicio percibida es o no satisfactoria.

El estándar GPRS [81] contempla, en su definición de calidad de servicio, varios parámetros de calidad. Este conjunto de parámetros no son suficientemente adecuados para evaluar la transmisión de vídeo en el presente trabajo. En primer lugar, los valores indicados para cada parámetro corresponden a valores medidos extremo a extremo (como el retardo); sin embargo, el interés de este trabajo se centra sobre una técnica de gestión de recursos radio, cuyo ámbito de aplicación no es todo el tramo de red que atraviesa la información desde la fuente hasta el receptor sino tan sólo la interfaz radio. Por otra parte, el tipo de servicio aquí considerado posee unos requisitos de calidad de servicio más exigentes que los habituales para otros servicios de datos, razón por la cual es preferible emplear como referencia otros parámetros más específicos del tipo de aplicación considerada, es decir, vídeo H.263.

Una medida de calidad ampliamente empleada en estudios relacionados con la transmisión de vídeo es la relación PSNR (*Peak Signal-to-Noise Ratio*). El valor concreto de este parámetro depende de multitud de factores tales como la resolución de vídeo (sub-QCIF, QCIF, CIF, 4CIF o 16 CIF), la tasa de tramas de vídeo por segundo o el tamaño de las tramas (cuanto mayor sean las tramas menor debe ser el valor de PSNR para ofrecer una misma calidad al usuario). Por ello, esta medida es subjetiva y su valor absoluto no resulta realmente representativo de la calidad experimentada en la transmisión de vídeo. Para evaluar de forma objetiva la calidad de una transmisión de vídeo, lo que verdaderamente interesa no es el valor absoluto de la relación PSNR obtenido en el receptor sino la degradación en la relación PSNR ocasionada por las pérdidas en el canal de transmisión, es decir, la diferencia entre la relación PSNR que se obtendría en el receptor en el supuesto caso de una transmisión ideal sin errores y la que realmente se obtiene, que resulta ser inferior a causa de los errores. Según se deriva de los estudios realizados en [82, 83], si la degradación de la relación PSNR es inferior a 1 dB, entonces la distorsión en la imagen de vídeo no es perceptible por el usuario, en cuyo caso se puede considerar que la calidad final de vídeo es aceptable. Como se demuestra en [82, 83], la degradación en la relación PSNR será inferior a un 1 dB siempre que el valor del BLER[34] experimentado en la transmisión sea inferior al 5%.

[34] A menudo, en la literatura suelen emplearse indistintamente los términos FER (*Frame Error Rate*) y PLR (*Packet Loss Ratio*) para referirse a lo que en este trabajo se designa mediante el término BLER (*BLock Error Rate*).

Es cierto que estos resultados se obtienen al emplear un sistema de vídeo en el cual los autores introducen ciertas mejoras; sin embargo, es la única referencia existente en la literatura en la que se proporciona una indicación similar para el caso en que se considera la transmisión de vídeo H.263. Por tanto, consideraremos esta medida como referencia en la evaluación de los algoritmos estudiados en este trabajo.

Por lo que respecta a las especificaciones técnicas, en [84] se señala el valor de la tasa de error que, desde el punto de vista del usuario, resulta tolerable para diversos tipos de aplicación. En el caso de vídeo se indica que para la transmisión en modo conversacional (tiempo real) el BLER máximo tolerable es del 1%, siendo del 2% para la transmisión en modo *streaming*. El estándar de vídeo considerado no se especifica, razón por la cual las indicaciones se consideran genéricas y, por tanto, aplicables al presente trabajo como referencia adicional en la evaluación de los algoritmos tratados.

Un último parámetro en función del cual puede expresarse la calidad requerida por una transmisión de vídeo es la proporción de bits erróneos en el receptor o BER. En lo referente a este parámetro, y dada la inexistencia de referencias en las cuales se indique el margen de BER máximo que una transmisión de vídeo H.263 puede tolerar para que su calidad pueda considerarse aceptable, se decidió finalmente considerar en este trabajo las indicaciones ofrecidas en el estándar [85] para transmisión de vídeo MPEG-4, donde se indica que para un BER igual a 10^{-6} la degradación no es perceptible por el usuario, un BER de 10^{-5} origina un ligero deterioro visible y un BER de 10^{-4} da lugar a varios artefactos perceptibles, siendo la aplicación práctica limitada cuando el BER excede el valor 10^{-3}. Es cierto que la tasa de errores admisible depende del códec concreto que se emplea. Aun así, cuando se evalúa la calidad de servicio en función de la capacidad de apreciación humana y no en función de otros parámetros relativos al sistema (como la relación PSNR, por ejemplo), es habitual que los resultados obtenidos sean aplicables a diversos códecs de vídeo. La especificación técnica [86] constituye un claro ejemplo de este caso.

4.4.- SUMARIO

En la introducción del presente capítulo se definió el concepto de servicio de tiempo real, caso analizado en este trabajo, poniendo de relieve el importante papel que juega el retardo en la calidad final experimentada por el usuario para este tipo de servicios. Puesto que para el trabajo aquí realizado se ha considerado la transmisión de vídeo usando el estándar H.263, en este capítulo se han descrito algunos principios básicos relativos a H.263 que resultan necesarios para los estudios sobre transmisión de vídeo descritos más adelante. También se ha pretendido dar una visión global, pero suficiente para los objetivos de este trabajo, del concepto de calidad de servicio. Puesto que los requisitos de calidad expuestos en las especificaciones GPRS no resultan lo suficientemente adecuados para la correcta evaluación de la transmisión de vídeo tal y como se estudia en el presente trabajo, el capítulo concluye con diversas indicaciones de calidad específicas del tipo de aplicación que se considera, es decir, transmisión de vídeo H.263, algunas de las cuales se han tomado de otros estándares de compresión de vídeo por no existir referencias específicas para H.263.

Capítulo 5

Modelos y Herramientas de Simulación

5.1.- INTRODUCCIÓN

Existen diversas formas de llevar a cabo el estudio de una investigación y evaluar los resultados [87]. Una primera forma consistiría en la implementación sobre un sistema real de telefonía móvil y la realización de medidas sobre dicho sistema, lo cual supone que todos los factores intervienen sobre los resultados finales y éstos son completamente fiables, permitiendo obtener una evaluación objetiva. Sin embargo, disponer de un sistema real para realizar experimentos supone un coste muy elevado y resulta complicado en la práctica cuando no imposible, haciendo poco atractiva esta alternativa.

Una segunda forma consiste en emplear modelos matemáticos analíticos. La ventaja de los modelos matemáticos reside en el hecho de que es posible establecer, mediante fórmulas matemáticas, una relación clara y directa entre los diversos parámetros que determinan el comportamiento de un sistema así como su grado de relevancia en la determinación de dicho comportamiento. Sin embargo, el método analítico requiere bastantes suposiciones y aproximaciones (lo cual limita su fiabilidad y exactitud) y no permite obtener una solución de forma sencilla. La creciente complejidad de los sistemas de comunicaciones móviles aumenta la dificultad de estudiar nuevas técnicas mediante métodos analíticos.

Una tercera forma consiste en la simulación por ordenador: los elementos y el funcionamiento de un sistema se modelan mediante software que se ejecuta en una computadora. Es cierto que este método necesita un elevado número de simulaciones para que los resultados obtenidos posean valor estadístico y sean representativos del comportamiento real de un sistema, lo cual requiere un elevado tiempo de simulación. Sin embargo, tiene la ventaja de que requiere muy pocas suposiciones en comparación con el análisis matemático (dependiendo de la complejidad del modelo de simulación), lo cual garantiza una mayor exactitud y fiabilidad de los resultados obtenidos y, por tanto, unos resultados más representativos de la eficacia de la técnica estudiada. Esta alternativa supone un buen compromiso entre coste, tiempo necesario, precisión y complejidad.

Los estudios realizados en este trabajo se llevan a cabo mediante simulación por ordenador, por lo que el presente capítulo estará dedicado a la presentación de los modelos empleados así como de las herramientas utilizadas para llevar a cabo dicha simulación.

El estudio por simulación de un sistema de comunicaciones móviles suele realizarse mediante simulaciones a dos niveles diferentes: a nivel de enlace y a nivel de sistema. En el nivel de sistema se consideran aspectos relativos al comportamiento de una red celular tales como, por ejemplo, la movilidad de los usuarios, sus relaciones de interferencia, la atenuación o el efecto sombra. Los resultados de la simulación a este nivel se expresan en términos de CIR medio. Sin embargo, no son completamente representativos de la calidad percibida en el enlace radio ya que no reflejan los efectos del desvanecimiento rápido debido al efecto multicamino.

Efectivamente, el desvanecimiento rápido origina variaciones abruptas en la relación CIR instantánea durante el proceso de transmisión. Es aquí donde aparece la simulación a nivel de enlace, en la cual se consideran aspectos de la transmisión radio como la codificación de canal, la modulación, y los efectos del canal de transmisión. A nivel de enlace, el comportamiento de la transmisión se modela a nivel de bit, lo cual sí permite tener en cuenta los efectos del desvanecimiento rápido. Así pues, la principal razón para desglosar la simulación en estos dos niveles es la diferente resolución de tiempo requerida: a nivel de sistema se consideran aspectos que varían en una escala de tiempos amplia[35] mientras que a nivel de enlace la escala temporal es más reducida (el bit), permitiendo así considerar los efectos del desvanecimiento rápido. Ambos simuladores trabajan de forma independiente y se enlazan entre sí mediante unas interfaces que permiten el análisis global del comportamiento del sistema.

5.2.- SIMULACIÓN A NIVEL DE ENLACE

Para la realización de las interfaces empleadas este trabajo se utilizó el mismo simulador a nivel de enlace que en [88]. En dicho simulador se implementan diversos aspectos relacionados con el proceso de transmisión radio. Además de la codificación de canal, también se implementa el formato y conformado de una ráfaga normal así como el proceso de modulación GMSK (en el transmisor), los efectos del ruido térmico y la interferencia cocanal de otros emisores (en el canal de transmisión) y el filtrado pasobanda, la demodulación (con correlador y ecualizador) y la descomposición de ráfaga (en el receptor). En este apartado no se proporcionarán más detalles acerca de la implementación y funcionamiento del simulador a nivel de enlace ya que para este trabajo no se han realizado modificaciones en este simulador, utilizando en el estudio los resultados obtenidos en [88]. Para una descripción más detallada, consultar [88].

Sin embargo, sí resulta relevante qué resultados son los que proporciona esta simulación, cómo se obtienen y cómo son utilizados en la simulación a nivel de sistema. Una simulación a nivel de enlace se realiza empleando un esquema de codificación de canal determinado y para un valor de CIR, velocidad del terminal móvil y canal radio (entorno urbano, rural, etc.) concretos. Bajo estas condiciones, se simula la transmisión radio de un número de bloques RLC lo suficientemente elevado como para que los resultados sean estadísticamente representativos. Esta transmisión se estudia ráfaga a ráfaga, evaluando los errores de bit introducidos en cada una de ellas por efecto del ruido térmico y la interferencia de otros emisores cocanal. De esta forma es posible obtener unas estadísticas de error que caracterizan el proceso de transmisión en función de la calidad del canal. Así es como se ha obtenido la figura 3.5. En dicha figura, cada uno de los puntos que conforman las curvas BLER en función de la relación CIR es el resultado obtenido en una simulación a nivel de enlace bajo un entorno TU50 (*Typical Urban*, entorno urbano, donde el móvil se desplaza a 50 km/h) con la relación CIR y el esquema de codificación correspondiente a cada caso.

[35] Se considera como unidad de tiempo aquella en la que se produce un cambio a nivel de sistema como, por ejemplo, la actualización del valor medio de la relación CIR.

Una vez realizadas las simulaciones pertinentes y habiendo obtenido una correspondencia entre calidad del canal de transmisión (expresada, por ejemplo, en términos de relación CIR) y la tasa de error asociada, estos valores son almacenados y utilizados en la simulación a nivel de sistema. En esta simulación, para la transmisión de cada ráfaga el transmisor conocerá la calidad del enlace radio (CIR) y, consultando los resultados de la simulación a nivel de enlace, obtendrá un valor de tasa de error en función del cual se determinará si la información transmitida ha sido recibida con error o no en el receptor. La forma exacta en que interactúan estos dos niveles requiere el uso de interfaces, lo cual será descrito más adelante.

Para cerrar este apartado es conveniente recordar que el hecho de que la simulación a nivel de enlace se modele a nivel de bit permite tener en cuenta los efectos del desvanecimiento rápido (*fast fading*) durante la transmisión en la interfaz radio.

5.3.- SIMULACIÓN A NIVEL DE SISTEMA

Existen tres tipos principales de simulación a nivel de sistema: continua, Monte Carlo y de eventos discretos. En la simulación continua las variables mediante las cuales se estudia el comportamiento del sistema varían constantemente a lo largo del tiempo y, por tanto, han de ser controladas de forma continuada. En las simulaciones Monte Carlo el tiempo no juega un papel importante y el estado del sistema es evaluado en instantes de tiempo aleatorios e independientes. En aquellos casos en los que la variable temporal es un factor a considerar pero no se hace necesaria una evaluación continua del sistema se emplea una simulación de eventos discretos. En este tipo de simulaciones, el sistema es evaluado en aquellos instantes temporales en los cuales se produce un cambio en alguna de sus variables. Dicho cambio recibe el nombre de evento. Así pues, en un sistema de comunicaciones móviles un evento podría ser, por ejemplo, el cambio de posición de un móvil dentro de una celda, el inicio de la transmisión de información o la solicitud de una retransmisión de datos.

Puesto que *Link Adaptation* (LA) es una técnica que adapta el nivel de protección contra errores en los datos a transmitir en función de la calidad del enlace radio y la transmisión a nivel de sistema se modela a nivel de ráfaga, no es necesaria información de la calidad del canal de forma continuada, sino que dicha información (CIR) es requerida sólo una vez para cada ráfaga. Por este motivo, la simulación por eventos discretos es la más adecuada para el estudio a nivel de sistema de LA y la que ha sido considerada en el simulador empleado.

El simulador a nivel de sistema [88] se ha implementado utilizando el lenguaje C++ y la librería CNCL (*Communications Networks Class Library*). CNCL [89] es una librería de clases escrita en C++ pensada para proveer funciones de simulación comunes a cualquier red de comunicaciones. Incluye funcionalidades para la generación de números aleatorios, análisis estadístico de resultados, creación de objetos contenedores como listas, *arrays*, colas o pilas y la simulación controlada por eventos. El elemento principal de esta librería es el *scheduler*, entidad central de control encargada de ordenar los eventos en función de su instante de ejecución y su prioridad así como de entregarlos a los manejadores de eventos asociados para su gestión.

5.3.1.- ARQUITECTURA DEL SIMULADOR

El simulador a nivel de sistema utilizado para el análisis de LA en este trabajo ha sido implementado siguiendo la estructura mostrada en la figura 5.1. En los subapartados siguientes se da una descripción de los elementos que la integran.

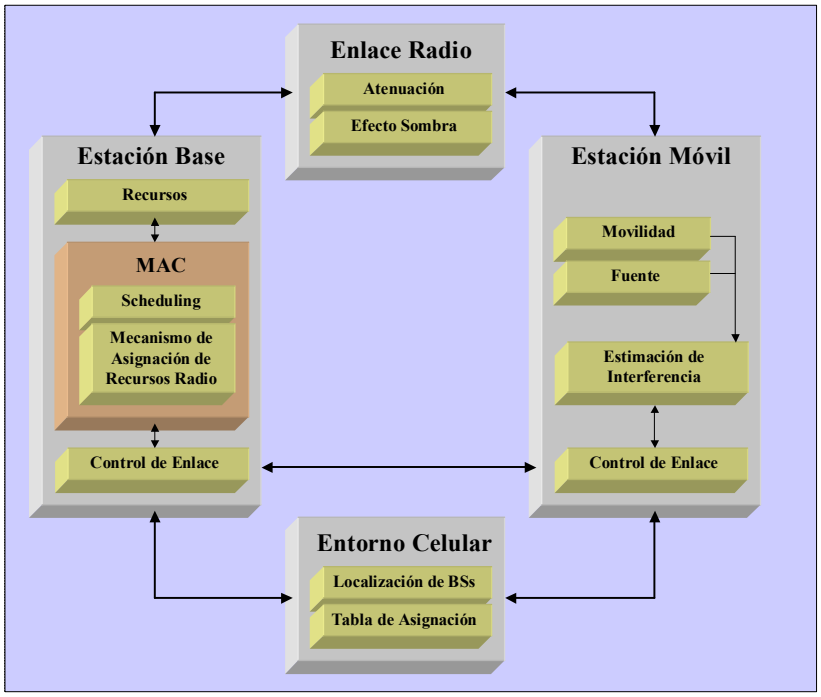


Figura 5.1.- Arquitectura del simulador a nivel de sistema

5.3.1.1.- ENTORNO CELULAR

En la entidad de *Entorno Celular* se almacena información relativa a la localización de las estaciones base así como una lista de los recursos asignados en cada una de ellas. Esta información, almacenada en el simulador en forma de tablas, resulta fundamental para la estimación de la calidad del canal de transmisión (CIR) por parte de cada uno de los terminales móviles que se hallan dentro del sistema simulado. En la estimación de la relación CIR interviene el nivel de señal que una estación móvil concreta recibe de la estación base transmisora y de las estaciones cocanal interferentes. En este cálculo suele ser habitual considerar únicamente, a efectos de interferencia, aquellas estaciones base cocanal más cercanas a la estación base a la cual se encuentra conectado el terminal móvil estudiado (ver figura 3.3), es decir, aquellas estaciones base cocanal cuya separación respecto a la estación base considerada es igual a la distancia de reutilización. Este conjunto de estaciones base recibe el nombre de primera corona. Sin embargo, tal y como se muestra en [90], el comportamiento dinámico de la relación CIR también se ve afectado por el nivel de señal recibido de las estaciones base que conforman la segunda corona, es decir, aquellas estaciones base cocanal alrededor de la estación base considerada cuya separación es (aproximadamente) igual a dos veces la distancia de reutilización. Así pues, y dado que el comportamiento de LA depende directamente de las variaciones en la calidad del canal de transmisión, en el simulador se han considerado, para los cálculos de la relación CIR, las estaciones base cocanal que conforman tanto la primera como la segunda corona.

5.3.1.2.- ENLACE RADIO

En la entidad *Enlace Radio* se modelan aquellos mecanismos del canal de transmisión que afectan a largo plazo a la señal transmitida, es decir, la atenuación y el efecto sombra. En el apartado 3.3 se vio que existe un tercer mecanismo, el efecto multicamino, que también afecta a la señal transmitida a través del medio radio. Ya se ha comentado anteriormente que estas variaciones se dan a corto plazo y por ello no han sido modeladas a nivel de sistema sino a nivel de enlace. No obstante, sus efectos también se tienen en cuenta en la simulación a nivel de sistema mediante el uso de interfaces entre ambos niveles que se verán más adelante.

Para estudiar el efecto de la atenuación, se ha implementado el modelo de propagación empírico Okumura-Hata [2, 23]. Este modelo fue obtenido por Hata en base a un amplio conjunto de medidas realizadas por Okumura en Japón. Según este modelo, la atenuación en dB viene dada por la expresión:

$$\begin{aligned} L[dB] &= 69.55 + 26.16 \cdot \log(f_c) - 13.82 \cdot \log(h_b) - \\ &- 3.2 \cdot \left[\log(11.75 h_m)\right]^2 + 4.97 + \left[44.9 - 6.55 \cdot \log(h_b)\right] \cdot \log(d) \end{aligned} \tag{5.1}$$

donde f_c es la frecuencia portadora en MHz, h_b es la altura de la estación base en metros, h_m es la altura de la estación móvil en metros y d es la distancia entre estación base y estación móvil expresada en kilómetros. El conjunto de medidas a partir del cual se determinó esta expresión se realizó bajo las siguientes condiciones:

a) Frecuencia portadora: $150\,MHz \le f_c \le 1500\,MHz$.

b) Altura de la estación base: $30\,m \le h_b \le 200\,m$.

c) Altura de la estación móvil: $1\,m \le h_m \le 10\,m$.

d) Distancia entre estación base y estación móvil: $1\,km \le d \le 20\,km$.

El sistema simulado está compuesto por celdas circulares cuyo radio es igual a 1 km, es decir, la distancia entre estación base y estación móvil es siempre igual o menor a 1 km. Sin embargo, tal y como se indica en [91], aunque el estudio original fue realizado para distancias entre estación base y estación móvil superiores a 1 km, la expresión (5.1) obtenida como resultado de dicho estudio sigue siendo válida y aplicable también para distancias inferiores a 1 km.

El efecto sombra ha sido modelado mediante un proceso aleatorio espacialmente correlado tal que el valor de las pérdidas debidas al efecto sombra que experimenta un terminal móvil en una posición concreta está correlado con el experimentado en una posición cercana. El valor numérico que representa las pérdidas por efecto sombra y que es restado al valor de la potencia transmitida se obtiene de forma aleatoria para cada posición del terminal móvil dentro de la celda. La correlación entre los valores relativos a posiciones cercanas observada en el ámbito real ha sido considerada en la modelización de este mecanismo de forma tal que los valores de pérdidas por efecto sombra para posiciones próximas, aun siendo números aleatorios, respetan esta correlación. Se omite una descripción más detallada acerca de la implementación por carecer aquí de relevancia. Para más detalles consultar [88].

5.3.1.3.- ESTACIÓN BASE

En el sistema implementado, la transmisión es simulada sólo en sentido *downlink* (estación base a estación móvil). En un sistema de comunicaciones móviles real es cierto que se transmite tanto en *downlink* como en *uplink*. Sin embargo, en el presente trabajo, y dado que en el ámbito real el flujo de información en *downlink* es mucho mayor que en *uplink*, se eligió estudiar el *downlink* por representar la situación más conflictiva y problemática (en *uplink* suelen transmitirse solicitudes para recibir páginas web, correos electrónicos o una película de vídeo, las cuales requieren una cantidad de recursos bastante menor que la propia información transmitida en *downlink*).

Por ello, es en la estación base donde se implementan las funciones de control de acceso al medio (MAC, *Medium Access Control*), responsables del control de la congestión o *scheduling* y de la asignación de canales radio. Dependiendo de la técnica de scheduling empleada, los paquetes de datos de un usuario sufrirán un retardo de cola (*queueing delay*[36]) concreto que repercutirá sobre el retardo global de la información transmitida. Sin embargo, el retardo relevante para este trabajo es el retardo de transmisión en la interfaz aire, por lo que el retardo de cola no es un factor importante y la elección de una técnica de scheduling u otra es indiferente. Es por ello que se decidió implementar la técnica FCFS (*First-Come First-Served*) por su enorme sencillez y simplicidad: las solicitudes de asignación de recursos a cada usuario y comienzo de la transmisión por la interfaz aire se satisfacen en el mismo orden en que surgen, sin mecanismos de reordenación en función de la prioridad o retardo máximo admisible.

En cuanto al mecanismo de asignación de recursos radio, aunque no forma parte de los factores estudiados en este trabajo, es cierto que juega un papel importante sobre el comportamiento de LA, ya que el criterio empleado a la hora de asignar recursos radio repercute sobre la calidad del canal de transmisión[37]. Se ha escogido la técnica de asignación de canales aleatoria RCA (*Random Channel Allocation*), en la cual a la hora de asignar un canal a un usuario simplemente se escoge uno aleatoriamente de entre los disponibles en ese momento. La justificación de esta elección reside en los buenos resultados ofrecidos por esta técnica [98] y en la sencillez de su implementación.

Por último, dentro de la entidad *Estación Base*, aparecen dos elementos más: *Recursos* y *Control de Enlace*. El primero no es más que una lista actualizada de los recursos radio tanto asignados como libres en la estación base, mientras que el segundo se encarga de supervisar diversos aspectos del funcionamiento de LA como, por ejemplo, la liberación del canal radio una vez finalizada la transmisión. Es en este elemento donde se implementa el algoritmo LA.

[36] Tiempo transcurrido desde que un paquete llega al búfer de transmisión hasta que comienza a ser transmitido.

[37] La asignación a un usuario de un canal con una calidad baja puede suponer un incremento de la interferencia sobre el resto de usuarios del sistema y, por tanto, una disminución de la calidad de transmisión experimentada por los demás terminales móviles. El comportamiento de LA depende de la calidad del enlace radio, por lo que la elección de la técnica de asignación canales repercute sobre el funcionamiento de LA.

5.3.1.4.- ESTACIÓN MÓVIL

En la entidad *Estación Móvil* también se implementa el *Control de Enlace*, análogo al implementado en la entidad *Estación Base* y encargado de las mismas labores. Otro elemento contemplado dentro de la *Estación Móvil* es la *Estimación de Interferencia*. En el simulador a nivel de sistema, cada estación móvil es la encargada de calcular la calidad de su propio canal de transmisión, expresada en términos de relación CIR. La relación CIR recordemos que es el cociente entre el nivel de potencia recibido de señal útil (C) y el nivel de potencia recibido de otras señales en la misma frecuencia (cocanales) denominadas interferentes (I), por lo que para su estimación es necesario estimar el valor de C así como el de I.

El valor de C vendrá dado por la potencia de transmisión de la estación base (recordemos que se simula el *downlink*) y por la distancia que separa la estación móvil de su estación base. Dicha distancia vendrá determinada por la ubicación física dentro del sistema de la estación base así como de la estación móvil. La situación de la estación base se obtiene consultando la tabla *Localización de BSs* almacenada en la entidad *Entorno Celular*, mientras que la posición de la estación móvil en cada instante vendrá determinada por el modelo de movilidad que se comenta más adelante. A partir de la distancia de separación entre estación base y estación móvil se puede determinar el valor de atenuación sufrido por la señal transmitida, mientras que dependiendo de la posición de la estación móvil dentro de la celda se obtendrá un valor concreto para las pérdidas por efecto sombra. Restando ambos valores a la potencia transmitida por la estación base se obtiene el nivel de señal útil (C) recibido por la estación móvil.

El valor de I vendrá dado por la potencia de transmisión de las estaciones base cocanal pertenecientes a la primera y segunda corona y por la distancia entre la estación móvil y cada una de estas estaciones base. La posición de cada elemento, así como el valor de la atenuación y de las pérdidas por efecto sombra se calcula por separado para cada estación base de forma análoga a como se realiza para obtener el valor de C. Una vez obtenido el valor de señal recibido de cada estación base interferente se suman todos ellos para obtener finalmente el valor de I.

El comportamiento de LA está determinado por la calidad experimentada en el enlace radio. Puesto que dicha calidad depende, entre otros muchos factores, de la posición relativa de la estación móvil y ésta varía con el tiempo en un sistema real, es necesaria la inclusión de un modelo de movilidad en el simulador. En dicho modelo [88], la posición inicial de cada estación móvil se establece de forma aleatoria de acuerdo a una distribución uniforme. A fin de poder integrar la movilidad (carácter continuo) dentro del simulador (controlado por eventos discretos), el movimiento del móvil viene determinado por una secuencia de posiciones aleatorias en instantes discretos de tiempo. La ubicación intermedia entre dos posiciones, en aquellos casos en que sea necesaria, se obtiene mediante interpolación lineal. La dirección del movimiento se establece añadiendo, en cada posición, un ángulo aleatorio respecto a la dirección previa. Dicho ángulo pertenece a una distribución normal con media igual a cero y cuya varianza depende de la velocidad del terminal móvil en el instante en cuestión. Este modelo de movilidad es bastante realista al demostrar ser consistente con datos reales [88].

5.3.2.- MODELOS DE TRÁFICO

En principio, la unidad de tráfico de fuente debería haber sido implementada en la entidad *Estación Base*, ya que la simulación se realiza en *downlink* y, por tanto, la estación base es el elemento transmisor dentro del sistema simulado. Al ser en la estación base donde se generan los paquetes de datos (cuyas características están determinadas por el modelo de tráfico empleado) es aquí donde hubiera resultado más lógico implementar el elemento *Fuente*, encargado de generar el tráfico. Sin embargo, y por motivos de simplicidad en la implementación del simulador, la unidad de tráfico se incluye en la entidad *Estación Móvil*. De esta forma, cada estación móvil genera el tráfico que recibe en lugar de ser la estación base la que se encarga de ello para cada una de las estaciones móviles. Esta forma de proceder, contraria a como sucede en la realidad, no afecta sin embargo al funcionamiento global del sistema simulado y permite un simulador con una estructura más ordenada.

Existen dos procedimientos habituales en la literatura para modelar el tráfico de una red de comunicaciones. Una primera aproximación consiste en monitorizar y capturar una sesión de tráfico en una red real. Posteriormente, durante la simulación, se reproduce la sesión previamente monitorizada. Una segunda aproximación consiste en intentar encontrar un modelo de distribución de probabilidad que describa con aceptable verosimilitud cada uno de los parámetros representativos del tráfico analizado. De esta forma es posible caracterizar un tráfico concreto de forma general mediante modelos analíticos.

Para la simulación a nivel de sistema se ha decidido emplear diversos tipos de tráfico: H.263, WWW (*World Wide Web*) y correo electrónico (mediante modelos que se basan en la segunda aproximación, es decir, modelos analíticos). Para los objetivos perseguidos en este trabajo, el comportamiento de interés es el del tráfico H.263, donde el retardo es un parámetro importante. Se decidió emplear otro tipo de tráfico a fin de evitar un sistema en el que sólo se transmite tráfico de un único tipo, situación muy poco realista. La elección del modelo de H.263 se justifica principalmente por la posibilidad de transmitir con una tasa binaria bastante reducida (16 kbps), cualidad que lo hace bastante adecuado a las velocidades binarias de transmisión que permiten los esquemas de codificación de GPRS, y por ser suficiente para los estudios de retardo aquí realizados. La navegación web y el correo electrónico son algunas de las aplicaciones más populares en las redes fijas de comunicaciones [92]. Puesto que esta tendencia parece confirmarse en redes de comunicaciones móviles, estos dos tipos de tráfico también han sido considerados en el análisis realizado en este trabajo. Además, estos dos modelos han sido ampliamente estudiados y utilizados en la literatura ([93] por ejemplo). Los modelos elegidos para navegación web y correo electrónico son modelos obtenidos a partir de medidas tomadas sobre redes fijas de comunicaciones. Sin embargo, tal y como se comenta en [88], no hay alternativa a esta elección al no existir modelos diseñados específicamente para redes de comunicaciones móviles. Además, las futuras tecnologías de tercera generación permitirán la comunicación no sólo a través de terminales móviles sino también empleado ordenadores portátiles, en cuyo caso estos modelos obtenidos a partir de redes fijas serían aplicables [88].

5.3.2.1.- TRÁFICO WORLD WIDE WEB (WWW)

Antes de explicar el modelo de tráfico WWW empleado es necesario señalar la diferencia entre los conceptos de *conexión TCP por objeto* y *conexión TCP por página* [88], ya que el modelado del tráfico WWW guarda relación con estos conceptos. Como es sabido, la navegación web hace uso del protocolo HTTP (*HyperText Transfer Protocol*) en la capa de aplicación, bajo la cual reside TCP (*Transmission Control Protocol*) como protocolo de la capa de transporte. Cuando un usuario solicita la descarga de una página web, está solicitando la transferencia de los diversos elementos u objetos que la integran: texto, imágenes, botones, listas desplegables, etc. Para la transferencia de estos objetos, el protocolo HTTP establece conexiones TCP. En las versiones HTTP/0.9 y HTTP/1.0 se establece una conexión TCP diferente para cada elemento de la página web transmitida, mientras que en la versión HTTP/1.1 se permite que todos los elementos de la página sean transmitidos utilizando una única conexión TCP. En el primer caso, cada una de las conexiones TCP utilizadas recibe el nombre de *conexión TCP por objeto*, mientras que en el segundo caso, la conexión utilizada para transferir la página web se denomina *conexión TCP por página*.

El modelo de tráfico WWW empleado en este estudio [94] está basado en el uso de conexiones TCP por objeto y su comportamiento se muestra en la figura 5.2.

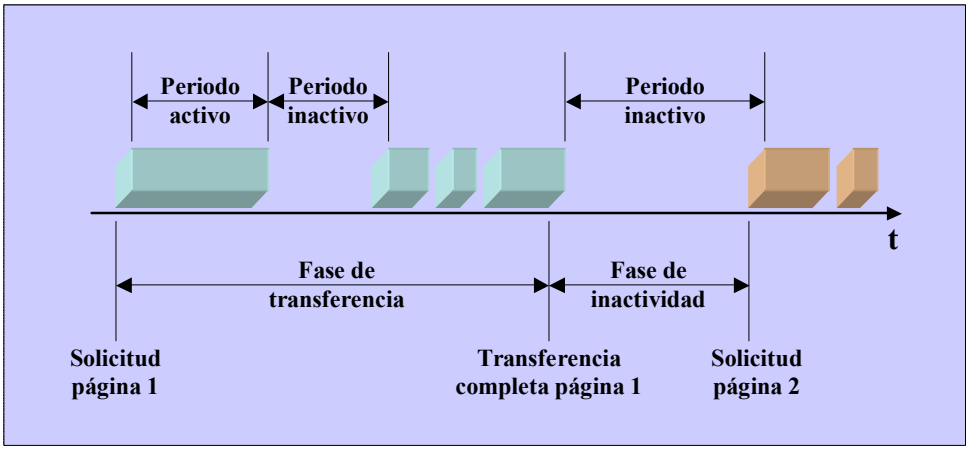


Figura 5.2.- Modelo de tráfico World Wide Web [88]

Durante una sesión HTTP un usuario visita varias páginas web. Al realizar la solicitud de una página web se inicia la fase de transferencia, que concluye cuando el usuario la ha recibido por completo. Entonces el usuario emplea cierto tiempo en leer la información contenida en la página (fase de inactividad) hasta que hace *clic* sobre un vínculo de la página o escribe una nueva dirección, solicitando una nueva página web e iniciando así su transferencia. Puesto que el modelo WWW utilizado se basa en el uso de conexiones TCP por objeto, durante la fase de transferencia existen periodos activos y periodos inactivos. Los periodos activos corresponden a los intervalos de tiempo durante los cuales se está transmitiendo un elemento de los que componen la página, mientras que los periodos inactivos corresponden a los intervalos de tiempo en los que se cierra una conexión TCP y se establece una nueva para transmitir un nuevo elemento de la página web.

La tabla 5.1 muestra los parámetros representativos del tráfico WWW, la distribución matemática mediante la cual puede ser descrito cada uno de ellos así como la función de distribución de probabilidad (PDF) y el valor de las constantes que en ella se aplican. El periodo activo en la fase de transferencia no aparece de forma explícita en la tabla ya que éste depende del tamaño del elemento u objeto a transmitir (además del ancho de banda disponible y de la calidad del enlace radio).

Parámetro representativo	Distribución matemática	Función de distribución de probabilidad (PDF)	Valor de las constantes
Tamaño de los elementos	Pareto	$f_x(x) = \frac{\alpha k^\alpha}{x^{\alpha+1}}$	k = 1000 α = 1.0
Periodo inactivo, Fase de transferencia	Weibull	$f_x(x) = \frac{bx^{b-1}}{a^b} e^{-\left(\frac{x}{a}\right)^b}$	a = 1.46 b = 0.382
Periodo inactivo, Fase de inactividad	Pareto	$f_x(x) = \frac{\alpha k^\alpha}{x^{\alpha+1}}$	k = 1 α = 1.5
Número de elementos por página web	Pareto	$f_x(x) = \frac{\alpha k^\alpha}{x^{\alpha+1}}$	k = 1 α = 2.43

Tabla 5.1.- Parámetros representativos del modelo de tráfico WWW [88]

5.3.2.2.- TRÁFICO DE CORREO ELECTRÓNICO

El modelo de tráfico para el correo electrónico [95] (mostrado en la figura 5.3) está basado, al igual que en el caso del tráfico WWW, en la combinación de periodos de actividad con periodos inactivos.

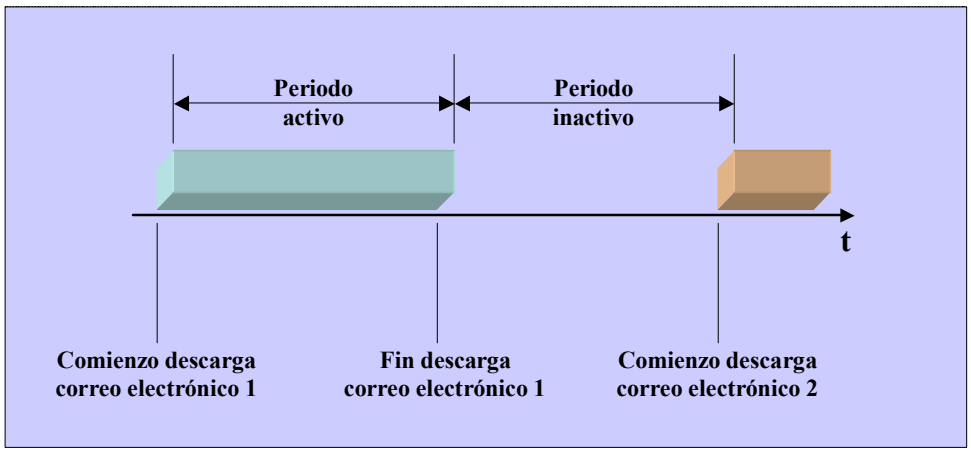


Figura 5.3.- Modelo de tráfico de correo electrónico [88]

Este modelo de tráfico asume que los correos se encuentran almacenados en un servidor de correo electrónico. Por tanto, cuando un usuario accede a su cuenta de correo a través del terminal móvil, se descarga solamente la cabecera de los correos almacenados en el servidor, en la cual se especifica el asunto. Si el usuario decide leer un correo concreto entonces es cuando se produce la descarga desde el servidor. Este proceso de transferencia corresponde al periodo activo. Desde que se termina de transferir un correo electrónico hasta que se solicita y comienza la transmisión del siguiente transcurre un intervalo de tiempo que recibe el nombre de periodo inactivo, tras el cual comienza un nuevo periodo de actividad en el que se produce la descarga del segundo correo electrónico y así sucesivamente.

La tabla 5.2 muestra los parámetros representativos del tráfico de correo electrónico, la distribución matemática mediante la cual puede ser descrito cada uno de ellos así como la función de distribución acumulativa (CDF) y el valor de las constantes que en ella se aplican. Nuevamente, el periodo activo no aparece de forma explícita en la tabla: éste depende del tamaño del correo electrónico a transmitir (además del ancho de banda disponible y la calidad del enlace radio). Observar que el tamaño del correo puede responder a dos distribuciones diferentes. Esto es así porque el modelado del tamaño del correo electrónico depende de si éste incorpora o no archivos adjuntos.

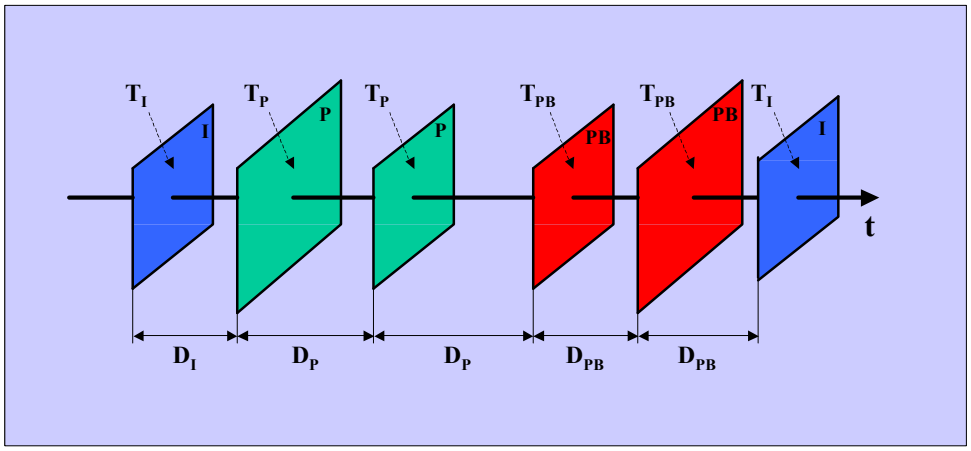
Parámetro representativo	Distribución matemática	Función de distribución acumulativa (CDF)	Valor de las constantes
Tamaño del correo electrónico	Weibull	$F_x(x) = \begin{cases} 1-\exp\left(-\exp\left(-k_1 x^{c_1}\right)\right) & F_x(x) \le 0.5 \\ 1-\exp\left(-\exp\left(-k_2 x^{c_2}\right)\right) & F_x(x) > 0.5 \end{cases}$	$k_1 = 17.64$ $k_2 = 3.61$ $c_1 = 2.04$ $c_2 = 0.37$
Periodo inactivo	Pareto	$F_x(x) = 1 - \left(\frac{k}{x}\right)^{\alpha}$	$k = 30$ $\alpha = 1.5$

Tabla 5.2.- Parámetros representativos del modelo de tráfico de correo electrónico [88]

5.3.2.3.- TRÁFICO DE VÍDEO H.263

En el estándar H.263 [71] 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[38] de vídeo que se transmite a través de la interfaz aire hacia el receptor. En el capítulo cuarto se explicó que, 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 5.4).

[38] Recordar que el concepto de trama de vídeo (imagen codificada) es diferente del de trama TDMA.



T_I: Tamaño trama tipo I
T_P: Tamaño trama tipo P
T_{PB}: Tamaño trama tipo PB
D_I: Duración trama tipo I
D_P: Duración trama tipo P
D_{PB}: Duración trama tipo PB

Figura 5.4.- Parámetros relevantes del tráfico H.263 [96]

El modelo de tráfico H.263 empleado en las simulaciones de este trabajo [96] se basa en el uso de una secuencia de vídeo real utilizada para su *parametrización*. 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 previamente es necesario codificar 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 [97]. 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, 64 kbps y 256 kbps así como con tasa binara variable (VBR, *Variable Bit-Rate*). En este caso se ha optado por una secuencia de una hora de la película Star Wars IV codificada a 16 kbps (esta tasa binaria es la que más se adecua a las tasas ofrecidas por los esquemas de codificación de canal disponibles en GPRS). Los detalles acerca de la configuración del codificador H.263 empleado se encuentran en [97].

A partir de estas tablas de valores disponibles en [97] en las que se especifica cada uno de los tipos de tramas generadas en la codificación H.263, su tamaño y su duración, el modelo de vídeo extrae ciertos valores de probabilidades y CDFs (*Cumulative Distribution Functions*) de parámetros 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 hace uso para ello de los datos relativos a la secuencia de video codificada.

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[39] que modela las transiciones entre los diversos tipos de tramas durante la transmisión de la secuencia de vídeo codificada.

[39] Una cadena de Markov pretende modelar un fenómeno concreto mediante un conjunto finito de estados discretos. Su funcionamiento se rige por la transición de un estado a otro. En su representación gráfica, los estados se encuentran relacionados entre sí mediante flechas. Cada flecha representa la transición de un estado a otro y sobre ella se indica la probabilidad de que dicha transición ocurra.

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 determinísticos, 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 [71], 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 5.5. Un estado representa la generación de una trama de tipo P y el otro representa la generación de una trama de tipo PB.

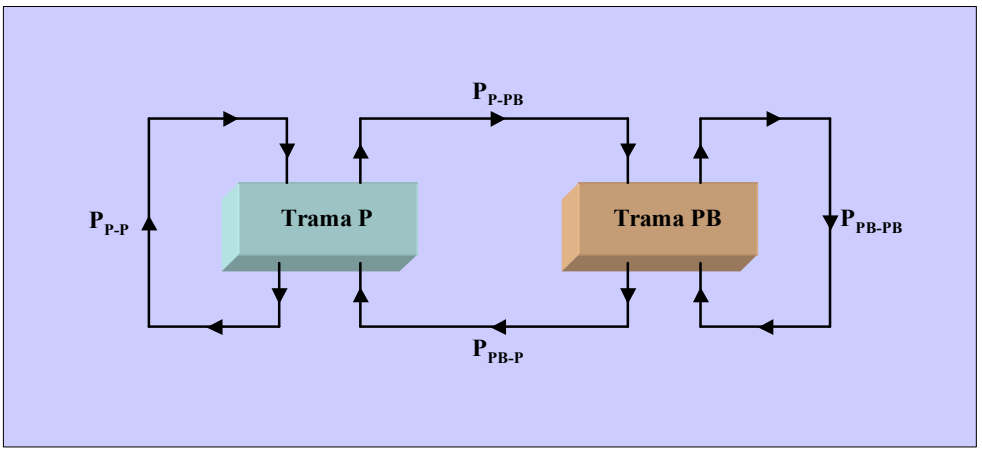


Figura 5.5.- Modelo de Markov para la transición entre tramas H.263 [96]

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 vendrá ya determinado 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 $P_{P\text{-}P}$, $P_{P\text{-}PB}$, $P_{PB\text{-}P}$ y $P_{PB\text{-}PB}$ contempladas en la cadena de Markov de la figura 5.5 pueden ser obtenidas a partir de los datos relativos a la secuencia de video 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.

Tras decidir el tipo de trama que se generará, el último paso consiste en obtener su tamaño y su duración. El proceso seguido hace uso de varias CDFs obtenidas a partir de los datos de la secuencia de vídeo real codificada (ver [96] para una descripción detallada) y respeta la correlación existente entre el tamaño de una trama y su duración [96]. Así pues, la secuencia de vídeo real se utiliza para la obtención no sólo de estas CDFs sino también del valor de las probabilidades $P_{X\text{-}Y}$ de la cadena de Markov de la figura 5.5 y ésta es la forma en que parametriza el modelo. La tabla 5.3 resume los parámetros representativos para la secuencia de vídeo empleada.

Parámetro representativo	Distribución matemática	Función de distribución de probabilidad (PDF)	Valor de las constantes
Tamaño de las tramas tipo I	Normal	$f_x(x) = \frac{1}{\sqrt{2\pi\sigma_I^2}} e^{-\frac{1}{2\sigma_I^2}(x-\mu_I)^2}$	μ_I = 1692.30 σ_I = 463080
Tamaño de las tramas tipo P	Normal	$f_x(x) = \frac{1}{\sqrt{2\pi\sigma_P^2}} e^{-\frac{1}{2\sigma_P^2}(x-\mu_P)^2}$	μ_P = 245.74 σ_P = 27783
Tamaño de las tramas tipo PB	Normal	$f_x(x) = \frac{1}{\sqrt{2\pi\sigma_{PB}^2}} e^{-\frac{1}{2\sigma_{PB}^2}(x-\mu_{PB})^2}$	μ_{PB} = 393.80 σ_{PB} = 38515
Duración de las tramas tipo I	Normal	$f_x(x) = \frac{1}{\sqrt{2\pi\sigma_I^2}} e^{-\frac{1}{2\sigma_I^2}(x-\mu_I)^2}$	μ_I = 830.00 σ_I = 114530
Duración de las tramas tipo P	Normal	$f_x(x) = \frac{1}{\sqrt{2\pi\sigma_P^2}} e^{-\frac{1}{2\sigma_P^2}(x-\mu_P)^2}$	μ_P = 140.10 σ_P = 8134
Duración de las tramas tipo PB	Normal	$f_x(x) = \frac{1}{\sqrt{2\pi\sigma_{PB}^2}} e^{-\frac{1}{2\sigma_{PB}^2}(x-\mu_{PB})^2}$	μ_{PB} = 178.21 σ_{PB} = 14890

Tabla 5.3.- Parámetros representativos del modelo de tráfico de vídeo H.263 [96]

Como puede verse, el modelo de vídeo empleado únicamente especifica, para cada trama generada, su tipo, su tamaño y su duración o tiempo al cabo del cual se generará la siguiente trama. El modelo no especifica ningún parámetro que determine si la transmisión de vídeo se realiza en tiempo real o en modo *streaming*, razón por la cual puede ser empleado para simular ambos tipos de transmisión. Así pues, el tipo de transmisión vendrá determinado por la implementación realizada en el simulador. La elegida corresponde a una transmisión en tiempo real y se describe a continuación.

Para transmitir una trama de vídeo es preciso, en primer lugar, proceder a la reserva de recursos y una vez que éstos son asignados se genera la trama, instante en el cual se introduce en el búfer del transmisor y comienza su transmisión a través de la interfaz aire. A partir de este momento pueden darse dos casos: que la trama de vídeo termine de transmitirse antes de que se genere la siguiente o que, por el contrario, se genere una nueva trama sin haber transmitido la anterior por completo. En este último caso, primero se procede a vaciar el búfer del transmisor para que la nueva trama, tras ser generada e introducida en el búfer, comience a ser transmitida en el mismo instante en que se genera. En este caso, y puesto que el usuario no deja de transmitir entre una trama y la siguiente, el canal no es liberado y se mantiene; tan sólo se procede a vaciar el búfer del transmisor y comenzar a transmitir la trama que se acaba de generar. En cambio, cuando una trama se termina de transmitir antes de que se genere la siguiente, entonces se procede a liberar los recursos que el usuario tenía asignados, ya que estos no volverán a ser necesarios hasta que llegue el instante en que se genere la siguiente trama, en el cual se volverán a reservar nuevos recursos y comenzará la transmisión.

Existe una excepción en la que, habiendo terminado de transmitir una trama de vídeo a tiempo, se prefiere mantener asignado el canal. Esto sucede cuando el tiempo restante hasta la generación de la siguiente trama es inferior al tiempo mínimo que se indica en la figura 5.6. Supongamos que el usuario del tercer time-slot termina de transmitir una trama de vídeo dentro de la trama TDMA *n*. Si antes de que comience el tercer time-slot de la trama *n*+1 se genera una nueva trama de vídeo para ese usuario, carece de sentido liberar el canal en la trama *n* y solicitar uno nuevo en la trama *n*+1, razón por la cual se prefiere mantener asignado el canal. Si tras haber terminado de transmitir la trama de vídeo el tiempo restante hasta la generación de la siguiente fuera superior al indicado en la figura 5.6, entonces sí se libera el canal, ya que en el instante en que comienza el tercer time-slot, el usuario no dispone de información para transmitir y es preferible liberarlo para que pueda ser aprovechado por otro usuario.

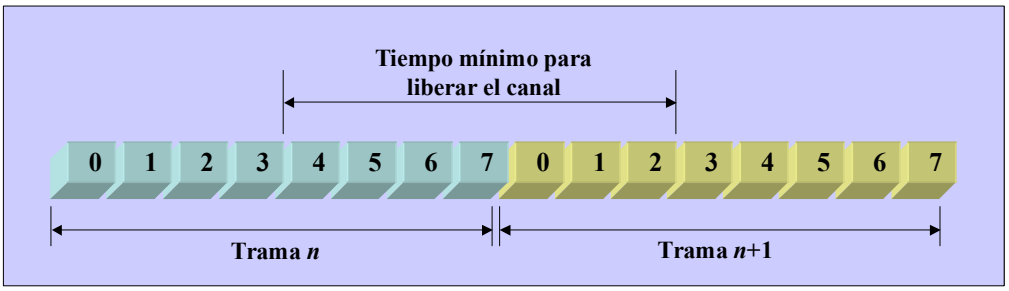


Figura 5.6.- Tiempo mínimo para liberar el canal

5.3.3.- UTILIZACIÓN DE ARQ

Los tipos de tráfico considerados responden a dos tipos de aplicaciones: sensibles al retardo (vídeo H.263) y no sensibles al retardo pero altamente sensibles a los errores de transmisión (navegación web y correo electrónico). En las aplicaciones no sensibles al retardo, éste no es un factor crítico mientras que la integridad de los datos transmitidos posee más relevancia. En estos casos es posible, y de hecho conveniente, emplear un mecanismo de retransmisión de datos ARQ. Sin embargo, en las aplicaciones sensibles al retardo, el uso de ARQ resulta bastante cuestionable. En el simulador a nivel de sistema se ha decidido utilizar ARQ para la transmisión de tráfico WWW y de correo electrónico mientras que para el tráfico de vídeo H.263 la retransmisión mediante ARQ se ha mantenido desactivada. La razón de esta decisión reside en el hecho de que la retransmisión de información supone un retardo adicional. En tráfico de vídeo el retardo es un factor clave, lo que significa que aumentar el retardo en la transmisión puede degradar considerablemente la calidad percibida por el usuario. En efecto, antes de que la secuencia de imágenes se reproduzca con lentitud debido al retardo, es preferible que la información llegue al receptor a su tiempo, aun cuando contenga algunos errores (el usuario apreciará una pequeña distorsión en una o algunas imágenes). Hay que tener en cuenta que en la decodificación de canal así como en la de fuente (H.263) suelen emplearse técnicas de corrección de errores que pueden incluso eliminar estos errores de transmisión.

5.3.4.- CONFIGURACIÓN DEL SIMULADOR

En la tabla 5.4 se muestra un resumen de la configuración empleada para el simulador a nivel de sistema.

Parámetro	Valor
Enlace estudiado	Downlink
Número de celdas cocanal en el sistema	25
Estaciones base cocanal consideradas	1ª y 2ª corona
Tamaño de clúster	4
Radio de las celdas	1 km
Sectorización	120º (3 sectores)
Portadoras por sector	2
Canales de tráfico de datos por sector	2 x 8 = 16
Número de canales por usuario	1
Usuarios por sector	12
Distribución del tráfico	Vídeo H.263: 50% (6 usuarios/sector) Navegación web: 25% (3 usuarios/sector) Correo electrónico: 25% (3 usuarios/sector)
Asignación de canales radio	RCA, FCFS
Movilidad	Sí implementada
Handover	No implementado
Control de potencia	No implementado
Salto en frecuencia	No implementado
Modelo de propagación (atenuación)	Okumura-Hata
Modelo de efecto sombra	Ley log-normal, desviación estándar = 6 dB, distancia de decorrelación = 20m
Entorno de simulación	Urbano (*Typical Urban*)
Velocidad del terminal móvil	50 km/h
ARQ	Sí implementado para WWW y correo No implementado para vídeo H.263
Tamaño de la ventana ARQ	64 bloques RLC
Periodo de confirmación ARQ	16 bloques RLC
Periodo de actualización de LA	60 ms

Tabla 5.4.- Resumen de la configuración del simulador

El sistema simulado está compuesto de 25 celdas organizadas según una estructura de clúster compuesta por 4 celdas. Cada celda posee un radio de 1 km (sistema macrocelular) y está compuesta por 3 sectores de 120º cada uno. En cada sector existen dos frecuencias portadoras. Puesto que en cada frecuencia se transmiten 8 time-slots, existe un total de 16 canales disponibles por sector. Estos 16 canales se dedican exclusivamente a la transmisión de paquetes de datos (PDTCH), ignorando en esta simulación canales de señalización ya que, como se ha comentado en más de una ocasión a lo largo de esta memoria, el presente trabajo centra su atención sobre la transmisión de datos en la interfaz aire careciendo de interés cualquier proceso de señalización.

La carga de información transmitida en el sistema puede controlarse variando en el simulador el número de usuarios por sector [98]. Las pruebas realizadas indican que para la cantidad de usuarios elegida para cada servicio (ver tabla 5.4) la ocupación media del ancho de banda es del 48%. Así pues, el sistema simulado soporta una carga de tráfico media.

La asignación de canales radio se realiza mediante RCA aplicando un mecanismo FCFS (justificación en el apartado 5.3.1.3). El número máximo de canales que se asigna a cada usuario es uno. Esta decisión resulta comprometida debido a que, analizando la velocidad binaria que requiere cada una de las tramas de la secuencia de vídeo (tamaño de la trama dividido entre tiempo disponible para su transmisión antes de que llegue la siguiente trama) se puede apreciar que, si bien la tasa binaria media de la secuencia es de 16 kbps, existen tramas que requieren una tasa menor pero también existen otras que requieren una tasa mucho mayor. En el anexo B se demuestra que, fijando como criterio que la calidad se considera aceptable si como mínimo el 95% de las tramas de vídeo se recibe sin retardo, entonces el empleo de un único time-slot por usuario resultaría insuficiente incluso en el hipotético caso de que la calidad del canal de transmisión fuera tal que permitiera utilizar CS-4 de forma continua. Son varios los estudios realizados [68, 70] en los que se afirma la imposibilidad de transmitir vídeo sobre GPRS empleando un único time-slot. De hecho, en ellos se expone que una de las características más atractivas de GPRS para la transmisión de vídeo es, precisamente, su capacidad de funcionamiento en modo multislot.

Sin embargo, las simulaciones de este trabajo se realizan empleando siempre un único time-slot por usuario debido a los siguientes motivos. En primer lugar, lo que aquí se pretende estudiar no es el hecho de que se consiga o no la transmisión de la secuencia de vídeo sin retardo (hecho que en el ámbito de un sistema real depende no sólo del algoritmo LA empleado sino también de otros muchos factores como, por ejemplo, el *scheduling*, la carga del sistema o la potencia de transmisión). Lo que en el presente trabajo se pretende es analizar, de forma comparativa, los diversos algoritmos de LA estudiados y examinar cuáles pueden contribuir en mayor medida a la mejora de la transmisión de datos sensibles al retardo a través de la interfaz aire. Es cierto que dicho estudio podría haberse realizado igualmente empleando tanto uno como varios time-slots por usuario. Sin embargo, en [99, 100] se demuestra la existencia de una potencial correlación entre la calidad de canal existente en los diversos time-slots que son asignados a un mismo usuario dentro de una trama TDMA cuando se opera en modo multislot. Por lo tanto, sería necesario en este caso elaborar un modelo adecuado que considerara esta correlación e implementarlo en el simulador a nivel de enlace.

La simulación de una transmisión multislot sin implementar este modelo no resulta totalmente correcta y, puesto que lo que aquí se pretende es realizar un estudio comparativo de los diversos algoritmos LA y para ello resulta suficiente la transmisión con un único time-slot, finalmente se decidió emplear un time-slot por usuario en la simulación, aun siendo conscientes de que la transmisión con una calidad aceptable realmente requiere más de un time-slot por usuario.

Como se comentó en el apartado 5.3.1.4, en el simulador se implementa un modelo de movilidad. No se han implementado, en cambio, modelos de *handover*, control de potencia ni salto en frecuencia. El *handover* no se ha considerado por no ser necesario al simular celdas cocanal en este trabajo. Como consecuencia, en el modelo de movilidad, cuando un usuario llega al borde de la celda vuelve hacia atrás de tal forma que el número de usuarios por celda permanece constante. El control de potencia no se ha implementado porque la variación de la potencia de transmisión supone la variación de la relación CIR experimentada y, por tanto, de la calidad del canal de transmisión. Puesto que el comportamiento de LA se basa en la calidad experimentada en el enlace radio, el uso de un mecanismo de control de potencia afecta al comportamiento de LA, pudiendo dar lugar a una estimación incorrecta de su rendimiento. Igualmente, el uso de salto en frecuencia implica que el usuario cambia constantemente de time-slot, cada uno de los cuales tendrá una calidad de transmisión (relación CIR) determinada. El resultado es que, a efectos prácticos, el algoritmo LA percibe un canal de transmisión cuya calidad varía de forma altamente aleatoria, dando lugar a que su comportamiento también lo sea (esta elevada aleatoriedad hace que sea incluso preferible no emplear LA en estos casos). Por ello, para el estudio de LA se han desactivado tanto el control de potencia como el salto en frecuencia.

En cuanto a ARQ, ya se ha dicho que éste no se implementa para H.263 pero sí para el resto de servicios, empleando un periodo de confirmación y una ventana de transmisión de 16 y 64 bloques RLC respectivamente. El periodo de confirmación se ha establecido en 16 bloques RLC de acuerdo con [43], es decir, el receptor envía un mensaje de confirmación por cada 16 bloques RLC recibidos. El tamaño de la ventana ARQ se ha establecido en 64 bloques RLC siguiendo las especificaciones de GPRS [101]. En cuanto al esquema de codificación empleado para realizar la retransmisión de datos, resulta obligatorio en el estándar GPRS que ésta se realice empleando el mismo esquema de codificación de canal que se utilizó para realizar la transmisión original. Esta restricción se debe al hecho de que la cantidad de datos transmitida en un bloque RLC que se codifica empleando un determinado esquema de codificación de canal no se puede acomodar en el campo de datos de ningún otro esquema de codificación existente en el estándar GPRS. Este problema ha sido solucionado en el sistema EDGE (*Enhanced Data-rates for Global Evolution*) [60], donde los esquemas de codificación-modulación se agrupan dentro de familias en las que el campo de datos de un esquema de codificación es múltiplo entero respecto a otros. Aun cuando este trabajo se basa en el uso de la interfaz aire de GPRS, en el simulador se ha permitido el cambio de esquema de codificación para realizar las retransmisiones con la intención de no limitar la generalidad de los resultados obtenidos al restringirlos a las particularidades de la interfaz radio de un sistema concreto. Una restricción adicional particular de GPRS es que el cambio de esquema de codificación de canal durante el proceso de transmisión sólo puede tener lugar entre bloques RLC pertenecientes a diferentes tramas LLC. Nuevamente, esta otra restricción ha sido ignorada por los mismos motivos que se han ignorado las restricciones para la retransmisión de información errónea.

Para fijar el periodo de actualización de LA se consideraron cuatro posibles valores candidatos a partir del estudio realizado en [56]: 20 ms, 60 ms, 100 ms y 200 ms. Para una carga media del sistema (como es el caso), un periodo de 20 ms no conviene [56]. Puesto que los resultados ofrecidos por los otros tres valores son similares, la decisión se toma analizando la secuencia real empleada por el modelo de vídeo. Puede observarse que el 31.5% de las tramas de vídeo tiene una duración inferior a 100 ms, lo que significa que para un periodo de 100 ms casi la tercera parte de las tramas de la secuencia de vídeo se transmitiría empleando el mismo esquema de codificación, algo poco deseable en el estudio de una técnica dinámica como LA. Para un periodo de 60 ms la cifra anterior se reduce a tan sólo el 3.3%, haciendo preferible, pues, un periodo de actualización de 60 ms.

5.4.- INTERFACES ENTRE SIMULACIÓN A NIVEL DE ENLACE Y NIVEL DE SISTEMA

Como se comentó en la introducción de este capítulo, las simulaciones de un sistema de comunicaciones móviles se desglosan en dos niveles diferentes: simulación a nivel de enlace y simulación a nivel de sistema. Ya se dijo que la principal razón para proceder de esta manera reside en la diferente resolución temporal requerida en ambos niveles. También es cierto que la simulación a nivel de enlace requiere un coste computacional significativamente más elevado y he aquí otro motivo adicional para proceder a esta separación entre niveles de simulación [88].

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 en lo sucesivo). 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) y su correspondiente tasa de error (expresada, por ejemplo, en términos de BLER).

Tal y como se discute en [102], el uso de una determinada interfaz u otra entre ambos niveles de simulación repercute sobre la precisión y la exactitud en el estudio de LA. A continuación se describirán dos interfaces entre ambos niveles. La primera de ellas trabaja a nivel de bloque RLC y es la que se utiliza habitualmente; la segunda [88] funciona a nivel de ráfaga y permite un modelado más preciso de los errores de transmisión en el enlace radio al capturar los efectos del desvanecimiento rápido (*fast fading*) y posibilitar la introducción de sus efectos sobre el proceso de transmisión en la simulación a nivel de sistema.

El simulador empleado en el presente trabajo hace uso de la interfaz a nivel de ráfaga por ofrecer resultados más exactos y precisos [102]. Sin embargo, también se presentará en este capítulo la interfaz a nivel de bloque RLC por ser utilizada en el capítulo siguiente para la presentación de los diversos algoritmos LA.

5.4.1.- INTERFAZ A NIVEL DE BLOQUE RLC

En el apartado 5.3 se explicó que el simulador a nivel de sistema modela el proceso de transmisión a nivel de ráfaga. Cada vez que se procede a la transmisión de una ráfaga de datos, se estima la calidad del canal de transmisión (relación CIR) en función de la posición de la estación móvil, de la estación base transmisora y de las estaciones base cocanal interferentes, de forma que la transmisión de cada ráfaga en la simulación a nivel de sistema lleva asociado un valor concreto de relación CIR. Los valores de la relación CIR correspondientes a las cuatro ráfagas que componen un bloque RLC determinado se promedian, dando lugar a un valor medio de la relación CIR para dicho bloque RLC. Este valor medio de la relación CIR obtenido en la simulación a nivel de sistema se convierte a un valor de BLER utilizando la figura 5.7 (la conversión depende del esquema de codificación de canal empleado para transmitir el bloque RLC).

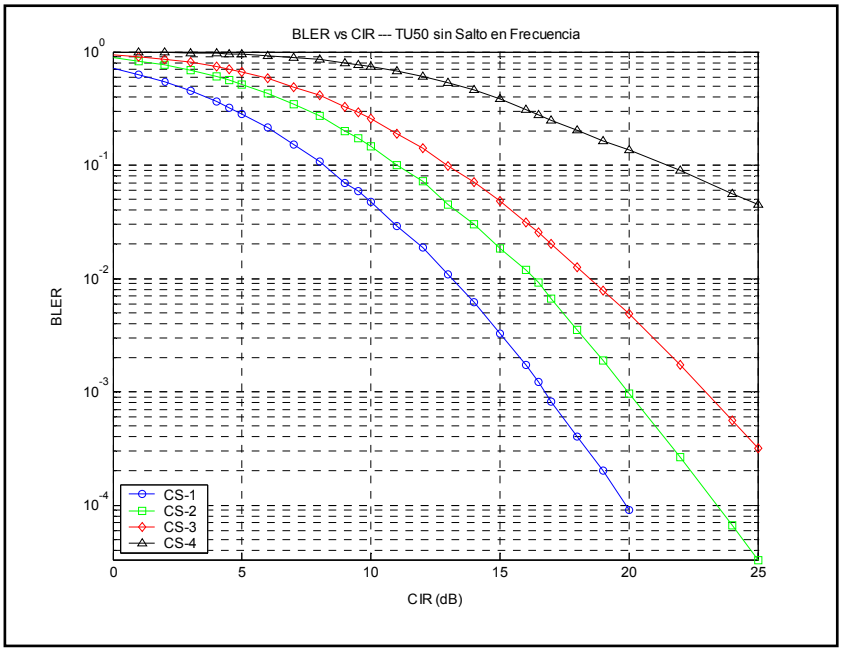


Figura 5.7.- Interfaz a nivel de bloque RLC [103]

Es esta figura la que constituye la interfaz (LUT) a nivel de bloque RLC entre ambos niveles de simulación. Recordemos que los puntos de esta figura fueron obtenidos a partir de varias simulaciones a nivel de enlace (variando el esquema de codificación empleado y la relación CIR en cada una, siempre en un entorno TU50, el utilizado para la simulación a nivel de sistema).

Mediante este procedimiento se obtiene, así pues, un valor de BLER para cada bloque RLC transmitido. Para determinar, en función de este valor, si el bloque ha sido transmitido con error o no en la simulación a nivel de sistema, se genera un número aleatorio x_o a partir de una variable aleatoria uniformemente distribuida entre 0 y 1. Si x_o es menor que el valor del BLER obtenido, entonces el bloque RLC se considera que ha sido recibido con error, mientras que si x_o es mayor, entonces se establece que el bloque RLC ha sido correctamente recibido por la estación móvil.

5.4.2.- INTERFAZ A NIVEL DE RÁFAGA

En la interfaz a nivel de bloque RLC, el BLER asociado a la transmisión de un bloque RLC se obtenía a partir de la relación CIR media experimentada en la transmisión de dicho bloque. Sin embargo, el uso de codificación convolucional así como de entrelazado origina que los errores de transmisión en bloques RLC dependan, no sólo de la calidad media experimentada en el canal, sino también de su distribución sobre las cuatro ráfagas empleadas para la transmisión de dicho bloque RLC [102]. La interfaz presentada a continuación [103], empleada para las simulaciones de este trabajo, tiene en cuenta la distribución de la calidad en las cuatro ráfagas de un bloque RLC, permitiendo así capturar los efectos del desvanecimiento rápido e incluirlos en la simulación a nivel de sistema.

La interfaz a nivel de ráfaga consta de dos grupos de LUTs que denominaremos LUT-1 y LUT-2. Las tablas LUT-1 requieren como entrada el valor de la relación CIR media experimentada por una ráfaga durante su transmisión (en la simulación a nivel de sistema). Existe una tabla LUT-1 para cada valor de la relación CIR. Puesto que cada una de estas tablas se obtiene como resultado de una simulación a nivel de enlace, el número de tablas LUT-1 es finito. Por ello, el valor de la relación CIR para una ráfaga se aproxima a uno de los valores de la relación CIR disponibles. Para dicho valor de relación CIR existe una tabla LUT-1 con un aspecto similar al de la mostrada en la figura 5.8 (correspondiente a una relación CIR de 6 dB).

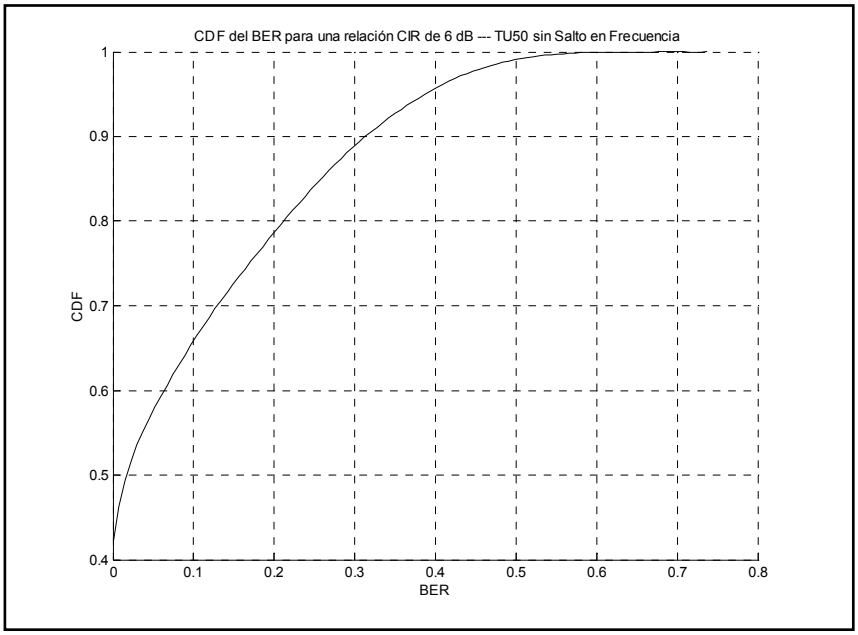


Figura 5.8.- Interfaz a nivel de ráfaga – LUT-1 [103]

A partir de la tabla LUT-1 correspondiente y de un proceso aleatorio se obtiene un valor determinado de BER. El interés de este procedimiento radica en la capacidad de capturar el efecto del desvanecimiento rápido sobre la relación CIR, ya que ahora, a diferencia de la interfaz a nivel de bloque RLC, para una misma relación CIR media se pueden obtener diferentes valores para la tasa de error (éste es, precisamente, el efecto del desvanecimiento rápido sobre la relación CIR).

El procedimiento se repite para cada una de las cuatro ráfagas que componen un bloque RLC. Es decir, para cada ráfaga se tiene un valor de relación CIR concreto; este valor se aproxima a uno de los valores de relación CIR disponibles y se utiliza la tabla LUT-1 asociada. Mediante un proceso aleatorio y haciendo uso de dicha tabla LUT-1 se obtiene un valor determinado de BER. Tras haber repetido el proceso para cada una de las cuatro ráfagas de un bloque RLC, se obtienen cuatro valores de BER[40], a partir de los cuales es posible determinar un valor de BER medio así como su desviación estándar. A partir de estos dos valores de BER y haciendo uso de las tablas LUT-2 se obtiene un valor de BLER. En la figura 5.9 se muestra la representación gráfica de los valores de la tabla LUT-2 correspondiente al esquema de codificación CS-1 en un entorno TU50 (existe una tabla LUT-2 para cada esquema de codificación y para diversas condiciones de simulación).

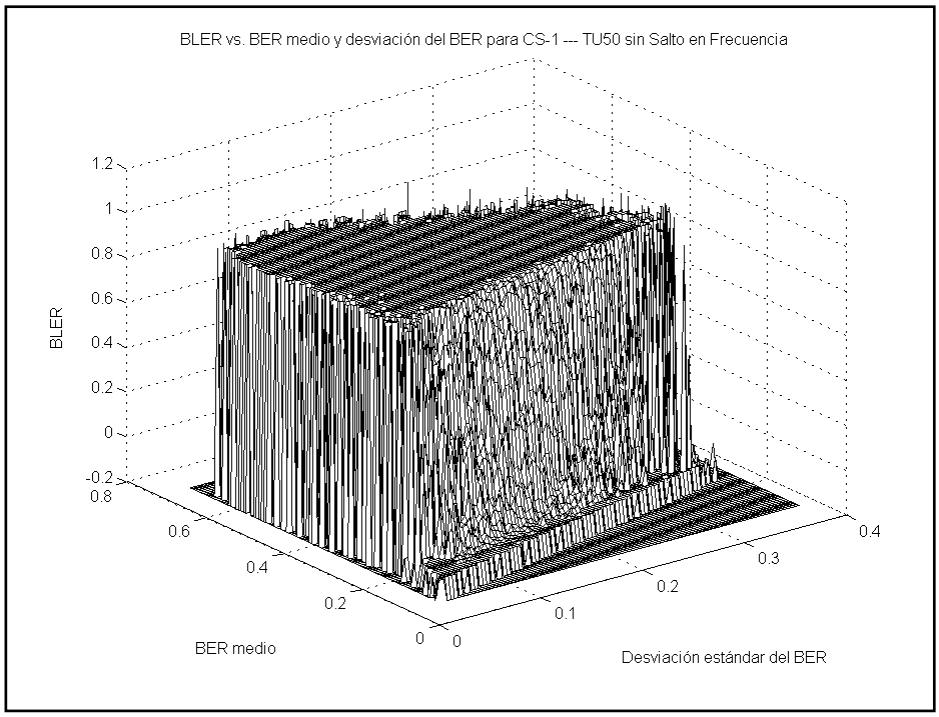


Figura 5.9.- Interfaz a nivel de ráfaga – LUT-2 [103]

Mediante este procedimiento se vuelve a obtener, igual que en el apartado anterior, un valor de BLER para cada bloque RLC transmitido, aunque su obtención se realiza a través de un procedimiento diferente, más elaborado, que permite considerar a nivel de sistema los efectos del desvanecimiento rápido. A partir de aquí, la forma en que se determina si el bloque RLC se recibe o no con error es idéntica a la explicada en el apartado anterior. Observar que en la figura 5.9 aparecen valores de BLER negativos. Esto, obviamente, es imposible, ya que el BLER, por definición, puede tomar valores sólo entre 0 y 1. Estos valores negativos simplemente representan aquellos puntos correspondientes a una combinación de BER medio y desviación estándar del BER que no se llegó a dar en la simulación a nivel de enlace realizada para obtener la figura.

[40] Observar que el BER es independiente del esquema de codificación empleado. Ello se debe a que dicho parámetro se extrae a la salida del ecualizador en el receptor, antes de proceder a la decodificación de canal.

5.5.- SUMARIO

En el presente capítulo se han explicado los modelos y herramientas de simulación empleados en la evaluación de los algoritmos LA estudiados. La simulación se ha desglosado en dos niveles: a nivel de enlace y a nivel de sistema.

Debe quedar claro que la implementación del simulador utilizado no forma parte de este trabajo[41] sino que, simplemente, representa una herramienta ya existente, necesaria para el estudio de LA, sobre la que se han realizado aquellas modificaciones (en el simulador a nivel de sistema) necesarias para incorporar la transmisión de vídeo en las simulaciones, labor que sí compete al trabajo aquí realizado. En lo relativo al simulador a nivel de enlace tan sólo se han descrito brevemente aquellos aspectos básicos que resultan necesarios para entender el funcionamiento global del proceso de simulación. La atención se ha centrado, pues, en la simulación a nivel de sistema. Sobre este simulador se han modificado tan sólo aquellos aspectos necesarios para los objetivos de este trabajo (implementación de los algoritmos LA estudiados, integración del modelo de vídeo en el simulador y desactivación de ARQ para vídeo).

Así pues, tras unos breves comentarios acerca del simulador a nivel de enlace, se ha procedido a describir con algo más de detalle el simulador a nivel de sistema: las entidades que componen su arquitectura (*Entorno Celular, Enlace Radio, Estación Base y Estación Móvil*), los modelos de tráfico implementados (navegación web, correo electrónico y vídeo H.263) y la utilización de ARQ. La descripción del simulador a nivel de sistema concluye con un resumen de su configuración en el que se explican algunas decisiones no tratadas anteriormente. El capítulo finaliza con la presentación de dos interfaces entre las simulaciones a ambos niveles. La primera de ellas (a nivel de bloque RLC) es la empleada más habitualmente; la segunda (a nivel de ráfaga) es más elaborada y permite considerar a nivel de sistema los efectos del desvanecimiento rápido. Esta segunda interfaz, más precisa, es la que se ha utilizado en las simulaciones de este trabajo.

[41] El simulador utilizado consta de un total de 13129 líneas de código repartidas en 39 clases.

Capítulo 6

Estudio y Diseño de Link Adaptation

6.1.- INTRODUCCIÓN

En los cuatro capítulos anteriores, desde el segundo hasta el quinto, se trataron diversos temas vinculados al estudio aquí realizado, tales como la descripción del sistema GPRS, de la técnica de gestión de recursos radio *Link Adaptation*, del concepto de calidad de servicio, del estándar de codificación de vídeo H.263 o de los modelos y herramientas de simulación aquí empleados. Este capítulo se dedicará al estudio de algunos algoritmos *Link Adaptation* (LA) ya propuestos por otros autores así como al diseño de otros nuevos que constituyen la principal aportación de este trabajo.

En primer lugar se presentan dos algoritmos propuestos por otros autores cuyo rendimiento será tomado como referencia para comparar los resultados ofrecidos por los nuevos algoritmos aquí propuestos. El primero de ellos es el algoritmo utilizado en el capítulo tercero para la explicación del concepto de LA, el cual pretende maximizar el throughput global del sistema. Este algoritmo es el empleado más habitualmente [42 – 44]. El segundo, propuesto en [104], no busca maximizar el throughput sino que su principal objetivo consiste en asegurar una transmisión con un nivel de calidad aceptable, manteniendo el BLER por debajo de un determinado valor por encima del cual se considera que la calidad experimentada por el usuario no es aceptable.

Es cierto que si el BLER experimentado por un usuario mejora, entonces el throughput de dicho usuario también mejorará. Sin embargo, cuando consideramos un sistema real en el cual interviene multitud de usuarios, entonces el throughput global del sistema y la tasa global de errores son dos factores comprometidos, tal y como se explica en [105]. En este contexto, el primero de los algoritmos que aquí se propone tiene como principal objetivo ofrecer un compromiso entre ambos factores, es decir, proporcionar unas prestaciones intermedias a las ofrecidas por los dos algoritmos anteriores, en los que throughput y BLER son considerados de forma aislada, es decir, el diseño se realiza considerando tan sólo uno de ellos e ignorando el otro. En este sentido, el primer algoritmo propuesto resulta novedoso en la medida en que el diseño se lleva a cabo teniendo en cuenta ambos factores.

Frente al criterio tradicional de maximizar el throughput se han propuesto, para servicios multimedia de tipo audio y vídeo, otros criterios en los que la decisión del esquema de codificación se basa en criterios de calidad. Algunos ejemplos son [104] (donde se intenta preservar el BLER) o [106] (donde se intenta maximizar la calidad final percibida por el usuario expresada en términos de relación PSNR[42]). Sin embargo, hasta la fecha, en la literatura no se ha encontrado ninguna propuesta donde el criterio seguido por LA tenga en cuenta el retardo, factor clave para este tipo de servicios. En algunos artículos [107] se afirma que para servicios en los que el retardo es importante, el throughput a largo plazo no cuantifica la calidad de servicio y, por ello, se sugiere la elección del esquema de codificación considerando el tamaño del paquete a transmitir [107, 108], ya que el retardo es proporcional a éste. Sin embargo, no se define de forma clara y concisa un algoritmo LA basado en criterios de retardo. El segundo algoritmo propuesto en este trabajo resulta innovador al basar su diseño no sólo en criterios de throughput, como ha venido siendo habitual, sino también en criterios de retardo.

[42] La relación PSNR (*Peak Signal to Noise Ratio*) o relación señal a ruido de pico es una medida de la calidad de imagen de vídeo que, aunque subjetiva, se emplea a menudo en estudios relacionados con la transmisión de vídeo.

6.2.- ALGORITMOS LINK ADAPTATION

Como se recordará de las explicaciones del capítulo tercero, la base de LA consiste en evaluar la calidad del enlace radio y seleccionar el esquema de codificación que resulta óptimo en dichas condiciones, de acuerdo a un criterio predefinido. A continuación se presentarán diversos criterios que pueden ser utilizados para realizar la elección y se ilustrará la forma en que estos criterios se materializan en la definición del algoritmo LA. Para ello se utilizarán primero las interfaces a nivel de bloque RLC mostradas en el capítulo anterior, ya que de esta forma la idea se muestra de forma más clara y la comprensión del procedimiento resulta más sencilla. A continuación, se repetirá el proceso pero empleando las interfaces a nivel de ráfaga, que son las que se implementan en el simulador utilizado en este trabajo ya que permiten la obtención de resultados más precisos y realistas.

6.2.1.- ALGORITMO BASADO EN THROUGHPUT

Este algoritmo tiene como principal objetivo maximizar el throughput global del sistema, es decir, maximizar la cantidad de bits correctamente transmitidos por unidad de tiempo. A continuación se ilustra, empleando las interfaces a nivel de bloque RLC, el procedimiento seguido para obtener los umbrales de decisión en función de los cuales se selecciona el esquema de codificación a emplear.

En el capítulo tercero se definió el throughput de cada esquema de codificación *CS-i* en función de su velocidad binaria $R_{CS\text{-}i}$ y de su BLER según la expresión

$$Throughput_{CS-i} = R_{CS-i} \cdot \left[1 - BLER_{CS-i}(CIR)\right] \tag{6.1}$$

Si para cada uno de los cuatro esquemas de codificación se sustituye en la expresión (6.1) $R_{CS\text{-}i}$ por su valor correspondiente (9.05 kpbs para CS-1, 13.4 kbps para CS-2, 15.6 kbps para CS-3 y 21.4 kbps para CS-4) y $BLER_{CS\text{-}i}(CIR)$ por los valores que se muestran en la figura 6.1, entonces se obtienen las curvas de la figura 6.2, en las que se representa el throughput que proporciona cada esquema de codificación en función de la calidad del canal de transmisión expresada en términos de relación CIR.

En la figura 6.2 se puede apreciar claramente que existen regiones en las que cada esquema de codificación proporciona un throughput superior al ofrecido por el resto, consecuencia directa del diferente grado de protección frente a errores que posee cada uno de ellos. Puesto que el objetivo es maximizar el throughput, entonces, según la figura 6.2, para cada valor de la relación CIR se debe emplear aquel esquema de codificación que proporcione el mayor throughput. Es decir, CS-1 se debe emplear para valores de relación CIR inferiores a 5 dB, CS-2 y CS-3 se deben emplear cuando la relación CIR tome valores entre 5 dB y 9.5 dB o entre 9.5 dB y 16.5 dB, respectivamente, mientras que CS-4 ha de emplearse si el valor de la relación CIR supera los 16.5 dB.

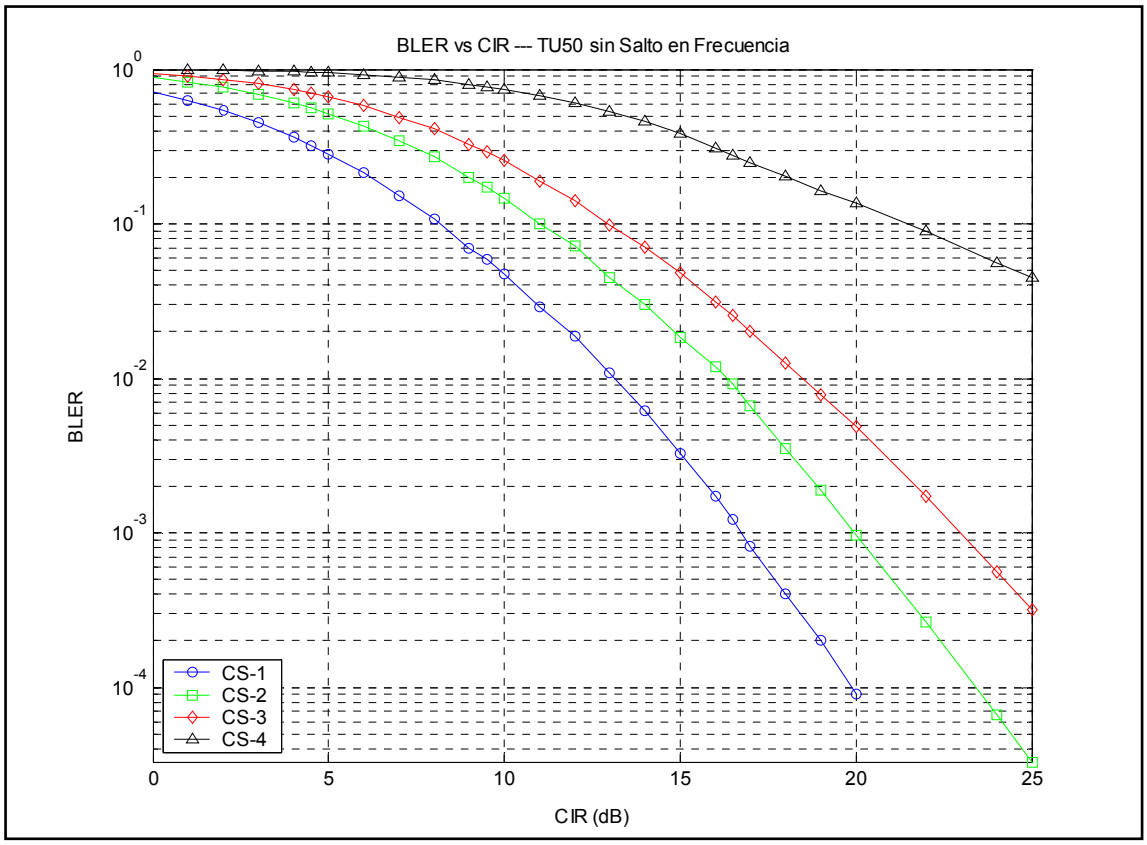


Figura 6.1.- BLER vs CIR

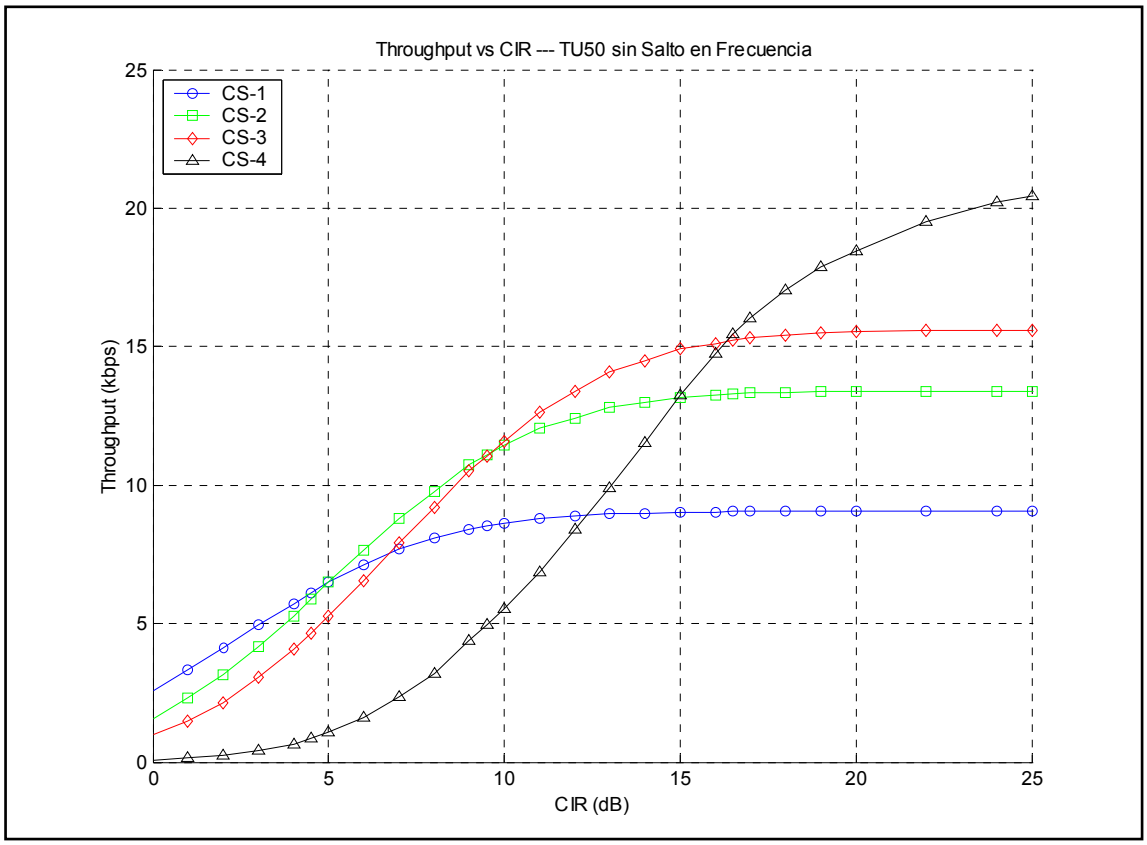


Figura 6.2.- Throughput vs CIR

La forma en que se implementa el algoritmo mediante las interfaces a nivel de ráfaga es similar a como se hace con las interfaces a nivel de bloque RLC. La figura 6.3 muestra, para cada esquema de codificación, las curvas LUT–2 de dichas interfaces, es decir, las curvas que muestran el BLER en función de la calidad del canal expresada, en este caso, en términos de BER medio y desviación estándar del BER (ambos valores se obtenían a partir de la relación CIR empleando las curvas LUT–1, apartado 5.4.2). Aparentemente, las curvas son iguales, pero se diferencian en la transición que tiene lugar para valores bajos de BER medio. Aplicando la expresión (6.1) a estas curvas con el valor de $R_{CS\text{-}i}$ correspondiente, se obtienen las curvas mostradas en la figura 6.4, en las que se representa el throughput en función del BER medio y la desviación del BER para cada esquema de codificación. El comportamiento del throughput respecto al BER medio que se observa en estas curvas recuerda al comportamiento del throughput respecto a la relación CIR que se apreciaba en la figura 6.2, excepto para el caso del esquema de codificación CS-4, el cual sólo proporciona un throughput distinto de cero cuando no tiene lugar ningún error en la transmisión, es decir, cuando el BER medio es cero y su desviación también. La diferencia respecto al caso anterior está en que ahora el throughput no depende sólo de la calidad media experimentada en el canal de transmisión sino también de sus variaciones instantáneas durante la transmisión de un bloque RLC (recordar que la interfaz a nivel de ráfaga permitía capturar los efectos del desvanecimiento rápido o *fast fading*). Representado todas las curvas sobre la misma gráfica (figura 6.5) se puede observar que, al igual que en el caso anterior, existen regiones en las que un esquema de codificación proporciona un throughput superior al ofrecido por el resto. Para cada combinación de BER medio y desviación estándar del BER se empleará aquel esquema de codificación que proporcione el mayor throughput.

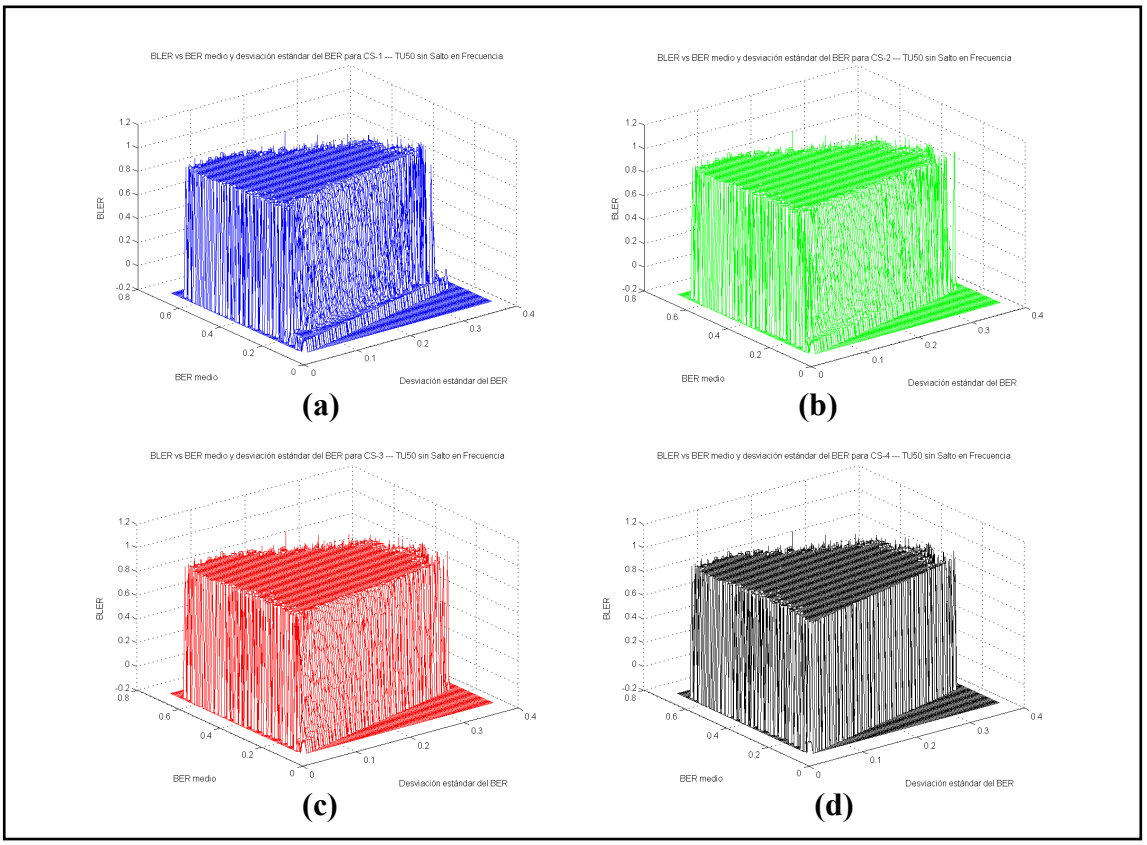


Figura 6.3.- BLER vs BER medio y desviación del BER para: (a) CS-1, (b) CS-2, (c) CS-3, (d) CS-4

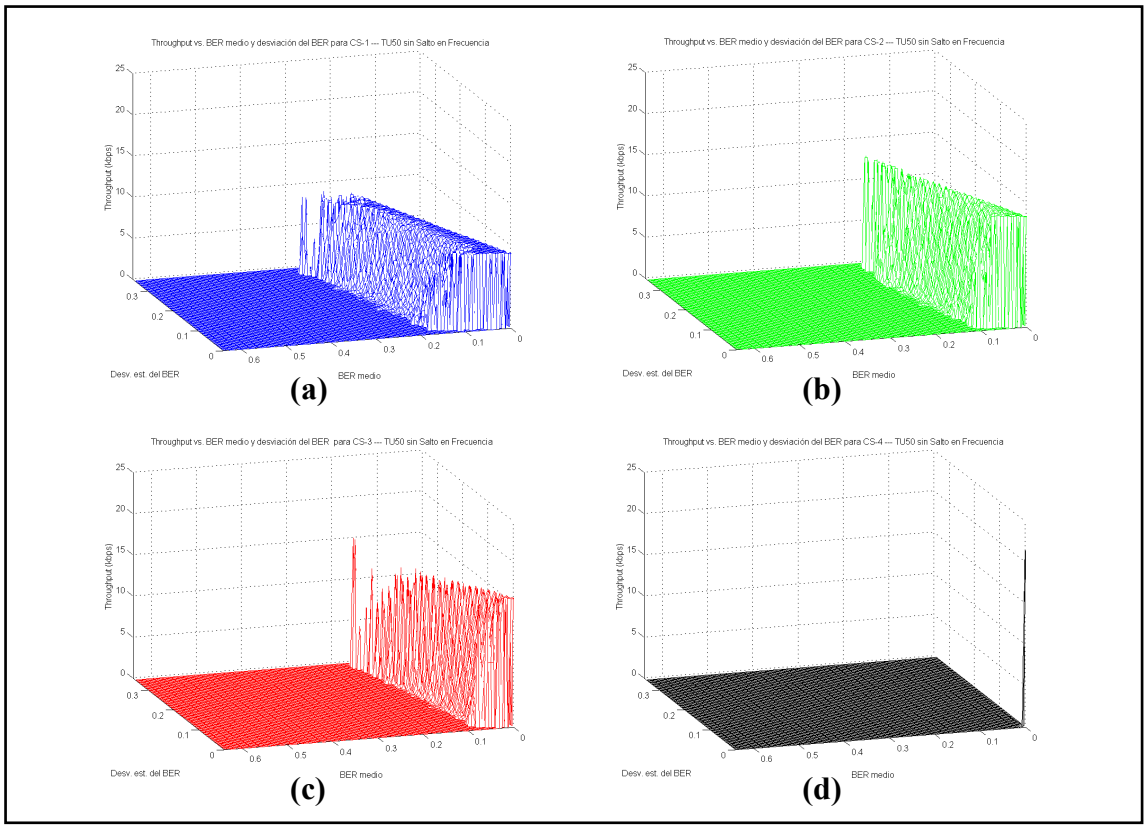


Figura 6.4.- Throughput vs BER medio y desviación del BER para: (a) CS-1, (b) CS-2, (c) CS-3, (d) CS-4

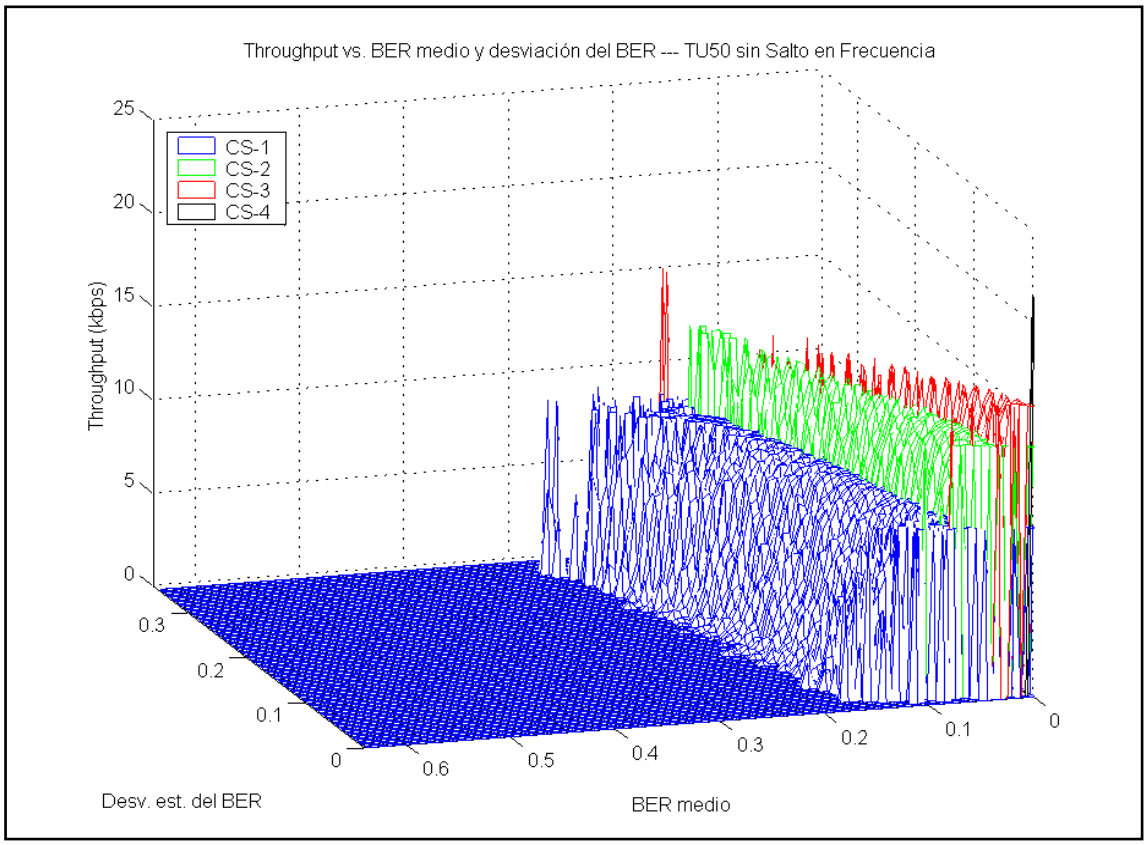


Figura 6.5.- Throughput vs BER medio y desviación del BER

La figura 6.6 muestra el *mapa de decisión* para este algoritmo (que no es más que la vista superior de la figura 6.5) a partir del cual es inmediato determinar el esquema de codificación que proporciona el mayor throughput para cada combinación de BER medio y desviación estándar del BER. Cada vez que haya que tomar una decisión en tiempo de simulación acerca del esquema de codificación a emplear, se determinará la calidad del canal (según se describió en el apartado 5.4.2) y, a partir de los valores del BER medio y de su desviación estándar se establecerá, empleando la figura 6.6, qué esquema de codificación es el que conviene emplear.

Puede apreciarse en la figura 6.6 que la región en la que CS-4 proporciona el mayor throughput corresponde tan sólo a un punto de la gráfica, en el cual el BER medio, como se ha dicho anteriormente, es nulo y, consecuentemente, también lo es la desviación estándar del BER. La razón de este hecho es que el esquema de codificación CS-4 es tan sumamente vulnerable frente a errores de transmisión que no resulta óptimo salvo que la transmisión se realice sin errores.

Por último, señalar que en el mapa de decisión se puede apreciar una sustancial cantidad de puntos en los que está indefinido el esquema de codificación a emplear. Ello se debe a que durante las simulaciones a nivel de enlace realizadas para obtener las curvas de la figura 6.3, estas combinaciones de BER medio y desviación del BER no se llegaron a producir (recordar del apartado 5.4.2 que en las curvas de la figura 6.3 estos puntos aparecen señalados con un valor negativo). Las combinaciones más habituales son aquellas con valores bajos de BER medio y de desviación del BER.

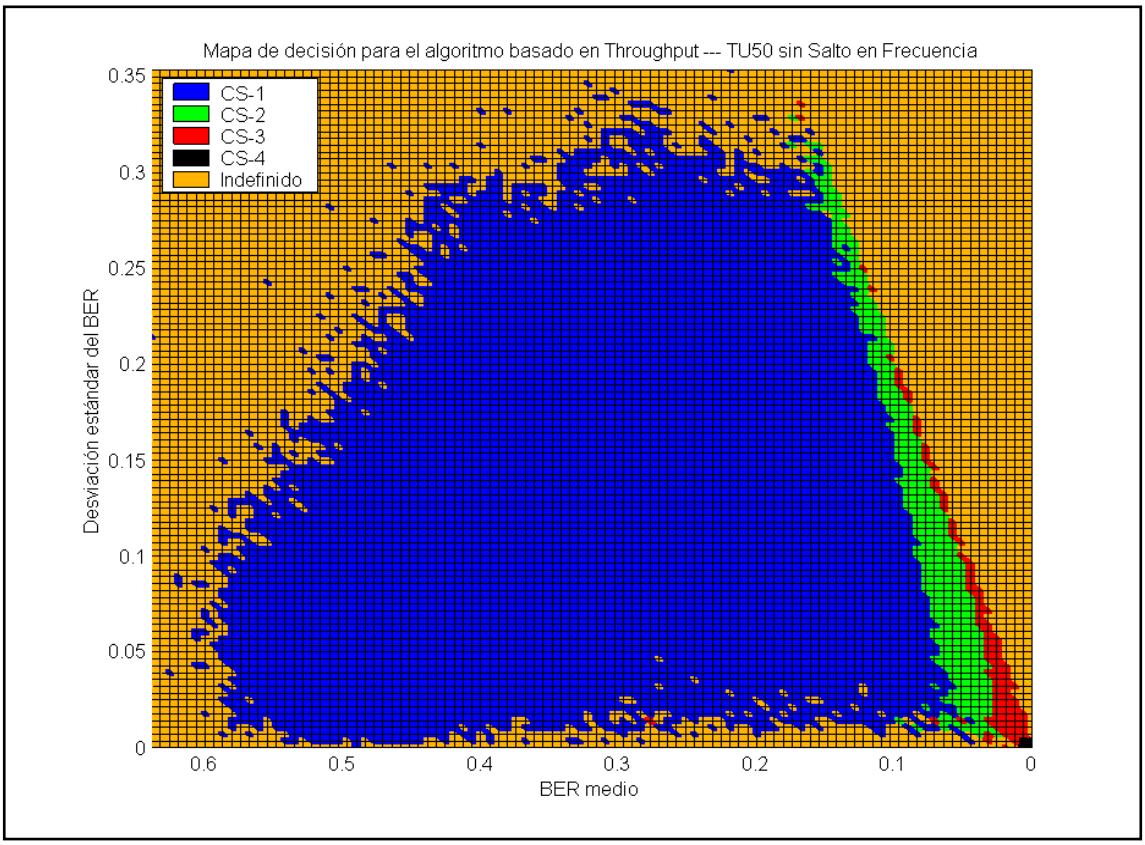


Figura 6.6.- Mapa de decisión para el algoritmo basado en throughput

6.2.2.- ALGORITMO BASADO EN BLER

El principal objetivo perseguido por este algoritmo no consiste en maximizar el throughput, como era el caso del algoritmo anterior, sino en preservar el valor del BLER experimentado por debajo de un determinado valor máximo. Este valor máximo se establece en función de la tasa máxima de errores que se considera aceptable para que la calidad experimentada por el usuario del servicio sea aceptable y depende del tipo de aplicación concreta: vídeo H.263, vídeo MPEG, audio, etc. La forma en que se establecen los umbrales de decisión empleando las curvas de las interfaces a nivel de bloque RLC se ilustra en la figura 6.7, en la cual se muestra el comportamiento del BLER en función de la relación CIR. Supongamos que para un determinado tipo de aplicación la tasa máxima de errores tolerable para que la calidad experimentada por el usuario sea aceptable corresponde a un BLER del 5%. En este caso, trazando una línea recta horizontal sobre la figura 6.7 para una ordenada igual a 0.05 (correspondiente a un BLER del 5%) se obtienen los umbrales de decisión. En la figura 6.7 se puede observar que CS-2 proporciona un BLER inferior al 5% cuando la relación CIR toma valores por encima de los 13 dB, CS-3 lo hace si la relación CIR toma valores superiores a los 15 dB mientras que CS-4 lo hace para valores por encima de los 24.5 dB. Estos tres valores serán los tres umbrales de decisión en función de los cuales se decidirá el esquema de codificación a emplear según la calidad del enlace radio. Observar que este criterio escoge aquel esquema de codificación que, proporcionando un BLER inferior al máximo aceptable, ofrece el mayor throughput.

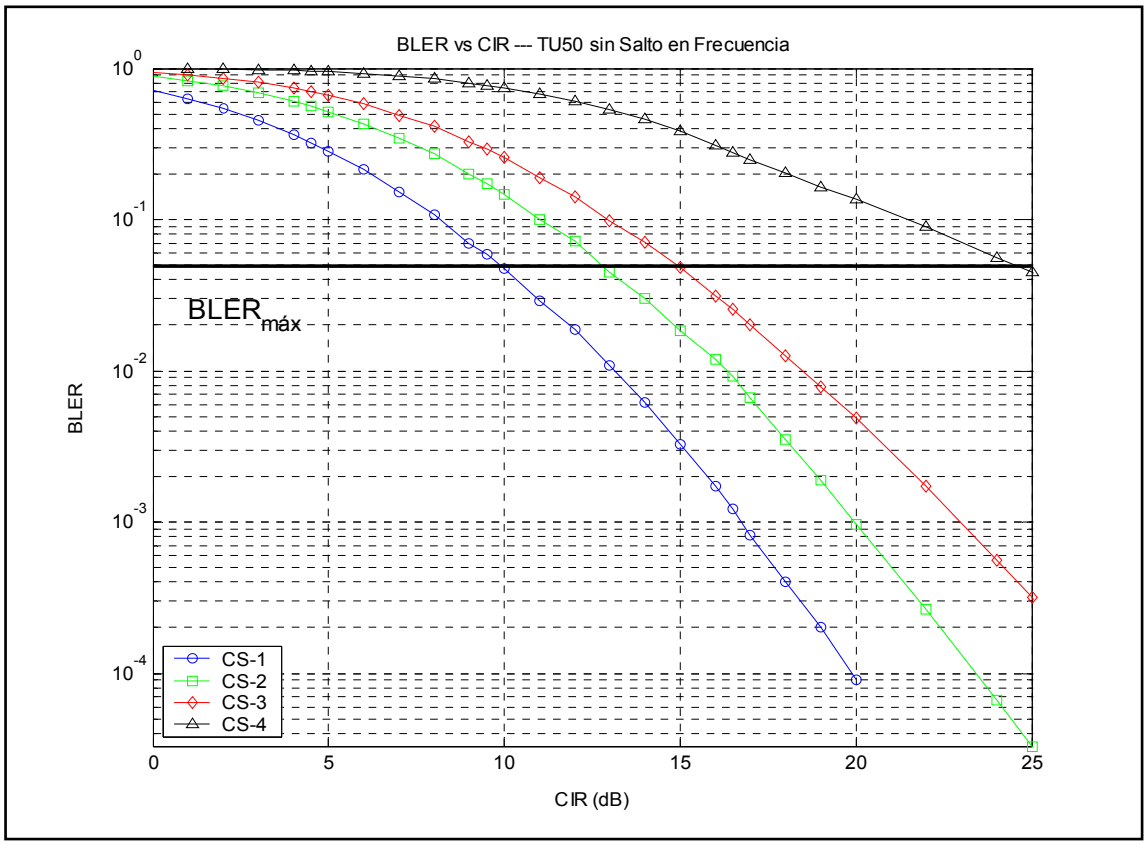


Figura 6.7.- BLER vs CIR

La implementación de este algoritmo mediante las interfaces a nivel de ráfaga se lleva a cabo a partir de las curvas mostradas en la figura 6.3. El procedimiento es análogo al mostrado en la figura 6.7: consiste en representar sobre los mismos ejes las curvas mostradas en la figura 6.3 y, haciendo uso de una superficie plana a la altura del valor del BLER deseado (5% por ejemplo), seccionar estas curvas (se omite esta gráfica por resultar bastante confusa). Realizando una proyección vertical del BLER sobre esta superficie se obtiene el mapa de decisión mostrado en la figura 6.8, el cual determina el esquema de codificación a emplear en función del BER medio y de su desviación.

Puede apreciarse que el aspecto del mapa de decisión que se obtiene con este algoritmo tiene un aspecto bastante similar al obtenido con el algoritmo basado en throughput. En primer lugar, los puntos correspondientes a combinaciones de BER medio y desviación del BER para los que no está definido el esquema de codificación a emplear son exactamente los mismos que en el caso anterior. Recordar que estas combinaciones no ocurrieron en las simulaciones a nivel de enlace realizadas para obtener las curvas de la figura 6.3. Puesto que para la definición de ambos algoritmos se utilizan estas curvas, ambos tienen estos puntos peculiares en común. En segundo lugar, se vuelve a observar que el esquema de codificación CS-4 se emplea sólo cuando el BER medio (y su desviación estándar) vale cero. De hecho, tal y como se observará más adelante, ésta es una característica común a cualquier algoritmo definido mediante las interfaces a nivel de ráfaga. La principal diferencia entre ambos mapas de decisión está en la zona derecha inferior, correspondiente, como ya se ha dicho anteriormente, a las combinaciones más frecuentes de BER medio y desviación estándar del BER.

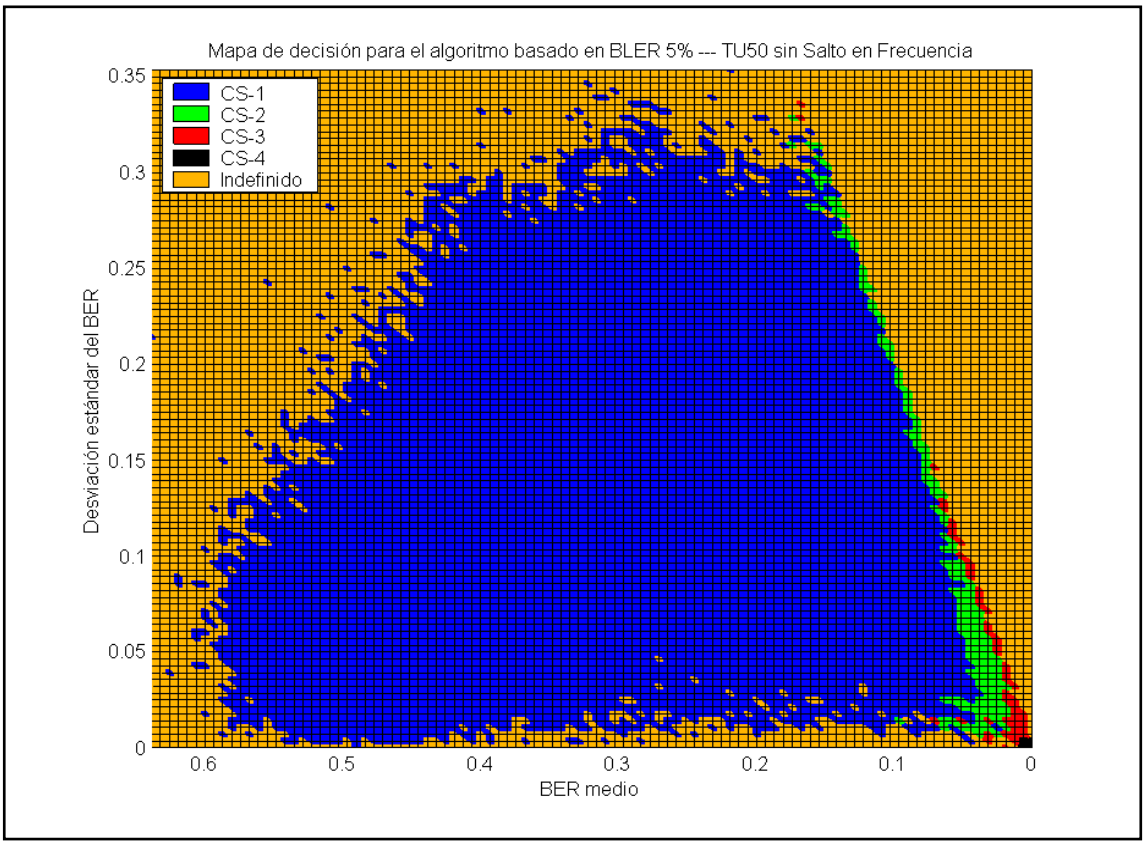


Figura 6.8.- Mapa de decisión para el algoritmo basado en BLER (para un BLER del 5%)

6.2.3.- PRIMER ALGORITMO PROPUESTO

El diseño de este algoritmo, a diferencia de los dos anteriores, se lleva a cabo teniendo en cuenta tanto el throughput como el BLER. Lo que con ello se pretende es lograr un algoritmo mediante el cual sea posible alcanzar un compromiso entre las prestaciones ofrecidas por ambos. En principio, podría parecer paradójico afirmar que existe un compromiso entre BLER y throughput, es decir, que la mejora de uno supone el empeoramiento del otro ya que, a partir de la expresión (6.1), se puede afirmar que una disminución de la tasa de errores lleva asociado un incremento del throughput. Sin embargo, esta afirmación es cierta al considerar una transmisión individual. En un sistema de comunicaciones móviles, en el que interviene multitud de usuarios, lo que importa no es el throughput o la tasa de errores de un usuario particular sino el throughput o la tasa de errores global del sistema. En este caso, el throughput y la tasa de errores sí son dos parámetros comprometidos [108]. La justificación es la que se describe a continuación.

Puede apreciarse en las figuras 6.2 y 6.7 que el criterio de preservar el BLER bajo un determinado valor proporciona unos umbrales de decisión superiores a los proporcionados por el criterio de maximizar el throughput. Según esta observación, con el algoritmo basado en BLER se requiere una mejor calidad del canal (mayor relación CIR) respecto al algoritmo basado en throughput para poder emplear los esquemas de codificación con mayor velocidad binaria. Por ello, cabe esperar que en el algoritmo basado en BLER se utilicen más frecuentemente aquellos esquemas de codificación con menor tasa binaria (y mayor protección contra errores), ofreciendo de esta forma un menor throughput global en el sistema (y una menor tasa de errores). Análogamente, el algoritmo basado en throughput muestra unos umbrales de decisión menores, según lo cual parece lógico suponer que los esquemas de codificación con mayor velocidad binaria se utilicen de forma más frecuente que en el algoritmo basado en BLER, lo cual proporciona un throughput global más elevado a cambio de una peor tasa de errores. Además, también se puede observar que el valor de los umbrales de decisión que proporciona el algoritmo basado en BLER aumenta conforme el BLER requerido por la aplicación disminuye, según lo cual cabe esperar que conforme mejora el BLER global, el throughput global del sistema empeore de forma paralela.

Aunque la explicación anterior se basa en las figuras obtenidas mediante las interfaces a nivel de bloque RLC, el comportamiento descrito se observa igualmente en aquellas obtenidas mediante las interfaces a nivel de ráfaga. Comparando las figuras 6.6 y 6.8 se puede apreciar, observando la región correspondiente a las combinaciones más habituales de BER medio y desviación estándar del BER (región derecha inferior), que para el algoritmo basado en BLER se requiere un BER medio más bajo que en el algoritmo basado en throughput para poder emplear los esquemas de codificación con mayor velocidad binaria. Igualmente, en el algoritmo basado en throughput resulta suficiente un menor BER medio respecto al algoritmo basado en BLER para poder emplear los esquemas de codificación con mayor velocidad binaria. Requerir un BER medio más bajo es sinónimo de requerir una mayor relación CIR o, en definitiva, una mejor calidad del canal de transmisión (y vicerversa), por lo que el razonamiento basado en las figuras 6.2 y 6.7 también es aplicable a las figuras 6.6 y 6.8.

Debe quedar claro que el razonamiento anterior no es en sí una demostración de que throughput y BLER son dos factores comprometidos a nivel de sistema, sino tan sólo una explicación intuitiva de dicho fenómeno. Sin embargo, y anticipando eventos, los resultados obtenidos en [104] y en las simulaciones realizadas sí demuestran que, efectivamente, el compromiso entre ambos factores realmente existe.

Así pues, lo que se pretende con este primer algoritmo propuesto es buscar un compromiso entre ambos factores. En un primer intento por establecer un compromiso, parece lógico evaluar la expresión

$$\frac{Throughput}{BLER} \tag{6.2}$$

en función de la relación CIR, en la cual intervienen ambos factores. La razón de evaluar esta expresión se debe a que interesa que el throughput sea lo más elevado posible, pero también conviene que el BLER se mantenga en valores cuanto más bajos mejor. Por ello, interesará escoger aquel esquema de codificación que, para cada valor de la relación CIR, maximice la expresión (6.2). La idea, pues, es evaluar este cociente para los diversos valores de la relación CIR y observar, en cada caso, qué esquema de codificación proporciona el mejor compromiso. A partir de las figuras 6.1 y 6.2 resulta inmediato obtener la figura 6.9, en la cual se representa, para cada uno de los cuatro esquemas de codificación, la expresión (6.2) en función de la relación CIR.

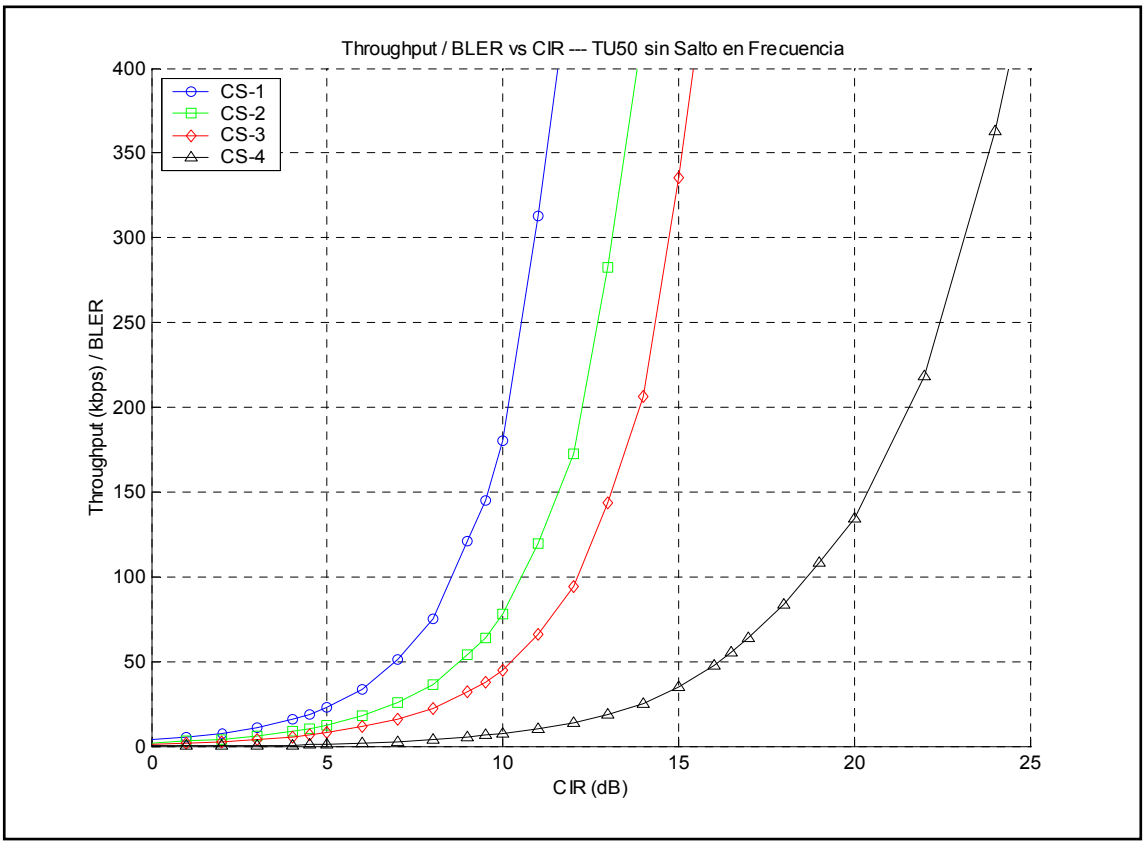


Figura 6.9.- Throughput / BLER vs CIR

La figura 6.9 muestra algo bastante peculiar ya que, según se puede observar, el mejor compromiso entre throughput y BLER siempre lo proporciona el esquema de codificación CS-1, independientemente de cuál sea la calidad del canal de transmisión. Observando las figuras 6.1 y 6.2 resulta sencillo encontrar una explicación para este comportamiento. Se puede apreciar que para valores bajos de la relación CIR, el throughput proporcionado por cada esquema de codificación es prácticamente nulo y el BLER es prácticamente igual a uno, por lo que para este rango de valores de CIR, el cociente de la expresión (6.2) vale aproximadamente cero. Conforme aumenta el valor de la relación CIR, para valores medios, el throughput comienza a aumentar al mismo tiempo que el BLER comienza a disminuir, lo cual explica que las curvas comiencen a crecer. Cuando la relación CIR toma valores altos, el throughput aumenta ligeramente hasta que se estabiliza en un valor constante (igual a la velocidad binaria del esquema de codificación en cuestión) mientras que el BLER sigue decreciendo y toma valores muy pequeños. En este rango de valores altos de relación CIR, el BLER proporcionado por cualquier esquema de codificación es tan pequeño que el valor de la expresión (6.2) toma valores cada vez más y más grandes a medida que la relación CIR aumenta.

Así pues, la principal conclusión que se extrae de estas observaciones es que, si bien es cierto que las curvas de la figura 6.9 son representativas del comportamiento tanto del throughput como del BLER, también es cierto que no lo son, en igual medida, a lo largo de todo el rango de valores de la relación CIR. Efectivamente, para valores altos de la relación CIR, el throughput toma valores constantes pero las curvas de la figura 6.9 se disparan hacia valores cada vez más elevados debido a que el BLER toma valores muy cercanos a cero. En esta región de la gráfica, la evolución del BLER es un factor claramente dominante frente al throughput, razón por la cual, para valores altos de la relación CIR, la forma de las curvas de la figura 6.9 viene determinada, principalmente, por el comportamiento del BLER.

Recordar que lo que se pretende con este algoritmo es obtener un compromiso entre throughput y BLER, para lo cual se precisa que ambos factores sean considerados de una forma más o menos equitativa. Por ello, la expresión (6.2) no resulta de utilidad para el objetivo perseguido por este algoritmo, ya que las curvas a las que da lugar, si bien es cierto que para valores bajos y medios de la relación CIR son representativas del comportamiento tanto del throughput como del BLER, para valores elevados de la relación CIR lo son, principalmente, del comportamiento del BLER. Sería deseable, en esta situación, que para valores altos de la relación CIR la importancia de los valores del BLER en la expresión (6.2) fuera menor para que, de esta forma, las curvas obtenidas fueran representativas, no sólo del comportamiento del BLER, sino también del comportamiento del throughput a lo largo de todo el rango de valores de interés de la relación CIR, ya que en este algoritmo ambos factores son relevantes.

En caso de que en la determinación del algoritmo uno de los dos parámetros debiera ser dominante frente al otro para valores altos de relación CIR, sería preferible que éste fuera el throughput y no el BLER. En efecto, para valores elevados de relación CIR, los valores del BLER son, en general, tan reducidos que la disminución del BLER en este rango de valores no posee tanta relevancia como la que puede llegar a tener un aumento del throughput; en este caso, el throughput debería ser el factor dominante en la elección del esquema de codificación de canal y no el BLER.

Por lo tanto, siguiendo el planteamiento anterior, una posibilidad consistiría en reemplazar la expresión (6.2) por la siguiente

$$\frac{Throughput}{f(BLER)} \tag{6.3}$$

donde *f*(BLER) representa una función matemática que, aplicada sobre el BLER, permita conseguir el efecto buscado de reducir su importancia en la expresión (6.3) para valores altos de la relación CIR. Esta reducción debe ser tal que, para este rango de valores de CIR, throughput y BLER tengan un grado de relevancia más o menos similar en la determinación del comportamiento de las curvas obtenidas mediante la expresión (6.3). En caso de que no fuera posible encontrar una función *f*(BLER) tal que las curvas obtenidas estuvieran determinadas en todo el rango de valores de la relación CIR de una forma más o menos equitativa por ambos factores, sería preferible, como se ha comentado anteriormente, que fuera el throughput y no el BLER el que determinara el comportamiento para valores altos de la relación CIR, si bien para valores medios sigue siendo deseable que, al igual que en la expresión (6.2), ambos parámetros tengan un grado de relevancia aproximadamente similar, ya que en este caso ninguno de ellos puede ser considerado de mayor trascendencia que el otro.

Llegados a este punto, surge el problema de encontrar una función *f*(BLER) que resulte adecuada para estos propósitos. Hay que tener en cuenta que este algoritmo queda abierto a la posibilidad de múltiples variantes, las cuales pueden ser fácilmente derivadas a partir del planteamiento inicial mediante la elección de una determinada función *f*(BLER) u otra. Lo que aquí se pretende no es determinar qué función *f*(BLER) resulta óptima u ofrece mejores resultados que las demás, sino encontrar una función *f*(BLER) que, siendo sencilla, cumpla los requisitos anteriormente citados y permita, de esta forma, conseguir un algoritmo mediante el cual sea posible alcanzar el objetivo principal: establecer un compromiso entre throughput y BLER[43].

Siguiendo estas pautas, se probaron varias funciones *f*(BLER) sencillas y se estudió su comportamiento en función de la relación CIR. Algunas de ellas se muestran en la figura 6.10, en la que se representa *f*(BLER) así como la expresión (6.3) en función de la relación CIR. En las figuras 6.10a, la representación se realiza considerando la expresión (6.2), es decir, *f*(BLER) = BLER. Estas figuras son las mismas que la 6.1 (con el eje de ordenadas en magnitud lineal en lugar de logarítmica) y la 6.9. En ellas se puede apreciar, como se ha discutido anteriormente, que para valores elevados de la relación CIR, el BLER es el factor que determina el comportamiento de las curvas obtenidas mediante la expresión (6.3). Para *f*(BLER) = $BLER^m$, con *m* positivo (6.10b), las curvas muestran un comportamiento similar y el BLER sigue siendo el factor dominante. En el caso de *f*(BLER) = $BLER^m$, pero con valores de *m* negativos (6.10c), el BLER es, una vez más, el factor determinante para valores altos de la relación CIR. Si en las curvas anteriores la expresión (6.3) crecía a gran velocidad para valores CIR elevados porque *f*(BLER) tomaba valores pequeños, en este caso la expresión (6.3) se hace prácticamente cero porque *f*(BLER) toma valores elevados respecto a los del throughput, haciendo que las curvas obtenidas mediante la expresión (6.3) sigan estando determinadas, para valores elevados de la relación CIR, por el comportamiento del denominador.

[43] La búsqueda de otras funciones *f*(BLER) alternativas y el estudio de los resultados ofrecidos por diversas variantes del algoritmo quedan propuestos como posible trabajo futuro.

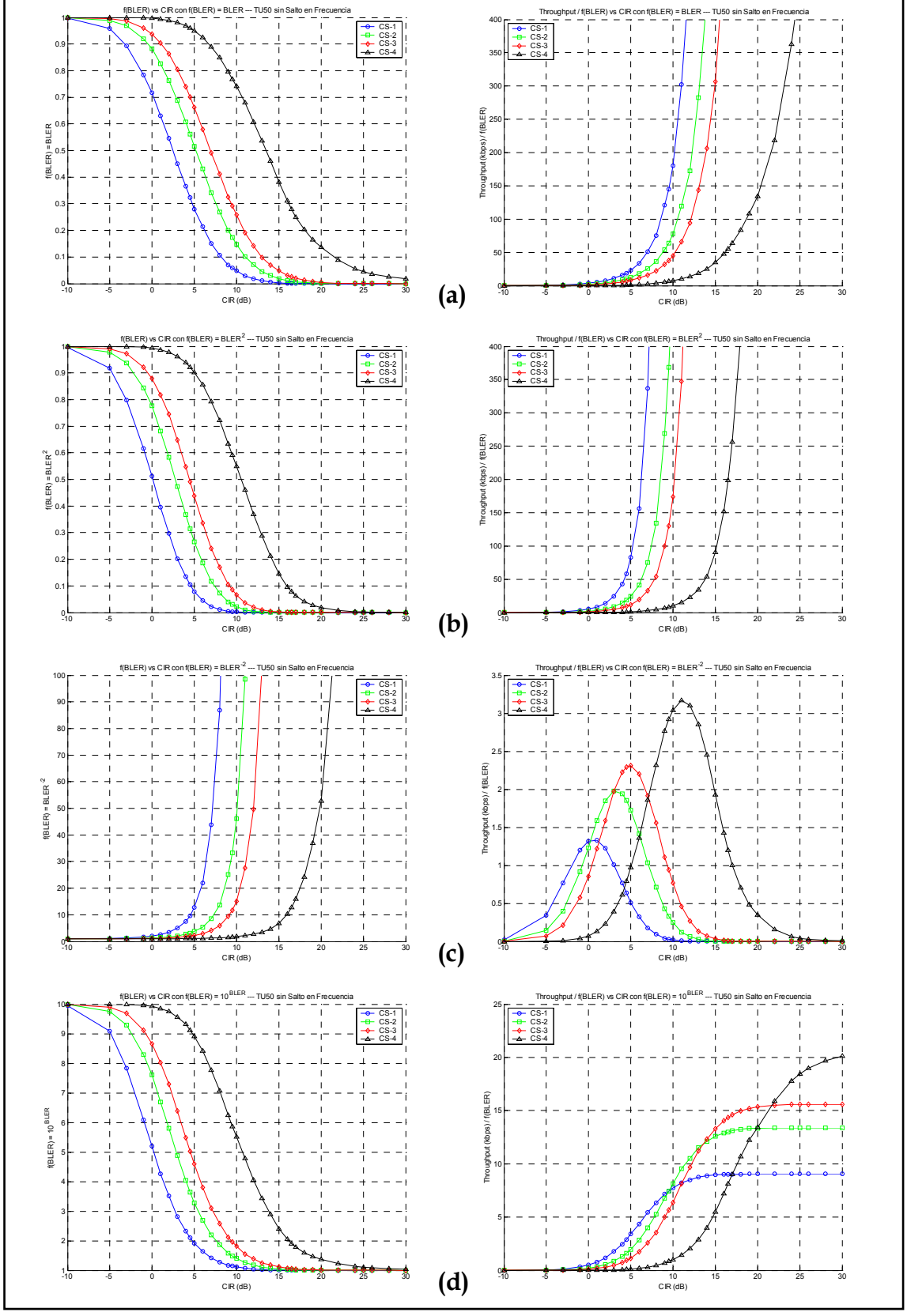


Figura 6.10.- f(BLER) vs CIR y Throughput / f(BLER) vs CIR para:
(a) f(BLER) = BLER, (b) f(BLER) = BLER^m (m = 2), (c) f(BLER) = BLER^m (m = -2), (d) f(BLER) = k^{BLER} (k = 10)

Sin embargo, cuando se considera $f(\text{BLER}) = k^{\text{BLER}}$ (figura 6.10d) se aprecian dos hechos que resultan bastante llamativos e interesantes. El primero de ellos es que, si se comparan las figuras 6.10a y 6.10d, se observa que al aplicar esta función f(BLER) sobre el BLER, las curvas obtenidas reproducen de forma fiel el comportamiento original del BLER frente a la relación CIR, a diferencia de $f(\text{BLER}) = \text{BLER}^m$ (con m negativo) o de otras funciones f(BLER) que, aunque aquí no aparezcan, también fueron estudiadas. El segundo de ellos es que, aunque la representación de la expresión (6.3) en función de la relación CIR empleando esta función f(BLER) pudiera llevar a pensar, dada la similitud con las curvas de la figura 6.2, que el comportamiento de las curvas obtenidas está determinado principalmente por el throughput a lo largo de todo el rango de valores de relación CIR, en realidad no es así, ya que la función hallada posee las características que, según se comentó más arriba, son deseables para la función f(BLER) buscada, tal y como se discute a continuación.

Al considerar $f(\text{BLER}) = k^{\text{BLER}}$ la expresión (6.3) queda como

$$\frac{Throughput}{k^{BLER}} \tag{6.4}$$

en la cual se pueden apreciar varias cosas. En primer lugar, para valores bajos de la relación CIR, el throughput es prácticamente nulo y el BLER vale aproximadamente uno, por lo que $f(\text{BLER}) = k$. En este caso la expresión (6.4) vale cero, tal y como se puede apreciar en la figura 6.10d. Cuando la relación CIR toma valores elevados, entonces el BLER vale aproximadamente cero, por lo que $f(\text{BLER}) \approx 1$ y el valor de la expresión (6.4) es aproximadamente igual al valor del throughput. En efecto, en la figura 6.10d se aprecia claramente que en este rango de valores de la relación CIR, el valor de la expresión (6.4) para cada esquema de codificación tiende hacia el valor de su velocidad binaria. Vemos, pues, que para valores elevados de la relación CIR las curvas obtenidas mediante la expresión (6.4) están determinadas, principalmente, por el comportamiento del throughput. Ésta es una de las características que se señalaron como convenientes anteriormente ya que, cuando la calidad del canal es buena, la tasa de errores proporcionada en general por los esquemas de codificación es bastante baja y el throughput juega un papel más importante. Sin embargo, para valores medios de la relación CIR, ninguno de estos dos factores puede ser considerado de mayor relevancia que el otro debido a que ni el throughput es lo suficientemente elevado como para que el BLER se pueda considerar más importante, ni el BLER resulta lo suficientemente reducido como para que el throughput se pueda considerar de mayor trascendencia. Por ello, es conveniente que, para valores medios de la relación CIR, ambos factores determinen, de una forma más o menos similar, el comportamiento de las curvas obtenidas. Esta otra característica deseable se observa que también tiene lugar con la expresión (6.4), ya que el throughput toma valores intermedios entre cero y la velocidad binaria del esquema de codificación considerado y f(BLER) toma valores intermedios dentro del intervalo $[1, k]$ para valores medios de la relación CIR.

Así pues, sin poder afirmar que $f(\text{BLER}) = k^{\text{BLER}}$ sea la función f(BLER) óptima o la que ofrece mejores resultados que otras, sí es posible concluir que resulta adecuada para las necesidades de este algoritmo ya que, en principio, promete lograr el objetivo propuesto: lograr un compromiso entre throughput y BLER a nivel de sistema.

La afirmación anterior se basa en la comparación de la representación gráfica de la expresión (6.4) en función de la relación CIR (figura 6.10d) con las figuras 6.2 y 6.7, donde se aprecia que los umbrales de decisión obtenidos con este nuevo algoritmo toman valores intermedios entre los proporcionados por los otros dos (el basado en throughput y el basado en BLER).

Umbral de decisión	Algoritmo basado en throughput	Primer algoritmo propuesto (k = 10)	Algoritmo basado en BLER (5%)
CS-1/CS-2	5 dB	9.5 dB	13 dB
CS-2/CS-3	9.5 dB	13.5 dB	15 dB
CS-3/CS-4	16.5 dB	21.5 dB	24.5 dB

Tabla 6.1.- Comparativa de los umbrales de decisión proporcionados por cada algoritmo

En la tabla 6.1 se aprecia que los umbrales de decisión del algoritmo basado en throughput son más pequeños que los análogos para el algoritmo basado en BLER. Tal y como se discutió al inicio de este apartado, éste es el motivo por el cual cabe esperar que el algoritmo basado en throughput proporcione un throughput global mejor que el algoritmo basado en BLER aunque una mayor tasa de errores. Análogamente, se espera que el algoritmo basado en BLER proporcione una mejor tasa de errores global pero un throughput global menor. Según este razonamiento, el hecho de que este algoritmo que se propone proporcione unos umbrales de decisión intermedios a los que se obtienen con los otros dos algoritmos hace suponer que el throughput global conseguido con este algoritmo será menor que el alcanzado con el algoritmo basado en throughput, aunque mayor que el proporcionado por el algoritmo basado en BLER. Por otra parte, la tasa global de errores debería ser mayor que la lograda con el algoritmo basado en BLER, pero menor que la observada empleando el algoritmo basado en throughput. Esta es la razón por la que el presente algoritmo, en principio, promete conseguir los fines propuestos.

Con la finalidad de analizar la evolución de los umbrales de decisión al variar el valor del parámetro k en la expresión (6.4), la figura 6.11 muestra la representación de dicha expresión en función de la relación CIR para diversos valores de k, exactamente, para k = 1, 2, 5, 10, 20, 50, 100 y 1000. Se puede apreciar que el valor de estas curvas es similar a las del algoritmo basado en throughput (figura 6.2) para valores bajos y altos de la relación CIR, independientemente del valor que tome el parámetro k (en el caso particular k = 1 las curvas son idénticas a las del algoritmo basado en throughput). La diferencia entre estas curvas está, por tanto, en la región correspondiente al rango de valores intermedios de la relación CIR (cuando $k \neq 1$), que precisamente corresponde a la región de la gráfica donde se establecen los umbrales de decisión. Puede apreciarse que el valor de estos umbrales, determinados por los puntos de cruce entre las diversas curvas, aumenta conforme lo hace el valor de k (ver la tabla 6.2). Esto significa que este algoritmo no sólo promete la consecución del buscado compromiso entre throughput y BLER sino que, además, permite modificar el grado de dicho compromiso con tan sólo variar el valor del parámetro k. Un aumento del valor de k supondría dar una mayor importancia a la tasa de errores, mientras que una reducción implicaría otorgar una mayor relevancia al throughput en el diseño del algoritmo.

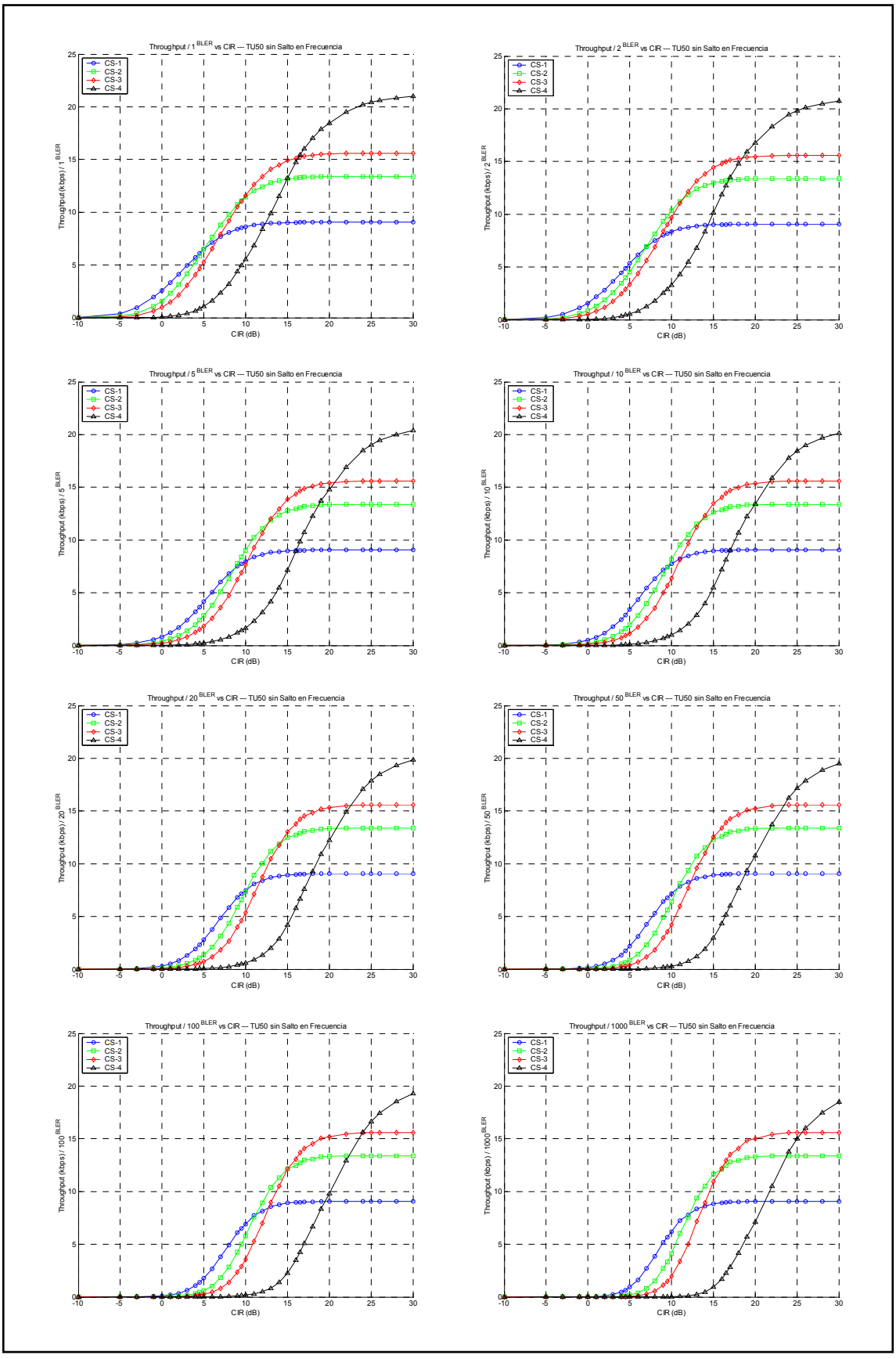


Figura 6.11.- Throughput / k^{BLER} vs CIR para k = 1, 2, 5, 10, 20, 50, 100 y 1000

Umbral	$k=1$	$k=2$	$k=5$	$k=10$	$k=20$	$k=50$	$k=100$	k=1000
CS-1/2	5 dB	7 dB	8.5 dB	9.5 dB	10 dB	10.5 dB	11 dB	12 dB
CS-2/3	9.5 dB	11.5 dB	13 dB	13.5 dB	14 dB	14.5 dB	15 dB	16 dB
CS-3/4	16.5 dB	18.5 dB	20.5 dB	21.5 dB	22.5 dB	23.5 dB	24 dB	25.5 dB

Tabla 6.2.- Umbrales de decisión del primer algoritmo propuesto para varios valores de k

Llegados a este punto, se podría objetar que si el compromiso entre ambos factores parece lograrse por el hecho de que los umbrales de decisión obtenidos toman valores intermedios a los de los umbrales proporcionados por los otros dos algoritmos, entonces dicho compromiso también podría alcanzarse fijando, de forma arbitraria, los umbrales de decisión en valores intermedios. Análogamente, también se podría variar su valor de forma arbitraria, lo cual también variaría el grado de compromiso entre throughput y BLER. Sin embargo, aunque esta forma de proceder podría conseguir variar el grado de compromiso, carece por completo de fundamento o razonamiento lógico alguno que la justifique. En cambio, este algoritmo se define considerando una expresión en la que tanto el throughput como el BLER intervienen a la hora de fijar los umbrales de decisión.

Hasta aquí la presentación de este algoritmo se ha realizado haciendo uso de las interfaces a nivel de bloque RLC. Si se representan las gráficas análogas a las de las figuras 6.9, 6.10 y 6.11 pero empleando las interfaces a nivel de ráfaga se observa que, como cabía esperar, el comportamiento de las curvas obtenidas mediante las interfaces a nivel de ráfaga frente al BER medio es similar al observado en las curvas obtenidas mediante las interfaces a nivel de bloque RLC frente a la relación CIR. Por ello no se repetirá todo el razonamiento partiendo desde el principio con las interfaces a nivel de ráfaga sino que, en su lugar, se presentarán directamente las gráficas más significativas en la definición de este algoritmo.

La figura 6.12 muestra la representación del throughput y de la expresión (6.4) para k = 10 frente al BER medio y la desviación estándar del BER. Aparentemente no existen grandes diferencias entre ambas; la figura 6.13 resulta más aclaradora.

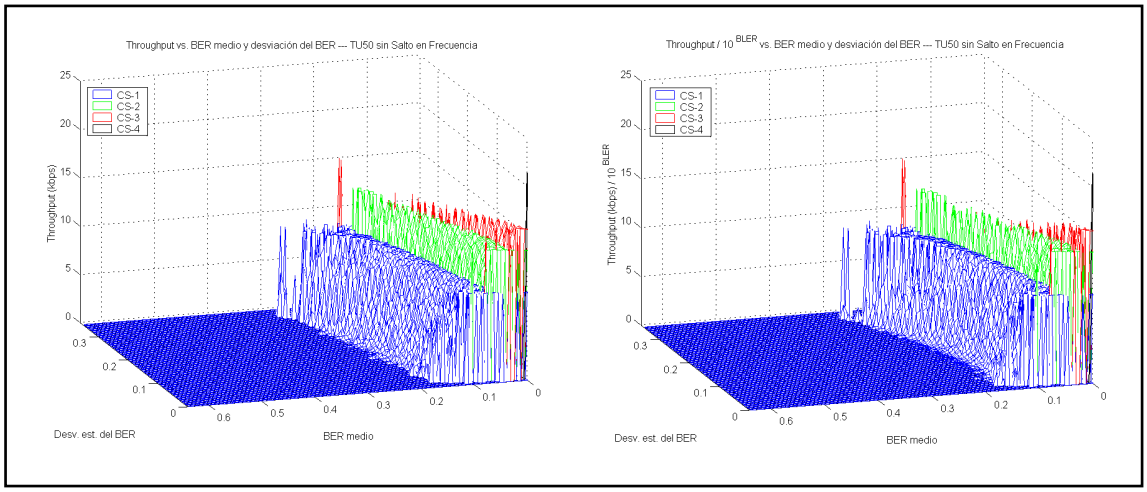


Figura 6.12.- (a) Throughput vs BER medio y desviación del BER,
(b) Throughput / 10^{BLER} vs BER medio y desviación del BER

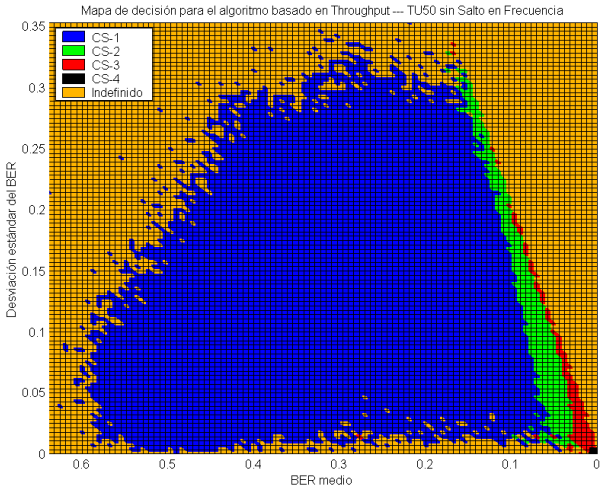
En la figura 6.13 se muestran los mapas de decisión para los tres algoritmos LA vistos hasta ahora: el algoritmo basado en throughput, el algoritmo basado en BLER (para un BLER del 5%) y el algoritmo que se acaba de presentar (para k = 10). Puede observarse que conforme se pasa de la figura 6.13a a la 6.13b y a la 6.13c, las regiones correspondientes a combinaciones de BER medio y desviación del BER en las que se utiliza CS-1 es cada vez más amplia, mientras que las regiones correspondientes a CS-2 y CS-3 cada vez son más reducidas (recordar que con estas interfaces, para CS-4 se obtiene siempre un único punto). Por lo tanto, se puede apreciar que el algoritmo propuesto requiere un BER medio más bajo (equivalente a una mayor relación CIR en el caso de las interfaces a nivel de bloque RLC o, en definitiva, a una mejor calidad del enlace radio) que el algoritmo basado en throughput para poder emplear los esquemas de codificación de mayor velocidad binaria. Además, el algoritmo basado en BLER requiere valores de BER medio todavía más bajos para poder utilizar estos esquemas de codificación. Este comportamiento de las gráficas obtenidas mediante las interfaces a nivel de ráfaga es equivalente al observado en las obtenidas mediante las interfaces a nivel de bloque RLC (mostrado en la tabla 6.1), donde los umbrales de decisión proporcionados por este algoritmo propuesto tomaban valores intermedios a los de los otros dos algoritmos.

Sería interesante, con fines comparativos, disponer en este caso de una tabla similar a la tabla 6.1; sin embargo, con las gráficas obtenidas a partir de las interfaces a nivel de ráfaga, el concepto de umbral de decisión resulta un tanto ambiguo. Aun así, resulta posible realizar una comparación cuantitativa entre los mapas de decisión. Para ello se tiene en cuenta el hecho de que cada mapa de decisión, tal y como se introduce en el simulador, es una matriz de 100 × 100 valores numéricos (0 para combinación desconocida de BER medio y desviación del BER, 1 para CS-1, 2 para CS-2, 3 para CS-3 y 4 para CS-4). La comparación podría realizarse en función del número de puntos correspondientes a cada esquema de codificación en el mapa de decisión del algoritmo (cada punto representa una combinación del BER medio y la desviación estándar del BER). En principio, esta comparación no es determinante ya que las combinaciones correspondientes a valores bajos tanto del BER medio como de la desviación estándar del BER son más probables que el resto. Sin embargo, es precisamente en esta región donde se producen la mayoría de los cambios de un mapa de decisión a otro (tal y como se puede apreciar en la figura 6.13), permaneciendo el resto del mapa de decisión prácticamente igual. Por esta razón, esta forma de comparar los algoritmos puede resultar bastante representativa.

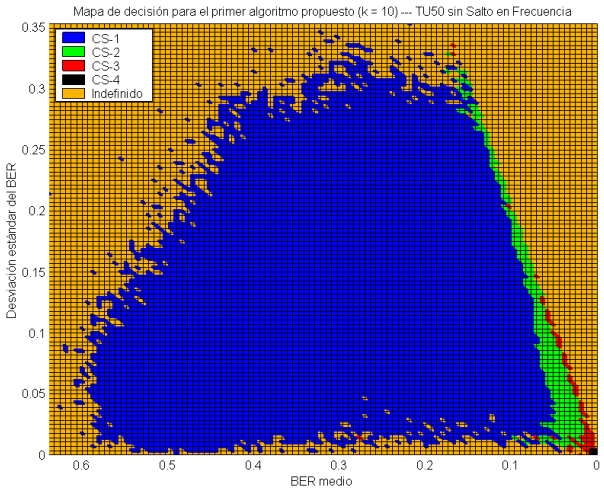
La tabla 6.3 muestra el número de puntos que corresponde a cada esquema de codificación de canal para los tres algoritmos vistos hasta ahora.

Esquema de codificación	Algoritmo basado en throughput	Primer algoritmo propuesto (k = 10)	Algoritmo basado en BLER (5%)
CS-1	5398	5517	5589
CS-2	299	221	151
CS-3	93	52	50
CS-4	1	1	1
Indefinido	4209	4209	4209

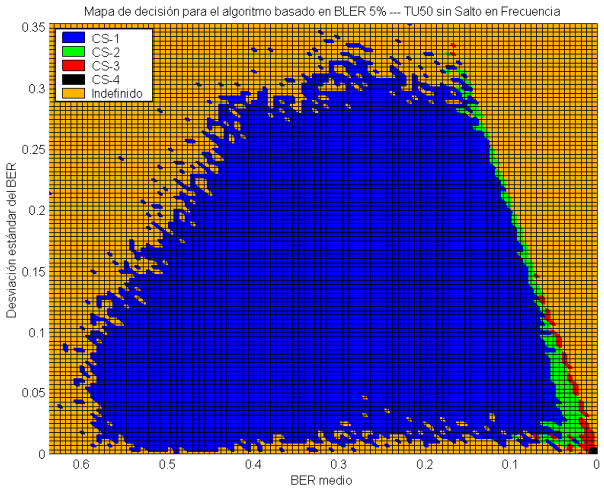
Tabla 6.3.- Número de puntos para cada esquema de codificación en los mapas de decisión



(a)



(b)



(c)

Figura 6.13.- Mapa de decisión para: (a) Algoritmo basado en throughput, (b) Primer algoritmo propuesto ($k = 10$), (c) Algoritmo basado en BLER (5%)

En la tabla 6.3 se puede apreciar que, para los tres algoritmos, CS-4 se utiliza siempre para una única combinación de BER medio y desviación estándar del BER. Esta combinación es siempre la misma para cualquier algoritmo y corresponde a un BER medio igual a cero y una desviación estándar del BER también igual a cero. Otro rasgo común a cualquier algoritmo definido mediante estas interfaces a nivel de ráfaga es el número de puntos (4209) correspondientes a combinaciones de BER medio y desviación estándar del BER que no llegaron a ocurrir durante las simulaciones a nivel de enlace realizadas para obtener las curvas de la figura 6.3.

Centrando la atención sobre la cantidad de puntos del mapa de decisión en que se utiliza cada esquema de codificación, se puede observar que el algoritmo para el que el número de puntos correspondientes a CS-1 es mayor es, tal y como cabía esperar, el algoritmo basado en BLER, ya que es en este algoritmo donde se espera que se utilicen más frecuentemente los esquemas de codificación con mayor protección ante errores. El número de puntos de CS-1 en el algoritmo basado en throughput es menor que en el algoritmo basado en BLER y, para el algoritmo propuesto, está en un valor intermedio. Puesto que el número de puntos en cada mapa de decisión es constante, el incremento del número de puntos correspondientes a CS-1 conforme se pasa del algoritmo basado en throughput al algoritmo propuesto y al algoritmo basado en BLER, va acompañado de una reducción del número de puntos en los que se emplea CS-2 y CS-3. Es decir, se intenta preservar la tasa de errores (algoritmo basado en BLER) aumentando el número de puntos en que se emplea el esquema de codificación CS-1 (que es el que mayor protección frente a errores proporciona) y reduciendo el número de puntos para CS-2 y CS-3 (que proporcionan menor protección). Igualmente, se intenta lograr un mayor throughput reduciendo el número de puntos para CS-1 (que es el que proporciona una menor velocidad binaria) y aumentando el número de puntos para CS-2 y CS-3 (que ofrecen una mayor velocidad binaria). Según estas observaciones, se puede afirmar que cuando se da mayor importancia al throughput, el número de puntos correspondientes a CS-1 se reduce para incrementar así el número de puntos para CS-2 y CS-3; de forma análoga, cuando se da mayor importancia a la tasa de errores, el número de puntos de CS-1 aumenta a costa de reducir el número de puntos asociados a CS-2 y CS-3. Los datos mostrados en la tabla 6.4 ratifican esta afirmación.

Esquema	$k = 1$	$k = 2$	$k = 5$	$k = 10$	$k = 20$	$k = 50$	$k = 100$	k=1000
CS-1	5398	5444	5495	5517	5535	5550	5559	5582
CS-2	299	268	236	221	209	199	193	174
CS-3	93	78	59	52	46	41	38	34
CS-4	1	1	1	1	1	1	1	1
Indefinido	4209	4209	4209	4209	4209	4209	4209	4209

Tabla 6.4.- Número de puntos para cada esquema de codificación para k = 1, 2, 5, 10, 20, 50, 100 y 1000

En efecto, un aumento del valor del parámetro k (equivalente, según se comentó anteriormente, a dar una mayor importancia a la tasa de errores) supone un aumento del número de puntos para CS-1 y una reducción del número de puntos para CS-2 y CS-3, mientras que la reducción de dicho parámetro (equivalente a considerar el throughput como factor más relevante) implica la reducción del número de puntos para CS-1 y el aumento del número de puntos para CS-2 y CS-3. Volvemos a observar, una vez más, que el presente algoritmo promete conseguir el objetivo inicialmente propuesto: establecer un compromiso entre throughput y BLER.

6.2.4.- SEGUNDO ALGORITMO PROPUESTO

En este apartado se presenta el segundo algoritmo propuesto en este trabajo, el cual se diseña considerando no sólo el throughput (como ha venido siendo habitual en la literatura) sino también otro factor adicional de especial relevancia para el caso que nos ocupa: el retardo. El retardo es, según la definición tradicional, el intervalo de tiempo transcurrido desde que un bloque de información[44] comienza a ser transmitido por el emisor hasta que éste ha sido completamente recibido por el receptor (retardo de transmisión). En una transmisión de datos en tiempo real, otro retardo importante y que afecta de forma crucial a la calidad de servicio experimentada por el usuario es el que se define como el exceso de tiempo incurrido en la transmisión de cada bloque de información (en lo sucesivo se denominará *retardo de trama*, para evitar confusiones). A continuación se describe de forma detallada la definición de retardo aquí considerada.

Supongamos que en el instante t_o se genera una trama de vídeo con un tamaño T y que al cabo de D unidades de tiempo se generará la siguiente trama de vídeo en el instante $t_o + D$. Ello implica que la trama generada en t_o debe haberse terminado de transmitir completamente en el instante $t_o + D$, instante en el cual debe comenzar a ser transmitida la siguiente trama. Si la calidad del canal da lugar a que la tasa binaria media a la cual se envía dicha trama de vídeo sea inferior a T/D, entonces en el instante $t_o + D$ la trama no se habrá terminado de transmitir completamente. Suponiendo que tras haber excedido el instante $t_o + D$ la trama de vídeo continúa transmitiéndose hasta haber completado su envío, entonces el instante en el cual se concluye la transmisión no será $t_o + D$, sino que será $t_o + D + \delta$, con $\delta > 0$. El parámetro δ es precisamente el retardo de trama que aquí se considera en la transmisión de la trama de vídeo y de su valor dependerá la calidad de servicio experimentada por el usuario.

Efectivamente, si $\delta > 0$, entonces la trama de vídeo no habrá llegado al receptor por completo en el instante en que debía haberlo hecho. El usuario que está viendo una secuencia de vídeo puede entonces experimentar fenómenos desagradables, en los que algunas de las imágenes se congelan durante breves intervalos de tiempo o incluso, si el retardo tiene lugar en cada trama de vídeo recibida, los movimientos se ralentizan. En cambio, si $\delta < 0$, no sólo habrá llegado la trama completa al receptor antes del instante debido sino que, además, los recursos que el usuario deja libre durante el tiempo que resta hasta que se genera la siguiente trama de vídeo pueden aprovecharse para incrementar la calidad de la transmisión de otros usuarios. En vídeo en tiempo real, si cuando se alcanza el instante $t_o + D$ no se ha transmitido por completo la trama de vídeo, la decisión tomada consiste en desechar el fragmento de trama que no ha dado tiempo a transmitir y comenzar a transmitir la siguiente trama, en lugar de seguir la transmisión de la trama hasta haber completado su envío. En este caso, el usuario no aprecia movimientos lentos sino distorsiones en la imagen de vídeo, las cuales pueden propagarse a lo largo de las siguientes imágenes de la secuencia.

[44] En el tipo de aplicación considerada en el presente trabajo (vídeo H.263) cada bloque de información está constituido por una trama de vídeo. Este algoritmo es igualmente aplicable a otro tipo de aplicaciones (vídeo MPEG o audio) en las que cada bloque de información lleve asociado un límite de tiempo para su envío.

Antes de continuar es necesario puntualizar que, cuando en el proceso de transmisión no se aplican retransmisiones de la información errónea, como es el caso, el retardo es función de la velocidad binaria del esquema de codificación empleado (independiente de la calidad del canal) y no del throughput que dicho esquema de codificación proporciona (dependiente de la calidad del canal). Recordar, en primer lugar, que la velocidad binaria se define como el cociente entre la cantidad de bits transmitidos (con o sin error) y el tiempo requerido para su transmisión, mientras que el throughput se define como la cantidad de bits correctamente transmitidos por unidad de tiempo. La velocidad binaria, por tanto, no depende de la calidad del canal de transmisión mientras que el throughput sí. Además, el retardo, tal y como se ha definido, dependerá, no de la cantidad de bits correctamente transmitidos por unidad de tiempo (throughput), sino de la cantidad total de bits (correctos o incorrectos) transmitidos por unidad de tiempo (velocidad binaria). En el supuesto caso de que se realizaran retransmisiones de la información errónea, entonces el retardo sí dependería del throughput, ya que cuanto menos información se envíe correctamente por unidad de tiempo serán necesarias más retransmisiones y, entonces, mayor será también el retardo. La razón de no considerar retransmisiones es, precisamente, mejorar el retardo a cambio, claro está, de un deterioro de la tasa de errores[45].

Así pues, queda claro que el retardo depende, no del throughput del esquema de codificación empleado, sino de su velocidad binaria (9.05 kbps para CS-1, 13.4 kbps para CS-2, 15.6 kbps para CS-3 y 21.4 kbps para CS-4). Teniendo en cuenta este hecho, la expresión analítica para el retardo de trama se halla como se describe a continuación. Para la transmisión sin retardo de una trama de vídeo con un tamaño T y una duración D se requiere una velocidad binaria igual a $V = T/D$. Supongamos que la transmisión de dicha trama se realiza íntegramente empleado el esquema de codificación de canal CS-i (i puede tomar los valores 1, 2, 3 ó 4) cuya velocidad binaria es $R_{CS\text{-}i}$. El tiempo que debe durar la transmisión sin retardo de la trama es $D = T/V$, mientras que el tiempo que realmente dura empleando el esquema de codificación CS-i es igual a $T/R_{CS\text{-}i}$. Por lo tanto, el retardo de trama incurrido durante la transmisión de esta trama de vídeo empleando el esquema de codificación CS-i vendrá dado por la expresión

$$Retardo\ de\ trama = \delta = \frac{T}{R_{CS-i}} - D \tag{6.5}$$

En esta expresión se puede apreciar que si la tasa binaria $R_{CS\text{-}i}$ ofrecida por el esquema de codificación CS-i considerado es menor que la tasa binaria T/D requerida por la trama de vídeo, entonces δ será positivo y habrá retardo de trama. En cambio, si $R_{CS\text{-}i}$ es mayor que T/D, entonces δ será negativo, lo que quiere decir que durante un tiempo $D - T/R_{CS\text{-}i}$ los recursos que estaban siendo utilizados quedarán libres y podrán ser aprovechados para mejorar la transmisión de otros usuarios. Otra observación a la expresión (6.5) es que la disminución del retardo de trama implica la utilización de un esquema de codificación con mayor velocidad binaria (y menor protección), lo que supone un incremento de la tasa de errores.

[45] Ante este dilema conviene tener en cuenta dos hechos. El primero es que el retardo es un parámetro crucial en la determinación de la calidad de servicio para una aplicación de vídeo. El segundo es que, independientemente de la existencia de técnicas de corrección de errores aplicadas en la decodificación de canal, también se aplican técnicas de corrección en la decodificación de fuente, lo cual palia las consecuencias de los errores de transmisión.

La definición de este algoritmo se basa en el hecho de que aquel criterio que maximiza el throughput no es, necesariamente, el que minimiza el retardo [107]. A modo de ejemplo, supongamos un sistema con una carga de usuarios más o menos elevada en el cual el nivel de interferencia global es considerable, dando lugar a que, en general, se experimenten relaciones CIR más bien pobres. En este caso, aquellos esquemas de codificación con mayor protección pueden proporcionar un throughput mayor que los que tienen mayor velocidad binaria ya que, aun siendo menor la cantidad de bits de datos que se transmiten por unidad de tiempo, al ser más robustos pueden llegar a ofrecer una mayor cantidad de bits correctos por unidad de tiempo, es decir, un mayor throughput[46]. El criterio que proporcionaría el mayor throughput global del sistema sería aquél que, bajo estas condiciones, propiciara una mayor utilización de aquellos esquemas de codificación con mayor protección frente a aquellos con mayor velocidad binaria. Sin embargo, un uso más frecuente de aquellos esquemas de codificación con menor velocidad binaria da lugar a un mayor retardo, con lo cual puede apreciarse que el criterio que maximiza el throughput no minimiza, necesariamente, el retardo. Por ello tiene sentido plantearse el comportamiento, frente a la relación CIR, de una expresión como la (6.6).

$$\frac{Throughput}{Retardo\ de\ trama} = \frac{Throughput}{\frac{T}{R_{CS-i}} - D} = \frac{Throughput \cdot R_{CS-i}}{T - D \cdot R_{CS-i}} \tag{6.6}$$

En caso de que la maximización del throughput implicara la minimización del retardo carecería de sentido plantear la expresión (6.6). Sin embargo, puesto que esto no es cierto en general, puede resultar interesante evaluar el comportamiento de dicha expresión en función de la relación CIR y observar, en cada caso, qué esquema de codificación es el que proporciona la mejor relación. No obstante, esta expresión tiene un pequeño inconveniente y es que, dada la definición de retardo de trama, cabe la posibilidad de que en ocasiones éste tome valores negativos (de hecho, esto es lo que nos interesa) en cuyo caso el esquema de codificación óptimo no resultaría ser aquél que proporciona el mayor valor para la expresión (6.6), ni tampoco el que ofrece el valor más reducido. Para aquellos esquemas de codificación para los que el retardo de trama toma valores negativos interesa que éste, en valor absoluto, sea lo más elevado posible; en esta ocasión, lo que interesa es maximizar la expresión

$$Throughput \cdot \left| Retardo\ de\ trama \right| = Throughput \cdot \left(D - \frac{T}{R_{CS-i}} \right) \tag{6.7}$$

Según este planteamiento, el algoritmo se definiría calculando el retardo de trama para cada trama de vídeo y, en función de si el esquema de codificación considerado proporciona un valor positivo o negativo, evaluar una expresión u otra. Dependiendo de la tasa binaria T/D requerida por la trama de vídeo, el retardo de trama proporcionado por cada esquema de codificación será positivo o negativo en función de si su velocidad binaria es menor o mayor, lo cual implica que se dará el caso en que se comparen dos esquemas de codificación mediante dos expresiones distintas, la (6.6) y la (6.7). Esta afirmación es totalmente cierta pero, como se demostrará a continuación, incluso hasta resulta conveniente.

[46] En la figura 6.2 puede apreciarse, por ejemplo, que para valores bajos de relación CIR existe una región en la que el throughput proporcionado por CS-1, CS-2 y CS-3 es mayor que el ofrecido por CS-4.

Antes de proceder a la evaluación de las expresiones (6.6) y (6.7) en función de la relación CIR, conviene señalar que para varias tramas de vídeo con tamaños (T) y duraciones (D) diferentes pero que requieren todas ellas la misma velocidad binaria (T/D), el valor del retardo de trama proporcionado por la expresión (6.5) es diferente para cada una de ellas. A pesar de ello, si la tasa binaria (T/D) requerida es la misma, aun dando lugar a diferentes valores de retardo de trama, las expresiones (6.6) y (6.7) en función de la relación CIR proporcionan unas curvas que dan lugar a los mismos umbrales de decisión. Por ello, para el análisis de las expresiones (6.6) y (6.7), es suficiente considerar tan sólo la velocidad binaria requerida por la trama de vídeo sin necesidad de tener en cuenta los valores específicos de tamaño y duración. En el anexo C se presenta el análisis detallado que corrobora estas afirmaciones.

La evaluación de las expresiones (6.6) y (6.7) se realizará considerando tramas de vídeo que requieren velocidades binarias de 6, 11, 14, 18, 28 y 40 kbps. Estos valores se han escogido para que se produzcan todas las posibles combinaciones de retardos de trama positivos y negativos con los diferentes esquemas de codificación: para una velocidad binaria de 6 kbps todos los esquemas de codificación proporcionan retardos de trama negativos según la expresión (6.5); para 11 kbps, CS-1 proporciona un retardo de trama positivo y todos los demás esquemas de codificación, un retardo de trama negativo; para 14 kbps, CS-1 y CS-2 dan lugar a retardos de trama positivos mientras que CS-3 y CS-4 ofrecen retardos de trama negativos; para 18 kbps, CS-4 es el único esquema de codificación que proporciona un retardo de trama negativo y para 28 kbps todos los esquemas de codificación ocasionan un retardo de trama positivo. El caso especial correspondiente a una velocidad binaria de 40 kbps (que no llega a tener lugar en la secuencia de vídeo simulada) se ha considerado con la intención de mostrar con mayor claridad la tendencia peculiar de las curvas cuando todos los retardos de trama toman valores positivos.

La figura 6.14 muestra el comportamiento de la expresión (6.6) y/o la expresión (6.7), según corresponda para cada esquema de codificación en cada gráfica, en función de la relación CIR. En la figura 6.14a se ha considerado una velocidad binaria requerida por la trama de 6 kbps, lo cual origina que el retardo de trama ofrecido por todos los esquemas de codificación según la expresión (6.5) sea negativo. En la figura 6.14b se puede apreciar que la curva correspondiente a CS-1 adquiere valores muy pequeños respecto al resto de curvas; la razón que explica esto es que para la velocidad binaria considerada (11 kbps), CS-1 ofrece un retardo de trama positivo mientras que el resto de esquemas de codificación ofrecen un retardo de trama negativo. Por ello, sobre CS-1 se aplica la expresión (6.6) y sobre el resto se aplica la expresión (6.7). Puede apreciarse, pues, que la expresión (6.6) ofrece valores mucho más pequeños que la expresión (6.7). Observando ahora la figura 6.14c puede apreciarse que se repite el mismo fenómeno, ahora también para CS-2, ya que, para tramas que requieren una velocidad binaria de 14 kbps, CS-2 también proporciona en este caso un retardo de trama positivo. Lo mismo sucede en la figura 6.14d con CS-3; en este caso particular se observa que, puesto que CS-4 es el único esquema de codificación que proporciona un retardo de trama negativo, es también el único que posee relevancia. A partir de estas observaciones, se puede afirmar que la forma en que se aplican las expresiones (6.6) y (6.7) en este algoritmo da lugar a que, para una trama de vídeo dada, aquellos esquemas de codificación que proporcionan un retardo de trama positivo (y degradan la calidad de la transmisión de vídeo) pierdan relevancia de forma considerable frente a aquellos otros que ofrecen un retardo de trama negativo (y mejoran la calidad de la

transmisión de vídeo). Esta es la razón por la que se señaló anteriormente que la comparación desigual entre esquemas de codificación a que daba lugar la aplicación de las expresiones (6.6) y (6.7) resultaba conveniente.

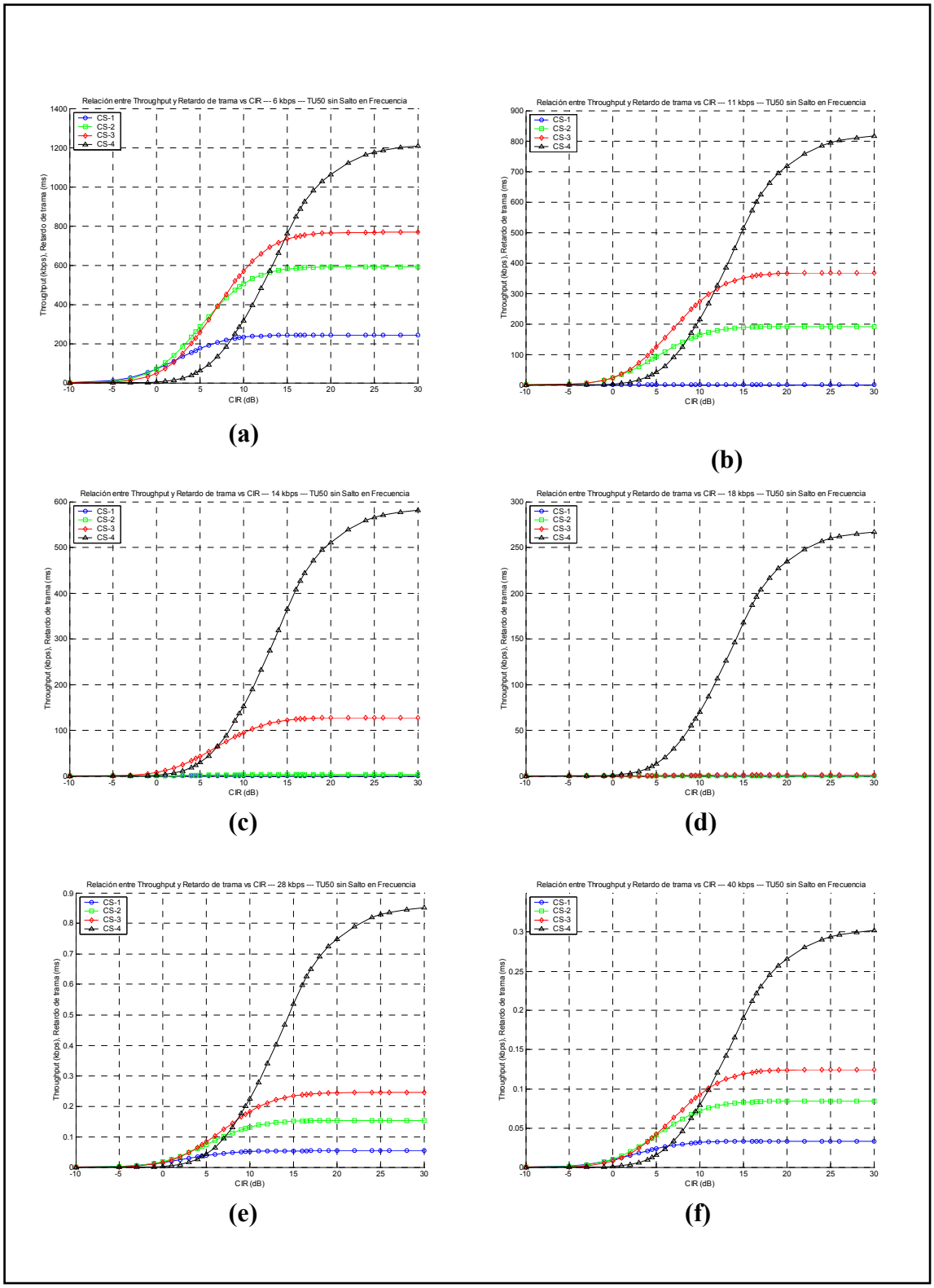


Figura 6.14.- Relación entre Throughput y Retardo de trama vs CIR para tramas que requieren diversas velocidades binarias: (a) 6 kbps, (b) 11 kbps, (c) 14 kbps, (d) 18 kbps, (e) 28 kbps, (f) 40 kbps

Otro hecho que se puede apreciar en la figura 6.14, y que resulta de bastante interés, es que conforme la velocidad binaria requerida por la trama de vídeo para ser transmitida sin retardo se incrementa, los umbrales de decisión proporcionados por el algoritmo son cada vez más bajos. Recordar que unos umbrales de decisión más bajos propiciaban un uso más frecuente de aquellos esquemas de codificación con mayor velocidad binaria, dando lugar a una mayor tasa binaria (y un menor retardo de trama) a cambio de una peor tasa de errores. Por lo tanto, se puede decir que éste es un algoritmo "inteligente" en la medida en que, para tramas que requieren una velocidad binaria mayor, proporciona unos umbrales de decisión más bajos con la finalidad de lograr una mayor tasa binaria y un menor retardo de trama a cambio de empeorar la tasa de errores. Cuando la trama de vídeo a transmitir requiere una velocidad binaria menor, entonces los umbrales de decisión toman valores más altos, lo cual da lugar a una mejora de la tasa de errores a cambio de una peor tasa binaria y un mayor retardo de trama; sin embargo, se puede permitir el empeoramiento de estos dos últimos factores ya que la tasa binaria que requiere la trama de vídeo para ser transmitida sin retardo es menor. Vemos, pues, que para cada trama de vídeo, el presente algoritmo proporciona unos umbrales de decisión diferentes que parecen buscar un compromiso entre tasa binaria (retardo de trama) y tasa de errores.

Frente a las dos características señaladas que muestra el presente algoritmo y que resultan bastante interesantes y deseables, también pueden apreciarse en la figura 6.14 otras dos características que, por el contrario, son más que indeseables. La primera de ellas, y probablemente la más evidente, es que los umbrales de decisión observados son sorprendentemente bajos. En efecto, ya para tramas que requieren una velocidad binaria de 6 kbps, donde la transmisión continua con CS-1 daría lugar a un retardo de trama negativo, los umbrales de decisión son incluso más bajos que los proporcionados por el algoritmo basado en throughput. Esto resulta contradictorio pues, en principio, sería lógico que esta situación ocurriera en aquellos casos en los que la trama de vídeo a transmitir necesitara una elevada velocidad binaria para ser transmitida sin retardo. Además, puede observarse en la figura 6.14d que, para una trama que requiera una velocidad binaria de 16 kbps[47], este algoritmo sugiere que se emplee durante toda la transmisión el esquema de codificación CS-4, independientemente de la calidad del canal cuando, en principio, podrían utilizarse todos los esquemas de codificación y el retardo de trama final experimentado podría ser nulo o incluso negativo siempre y cuando la velocidad binaria media experimentada durante la transmisión de la trama de vídeo fuera, respectivamente, igual o mayor que 16 kbps. Un segundo rasgo nada deseable que caracteriza la definición del presente algoritmo se puede deducir a partir de las figuras 6.14e y 6.14f. En ambas, el retardo de trama de cualquier esquema de codificación es positivo. Puede verse claramente que, para retardos de trama positivos, al incrementarse la velocidad binaria requerida por la trama de vídeo los umbrales de decisión también aumentan. Esta tendencia, además de contradecir la observada en las figuras anteriores, resulta indeseable: si para una trama que requiera 18 kbps se ha de emplear (según la figura 6.14d) el esquema de codificación CS-4 en cualquier situación con la intención de reducir el retardo, también debería hacerse para aquellas tramas que requieran velocidades binarias mayores, en lugar de sugerir, tal y como puede apreciarse en las figuras 6.14e y 6.14f, el empleo de otros esquemas de codificación.

[47] Aproximadamente el 50% de las tramas de la secuencia de vídeo empleada en las simulaciones requiere una velocidad binaria de 16 kbps para la transmisión sin retardo, según puede apreciarse en la figura B.1 del anexo B.

A continuación se discutirá el motivo que origina estos comportamientos tan extraños y se demostrará que la solución al problema, bien sencilla, permite subsanar estos inconvenientes al mismo tiempo que preserva las virtudes observadas.

En las expresiones (6.6) y (6.7), el valor considerado para R_{CS-i} fue, tal y como se comentó, el correspondiente a la velocidad binaria de cada esquema de codificación, es decir, 9.05 kbps para CS-1, 13.4 kbps para CS-2, 15.6 kbps para CS-3 y 21.4 kbps para CS-4. Estos valores corresponden a la velocidad binaria ofrecida por cada esquema de codificación suponiendo que cada usuario dispone de un único time-slot. Sin embargo, tal y como se demuestra en el anexo B, para que la secuencia de vídeo simulada se pueda transmitir con una calidad aceptable se requiere el empleo de más de un time-slot por usuario. Vemos, una vez más, que el algoritmo puede ser calificado como "inteligente" ya que, debido a que con un único time-slot por usuario la velocidad binaria de que se dispone resulta insuficiente para la transmisión sin retardo de un número considerable de tramas, los umbrales de decisión se establecen en valores considerablemente bajos con la esperanza de incrementar la velocidad binaria y reducir así el retardo de trama. Al ser insuficiente un time-slot por usuario, los umbrales de decisión toman valores sorprendentemente bajos, incluso para aquellas tramas que requieren una velocidad binaria reducida. Así pues, en base a estas observaciones es de esperar que si el diseño del algoritmo se realiza considerando un número de time-slots adecuado, los umbrales de decisión que se obtienen tomen valores más coherentes.

El problema ahora consiste en determinar el número de time-slots por usuario que son necesarios. Una posible solución consistiría en realizar el diseño considerando la indicación ofrecida por el anexo B. En dicho anexo se señala, a modo de guía, que el número de time-slots que se necesita por cada usuario varía en función del esquema de codificación empleado. El diseño podría realizarse bajo estas consideraciones; sin embargo, de cara a la implementación en un sistema real, esta decisión puede resultar equivocada por dos motivos. El primero es que el cambio a un esquema de codificación con menor velocidad binaria estaría condicionado por la disponibilidad de recursos en ese momento. El segundo es que, teniendo en cuenta que el periodo de actualización suele ser del orden de milisegundos, el número de cambios de esquema de codificación que se produce por segundo es considerable, ocasionando que la carga de señalización requerida por esta aproximación pueda resultar, muy probablemente, excesiva. Por ello es preferible que el número considerado de time-slots por usuario sea constante, independientemente del esquema de codificación que se emplee en cada momento. A partir de las indicaciones del anexo B se decidió considerar tres time-slots por usuario a la hora de realizar el diseño del algoritmo.

Bajo estas consideraciones, las expresiones (6.6) y (6.7) quedarían de la forma

$$\frac{Throughput}{Retardo\ de\ trama} = \frac{Throughput}{\frac{T}{N \cdot R_{CS-i}} - D} = \frac{Throughput \cdot N \cdot R_{CS-i}}{T - D \cdot N \cdot R_{CS-i}} \tag{6.8}$$

$$Throughput \cdot \left| Retardo\ de\ trama \right| = Throughput \cdot \left(D - \frac{T}{N \cdot R_{CS-i}} \right) \tag{6.9}$$

donde N representa el número de time-slots por usuario (3 en este caso).

En las expresiones (6.8) y (6.9) el valor del throughput, en principio, también se vería multiplicado por N[48], es decir,

$$Throughput_{CS-i} = N \cdot R_{CS-i} \cdot \left[1 - BLER_{CS-i}(CIR)\right] \tag{6.10}$$

sin embargo, el hecho de considerar o no el factor N en la expresión (6.10) no afecta a los umbrales de decisión obtenidos[49].

La evaluación de las expresiones (6.8) y (6.9) da lugar a las gráficas mostradas en la figura 6.15. En este caso, al realizarse el diseño considerando más de un time-slot por usuario, los valores seleccionados anteriormente de velocidad binaria para realizar el análisis no dan lugar a todas las posibles combinaciones de retardo de trama tal y como antes sucedía; sin embargo, sí resultan representativos de los posibles valores de velocidad binaria que pueden requerir las diversas tramas de vídeo que integran la secuencia de vídeo estudiada. Observando las gráficas de la figura 6.15 se observa que ahora los umbrales de decisión obtenidos no son tan bajos como en el caso anterior, debido a que el diseño se ha realizado considerando un número de time-slots por usuario más ajustado a las verdaderas necesidades de la secuencia de vídeo estudiada. Consecuentemente, los umbrales de decisión obtenidos son más coherentes. El segundo rasgo contradictorio que anteriormente se observó en las gráficas de la figura 6.14, es decir, el incremento de los umbrales de decisión conforme aumenta la velocidad binaria requerida por la trama de vídeo para ser transmitida sin retardo (cuando el retardo de trama proporcionado por todos los esquemas de codificación era positivo), no tiene lugar en este caso debido a que, tal y como puede deducirse, el valor mínimo de velocidad binaria para que el retardo de trama ofrecido por todos los esquemas de codificación sea positivo ha de ser $3 \times 21.4 = 64.2$ kbps. Este valor es muy superior al máximo observado en la secuencia de vídeo (del orden de los 32 kbps), por lo que con este nuevo diseño no se llega a producir este comportamiento indeseable.

Si bien las características indeseables que se observaban en el diseño inicial han desaparecido, aquellas que resultaban deseables se siguen manteniendo. Puede verse en la figura 6.15 que los umbrales de decisión, al igual que en el caso anterior, siguen decreciendo al aumentar la tasa binaria requerida por la trama de vídeo, hecho que se señaló anteriormente como deseable. Además, para una trama de vídeo dada, aquellos esquemas de codificación que proporcionan un retardo de trama positivo (y degradan la calidad de vídeo) pierden relevancia frente a aquellos cuyo retardo de trama es negativo (y mejoran la calidad), como se observa en las figuras 6.15e y 6.15f.

Las figuras 6.16 y 6.17 muestran las mismas curvas que la figura 6.15 pero para un diseño en el que se ha considerado 2 y 4 time-slots por usuario, respectivamente. Se aprecia que los umbrales de decisión de la figura 6.15 (3 time-slots por usuario) varían significativamente respecto a los de la figura 6.16 (2 time-slots por usuario), mientras que de la figura 6.17 (4 time-slots por usuario) a la 6.15 (3 time-slots por usuario) las

[48] En realidad, al aumentar el número de time-slots por usuario se incrementa el nivel de interferencia global en el sistema, propiciando un uso más habitual de los esquemas de codificación con menor velocidad binaria, por lo que al multiplicar por N el número de time-slots por usuario el throughput obtenido no es, en general, N veces mayor.

[49] Los umbrales de decisión son los puntos donde se cortan las curvas de cada esquema de codificación, es decir, donde toman valores iguales. Multiplicar por el mismo factor los valores de todas las curvas modificaría el valor de la ordenada donde se cortan las curvas, pero el valor de la abcisa, es decir, del umbral de decisión, seguiría igual.

diferencias son mucho menos acentuadas. Esto hace suponer que la mejora que cabe esperar al emplear 4 time-slots por usuario no compensa el incremento en la cantidad de recursos necesarios respecto al empleo de 3 time-slots por usuario, reforzando la idea de que la consideración de 3 time-slots por usuario constituye una buena elección.

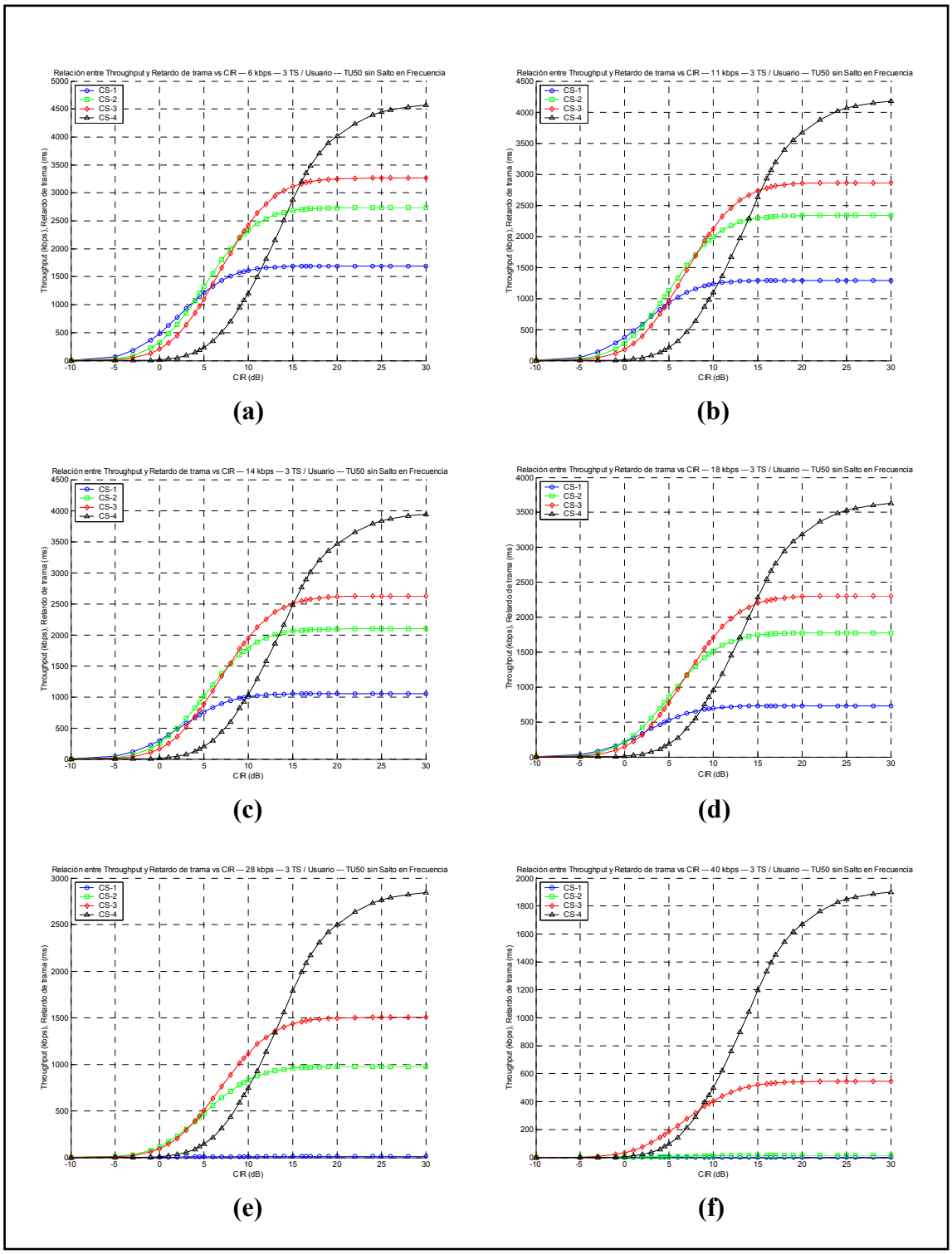


Figura 6.15.- Relación entre Throughput y Retardo de trama vs CIR, considerando 3 time-slots por usuario, para tramas que requieren diversas velocidades binarias:
(a) 6 kbps, (b) 11 kbps, (c) 14 kbps, (d) 18 kbps, (e) 28 kbps, (f) 40 kbps

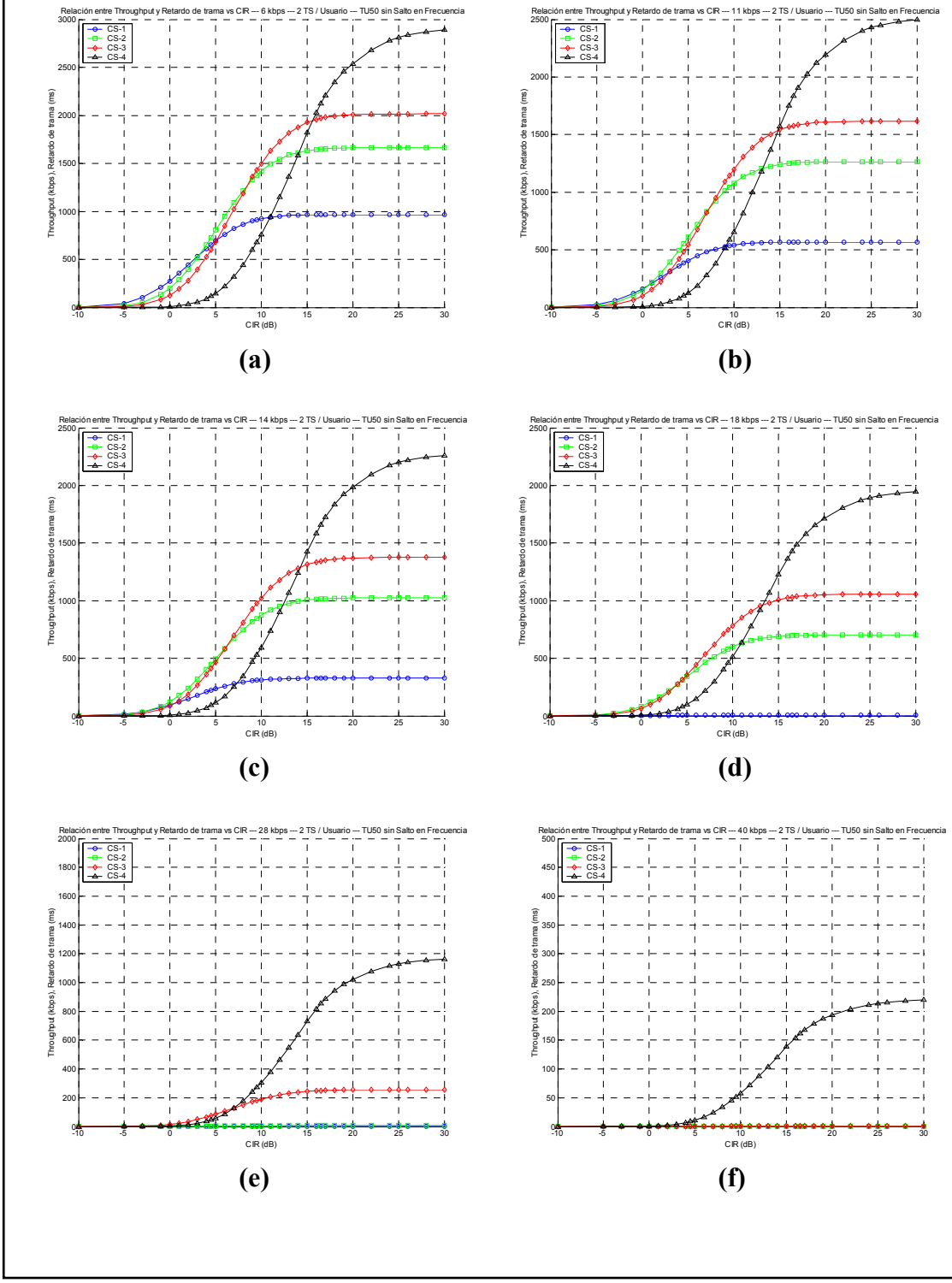


Figura 6.16.- Relación entre Throughput y Retardo de trama vs CIR, considerando 2 time-slots por usuario, para tramas que requieren diversas velocidades binarias:
(a) 6 kbps, (b) 11 kbps, (c) 14 kbps, (d) 18 kbps, (e) 28 kbps, (f) 40 kbps

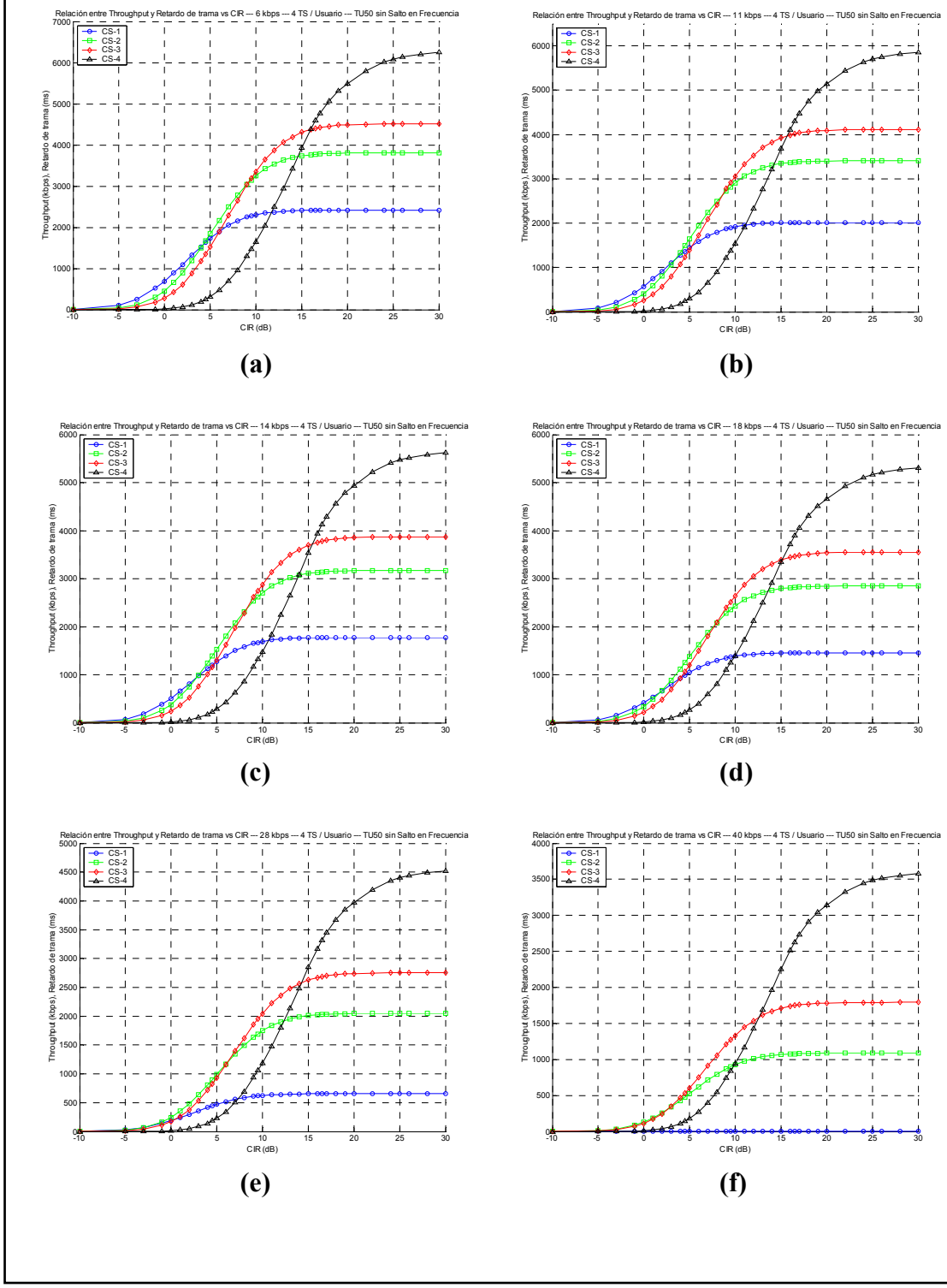


Figura 6.17.- Relación entre Throughput y Retardo de trama vs CIR, considerando 4 time-slots por usuario, para tramas que requieren diversas velocidades binarias:
(a) 6 kbps, (b) 11 kbps, (c) 14 kbps, (d) 18 kbps, (e) 28 kbps, (f) 40 kbps

Las explicaciones dadas hasta aquí han hecho uso de las interfaces a nivel de bloque RLC. El proceso a seguir empleando las interfaces a nivel de ráfaga es, como ya se podrá suponer, el mismo que el llevado a cabo con las interfaces a nivel de bloque RLC. La evaluación de las expresiones (6.8) y (6.9) con la curvas de las interfaces a nivel de ráfaga y con un valor de $N = 3$ da lugar a las curvas mostradas en la figura 6.18. En ellas se pueden apreciar las mismas tendencias que anteriormente se comentaron para las curvas de la figura 6.15. Conforme aumenta la velocidad binaria requerida por la trama de vídeo para su transmisión sin retardo, el BER medio necesario para emplear los esquemas de codificación con mayor velocidad binaria es menor, fomentando un mayor uso de estos esquemas con la finalidad de incrementar la tasa binaria y reducir el retardo de trama, a cambio de un empeoramiento de la tasa de errores. Al disminuir la tasa binaria requerida por la trama de vídeo se observa lo contrario. En cualquier caso, esta variación del BER medio frente a la velocidad binaria requerida por la trama de vídeo es equivalente a la observada para la relación CIR en el caso de la figura 6.15. Además, se observa nuevamente que para una trama dada, aquellos esquemas de codificación que dan lugar a un retardo de trama positivo pierden relevancia frente a los que ofrecen un retardo de trama negativo: en la figura 6.18e la tasa binaria (28 kbps) es superior a la proporcionada por CS-1 ($3 \times 9.05 = 27.15$ kbps) y, consecuentemente, la región correspondiente a CS-1 disminuye, mientras que en la figura 6.18f la tasa binaria considerada (40 kbps) da lugar a que CS-2 esté próximo a proporcionar un retardo de trama positivo ($3 \times 13.4 = 40.2$ kbps) y, por tanto, el fenómeno apreciado es el mismo.

Las observaciones que se acaban de realizar pueden apreciarse con una mayor claridad en la figura 6.19, donde se muestran los mapas de decisión de este algoritmo para las diferentes velocidades binarias de trama consideradas (con $N = 3$). La tabla 6.5 recoge el número de puntos que en el mapa de decisión de estas figuras corresponde a cada esquema de codificación. Los valores observados corroboran, una vez más, las observaciones realizadas.

Esquema de codificación	6 kbps	11 kbps	14 kbps	18 kbps	28 kbps	40 kbps
CS-1	5370	5340	5310	5245	5162	4683
CS-2	314	338	362	421	460	682
CS-3	106	112	118	124	168	425
CS-4	1	1	1	1	1	1
Indefinido	4209	4209	4209	4209	4209	4209

Tabla 6.5.- Número de puntos para cada esquema de codificación para diversas velocidades binarias

Al aumentar la velocidad binaria requerida por la trama de vídeo para su transmisión sin retardo se reduce el número de puntos correspondientes a CS-1 (con menor velocidad binaria) al mismo tiempo que se incrementa el número de puntos en los que se emplean CS-2 y CS-3 (con mayor velocidad binaria y capaces de ofrecer un menor retardo de trama). Si la velocidad binaria requerida por la trama de vídeo se reduce, entonces se observa el fenómeno contrario, ya que cuanto menor sea ésta se hace menos necesario emplear con mayor frecuencia aquellos esquemas de codificación con mayor velocidad binaria para la obtención de un bajo retardo de trama.

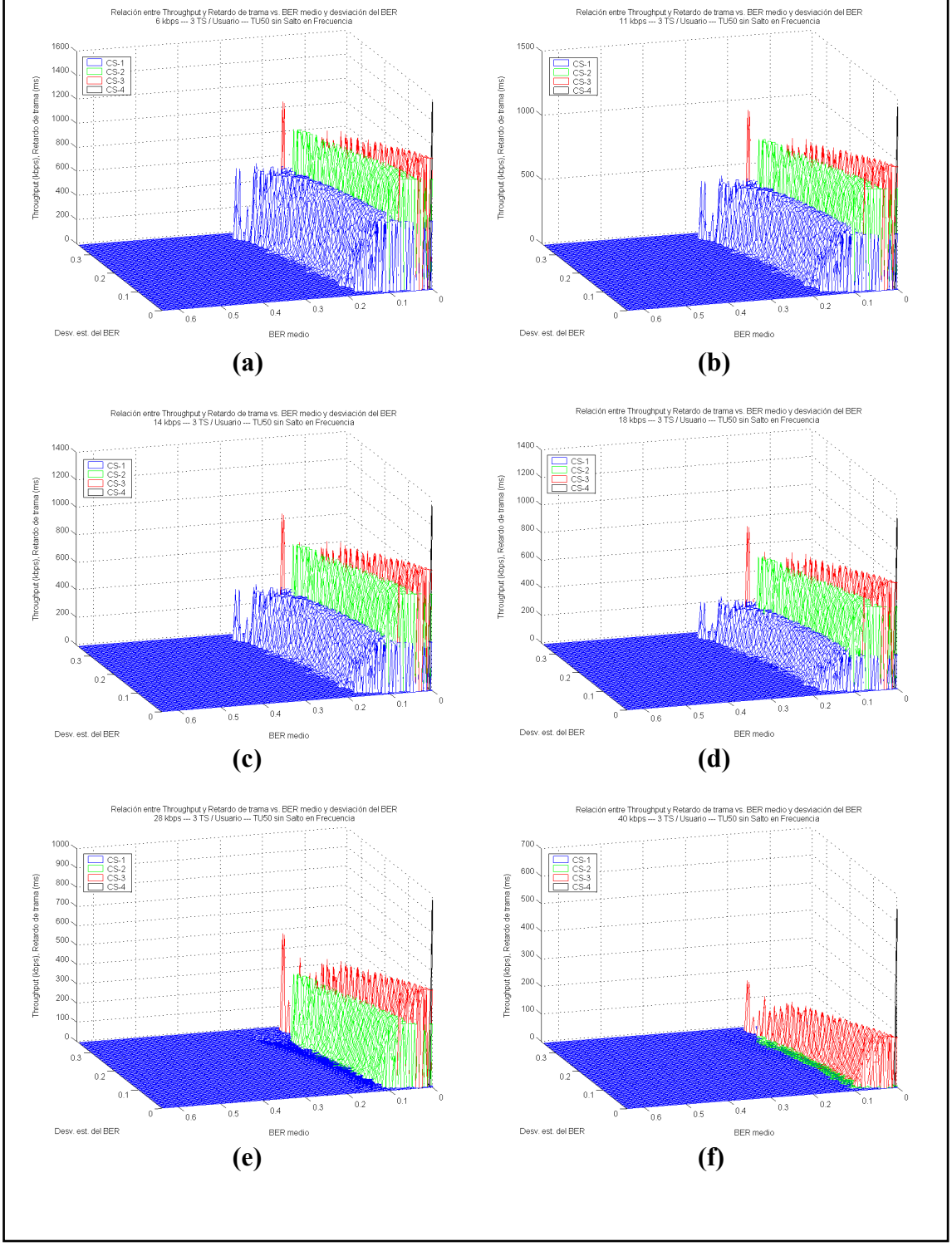


Figura 6.18.- Relación entre Throughput y Retardo de trama vs BER medio y desviación del BER, considerando 3 time-slots por usuario, para tramas que requieren diversas velocidades binarias: (a) 6 kbps, (b) 11 kbps, (c) 14 kbps, (d) 18 kbps, (e) 28 kbps, (f) 40 kbps

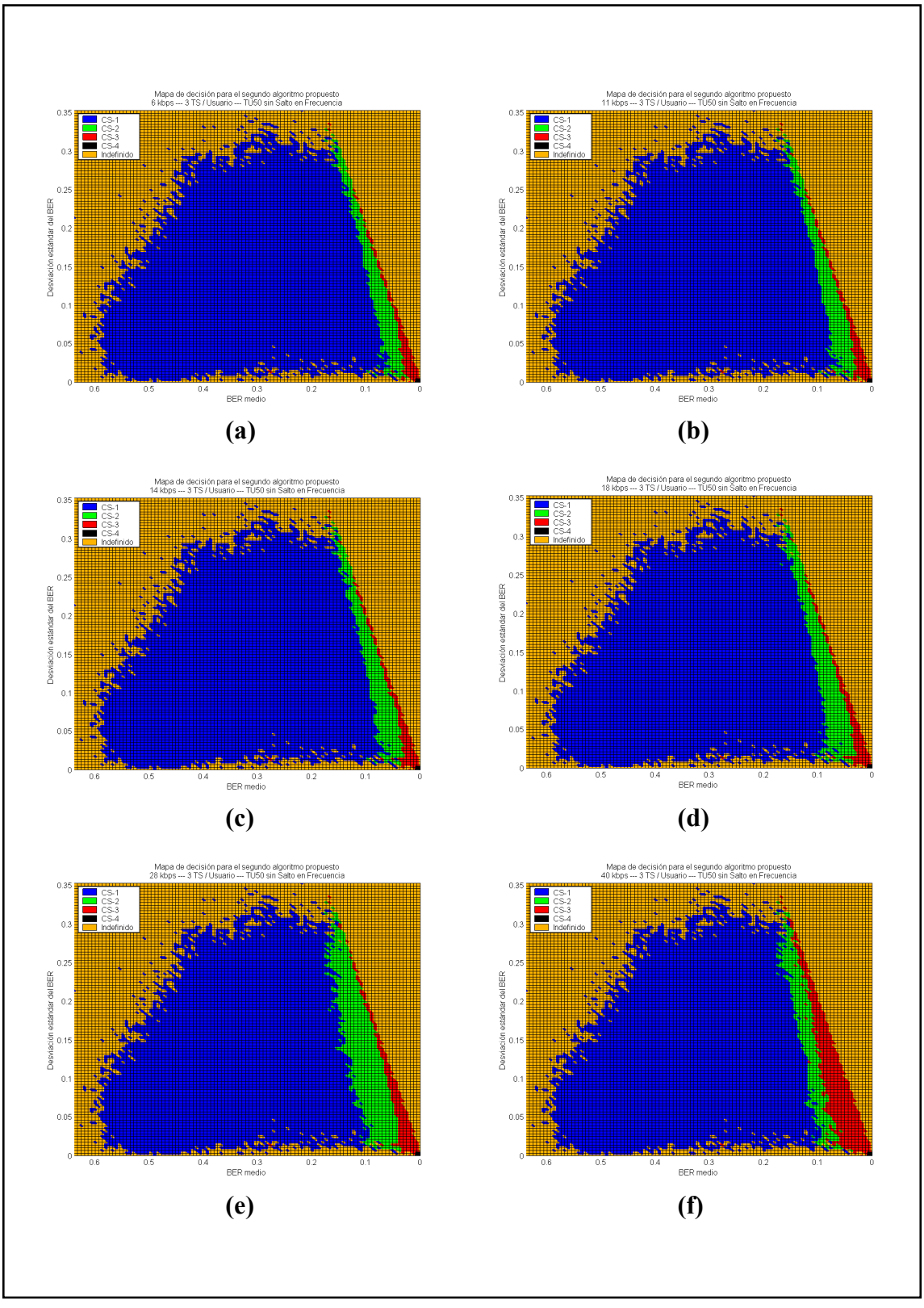


Figura 6.19.- Mapa de decisión del segundo algoritmo propuesto,
considerando 3 time-slots por usuario, para tramas que requieren diversas velocidades binarias:
(a) 6 kbps, (b) 11 kbps, (c) 14 kbps, (d) 18 kbps, (e) 28 kbps, (f) 40 kbps

Existe un caso bastante peculiar que puede darse en el ámbito práctico y que en la definición de este algoritmo no se ha considerado hasta ahora. Este caso corresponde a aquellas situaciones en las que para un esquema de codificación concreto, la trama de vídeo a transmitir posee un tamaño T y una duración D tales que el retardo de trama que se obtiene a partir de la expresión (6.5) es cero, lo cual ocurre cuando la velocidad binaria ofrecida por el esquema de codificación considerado coincide con la velocidad binaria que requiere la trama de vídeo para su transmisión sin retardo (T/D). Esta situación, que curiosamente no se ha llegado a observar ni una sola vez en las simulaciones realizadas, puede darse en el ámbito real y es necesario que se considere en la definición del algoritmo.

Cuando el retardo de trama empleando un esquema de codificación concreto es cero para una trama de vídeo, surge el problema de cuál de las dos expresiones en que se basa el algoritmo, la (6.8) y la (6.9), conviene aplicar. Aplicando la expresión (6.8) cuando el retardo de trama es nulo, el cociente obtenido tiende a infinito, lo cual significa que el esquema de codificación que proporciona un retardo de trama igual a cero para la trama a transmitir es el que conviene emplear, independientemente de la calidad del canal de transmisión y de que otros esquemas de codificación con mayor velocidad binaria sean capaces de ofrecer un retardo de trama negativo, lo cual no es correcto. Por otra parte, aplicando la expresión (6.9) cuando el retardo de trama es nulo, el valor obtenido es cero, lo que significa que el esquema de codificación que para una determinada trama de vídeo proporciona un retardo de trama nulo no se debe emplear nunca, independientemente de cuál sea la calidad del canal y de que otros esquemas de codificación con menor velocidad binaria, sí empleados en la transmisión, den lugar a retardos de trama positivos, lo cual tampoco es correcto. Por tanto, puede apreciarse que cuando el retardo de trama es cero no es correcto aplicar directamente ninguna de las dos expresiones en que se basa el presente algoritmo.

Cuando para un esquema de codificación se obtiene un retardo de trama nulo, entonces se promedia el retardo de trama obtenido para los esquemas de codificación con una velocidad binaria inmediatamente inferior (el retardo de trama será positivo) y superior (el retardo de trama será negativo). Si, por ejemplo, se obtiene un retardo de trama nulo para CS-2, entonces se promedia el retardo de trama de CS-1 (positivo) con el de CS-3 (negativo) y el resultado obtenido será el retardo de trama considerado para CS-2 a la hora de aplicar las expresiones (6.8) y (6.9). Cuando el retardo de trama para CS-3 sea cero, entonces se promedian los retardos de trama de CS-2 y CS-4. Puesto que el nuevo retardo de trama se obtiene promediando un valor positivo y otro negativo, podría pensarse que cabe la posibilidad de que el nuevo valor obtenido siguiera siendo cero. No obstante, a partir de la expresión (6.5) resulta sencillo comprobar que, si para un esquema de codificación el retardo de trama es cero, esta forma de obtener el nuevo retardo de trama da lugar a un valor no nulo. Cuando el retardo de trama sea nulo para CS-1 o para CS-4, entonces la solución adoptada consiste en realizar una pequeña aproximación en la que se considera que el retardo de trama obtenido es inversamente proporcional a la velocidad binaria del esquema de codificación (en la expresión 6.5 se ve que realmente no es así). Si empleando CS-2 (cuya velocidad binaria es conocida) se obtiene un determinado retardo de trama (también conocido), entonces, empleando CS-1 se tendrá un retardo de trama que, supuesto proporcional a la velocidad binaria, será distinto de cero, pues el retardo de trama de CS-2 también lo es. Para CS-4 el método es el mismo pero obteniendo el nuevo retardo de trama a partir del de CS-3. Esta pequeña aproximación permite solventar el problema de retardos de trama nulos.

6.3.- SUMARIO

El presente capítulo, dedicado al estudio y diseño de *Link Adaptation*, comienza con una descripción de dos algoritmos propuestos por otros autores. El primero de ellos es el algoritmo tradicional cuyo criterio se basa en maximizar el throughput a largo plazo (que ya fue utilizado en el capítulo tercero para la explicación del concepto de *Link Adaptation*). Se describe también un segundo algoritmo cuyo criterio se basa en intentar preservar la tasa de errores (expresada en términos de BLER) por debajo de un determinado valor, el cual se establece en función del BLER máximo tolerable para que la calidad de servicio experimentada por el usuario sea aceptable. A continuación se proponen dos nuevos algoritmos que constituyen la aportación más significativa del presente trabajo.

El primer algoritmo propuesto tiene, como principal objetivo, alcanzar un compromiso entre los dos algoritmos anteriores, intentando lograr unas prestaciones intermedias a las ofrecidas por ambos. Partiendo con este objetivo inicial, el diseño finaliza obteniendo un algoritmo en el que el grado de compromiso se puede variar, modificando un parámetro de diseño, en función de qué factor, throughput o tasa de errores, se considere de mayor relevancia respecto al otro.

El segundo de los algoritmos que aporta este trabajo se diseña teniendo en cuenta, no sólo el throughput, tal y como ha venido siendo habitual, sino también el retardo, parámetro trascendental para servicios de tiempo real. El principio de este planteamiento reside en el hecho de que el criterio de que maximiza el throughput no es, necesariamente, el que también minimiza el retardo. Lo que se pretende con este algoritmo es que la elección del esquema de codificación no esté determinada únicamente por el throughput sino que el retardo sea también un elemento influyente en la decisión final ya que, de hecho, constituye un factor igualmente importante para el tipo de servicio considerado. El resultado es un algoritmo en el cual el proceso de selección del esquema de codificación se adapta de forma dinámica a las características del flujo de datos transmitido.

Tras la presentación de estos dos nuevos algoritmos propuestos, en el capítulo siguiente se muestran y se analizan los resultados obtenidos en las simulaciones y se comparan con las prestaciones ofrecidas por los otros dos algoritmos, considerados a modo de referencia.

Capítulo 7

Análisis de Resultados

7.1.- INTRODUCCIÓN

Este capítulo estará dedicado a presentar y analizar los resultados obtenidos en las simulaciones de los cuatro algoritmos estudiados en este trabajo. Las prestaciones ofrecidas por los dos algoritmos propuestos por otros autores (el basado en throughput y el basado en BLER) serán tomadas como punto de referencia para la comparación de los resultados ofrecidos por los otros dos algoritmos que aquí se proponen. El análisis realizado será comparativo, observando el grado de mejora (o empeoramiento) de cada uno de los algoritmos analizados respecto a los demás y examinando si los algoritmos tratados (especialmente los que se proponen en el presente trabajo) permiten alcanzar el objetivo inicialmente establecido y para el que fueron diseñados.

7.2.- ALGORITMOS DE REFERENCIA

Antes de entrar en detalle en el análisis de los dos algoritmos propuestos, se compararán, en este apartado, los resultados de los dos algoritmos de referencia: el algoritmo basado en throughput y el algoritmo basado en BLER (para un BLER del 1%, 2% y 5%). En primer lugar, recordemos del capítulo anterior que el algoritmo basado en BLER requería una mejor calidad en el enlace radio que el algoritmo basado en throughput para utilizar los esquemas de codificación con mayor velocidad binaria. Se apuntó que este hecho debería fomentar, en principio, una utilización más frecuente, por parte del algoritmo basado en BLER, de aquellos esquemas de codificación con mayor protección ante errores, mientras que para el algoritmo basado en throughput cabía esperar lo contrario. La tabla 7.1, que recoge el porcentaje de utilización de cada esquema de codificación en la transmisión de vídeo H.263 para ambos algoritmos, muestra que estas suposiciones son correctas y se cumple el comportamiento esperado.

Algoritmo	CS-1	CS-2	CS-3	CS-4
BLER 1%	16.07 %	11.03 %	17.28 %	55.62 %
BLER 2%	13.67 %	9.85 %	19.94 %	56.54 %
BLER 5%	11.06 %	8.59 %	22.63 %	57.72 %
Throughput	6.63 %	7.14 %	26.41 %	59.82 %

Tabla 7.1.- Porcentaje de utilización de cada esquema de codificación en transmisión de vídeo

Cualquiera de las versiones del algoritmo basado en BLER utiliza con mayor frecuencia que el algoritmo basado en throughput aquellos esquemas de codificación con una mayor protección, es decir, CS-1 y CS-2 (especialmente el primero), mientras que en el algoritmo basado en throughput se emplean CS-3 y CS-4 (los esquemas de codificación con mayor velocidad binaria) de forma más habitual. Además, conforme el algoritmo basado en BLER exige una menor tasa de errores, se incrementa la calidad del canal requerida para emplear los esquemas con mayor tasa binaria, propiciando, tal y como se puede apreciar en la tabla 7.1, un uso más reducido de estos esquemas de codificación y, por tanto, un uso más frecuente de los codificadores más robustos.

El comportamiento que se observa en la tabla 7.1 permite explicar los resultados recogidos en la tabla 7.2, donde se muestra, para varios parámetros relevantes, el valor medio ofrecido por cada algoritmo así como el valor en el 95% y 99% de los casos, los cuales han sido obtenidos a partir de las CDFs mostradas en las figuras 7.1 a la 7.4[50].

Parámetro	Algoritmo	Medio	95%	99%
Throughput (kbps)	BLER 1%	15.85	9.50	7.05
	BLER 2%	16.04	9.68	7.07
	BLER 5%	16.24	9.92	7.22
	Throughput	16.56	10.46	7.53
BLER (%)	BLER 1%	9.93	21.91	36.87
	BLER 2%	10.25	22.77	37.52
	BLER 5%	10.60	23.73	38.43
	Throughput	11.18	25.90	39.73
BER (%)	BLER 1%	2.08	7.57	12.76
	BLER 2%	2.05	7.56	12.75
	BLER 5%	1.99	7.50	12.73
	Throughput	1.86	7.17	12.42
Retardo normalizado (ms/kbit)	BLER 1%	63.70	119.63	138.90
	BLER 2%	62.34	116.36	137.37
	BLER 5%	60.82	110.21	134.76
	Throughput	58.32	99.08	129.44

Tabla 7.2.- Throughput, BLER, BER y retardo normalizado para los algoritmos de referencia en vídeo

Se observa que el throughput ofrecido por el algoritmo basado en throughput es superior al ofrecido por cualquiera de los algoritmos basados en BLER, tal y como cabía esperar. Ello se debe a dos causas: una mayor utilización de los codificadores con mayor tasa binaria (que permite alcanzar una mayor velocidad binaria media y, por tanto, transmitir una mayor cantidad de bits por unidad de tiempo) y una disminución del BER o proporción de bits transmitidos con error. No sólo se transmiten más bits por unidad de tiempo empleando el algoritmo basado en throughput sino que, además, la proporción de bits con error es menor, dando lugar a que el throughput experimente un aumento respecto a los algoritmos basados en BLER, tanto en valor medio como en el 95% y 99% de los casos (tabla 7.2). Además, a medida que el algoritmo basado en BLER se vuelve más conservador, el decrecimiento del porcentaje de utilización de los codificadores con mayor velocidad binaria visto en la tabla 7.1 se ve acompañado por un deterioro del BER, lo cual explica la disminución del throughput que se observa en la tabla 7.2 así como en la figura 7.1. El algoritmo basado en throughput ofrece una mayor cantidad de bits correctos por unidad de tiempo pagando el precio de un BLER más alto. En efecto, tal y como se aprecia en la tabla 7.2 y en la figura 7.2, el BLER del algoritmo basado en throughput es el más elevado de todos, mientras que el BLER alcanzado por los algoritmos de BLER se reduce a medida que el algoritmo busca una menor tasa de bloques RLC erróneos[51].

[50] Para el throughput, estos valores corresponden al 1% y 5% de la CDF (throughput mínimo). Para el resto de parámetros de la tabla, al 95% y 99% de la CDF (BLER, BER o retardo normalizado máximos).

[51] Puede apreciarse que ninguno de los algoritmos basados en BLER permite alcanzar el valor de BLER para el que fueron diseñados. Esta cuestión será retomada y tratada, dada su relevancia, en un apartado posterior.

La reducción del BLER conforme el algoritmo basado en BLER se vuelve más exigente se debe a que el porcentaje de utilización de los esquemas de codificación con mayor tasa binaria (CS-3 y CS-4) decrece y se incrementa el grado de utilización de los codificadores más robustos (CS-1 y CS-2). Consecuentemente, decrece la proporción de bloques RLC recibidos con error, es decir, el BLER, tal y como se aprecia. La afirmación anterior parece estar en contradicción con el incremento del BER observado en la tabla 7.2 y en la figura 7.3: no resulta coherente, en principio, que utilizando de manera más frecuente los esquemas de codificación más robustos se incremente la proporción de bits recibidos con error. Una disminución del BLER sí resulta normal pero, aun así, tampoco parece lógico que éste disminuya cuando aumenta el BER.

En cuanto a la primera observación, hay que tener en cuenta que al utilizar con mayor frecuencia los esquemas de codificación más robustos, se transmite una menor cantidad de bits de datos por unidad de tiempo o, lo que es lo mismo, la transmisión de una misma cantidad de información requiere más tiempo. Si se utiliza con mayor frecuencia estos esquemas, entonces las transmisiones tendrán una duración media mayor, lo cual supondrá un incremento del nivel de interferencia global en el sistema y un deterioro en la calidad del canal de transmisión, existiendo la posibilidad de que este grado de deterioro experimentado sea mayor que el grado de robustez adquirido al emplear, de forma más habitual, los codificadores con menor tasa binaria. Es cierto que para una misma calidad del canal de transmisión, un uso más habitual de los codificadores más robustos reditúa en una reducción de la tasa de bits erróneos, pero si la calidad empeora en mayor medida de lo que aumenta el grado de robustez, entonces no se produce una reducción del BER sino, por el contrario, un incremento, lo cual explica la primera observación realizada.

Por lo que respecta a la segunda observación, es cierto que tanto el BER como el BLER son dos tasas de errores y que lo más lógico sería que el comportamiento de ambas fuera paralelo. No obstante, esto no ha de ser necesariamente así ya que el BLER se mide sobre bloques RLC, es decir, sobre grupos de bits obtenidos una vez realizado el proceso de decodificación de canal, en el que, como se vio en el capítulo segundo, se aplican técnicas de corrección de errores. Si los decodificadores son capaces de corregir los bits erróneos, el BLER o proporción de bloques RLC con error tras la decodificación de canal no tiene por qué comportarse de forma paralela al BER. De hecho, el uso más frecuente de codificadores con mayor redundancia, si bien puede ocasionar un mayor BER en el enlace radio por los motivos anteriormente comentados, también da lugar a que la proporción de bits erróneos a la salida del decodificador de canal (que no es el BER mostrado en la tabla 7.2) sea menor, lo cual explica que el BLER decrezca.

Tras analizar de forma cualitativa el comportamiento del throughput, BLER y BER para los dos algoritmos de referencia, en un análisis más cuantitativo se puede observar a partir de la tabla 7.2 que, en general, la mejora del BLER alcanzada por los algoritmos basados en BLER se obtiene a cambio de empeorar, en mayor proporción, el throughput global del sistema. Si comparamos, por ejemplo, el algoritmo basado en un BLER del 1% y el basado en throughput en valor medio y en el 95% de los casos, entonces el BLER mejora, respectivamente, un 1.25% y un 4% mientras que, de forma paralela, el throughput empeora en un 3.3% y 4.5%. Este es un dato a tener en cuenta a la hora de diseñar y aplicar estos algoritmos basados en mantener el BLER, en especial si la aplicación concreta, como es el caso, requiere una tasa de error bastante reducida.

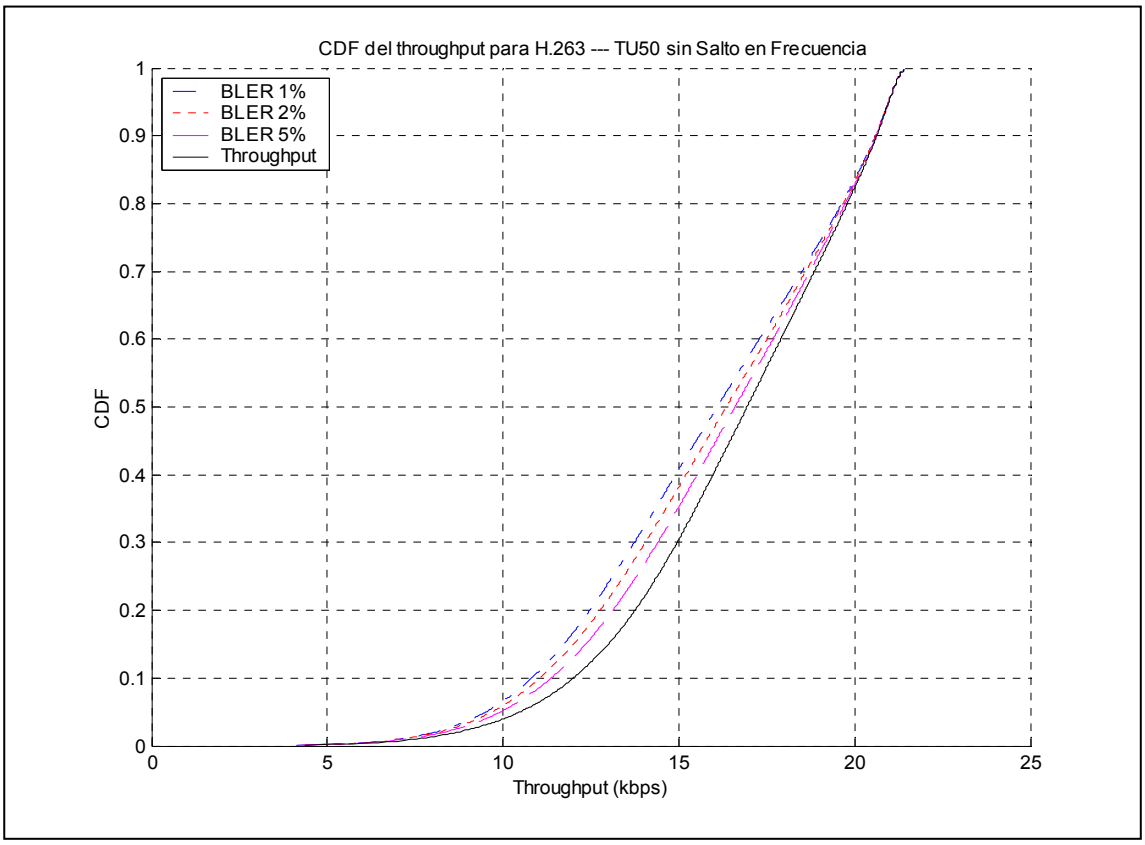


Figura 7.1.- CDF del throughput para los algoritmos de referencia en transmisión de vídeo

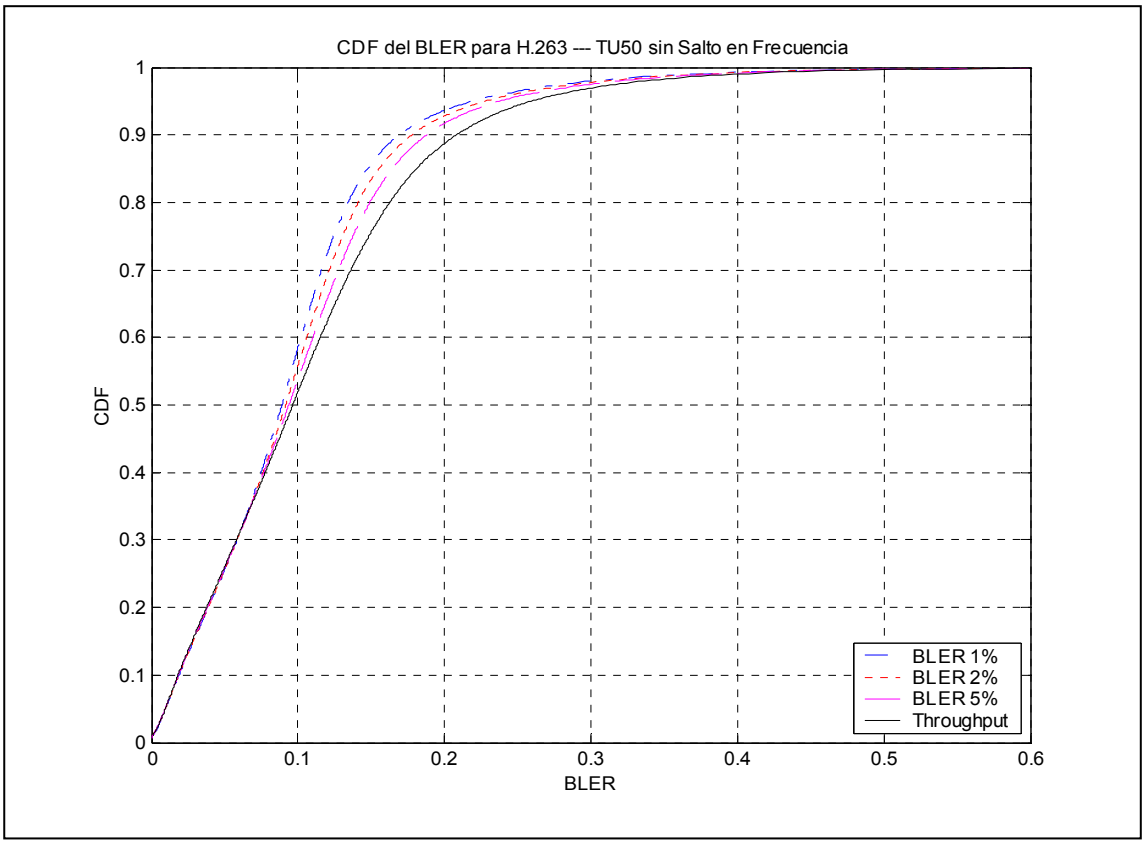


Figura 7.2.- CDF del BLER para los algoritmos de referencia en transmisión de vídeo

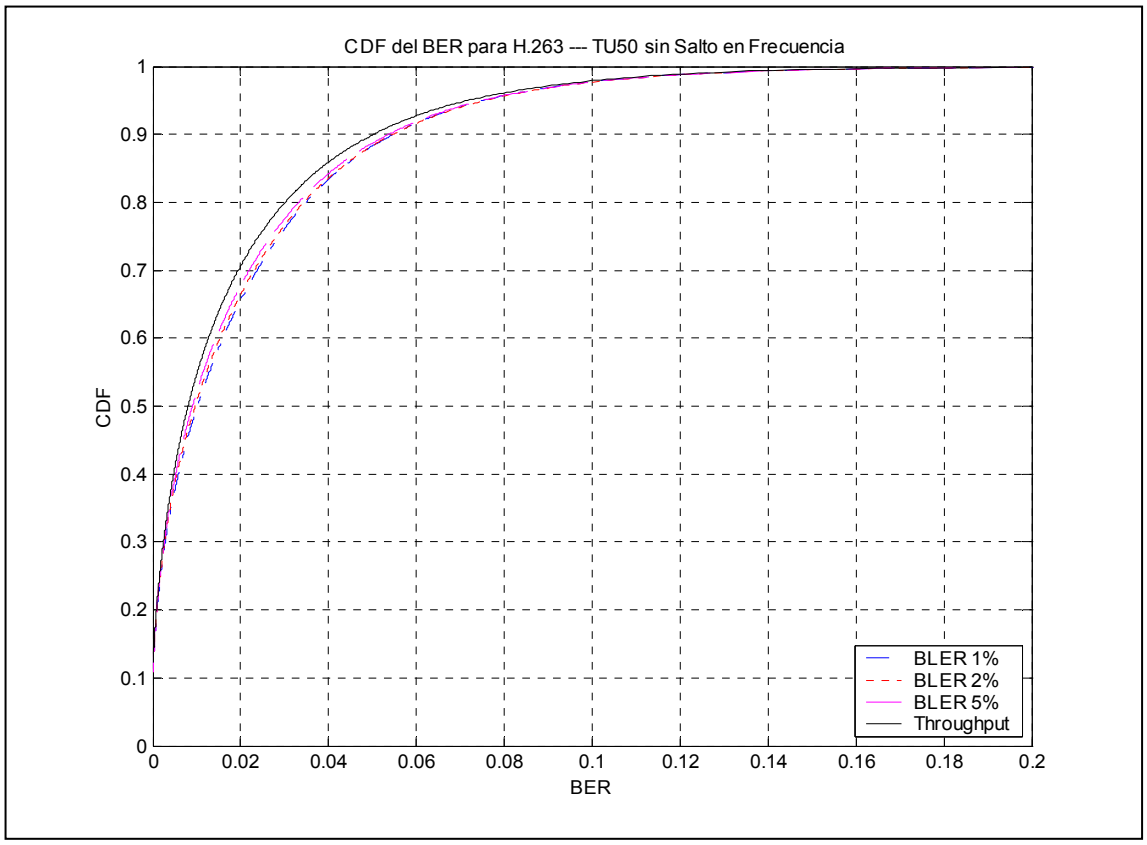


Figura 7.3.- CDF del BER para los algoritmos de referencia en transmisión de vídeo

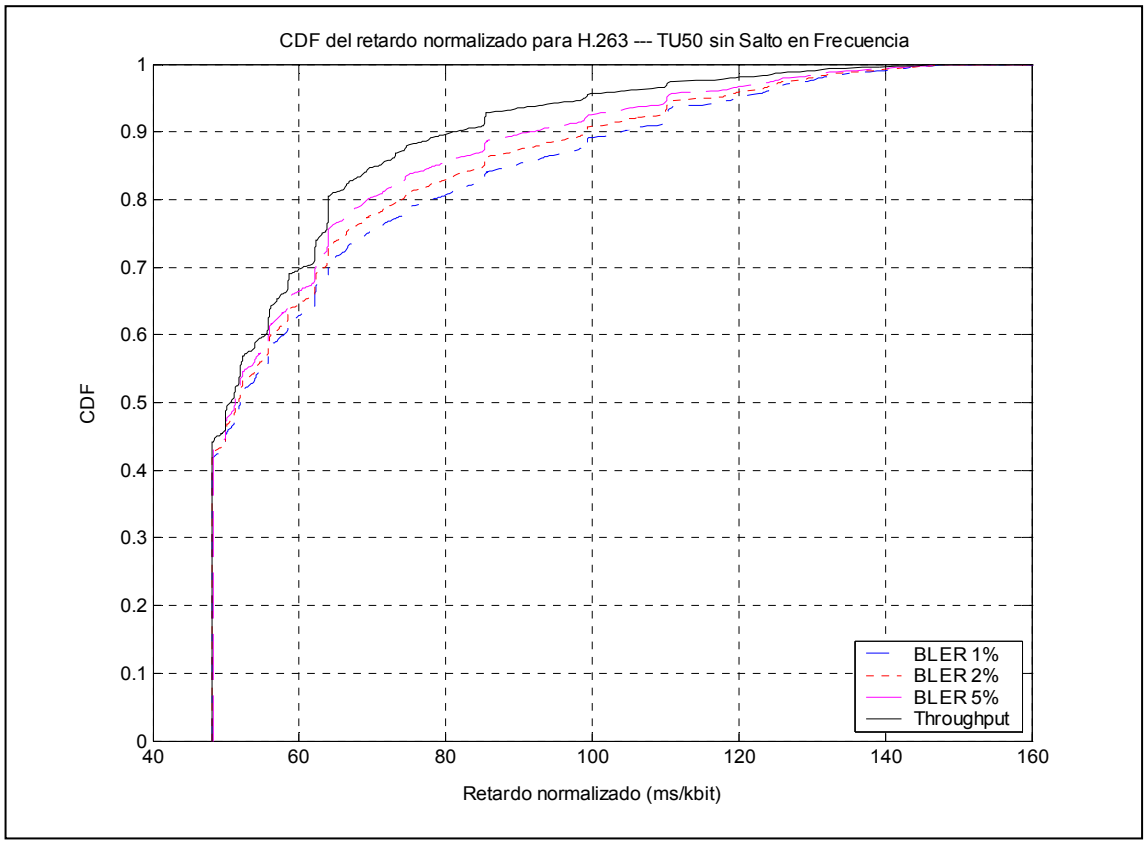


Figura 7.4.- CDF del retardo normalizado para los algoritmos de referencia en transmisión de vídeo

Un parámetro de vital importancia para el tipo de servicio considerado en este trabajo es el retardo, el cual se señaló que era el tiempo requerido para transmitir un determinado bloque de información (trama de vídeo en el caso particular que aquí se considera). Sin embargo, el retardo en sí no permite evaluar de forma objetiva la forma en que se comporta un determinado algoritmo, puesto que será mayor para aquellas tramas de vídeo con un mayor tamaño. Por ello, un parámetro más conveniente para el análisis es el retardo normalizado, el cual se define como el cociente entre el tiempo requerido para transmitir un determinado bloque de información y el tamaño de dicho bloque. Con este nuevo parámetro es posible comparar, de forma equitativa, el retardo para tramas con un tamaño diferente. El valor de este parámetro para los algoritmos de referencia se muestra en la tabla 7.2, así como en la figura 7.4[52].

El algoritmo basado en throughput, caracterizado por utilizar con una mayor frecuencia los codificadores de mayor velocidad binaria, es el que alcanza una mayor tasa binaria media y un menor retardo normalizado (llegándose a observar una mejora relativa del 17% en el 95% de los casos respecto al algoritmo basado en un BLER del 1%). En cambio, para el algoritmo basado en BLER, que utiliza más habitualmente los codificadores más robustos, la tasa binaria media es más reducida y, en consecuencia, el retardo normalizado, mayor. A medida que el algoritmo se vuelve más conservador, se emplean con mayor frecuencia todavía los codificadores de menor velocidad binaria y el retardo normalizado aumenta. Por tanto, resulta lógico, como muestra la tabla 7.3, que el porcentaje de tramas de vídeo transmitidas con retardo[53] sea mayor para el algoritmo basado en BLER y que este valor se incremente a medida que el algoritmo se vuelve más exigente, hasta suponer un empeoramiento relativo del 22% respecto al algoritmo basado en throughput.

Algoritmo	**Porcentaje de tramas de vídeo transmitidas**	
	Sin retardo	**Con retardo**
BLER 1%	62.33 %	37.67 %
BLER 2%	64.03 %	35.97 %
BLER 5%	65.97 %	34.03 %
Throughput	69.16 %	30.84 %

Tabla 7.3.- Porcentaje de tramas de vídeo transmitidas con y sin retardo para los algoritmos de referencia

[52] Observar que el retardo normalizado mínimo para una trama de vídeo nunca puede ser inferior a la inversa de la máxima velocidad binaria posible (21.4 kbps). En cambio, el retardo normalizado máximo sí puede ser superior a la inversa de la mínima velocidad binaria de un esquema de codificación (9.05 kbps), cuando en el último bloque RLC de la transmisión de una trama de vídeo no hay información suficiente para llenar todo el campo de datos, en cuyo caso se insertan bits de relleno, que no contribuyen a la velocidad binaria final. Teniendo en cuenta que la duración mínima de una trama de vídeo es de 40 ms (dos bloques RLC), el peor de los casos correspondería a transmistir, con CS-1, 182 bits de datos (ver tabla 2.1), en cuyo caso la velocidad binaria real sería 182 bits / 40 ms = 4.55 kbps y, el retardo normalizado máximo posible, de unos 220 ms / kbit.

[53] En lo sucesivo, el término *trama transmitida con retardo* se utilizará para referirse a aquellas tramas de vídeo cuyo plazo de tiempo disponible para su transmisión expira sin haber sido transmitidas por completo. El término *trama transmitida sin retardo* se empleará para referirse a aquellas que, por el contrario, sí son transmitidas íntegramente dentro del plazo de tiempo para ello disponible.

Sin embargo, es necesario tener en cuenta que el hecho de que el porcentaje de tramas sin retardo sea superior no es suficiente para poder afirmar que el rendimiento obtenido sea mejor ya que, como se ha visto, este incremento se debe a una reducción del retardo normalizado, consecuencia de un empleo más habitual de los codificadores con mayor velocidad binaria (y menor protección ante errores). Por ello, resulta más representativo comparar, no sólo el porcentaje de tramas sin retardo, sino también el porcentaje de tramas de vídeo que, además de haber llegado sin retardo al receptor, satisfacen un requisito concreto de calidad (ver apartado 4.3). En la tabla 7.4 se muestra esta comparativa.

Requisito	BLER 1%	BLER 2%	BLER 5%	Throughput
BLER ≤ 1%	41.66 %	42.53 %	43.41 %	44.73 %
BLER ≤ 2%	41.68 %	42.55 %	43.43 %	44.75 %
BLER ≤ 5%	41.91 %	42.78 %	43.66 %	44.98 %
BER ≤ 10^{-3}	42.69 %	43.51 %	44.42 %	45.93 %
BER ≤ 10^{-4}	37.59 %	38.30 %	39.10 %	40.40 %
BER ≤ 10^{-5}	36.62 %	37.31 %	38.08 %	39.33 %
BER ≤ 10^{-6}	36.52 %	37.21 %	37.99 %	39.23 %

Tabla 7.4.- Porcentaje de tramas de vídeo transmitidas sin retardo
y que cumplen un determinado requisito de calidad

Se puede apreciar que el algoritmo basado en throughput, que es el que obtiene el mayor porcentaje de tramas transmitidas sin retardo, es también el que alcanza una mayor proporción de tramas sin retardo con una determinada calidad (sobre un 3% más respecto al algoritmo basado en un BLER del 1% para los diversos objetivos de calidad señalados en la tabla). Este hecho no resulta extraño si se tiene en cuenta que, puesto que es el algoritmo que transmite una mayor proporción de tramas sin retardo, también debe ser el que transmita un mayor porcentaje de tramas sin retardo de todos los *tipos*, es decir, con un BLER inferior al 1%, 2% y 5% o un BER inferior a 10^{-3}, 10^{-4}, 10^{-5} y 10^{-6}. No obstante, si se consideran únicamente las tramas transmitidas sin retardo y, dentro de este conjunto, se examina el porcentaje que cumple estos mismos requisitos de calidad, los resultados ahora son bien distintos, como puede apreciarse en la tabla 7.5. En este caso, el algoritmo basado en BLER, si bien transmite un porcentaje de tramas sin retardo más bajo que el algoritmo basado en throughput, la proporción de este porcentaje que satisface un determinado requisito de calidad es mayor que en el caso del algoritmo basado en throughput y, de hecho, se incrementa a medida que el algoritmo busca un BLER más reducido, lo cual demuestra que el retardo y la tasa de errores son dos factores comprometidos a nivel de sistema y que mejorar uno implica deteriorar el otro.

Requisito	BLER 1%	BLER 2%	BLER 5%	Throughput
BLER ≤ 1%	66.85 %	66.42 %	65.81 %	64.67 %
BLER ≤ 2%	66.87 %	66.44 %	65.84 %	64.70 %
BLER ≤ 5%	67.24 %	66.80 %	66.19 %	65.04 %
BER ≤ 10^{-3}	68.50 %	67.95 %	67.34 %	66.41 %
BER ≤ 10^{-4}	60.31 %	59.82 %	59.27 %	58.41 %
BER ≤ 10^{-5}	58.75 %	58.27 %	57.73 %	56.87 %
BER ≤ 10^{-6}	58.60 %	58.12 %	57.58 %	56.73 %

Tabla 7.5.- De las tramas de vídeo transmitidas sin retardo, porcentaje que cumple un determinado requisito de calidad

Comparando con mayor detenimiento los valores que se recogen en las tablas 7.3 y 7.5 se aprecia que, si bien de las tramas transmitidas sin retardo, la proporción de tramas con una determinada calidad mejora en el algoritmo basado en BLER respecto al algoritmo basado en throughput, el porcentaje de mejora no es tan notable como lo es el de la cantidad de tramas que se transmiten sin retardo empleando el algoritmo basado en throughput. Comparando los valores de la tabla 7.5 correspondientes al algoritmo basado en throughput y al basado en un BLER del 5%, éste último mejora al primero en este parámetro aproximadamente en un 1% para los diferentes objetivos de BER y BLER mostrados. La mejora de los algoritmos basados en un BLER del 2% y 1% ronda, respectivamente, el 1.5% y 2%. Sin embargo, a partir de la tabla 7.3, se puede afirmar que el algoritmo basado en throughput consigue transmitir un 3%, un 5% e incluso hasta un 7% más de tramas sin retardo que en el caso de los algoritmos basados en un BLER del 5%, 2% y 1%, respectivamente.

Por lo tanto, lo que se concluye de todo este análisis es que se observa un compromiso entre cantidad de tramas transmitidas sin retardo y calidad de las tramas transmitidas sin retardo. El algoritmo basado en throughput consigue transmitir una mayor proporción de tramas sin retardo, si bien las tramas transmitidas sin retardo empleando el algoritmo basado en BLER, aun constituyendo un porcentaje menor, muestran una mejor calidad en un mayor porcentaje de veces, el cual se incrementa a medida que el algoritmo persigue un BLER más reducido. No obstante, el porcentaje de mejora en lo que se refiere a calidad en las tramas sin retardo obtenido al emplear el algoritmo basado en BLER no es tan significativo como el incremento en el porcentaje de tramas que se transmiten sin retardo con el algoritmo basado en throughput.

El análisis anterior muestra que el algoritmo basado en BLER es capaz de lograr una mejor calidad que el algoritmo basado en throughput para las tramas transmitidas sin retardo. No obstante, para el análisis de la calidad final de vídeo experimentada por el usuario, es necesario considerar también las tramas que se transmiten con retardo.

Algoritmo	Global		
	Medio	95 %	99 %
BLER 1%	36.90 %	67.94 %	78.70 %
BLER 2%	36.43 %	66.75 %	78.55 %
BLER 5%	35.89 %	66.15 %	78.23 %
Throughput	34.88 %	64.15 %	76.17 %

Tabla 7.6.- Porcentaje de información perdida en tramas de vídeo transmitidas con retardo

En efecto, lo que realmente afecta a la calidad final de vídeo (y de otro tipo de servicios de tiempo real), es el hecho de tener que descartar, es decir, dar por perdida en el receptor, parte de la información que constituye el flujo de datos original. El tener que descartar una mayor parte de la información ocasionará un mayor deterioro en la calidad final de servicio que experimenta el usuario, sin importar el hecho de que esta pérdida de información se deba a que se recibe con error o a que, simplemente, no se recibe. Por lo tanto, si bien es cierto que una reducción del retardo y, por tanto, de la información no recibida a tiempo, se consigue a cambio de empeorar la calidad de la que sí se recibe a tiempo, también es cierto que una reducción de la tasa de errores y, por tanto, de la cantidad de información recibida con error, se consigue pagando el precio de una mayor cantidad de información no recibida, tal y como se desprende de los análisis anteriormente realizados. Lo que habría que examinar en esta circunstancia es cuál de las dos estrategias anteriores da lugar a una mayor cantidad de información correcta finalmente recibida en el receptor y, por tanto, cuál es la que contribuye a una mejor calidad de servicio final para el usuario.

La tabla 7.6 y la figura 7.5 muestran que la cantidad de información descartada en tramas con retardo es mayor con el algoritmo basado en BLER que con el algoritmo basado en throughput y que se incrementa conforme el BLER perseguido se reduce, hecho que se debe a una mayor utilización, por parte del algoritmo basado en BLER, de aquellos codificadores más robustos, lo cual ofrece una menor velocidad binaria y, por tanto, una mayor cantidad de tramas transmitidas con retardo así como de información perdida en tramas con retardo. Es decir, no sólo se reciben más tramas con retardo empleando el algoritmo basado en BLER sino que, además, la cantidad de información que en ellas se ha de dar por perdida es también mayor.

A partir de este dato y de la comparación de las tablas 7.3 y 7.5 anteriormente realizada, lo que se plantea es que la mayor proporción de tramas con una calidad correcta dentro del grupo de tramas transmitidas sin retardo que logra el algoritmo basado en BLER respecto al basado en throughput, probablemente no compense el (mayor) porcentaje de incremento de tramas con retardo así como de información perdida en estas tramas, en cuyo caso la cantidad de información que finalmente se ha de descartar en el receptor sería mayor con este algoritmo que con el algoritmo basado en throughput. Esta observación hace cuestionable, al menos bajo las condiciones de operación estudiadas, el hecho de que los beneficios que supuestamente deberían obtenerse con este algoritmo, respecto al algoritmo basado en throughput, realmente se lleguen a conseguir para servicios en tiempo real.

En un servicio de tiempo no real el algoritmo sí puede resultar efectivo, pues el incremento del retardo no implica necesariamente una pérdida de información en este tipo de servicios, no sensibles al retardo. El resultado es que el receptor recibe toda la información original y, además, con una menor tasa de errores que la que ofrecería el algoritmo basado en throughput. No existe motivo aparente que impida considerar la aplicación de este algoritmo a servicios de tipo *streaming* que, aun siendo sensibles al retardo, lo son en menor medida que los de tiempo real, ya que la información se almacena en un búfer hasta que puede ser transmitida, en lugar de desechar la que se estaba transmitiendo anteriormente. Un incremento (moderado) del retardo puede ser aceptable en servicios *streaming* sin que ello suponga necesariamente que se haya de dar por perdida una mayor cantidad de información, lo que permitiría aplicar este algoritmo para lograr una reducción de la tasa de errores y una mejora de la calidad de servicio [104]. Sin embargo, en un servicio de tiempo real, un incremento del retardo sí supone la pérdida de una mayor cantidad de información, lo cual resulta equivalente a recibirla con error y darla por perdida en el receptor. Por lo tanto, lo que finalmente se concluye es que cabe la posibilidad de que la aplicación de este algoritmo a servicios en tiempo real no mejore la calidad final experimentada por el usuario respecto al algoritmo basado en throughput (que es lo que en definitiva persigue este algoritmo) sino que, por el contrario, sufra un deterioro a causa del excesivo retardo en el que se incurre, el cual no se ve compensado por la mejora en la calidad de las tramas sin retardo, con lo que la cantidad final de información correcta en el receptor resulta ser menor y, por tanto, la calidad de servicio final, peor[54].

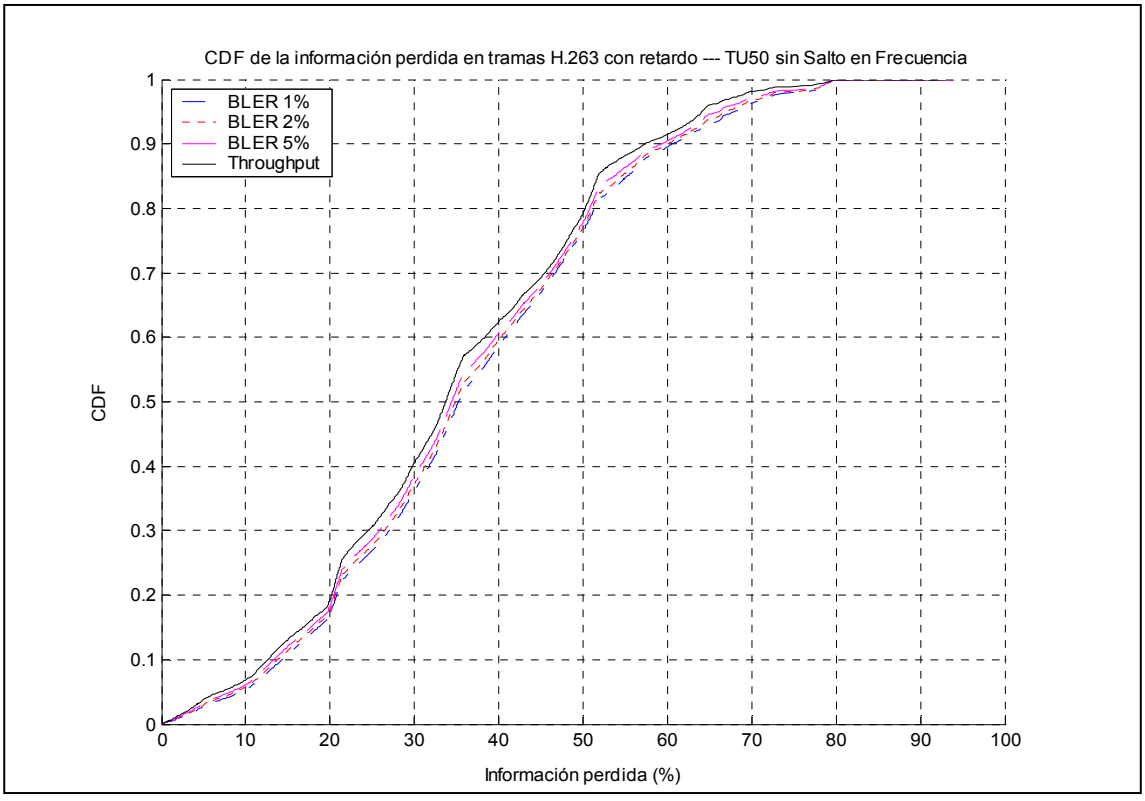


Figura 7.5.- CDF del porcentaje de información perdida en tramas de vídeo transmitidas con retardo

[54] El resultado final y concluyente dependerá del proceso de decodificación de vídeo en sí y de la eficacia de sus técnicas de corrección de errores, así como del estudio bajo otras condiciones de operación como, por ejemplo, en modo multislot, razón por la cual este hecho sólo puede señalarse aquí como una posibilidad.

7.2.1.- CAUSA DE LA INEFICACIA DEL ALGORITMO BASADO EN BLER

En el apartado anterior se pudo observar que el algoritmo basado en BLER, si bien permite reducir la tasa de bloques RLC recibidos con error respecto al algoritmo basado en throughput y muestra una clara tendencia de reducción a medida que busca un BLER más reducido, es incapaz de mantener su valor global por debajo del margen que se había propuesto en su diseño, según se observa en los resultados recogidos en la tabla 7.7.

Algoritmo	BLER medio
BLER 1%	9.93
BLER 2%	10.25
BLER 5%	10.60

Tabla 7.7.- Comparación del BLER obtenido para los diversos algoritmos basados en BLER

Este hecho, bastante sorprendente, se observa que también tiene lugar en los resultados obtenidos en [104], donde el diseño del algoritmo se realiza para un BLER del 15% y, tras las simulaciones realizadas, los resultados obtenidos muestran que el BLER medio final es del 23%. Si bien es cierto que existen diferencias entre el estudio aquí realizado y el de [104] (donde se utiliza control de potencia e interfaces a nivel de bloque RLC en el simulador, menos precisas que las aquí empleadas al no permitir la consideración de los efectos del desvanecimiento rápido en la simulación a nivel de sistema), puede apreciarse que el comportamiento cualitativo del algoritmo coincide en ambos casos, es decir, se produce una reducción del BLER respecto al algoritmo basado en throughput pero sin lograr el margen considerado en el diseño. Este hecho implica que para la obtención de la tasa de error buscada sería necesario sobredimensionar el diseño del algoritmo, es decir, considerar un BLER muy por encima del deseado. Aun así, este planteamiento no parece garantizar la consecución del objetivo, afirmación realizada en virtud a la escasa diferencia entre los algoritmos de BLER 1% y BLER 5%[55]. Las únicas garantías que ofrece este algoritmo son una reducción del BLER respecto al algoritmo basado en throughput (a penas apreciable en este caso) y su disminución a medida que el diseño se realiza para un margen de BLER más reducido.

Una explicación coherente para un BLER tan elevado es que el aumento en el porcentaje de utilización de los esquemas de codificación más robustos no resulta lo suficientemente elevado como para que la tasa de errores experimente el descenso que se pretende. Otra forma de verlo es que resulta insuficiente la reducción en el grado de utilización de los codificadores menos robustos. En la tabla 7.1 se aprecia que CS-4 es, con diferencia, el esquema de codificación más utilizado, incluso cuando el BLER que se persigue es considerablemente reducido. Una utilización tan elevada de CS-4 resulta comprensible cuando lo que se pretende es maximizar el throughput, pero no cuando el objetivo es una reducción del BLER.

[55] Reducir el valor de BLER tomado en el diseño del algoritmo en un 4%, supone una mejora de sólo un 0.7% en el BLER medio global del sistema, siendo esta mejora ligeramente superior, del 1.8% y del 1.6%, para el valor máximo que puede garantizarse, respectivamente, al 95% y 99% de los usuarios (tabla 7.2).

Según la afirmación anterior, la consecución del BLER buscado haría necesaria una reducción forzada de la elevada utilización del codificador CS-4, pues el algoritmo, tal y como ha sido definido, permite que CS-4 siga siendo el esquema de codificación más utilizado incluso cuando el BLER marcado como objetivo es tan bajo como un 1%. A continuación se muestran los resultados que corroboran esta afirmación, obtenidos al simular nuevamente los algoritmos de referencia modificando un parámetro relativo a la configuración de *Link Adaptation*.

El parámetro modificado es el esquema de codificación utilizado al inicio de cada nueva transmisión. Según se comentó en el capítulo tercero (apartado 3.5.4), al comienzo de cada nueva transmisión de datos no se dispone de información alguna sobre la calidad del enlace radio en función de la cual decidir el codificador a emplear. Una posible solución a este problema consiste en emplear siempre el mismo esquema de codificación, el cual se fija en tiempo de diseño en función de la estrategia seguida por el algoritmo. Otra posible solución consiste en emplear el esquema de codificación con el que se concluyó la transmisión anterior. Los resultados mostrados hasta ahora corresponden a este último caso. También se realizaron las simulaciones haciendo uso de la otra alternativa, en la que el codificador CS-2 se empleó siempre al inicio de cada nueva transmisión [42].

La tabla 7.8, que recoge el porcentaje de utilización de cada codificador para los dos algoritmos de referencia en ambos casos, muestra que la modificación realizada sobre este parámetro fuerza una reducción considerable en el porcentaje de uso del codificador CS-4, más significativa incluso que la lograda por el algoritmo basado en un BLER del 1% respecto al algoritmo basado en throughput (ver tabla 7.1).

Algoritmo	**CS-1**		**CS-2**		**CS-3**		**CS-4**	
	Último CS	CS-2	Último CS	CS-2	Último CS	CS-2	Último CS	CS-2
BLER 5%	11.06 %	11.04 %	8.59 %	25.85 %	22.63 %	19.73 %	57.72 %	43.38 %
Throughput	6.63 %	6.64 %	7.14 %	25.58 %	26.41 %	22.96 %	59.82 %	44.82 %

Tabla 7.8.- Porcentaje de utilización de cada esquema de codificación en transmisión de vídeo empleando el último codificador o CS-2 al inicio de cada nueva transmisión

Puede apreciarse que el porcentaje de utilización de CS-1 no varía de un caso al otro. Sin embargo, forzando la utilización de CS-2 al inicio de cada nueva transmisión de datos se consigue incrementar el porcentaje de utilización de este codificador en un 17% (aproximadamente) para ambos algoritmos a cambio de reducir el uso de CS-3 en un 3% y el de CS-4 en un 14%. Vemos pues, que empleando CS-2 al comienzo de cada nueva transmisión se consigue disminuir la utilización de los esquemas de codificación menos robustos, en especial, la de CS-4. Analizando la reducción en el BLER que este cambio supone y comparándola con la obtenida por el algoritmo de BLER respecto al de throughput, se refuerza la afirmación anterior de que la consecución del BLER que se pretende requiere forzar una reducción drástica del porcentaje de uso de CS-4.

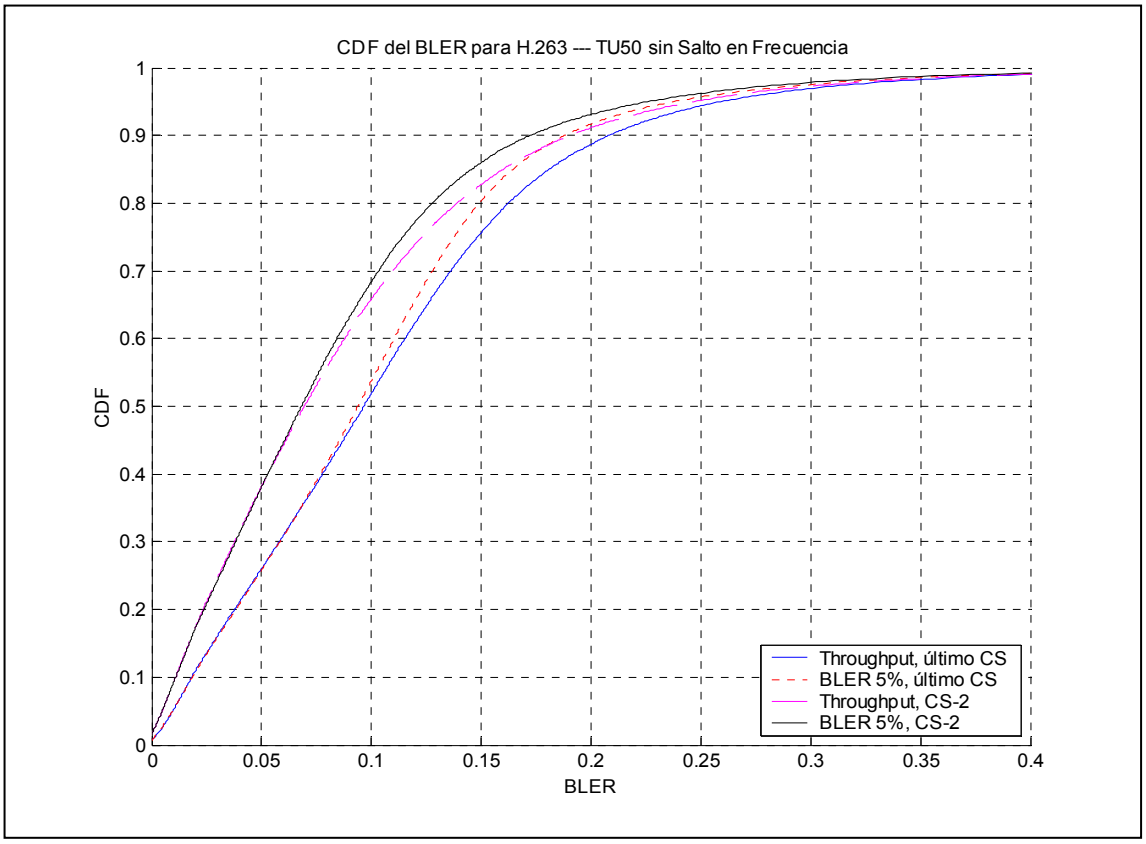


Figura 7.6.- Comparación de la CDF del BLER en transmisión de vídeo empleando el último codificador o CS-2 al inicio de cada nueva transmisión

En la figura 7.6 se puede apreciar, por una parte, que el BLER obtenido para un mismo caso siempre es mejor con el algoritmo basado en BLER que con el algoritmo basado en throughput, tendencia ya observada en el apartado anterior. El dato de interés que se observa es que si en un mismo algoritmo se pasa a utilizar CS-2 al inicio de cada nueva transmisión, es decir, se reduce el empleo de CS-4 de forma notable, entonces el BLER obtenido disminuye, lo cual demuestra que la obtención de tasas de errores tan elevadas es consecuencia del excesivo uso de CS-4. Es más, si se toma como referencia la curva del algoritmo basado en throughput en el cual se utiliza el último esquema de codificación al inicio de cada nueva transmisión, se observa que la mejora en cuanto a BLER que se obtiene al utilizar en su lugar CS-2 para el mismo algoritmo y, por tanto, al reducir de forma considerable el uso de CS-4, es tan significativa o incluso mayor que la que se consigue empleando el algoritmo basado en BLER (utilizando el último codificador en ambos casos)[56], lo cual demuestra la importancia que tiene la reducción del uso de CS-4 para la obtención de un descenso significativo del BLER[57]. En valores medios, este comportamiento se observa claramente: la reducción del BLER al emplear el algoritmo basado en BLER (último codificador) es de un 0.6%, mientras que utilizando CS-2 con el mismo algoritmo (el de throughput) la reducción es del 2%.

[56] Para un porcentaje de casos ligeramente inferior al 90%, la mejora en el BLER obtenida en ambos casos es la misma, mientras que para porcentajes menores se aprecia que la utilización de CS-2 al inicio de cada transmisión de datos proporciona una mayor reducción del BLER que la que supone el propio algoritmo de BLER eligiendo el último codificador empleado en la transmisión anterior.

[57] De hecho, para porcentajes inferiores al 50%, reducir el uso de CS-4 es el único hecho que rebaja el BLER.

Los resultados anteriores confirman que la excesiva utilización de CS-4 es el motivo de la obtención de un BLER tan elevado. Queda claro, pues, que la consecución del BLER deseado hace necesaria la implementación de un mecanismo adicional que se encargue de controlar el grado de utilización de CS-4, pues el algoritmo basado en BLER, tal y como se definió, es incapaz de lograr por sí solo una reducción significativa del grado de utilización de este codificador. Ante este hecho, surge la pregunta de por qué el algoritmo basado en BLER no consigue disminuir lo suficiente el uso de CS-4.

En primer lugar, resulta bastante llamativo el hecho de que este esquema de codificación se utilice tan a menudo, especialmente si se tiene en cuenta que la región correspondiente a CS-4 en los mapas de decisión de los diversos algoritmos estudiados es mínima. En efecto, recordemos que el mapa de decisión de un algoritmo en el simulador consta de un total de 10000 puntos, cada uno de los cuales corresponde a una combinación distinta de BER medio y desviación estándar del BER. En 4209 de estos 10000 puntos, el codificador no está definido por corresponder a combinaciones de BER medio y desviación del BER que no se llegaron a producir en las simulaciones a nivel de enlace. De los 5791 puntos restantes en el mapa de decisión en los que sí está definido el codificador a emplear, tan sólo uno de ellos, correspondiente a un BER medio y a una desviación estándar del BER igual a cero, es el que indica la utilización de CS-4. Se dijo que las combinaciones más habituales de estos dos valores de BER eran las correspondientes a puntos cercanos a CS-4; aun así surge la cuestión de si realmente resulta tan probable la elección de este codificador. La respuesta a esta pregunta es sí.

Observando la figura 5.8, en la que se representa la CDF del BER para una relación CIR de 6 dB, se puede observar que incluso para una calidad del canal que se puede considerar más bien pobre[58] existe una probabilidad superior al 40% de que no ocurra ningún bit erróneo en la transmisión (BER igual a cero) y, por tanto, de que se elija CS-4 como esquema de codificación. La modificación del mapa de decisión de un algoritmo a otro no modifica este hecho ya que, si bien es cierto que se observan puntos que para un algoritmo corresponden a un determinado codificador y para otro más conservador (en lo que a tasa de errores se refiere) corresponden a un codificador más robusto, también es cierto que el punto correspondiente a CS-4 es común a todos ellos y que, es más, existe una elevada probabilidad de que sea seleccionado. Esta observación explica la reducción del BLER experimentada, ya que al buscar un BLER más bajo se emplean más a menudo CS-1 y CS-2 al mismo tiempo que disminuye el porcentaje de utilización de CS-3. Sin embargo, a penas desciende el porcentaje de utilización de CS-4, el esquema de codificación menos robusto y el más utilizado con diferencia, lo cual explica que la variación del BLER obtenido de un algoritmo a otro sea tan poco significativa.

Esta elevada probabilidad de utilización de CS-4, no considerada en el diseño del algoritmo, explica que la tasa de errores obtenida finalmente resulte tan elevada y difiera tanto de la que se busca, independientemente del objetivo de BLER perseguido por el algoritmo. Esta observación confirma la necesidad de utilizar algún mecanismo que limite la utilización de CS-4, ya que el cambio de criterio de un algoritmo a otro tan sólo modifica el mapa de decisión, hecho que resulta insuficiente para asentar el BLER.

[58] Ya para el algoritmo basado en throughput y en interfaces a nivel de bloque RLC, una relación CIR de 6 dB está próxima al umbral de decisión entre CS-1 y CS-2.

7.3.- PRIMER ALGORITMO PROPUESTO

Una vez analizadas las prestaciones de los algoritmos de referencia y habiendo presentado sus resultados, se compararán con los ofrecidos por el primer algoritmo propuesto en el presente trabajo, cuyo principal objetivo es establecer un compromiso entre el throughput global y la tasa de errores global (expresada en términos de BLER) experimentados en el sistema. Los resultados analizados corresponden a los ofrecidos para varios valores de k (2, 5, 20 y 1000), parámetro de diseño cuya variación permitía, en principio, modificar el grado de compromiso establecido entre estos dos factores.

Según se expuso en el capítulo previo, al analizar el comportamiento observado en el mapa de decisión de este algoritmo conforme variaba el valor del parámetro k, cuando éste sufría una disminución resultaba suficiente una peor calidad en el canal de transmisión para poder emplear los esquemas de codificación con menor protección frente a errores, lo cual se señaló como razón por la cual cabía esperar un incremento en el porcentaje de utilización de los codificadores menos robustos. Un incremento de k ocasionaba el efecto contrario, por lo que debería fomentar, en principio, un uso más frecuente de los codificadores más robustos. Puesto que los valores de k considerados son todos superiores a uno[59], los codificadores más robustos deberían emplearse más a menudo que en el caso del algoritmo basado en throughput, mientras que los de mayor velocidad binaria deberían experimentar un descenso en su empleo.

En la tabla 7.9 se muestra el porcentaje de uso de cada esquema de codificación. Si se comparan los valores del algoritmo basado en throughput y del primer algoritmo propuesto, se aprecia que el comportamiento predicho se cumple. Sin embargo, cuando se incluyen en la comparación los algoritmos basados en BLER se observan algunas irregularidades.

Algoritmo	CS-1	CS-2	CS-3	CS-4
Throughput	6.63 %	7.14 %	26.41 %	59.82 %
Algoritmo 1, k = 2	7.27 %	8.77 %	24.67 %	59.29 %
Algoritmo 1, k = 5	8.31 %	9.50 %	23.60 %	58.59 %
Algoritmo 1, k = 20	9.16 %	10.59 %	22.03 %	58.22 %
Algoritmo 1, k = 10^3	11.02 %	12.06 %	19.56 %	57.36 %
BLER 5%	11.06 %	8.59 %	22.63 %	57.72 %
BLER 2%	13.67 %	9.85 %	19.94 %	56.54 %
BLER 1%	16.07 %	11.03 %	17.28 %	55.62 %

Tabla 7.9.- Porcentaje de utilización de cada esquema de codificación en transmisión de vídeo

[59] Recordar que este algoritmo para k = 1 era precisamente el algoritmo basado en throughput.

El porcentaje de utilización de CS-1 ofrece una tendencia homogénea, si bien el de CS-2 resulta algo extraño. Sin embargo, observando el porcentaje de uso de CS-3 y CS-4 (los codificadores más empleados) se observa un comportamiento bien definido y similar al de CS-1. La única excepción observada respecto a la tendencia que se predijo es que el algoritmo con k = 1000 parece ser más conservador que el propio algoritmo basado en un BLER del 5%. No obstante, este hecho no es contradictorio, ya que un incremento de k supone una reducción en el uso de los codificadores menos robustos; si el valor alcanzado por este parámetro es lo suficientemente elevado, como es el caso, resulta normal que la reducción obtenida en la utilización de estos codificadores sea incluso mayor que la que supone la aplicación del algoritmo basado en un BLER del 5%, con lo que, efectivamente, el algoritmo sería más conservador en cuanto a tasa de errores.

Si bien las tendencias apreciadas en el porcentaje de uso de cada esquema de codificación no coinciden plenamente con las que cabría esperar según las suposiciones realizadas en el capítulo sexto, este hecho no debe resultar excesivamente preocupante, pues los esquemas de codificación empleados en el 80% de las ocasiones (CS-3 y CS-4) muestran un comportamiento bien definido y acorde con el que se predijo, hecho que debería ser responsable de que el algoritmo se comporte según cabe esperar.

Parámetro	Algoritmo	Medio	95%	99%
Throughput (kbps)	Throughput	16.56	10.46	7.53
	Algoritmo 1, k = 2	16.49	10.36	7.47
	Algoritmo 1, k = 5	16.39	10.21	7.34
	Algoritmo 1, k = 20	16.33	10.08	7.32
	Algoritmo 1, k = 10^3	16.18	9.90	7.21
	BLER 5%	16.24	9.92	7.22
	BLER 2%	16.04	9.68	7.07
	BLER 1%	15.85	9.50	7.05
BLER (%)	Throughput	11.18	25.90	39.73
	Algoritmo 1, k = 2	10.98	25.21	39.30
	Algoritmo 1, k = 5	10.90	24.83	39.27
	Algoritmo 1, k = 20	10.71	24.27	38.58
	Algoritmo 1, k = 10^3	10.40	23.28	38.01
	BLER 5%	10.60	23.73	38.43
	BLER 2%	10.25	22.77	37.52
	BLER 1%	9.93	21.91	36.87
Retardo normalizado (ms/kbit)	Throughput	58.32	99.08	129.44
	Algoritmo 1, k = 2	58.88	99.35	130.47
	Algoritmo 1, k = 5	59.55	104.24	131.61
	Algoritmo 1, k = 20	60.10	109.11	132.26
	Algoritmo 1, k = 10^3	61.26	110.22	134.88
	BLER 5%	60.82	110.21	134.76
	BLER 2%	62.34	116.36	137.37
	BLER 1%	63.70	119.63	138.90

Tabla 7.10.- Comparación del throughput, BLER y retardo normalizado en transmisión de vídeo para el primer algoritmo propuesto

La tabla 7.10 recoge todos los valores relevantes relativos al throughput, BLER y retardo normalizado obtenidos para los algoritmos de referencia y para las diversas variantes estudiadas del primer algoritmo propuesto. Centrando la atención sobre los resultados referentes al throughput, se observa que el algoritmo basado en throughput es el que alcanza el resultado más elevado, tanto en valor medio como en el 95% y 99% de los casos. Los valores obtenidos mediante el empleo del algoritmo propuesto resultan ser intermedios a los ofrecidos por los dos algoritmos de referencia, lo que permite afirmar, por el momento, que sí que es capaz de lograr el compromiso buscado entre ambos, por lo menos, en lo referente al throughput. Además, el incremento del valor de k para el primer algoritmo propuesto supone una reducción del throughput, tal y como cabía esperar, lo cual es consecuencia directa de una mayor utilización de los esquemas de codificación con menor velocidad binaria a medida que k se incrementa. Esta tendencia se observa en los valores mostrados en la tabla 7.10 y, más claramente, en el detalle de la figura 7.7 que se ofrece en la figura 7.8. Un examen detallado de estas figuras muestra que el throughput que se obtiene mediante el primer algoritmo que se propone para k = 1000 resulta ser menor que el ofrecido por el algoritmo basado en un BLER del 5% (aunque superior al obtenido por el diseñado para mantener el BLER por debajo del 2%). Esto confirma la observación realizada anteriormente, según la cual este valor de k ocasiona que el primer algoritmo propuesto sea más conservador que el propio algoritmo basado en un BLER del 5% al permitir un uso más restringido de los esquemas de codificación menos robustos (no tan acentuado como el algoritmo basado en un BLER del 2%).

Por lo que respecta al BLER global que se experimenta en el sistema, se puede apreciar que el BLER más reducido lo sigue proporcionando el algoritmo basado en un BLER del 1% mientras que el más elevado sigue correspondiendo al algoritmo basado en throughput. Por tanto, el valor del BLER global que se obtiene al aplicar este nuevo algoritmo ofrece resultados intermedios. Tomando como referencia los resultados del algoritmo basado en throughput, un análisis con mayor detenimiento de la tabla 7.10 y las figuras 7.8 y 7.10 (en la que se muestra un detalle de la figura 7.9), muestra que el algoritmo con k = 2 mejora ligeramente el BLER ofrecido por el algoritmo basado en throughput (equivalente a k = 1) a cambio de un pequeño descenso en el throughput, hecho lógico considerando el compromiso existente entre ambos factores. Un aumento del valor de este parámetro hasta k = 5 origina una mejora en el BLER todavía mayor que la obtenida con k = 2 y un descenso del throughput también más acentuado. Conforme se continúa incrementando el valor de k y, por tanto, se sigue fomentando una mayor utilización de los codificadores más robustos, el descenso del BLER se ve acompañado por un aumento del throughput, hasta que para k = 1000 el algoritmo restringe la utilización de los codificadores menos robustos en mayor proporción que el algoritmo basado en un BLER del 5%, lo cual explica que el BLER obtenido para este valor de k sea menor que para el algoritmo basado en BLER (y el throughput, mayor). Lo que finalmente se concluye a partir de este examen es que el algoritmo logra de forma eficiente el objetivo marcado: no sólo es capaz de establecer el compromiso buscado entre throughput global y BLER global del sistema sino que además permite modificar, de manera sencilla, el grado de compromiso entre ambos factores[60].

[60] La escasa diferencia que se observa entre los resultados de los diferentes algoritmos de la tabla 7.10 se debe, según se comentó en el apartado anterior, a que el elevado empleo de CS-4 es el factor común a todos ellos.

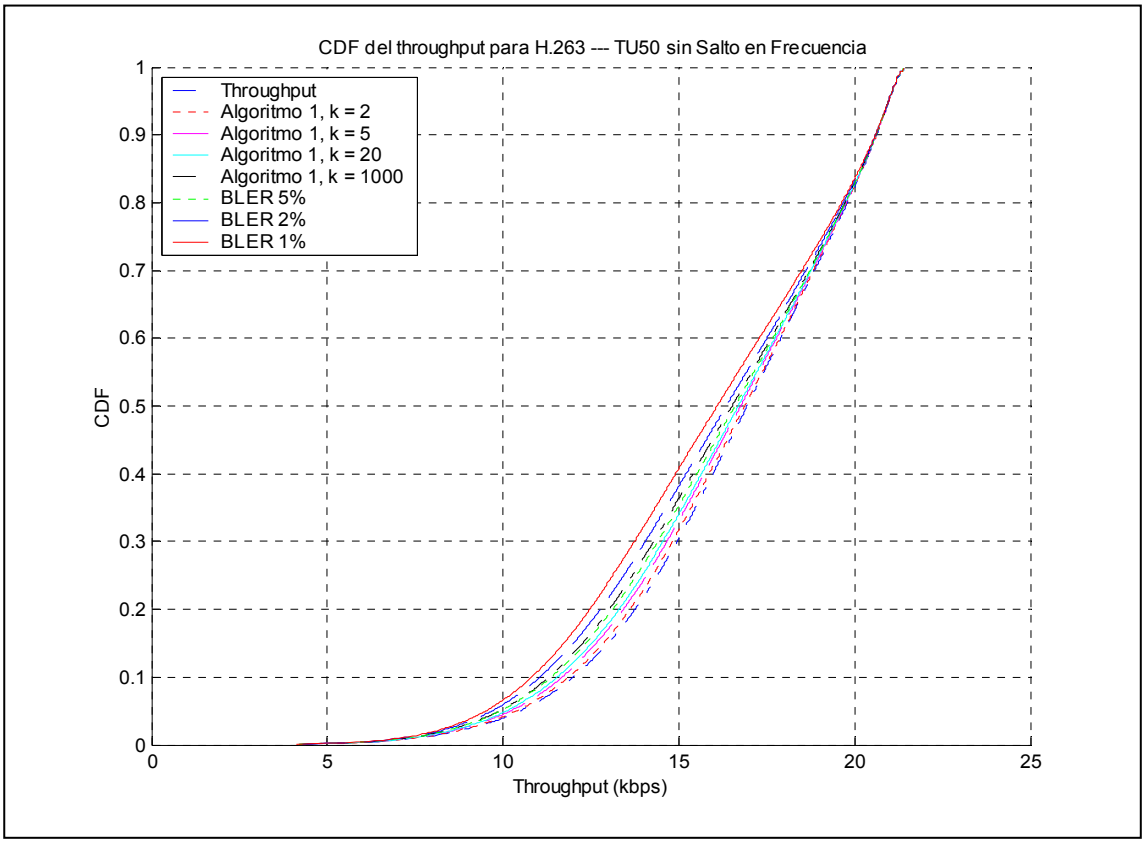


Figura 7.7.- CDF del throughput para el primer algoritmo propuesto en transmisión de vídeo

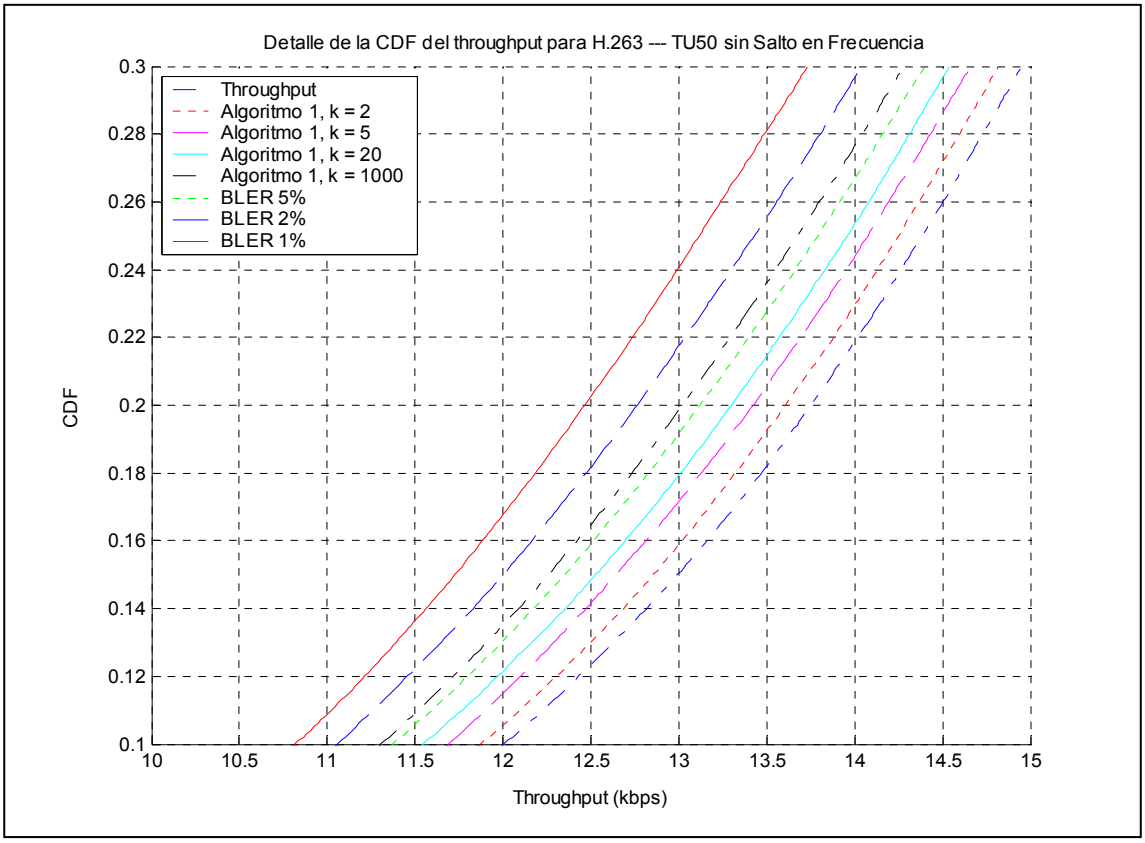


Figura 7.8.- Detalle de la CDF del throughput para el primer algoritmo propuesto en transmisión de vídeo

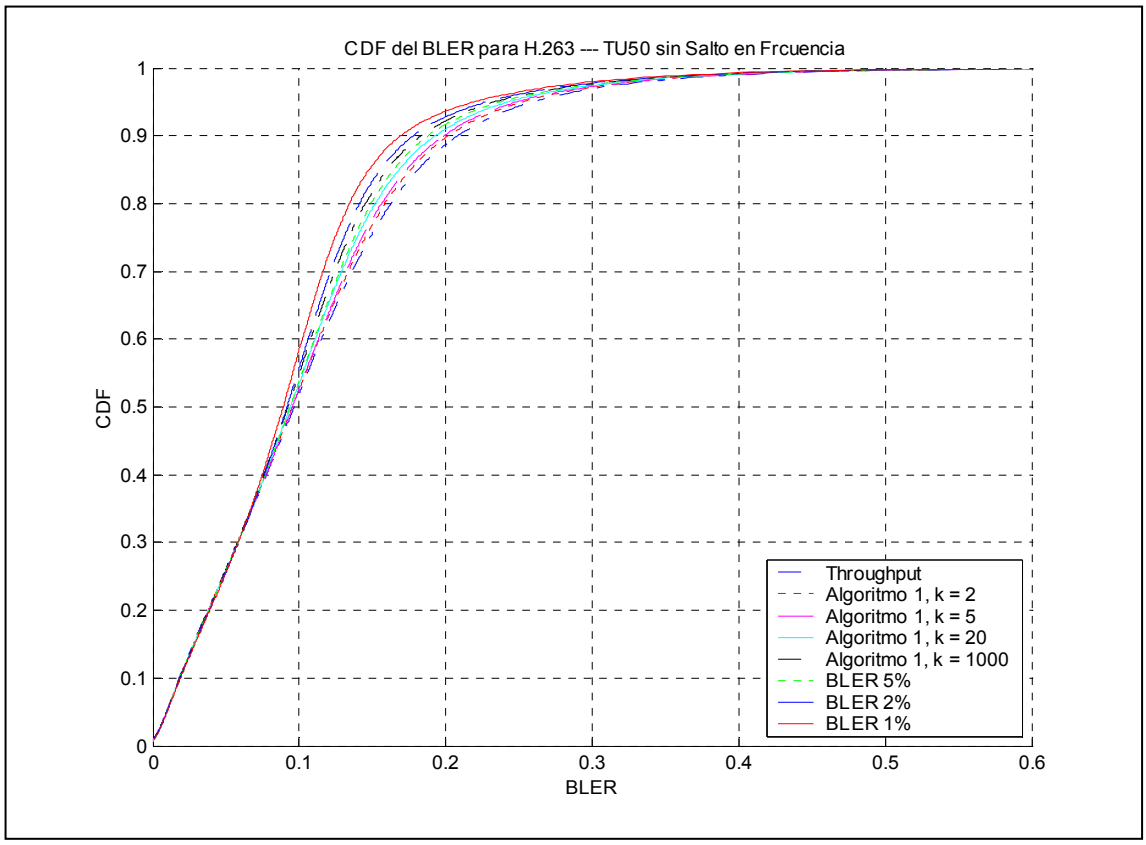


Figura 7.9.- CDF del BLER para el primer algoritmo propuesto en transmisión de vídeo

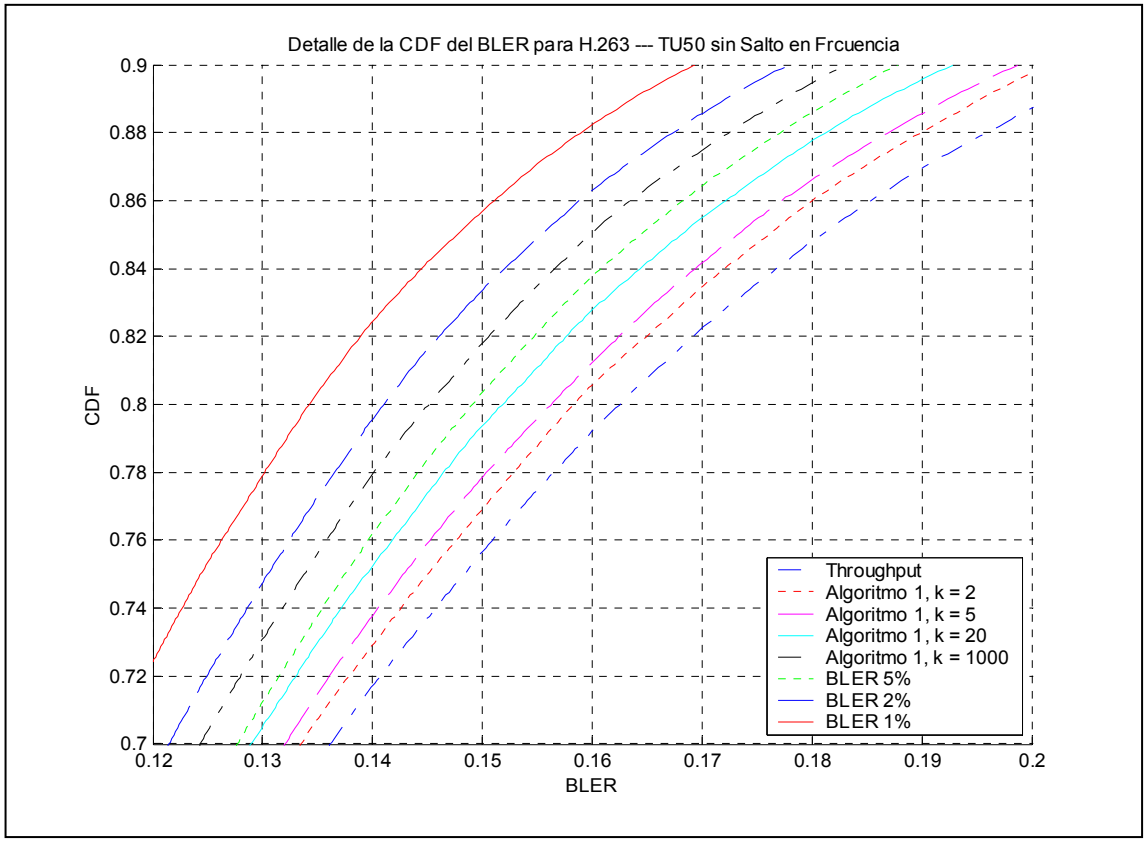


Figura 7.10.- Detalle de la CDF del BLER para el primer algoritmo propuesto en transmisión de vídeo

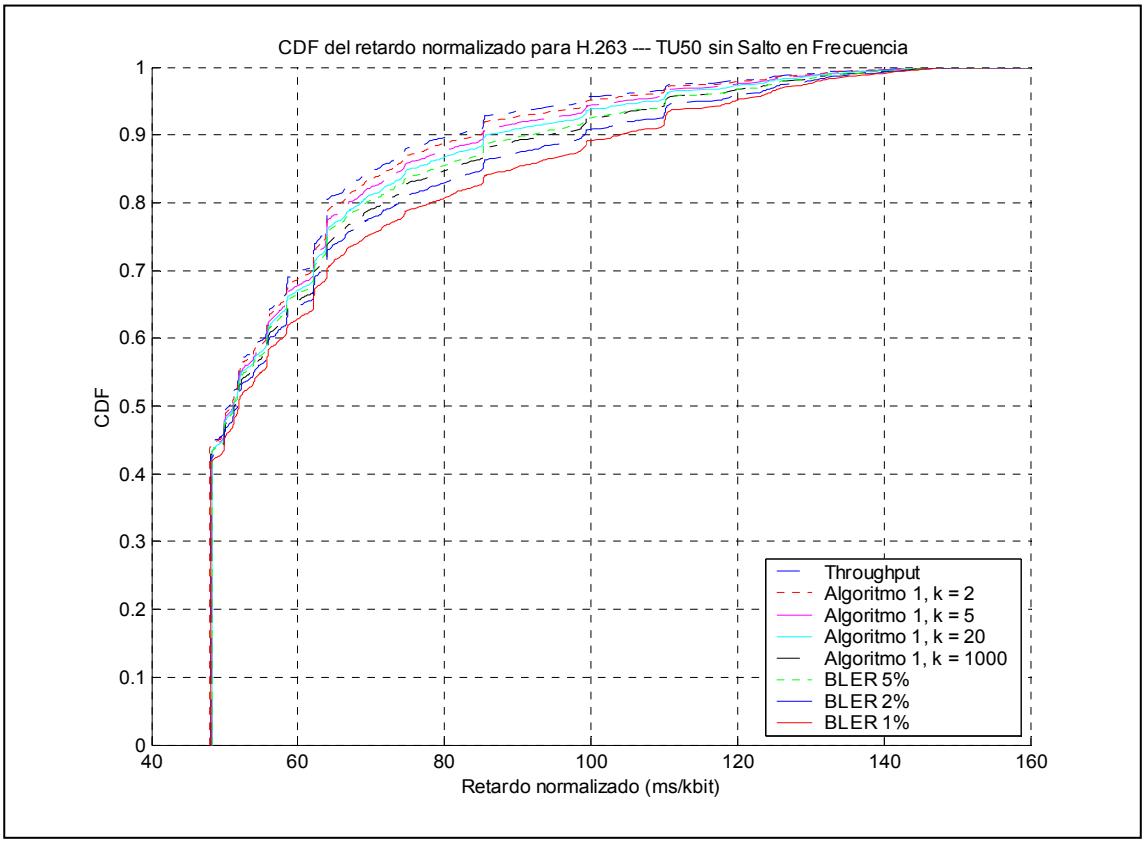


Figura 7.11.- CDF del retardo normalizado para el primer algoritmo propuesto en transmisión de vídeo

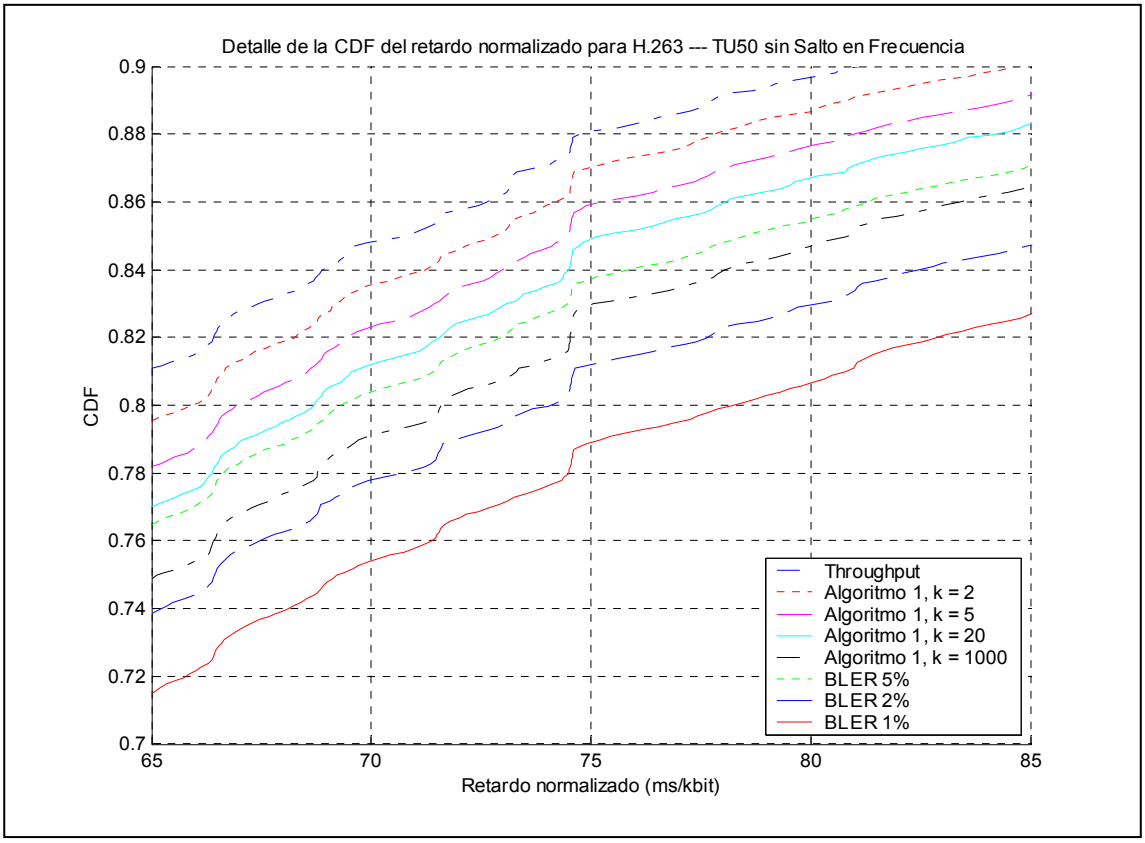


Figura 7.12.- Detalle de la CDF del retardo normalizado para el primer algoritmo propuesto en vídeo

Al analizar los algoritmos de referencia se pudo comprobar que una mejora del BLER siempre llevaba asociado un deterioro del throughput así como del retardo que se experimentaba durante la transmisión y viceversa. Puesto que el primer algoritmo propuesto demuestra ser capaz de modificar el compromiso existente entre throughput y BLER, no resulta extraño que también varíe, de manera paralela, el compromiso existente entre BLER y retardo a nivel de sistema. En efecto, la tabla 7.10 muestra que el incremento del parámetro k no sólo conlleva una reducción del throughput y una mejora del BLER sino que, además, conforme aumenta su valor (y por lo tanto, se reduce la utilización de los codificadores capaces de ofrecer una mayor tasa binaria), se incrementa el retardo (normalizado) que experimentan las tramas de vídeo durante su transmisión. Se aprecia, una vez más, que el algoritmo correspondiente a k = 1000 es más conservador que el algoritmo basado en un BLER del 5%. La CDF correspondiente se muestra en la figura 7.11; el detalle de la figura 7.12 permite apreciar claramente el comportamiento señalado.

Al aumentar el valor del parámetro k, y como consecuencia de un incremento del retardo normalizado, el algoritmo que se propone ocasionará un mayor número de tramas de vídeo que son transmitidas con retardo, es decir, que no llegan al receptor antes de que el plazo de tiempo prefijado expire, hecho que queda patente en los resultados mostrados en la tabla 7.11.

Algoritmo	**Porcentaje de tramas de vídeo transmitidas**	
	Sin retardo	**Con retardo**
Throughput	69.16 %	30.84 %
Algoritmo 1, k = 2	68.37 %	31.63 %
Algoritmo 1, k = 5	67.48 %	32.52 %
Algoritmo 1, k = 20	66.72 %	33.28 %
Algoritmo 1, k = 1000	65.23 %	34.77 %
BLER 5%	65.97 %	34.03 %
BLER 2%	64.03 %	35.97 %
BLER 1%	62.33 %	37.67 %

Tabla 7.11.- Porcentaje de tramas de vídeo transmitidas con y sin retardo para el 1er. algoritmo propuesto

Un descenso en el valor de este parámetro, por tanto, ocasionará el efecto contrario, es decir, un incremento en el porcentaje de tramas de vídeo que llegan al receptor antes de que expire el plazo de tiempo disponible para su transmisión. Esta mejora se verá acompañada, de manera similar a como se pudo apreciar en la tabla 7.4, por un incremento general en el porcentaje de tramas que, además de transmitirse sin retardo, satisfacen diversos requisitos de calidad de servicio expresada en términos de BLER y BER. Este hecho, mostrado en la tabla 7.12, muestra que el porcentaje de tramas sin retardo con una tasa de errores aceptable y, por tanto, capaces de contribuir a una mejora de la calidad de servicio, aumenta al disminuir el valor del parámetro k.

Requisito	Throughput	Alg. 1, k = 2	Alg. 1, k = 5	Alg. 1, k = 20	Alg. 1, k = 1000	BLER 5%	BLER 2%
BLER ≤ 1%	44.73 %	44.44 %	44.03 %	43.78 %	43.18 %	43.41 %	42.53 %
BLER ≤ 2%	44.75 %	44.46 %	44.04 %	43.79 %	43.20 %	43.43 %	42.55 %
BLER ≤ 5%	44.98 %	44.70 %	44.28 %	44.03 %	43.43 %	43.66 %	42.78 %
BER ≤ 10^{-3}	45.93 %	45.53 %	45.03 %	44.77 %	44.12 %	44.42 %	43.51 %
BER ≤ 10^{-4}	40.40 %	40.06 %	39.63 %	39.39 %	38.84 %	39.10 %	38.30 %
BER ≤ 10^{-5}	39.33 %	39.00 %	38.59 %	38.36 %	37.83 %	38.08 %	37.31 %
BER ≤ 10^{-6}	39.23 %	38.90 %	38.50 %	38.26 %	37.74 %	37.99 %	37.21 %

Tabla 7.12.- Porcentaje de tramas de vídeo transmitidas sin retardo y que cumplen un determinado requisito de calidad

Requisito	Throughput	Alg. 1, k = 2	Alg. 1, k = 5	Alg. 1, k = 20	Alg. 1, k = 1000	BLER 5%	BLER 2%
BLER ≤ 1%	64.67 %	65.01 %	65.25 %	65.61 %	66.20 %	65.81 %	66.42 %
BLER ≤ 2%	64.70 %	65.03 %	65.27 %	65.63 %	66.23 %	65.84 %	66.44 %
BLER ≤ 5%	65.04 %	65.38 %	65.62 %	65.98 %	66.59 %	66.19 %	66.80 %
BER ≤ 10^{-3}	66.41 %	66.60 %	66.74 %	67.10 %	67.64 %	67.34 %	67.95 %
BER ≤ 10^{-4}	58.41 %	58.60 %	58.73 %	59.04 %	59.55 %	59.27 %	59.82 %
BER ≤ 10^{-5}	56.87 %	57.05 %	57.20 %	57.49 %	58.00 %	57.73 %	58.27 %
BER ≤ 10^{-6}	56.73 %	56.91 %	57.05 %	57.35 %	57.86 %	57.58 %	58.12 %

Tabla 7.13.- De las tramas de vídeo transmitidas sin retardo, porcentaje que cumple un determinado requisito de calidad

En el apartado anterior se puso de relieve que el mayor número de tramas de vídeo sin retardo que obtenía el algoritmo basado en throughput, se lograba a cambio de una peor calidad global en las tramas que llegaban sin retardo al receptor, aspecto en el cual era superado por el algoritmo basado en BLER. Además, cuando el algoritmo basado en BLER perseguía una tasa de errores más reducida, aumentaba el número de tramas con retardo pero mejoraba la calidad de las que se transmitían sin retardo.

A partir de esta observación, y aplicando este mismo razonamiento al primer algoritmo propuesto, lo lógico sería que al aumentar el valor de k, es decir, conforme el algoritmo se vuelve más exigente en cuanto a tasa de errores, se observara la misma tendencia que al perseguir un BLER más reducido.

La tabla 7.13 demuestra que esta tendencia se cumple (por los mismos motivos que se explicaron al analizar los algoritmos de referencia), ya que se observa que un incremento del valor del parámetro *k*, que supone que el algoritmo reduzca el valor del BLER obtenido, ocasiona que la calidad de las tramas recibidas sin retardo mejore, a cambio de aumentar el porcentaje de tramas que se transmiten con retardo y, según puede apreciarse en las figuras 7.13 y 7.14, de ocasionar que la cantidad de información perdida en estas tramas sea también mayor. Por los mismos motivos por los cuales se cuestionó la utilidad del algoritmo basado en BLER para mejorar la calidad de servicio final de un servicio de tiempo real (al menos, bajo las condiciones estudiadas), es cuestionable para este algoritmo la conveniencia de realizar el diseño considerando para el parámetro *k* valores excesivamente elevados.

Para finalizar, y resumiendo los resultados observados, es posible afirmar que el primer algoritmo que se propone en este trabajo consigue establecer un compromiso entre el throughput global y el BLER global a nivel de sistema, objetivo para el cual fue diseñado. Además, según se ha podido comprobar, permite modificar el grado de compromiso establecido entre ambos factores (variando el valor del parámetro *k*) en función de cuál de los dos se considere de mayor relevancia respecto al otro. Debido a que a nivel de sistema existe un compromiso no sólo entre el throughput y el BLER, sino también entre éste último y el retardo, la variación del compromiso establecido entre los dos primeros implica también la variación del compromiso existente con el retardo global del sistema.

En los resultados mostrados en este apartado se muestra que un incremento del valor de *k* es equivalente a otorgar una mayor importancia al BLER global del sistema, cuya reducción se logra a cambio de disminuir el throughput y degradar el retardo que sufren las tramas de vídeo. En consecuencia, un incremento del valor de *k* implica un incremento del porcentaje de tramas transmitidas con retardo y de la cantidad total de información que se ha de descartar en el receptor. A cambio, el beneficio obtenido con el aumento de *k* es una mejora apreciable en la calidad de las tramas de vídeo que llegan al receptor sin retardo. La variación del valor de *k* repercute sobre otro factor que, aunque hasta ahora no se ha mencionado, resulta relevante en el estudio del comportamiento de *Link Adaptation*: el número de cambios de esquema de codificación por segundo. Este parámetro, si bien depende del periodo de actualización usado, al ser comparado entre varios algoritmos para los que éste se mantiene constante, proporciona una estimación de la carga de señalización inducida sobre el sistema por el empleo de cada uno de los diferentes algoritmos. Puede apreciarse que, a medida que aumenta el valor de *k*, se incrementa el número de cambios de codificador por segundo. Ello se debe a que a medida que el algoritmo se vuelve más conservador, el número de veces en las que se emplea un esquema de codificación más robusto de lo requerido por la calidad del enlace radio se incrementa, con lo cual tiene lugar un mayor número de cambios. Un descenso del valor de *k*, no sólo aumenta el throughput sino que también reduce la carga de señalización del sistema, al precio ya conocido.

Throughput	**Alg. 1, k = 2**	**Alg. 1, k = 5**	**Alg. 1, k = 20**	**Alg. 1, k = 1000**	**BLER 5%**	**BLER 2%**
4.69	4.84	4.94	5.03	5.15	5.02	5.16

Tabla 7.14.- Número medio de cambios de esquema de codificación por segundo

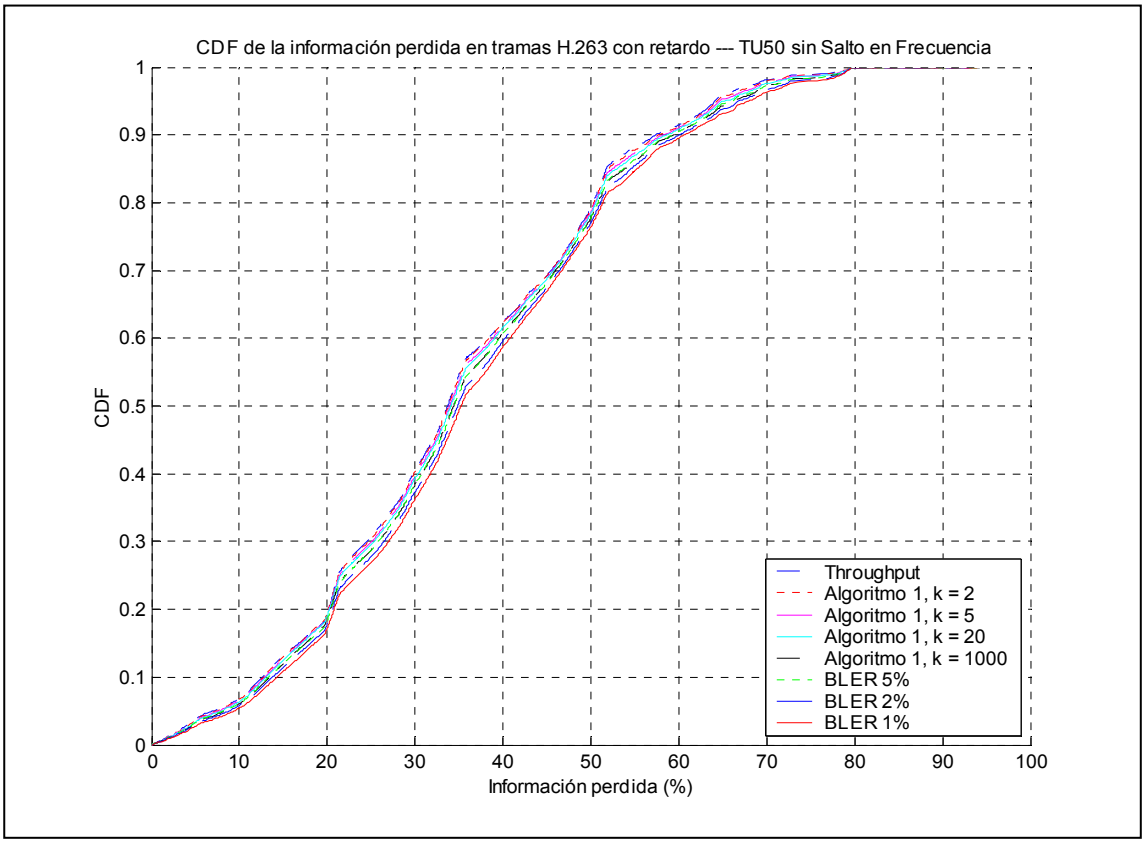


Figura 7.13.- CDF del porcentaje de información perdida en tramas de vídeo con retardo

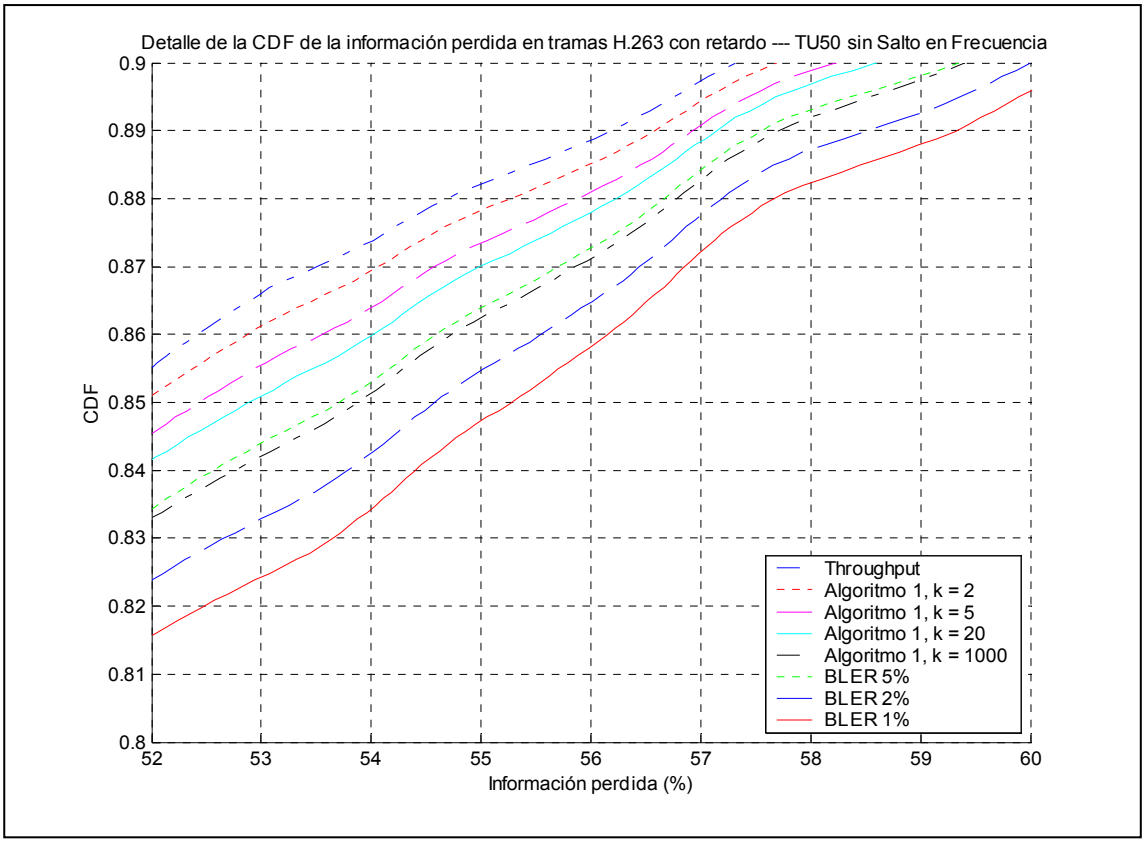


Figura 7.14.- Detalle de la CDF del porcentaje de información perdida en tramas de vídeo con retardo

7.4.- SEGUNDO ALGORITMO PROPUESTO

Frente los algoritmos basados en throughput y en BLER, cuyos criterios estaban basados, respectivamente, en la mejora de la cantidad de bits correctos transmitidos por unidad de tiempo y de la tasa de bloques RLC recibidos con error, el segundo de los algoritmos que se proponen en el presente trabajo establece un criterio de selección del esquema de codificación de canal en el que interviene, no sólo el throughput, sino también el retardo (reflejado en el concepto de retardo de trama, empleado en el diseño del algoritmo en el capítulo anterior).

A diferencia de estos dos algoritmos, que mostraban una clara predilección por el empleo de codificadores más robustos o con mayor velocidad binaria en función del objetivo perseguido, este segundo algoritmo no es propenso a fomentar el uso de un determinado grupo de codificadores. Recordemos que una diferencia significativa de este algoritmo en comparación con los anteriores es que se caracterizaba por exhibir un comportamiento dinámico en el cual el empleo más frecuente de unos codificadores u otros no era permanente sino que venía determinado por el tamaño y la duración de la trama de vídeo a transmitir en cada situación o, más exactamente, por la tasa binaria que requería para su transmisión sin retardo así como por la cantidad de recursos por usuario de los que se disponía. Por lo tanto, el que se fomente el empleo de un conjunto concreto de codificadores no vendrá dado tanto por el propio algoritmo como por los requisitos de la secuencia de vídeo transmitida y la cantidad de recursos disponibles.

Este hecho explica la variación observada en la tabla 7.15 en el porcentaje de utilización de cada esquema de codificación en función del número de time-slots por usuario (TS/u) que se considera en el diseño del algoritmo. Mientras que la secuencia de vídeo es la misma en todos los casos, es decir, la cantidad de recursos que requiere el flujo de información para una transmisión adecuada permanece constante, al reducir el número de time-slots (considerado) por usuario, aumenta el porcentaje de ocasiones en que la velocidad binaria que se puede ofrecer es menor que la que necesita la trama de vídeo para ser transmitida sin retardo, por lo cual el algoritmo propicia en estos casos que se utilicen con mayor frecuencia los codificadores de mayor tasa binaria. Si el diseño se realiza considerando un mayor número de canales por usuario, entonces el algoritmo se comporta de forma acorde, es decir, no fomenta de forma tan acentuada el uso de los codificadores con mayor tasa binaria, tal y como se vio en el capítulo sexto.

Algoritmo	CS-1	CS-2	CS-3	CS-4
Algoritmo 2, 1 TS/u	3.37 %	2.92 %	27.98 %	65.73 %
Algoritmo 2, 2 TS/u	4.41 %	6.38 %	28.11 %	61.10 %
Algoritmo 2, 3 TS/u	5.21 %	6.89 %	27.35 %	60.55 %
Algoritmo 2, 4 TS/u	5.75 %	6.92 %	27.12 %	60.21 %
Throughput	6.63 %	7.14 %	26.41 %	59.82 %
BLER 5%	11.06 %	8.59 %	22.63 %	57.72 %

Tabla 7.15.- Porcentaje de utilización de cada esquema de codificación en transmisión de vídeo

Se hace necesario, en este punto, realizar un pequeño inciso para dejar claro un aspecto relativo al número de time-slots por usuario. El número de time-slots de los que realmente dispone un usuario durante la transmisión de vídeo[61], sí bien determina los resultados finales obtenidos, no influye, en cambio, sobre el modo de comportarse del algoritmo. En efecto, el algoritmo *desconoce* la cantidad de recursos utilizada para cada usuario en la simulación; sí *conoce*, en cambio, el número de time-slots por usuario que se le indican en el diseño y, en función de éste, el algoritmo se comportará de una forma u otra.

Para cada trama de vídeo a transmitir, el algoritmo obtendrá la tasa binaria requerida para que sea transmitida dentro del tiempo preestablecido[62] y la comparará con la velocidad binaria que cada esquema de codificación es capaz de ofrecer, lo cual dependerá del número de time-slots por usuario que se haya considerado en el diseño. En función de esta comparación, el algoritmo generará un mapa de decisión en el que se fomenta un mayor uso de aquellos esquemas de codificación que son capaces de ofrecer una velocidad binaria igual o mayor a la requerida por la trama y, por tanto, que son capaces de conseguir su transmisión sin retardo, según pudo apreciarse en el capítulo sexto. Si se indica un número de time-slots inferior al que requiere la secuencia de vídeo, entonces habrá un mayor porcentaje de ocasiones en las que la velocidad binaria que se puede ofrecer no sea suficiente para transmitir la trama sin retardo, con lo que el algoritmo generará para esta trama un mapa de decisión en el que se fomenta una mayor utilización de los codificadores con mayor velocidad binaria, que es lo que se observa en la tabla 7.15.

Evidentemente, el diseño coherente es aquél en el que se considera el número de time-slots por usuario de los que realmente se dispone, ya que de esta forma el algoritmo se comportará de forma acorde a la realidad. Sin embargo, esto no impide que, a pesar de simular la transmisión de vídeo con un único time-slot por usuario, el algoritmo se diseñe considerando otros valores y se estudie su comportamiento. Cabe esperar, bajo estas circunstancias, que si el algoritmo se comporta *considerando* que cada usuario dispone de un número de time-slots mayor del que realmente tiene lugar, entonces no será tan eficaz en la consecución del objetivo como en aquellos casos en los que el diseño se realiza empleando el número de canales por usuario utilizados en la transmisión real, hecho que se corroborará en los resultados.

Por ello, la comparación adecuada de este algoritmo con los algoritmos de referencia es aquella en la que se considera el algoritmo diseñado con un time-slot por usuario, que corresponde a la cantidad de recursos que realmente se dedica a cada uno en la simulación. Sin embargo, también se presentarán los resultados obtenidos al simular los algoritmos diseñados para dos, tres y hasta cuatro time-slots por usuario en un sistema en el que cada usuario tan sólo dispone de uno. El motivo por el cual se hace esto es para mostrar que el presente algoritmo, incluso cuando se diseña para cuatro time-slots por usuario, supera en varios aspectos a los algoritmos de referencia.

[61] En el capítulo quinto se expusieron los motivos de emplear un único time-slot por usuario en la simulación cuando, según se vio en el anexo B, era necesario un número más elevado.

[62] Igual al cociente entre su tamaño y su duración.

Parámetro	Algoritmo	Medio	95%	99%
Throughput (kbps)	Algoritmo 2, 1 TS/u	16.94	11.33	8.09
	Algoritmo 2, 2 TS/u	16.69	10.74	7.57
	Algoritmo 2, 3 TS/u	16.66	10.67	7.63
	Algoritmo 2, 4 TS/u	16.61	10.57	7.51
	Throughput	16.56	10.46	7.53
	BLER 5%	16.24	9.92	7.22
BLER (%)	Algoritmo 2, 1 TS/u	12.20	28.91	42.71
	Algoritmo 2, 2 TS/u	11.70	28.18	42.36
	Algoritmo 2, 3 TS/u	11.37	27.00	40.72
	Algoritmo 2, 4 TS/u	11.35	26.72	40.73
	Throughput	11.18	25.90	39.73
	BLER 5%	10.60	23.73	38.43
Retardo normalizado (ms/kbit)	Algoritmo 2, 1 TS/u	55.36	85.32	119.59
	Algoritmo 2, 2 TS/u	57.01	85.73	124.35
	Algoritmo 2, 3 TS/u	57.52	92.22	125.88
	Algoritmo 2, 4 TS/u	57.83	95.59	127.46
	Throughput	58.32	99.08	129.44
	BLER 5%	60.82	110.21	134.76

Tabla 7.16.- Comparación del throughput, BLER y retardo normalizado en transmisión de vídeo para el segundo algoritmo propuesto

El primero de ellos es, como puede apreciarse en la tabla 7.16, el throughput. En efecto, el segundo algoritmo propuesto no sólo es capaz de ofrecer un throughput más elevado que el alcanzado por el propio algoritmo basado en throughput, tanto en valor medio como en el 95% y 99% de los casos, sino que sigue siéndolo cuando el diseño se realiza considerando cuatro time-slots (salvo un caso puntual). Al comparar los valores de la tabla referentes al algoritmo basado en throughput y al segundo que se propone diseñado con 1 time-slot por usuario, se observa que la mejora relativa es del 2.3% en valor medio y del 8.3% y 7.4% en el 95% y 99% de los casos, respectivamente. Respecto a algoritmos más robustos, como el basado en un BLER del 5%, la diferencia es todavía más significativa, del 4.3%, 14.2% y 12% para estos mismos valores, respectivamente.

La razón de esta mejora hay que buscarla en el aumento del porcentaje de uso de los esquemas de codificación con mayor velocidad binaria, en especial CS-4, cuyo incremento respecto al algoritmo basado en throughput es del 6%. Sin embargo, este aumento, que en comparación con los observados hasta ahora es bastante considerable, no justifica por sí sólo que el throughput también aumente, ya que la protección ante errores de estos codificadores es menor y el transmitir más bits por unidad de tiempo (aumentar la velocidad binaria) no supone un incremento del throughput si la tasa de errores aumenta, lo cual es más probable al emplear con mayor frecuencia un esquema de codificación poco robusto. El dato que explica este aumento del throughput es el que se muestra en la tabla 7.17, donde se recoge el porcentaje de ocasiones en las que cada uno de estos dos algoritmos elige el esquema de codificación óptimo[63].

[63] Al comienzo de un periodo de actualización, el algoritmo selecciona un determinado codificador. Una vez concluido éste, si el codificador empleado coincide con el que el algoritmo hubiera escogido a partir de la calidad del enlace radio experimentada durante este periodo, entonces se ha utilizado el codificador óptimo.

En la tabla 7.17 se puede apreciar que la diferencia en este parámetro no es excesivamente significativa (del 1.3% como máximo) para el algoritmo en cuyo diseño se consideran dos, tres y cuatro time-slots por usuario respecto al algoritmo basado en throughput (tampoco lo es la del throughput para estos números de time-slots). Sin embargo, para el diseño que se realiza considerando un único time-slot por usuario, la diferencia es notablemente mayor, del 5.2% exactamente.

Algoritmo	Porcentaje de elección del codificador óptimo
Throughput	65.99 %
Algoritmo 2, 4 TS/u	66.38 %
Algoritmo 2, 3 TS/u	66.63 %
Algoritmo 2, 2 TS/u	67.30 %
Algoritmo 2, 1 TS/u	71.17 %

Tabla 7.17.- Porcentaje de utilización del codificador óptimo

Éste es el motivo que explica que el segundo algoritmo propuesto sea capaz de alcanzar un throughput más elevado que el que logra el propio algoritmo basado en la maximización del throughput: no sólo se usan de forma más frecuente los esquemas de codificación con mayor velocidad binaria, sino que también el porcentaje de veces en los que la elección de estos codificadores resulta ser óptima también es mayor. El hecho de transmitir más bits por unidad de tiempo y de que el algoritmo acierte en un mayor porcentaje de ocasiones a seleccionar estos codificadores de forma correcta, ocasiona un aumento en la cantidad de bits correctos transmitidos por unidad de tiempo o, lo que es lo mismo, del throughput, el cual viene acompañado por una reducción relativa del 16% en el número de cambios de codificador por segundo para el diseño de 1 TS/u.

Conforme el diseño se realiza con un mayor número de time-slots, el porcentaje de utilización de estos codificadores se reduce (por los motivos ya explicados) y el throughput también, según se puede apreciar en la tabla 7.16 así como en el detalle de la figura 7.15 mostrado en la figura 7.16. Aun así, el algoritmo diseñado considerando cuatro time-slots por usuario sigue ofreciendo, en general, un throughput mayor que el alcanzado por el algoritmo cuyo criterio consiste en la maximización el throughput.

En todos los resultados mostrados anteriormente se pudo ver que la mejora del throughput a nivel de sistema se veía acompañada por un empeoramiento del BLER y, como cabe esperar, este caso no constituye una excepción. A medida que el throughput que se obtiene aumenta, el BLER aumenta también. Este comportamiento se observa que tiene lugar cuando desciende el número de time-slots por usuario considerado en el diseño, lo cual se debe, como se explicó anteriormente, a que este hecho fomenta un mayor empleo de los codificadores menos robustos, con lo que la tasa de errores crece. Las figuras 7.17 y 7.18 muestran de una forma más clara esta tendencia. En la tabla 7.16 se aprecia que el deterioro del BLER originado por el algoritmo diseñado para un time-slot por usuario es del 1%, 3% y 3% para el valor medio y en el 95% y 99% de los casos.

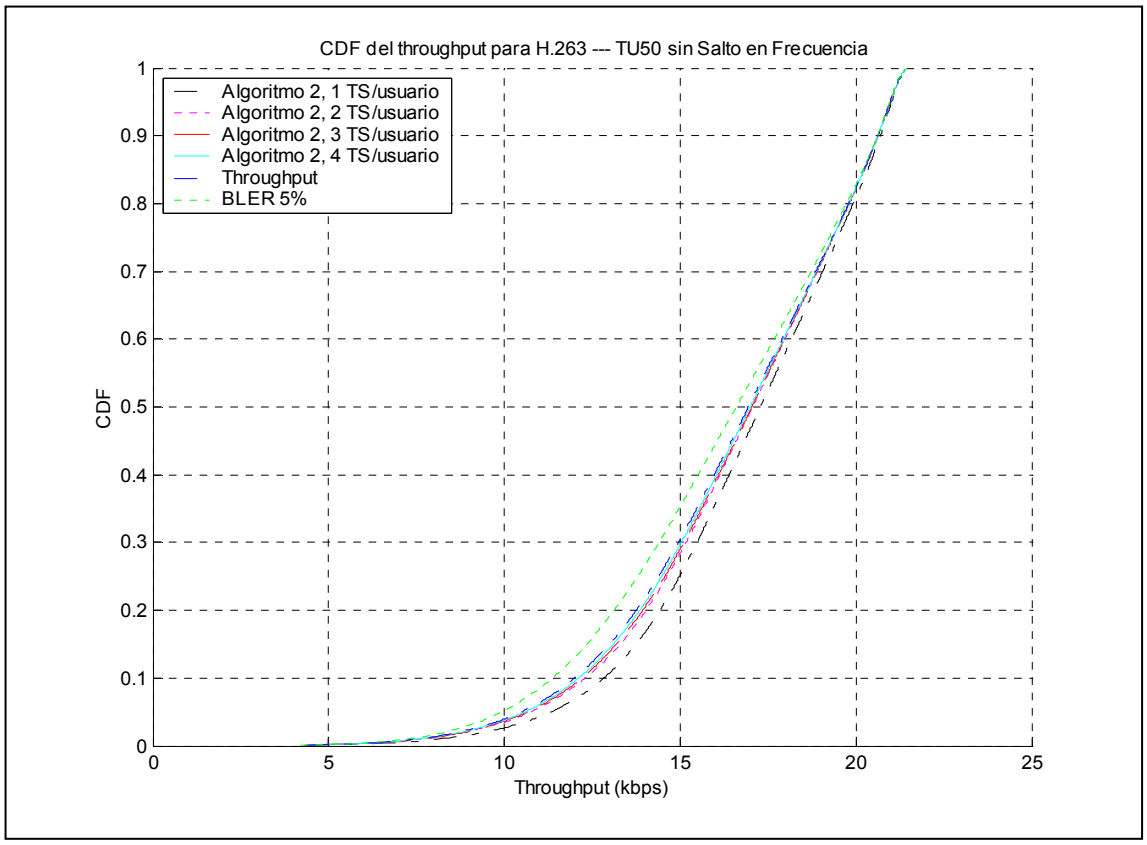


Figura 7.15.- CDF del throughput para el segundo algoritmo propuesto en transmisión de vídeo

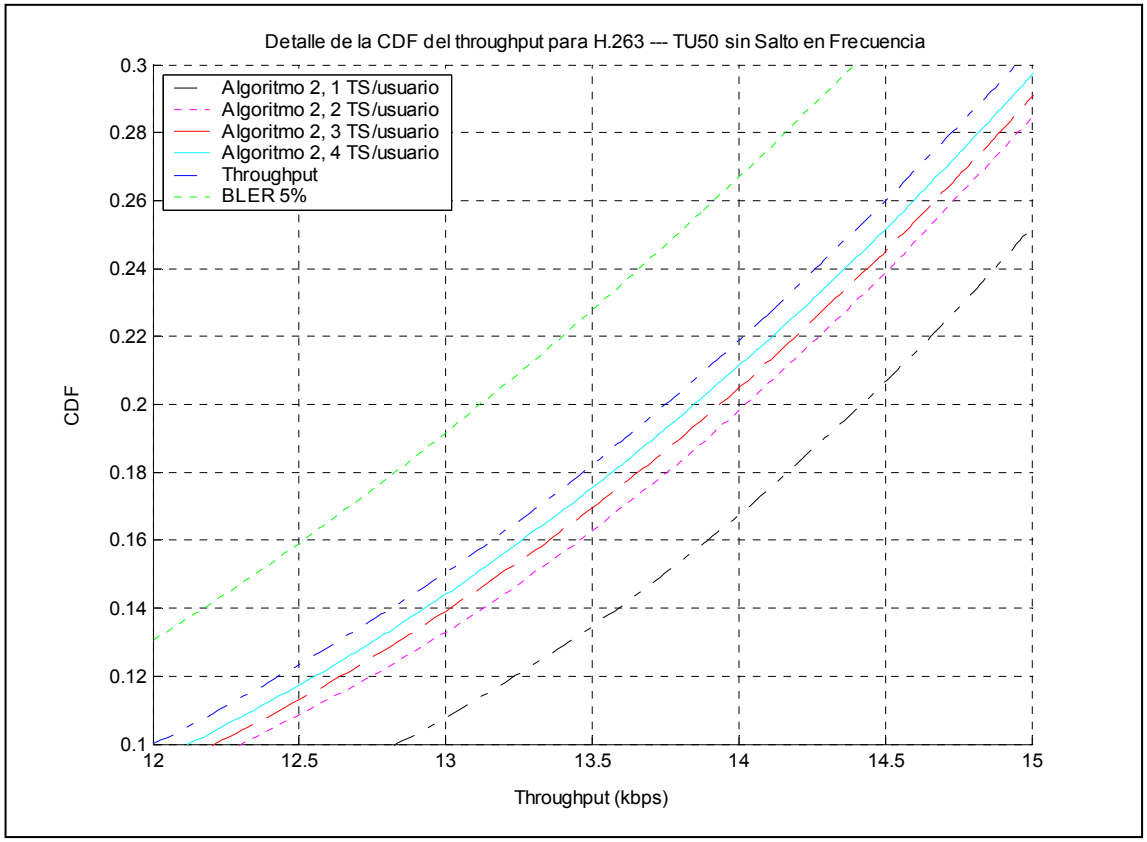


Figura 7.16.- Detalle de la CDF del throughput para segundo algoritmo propuesto en transmisión de vídeo

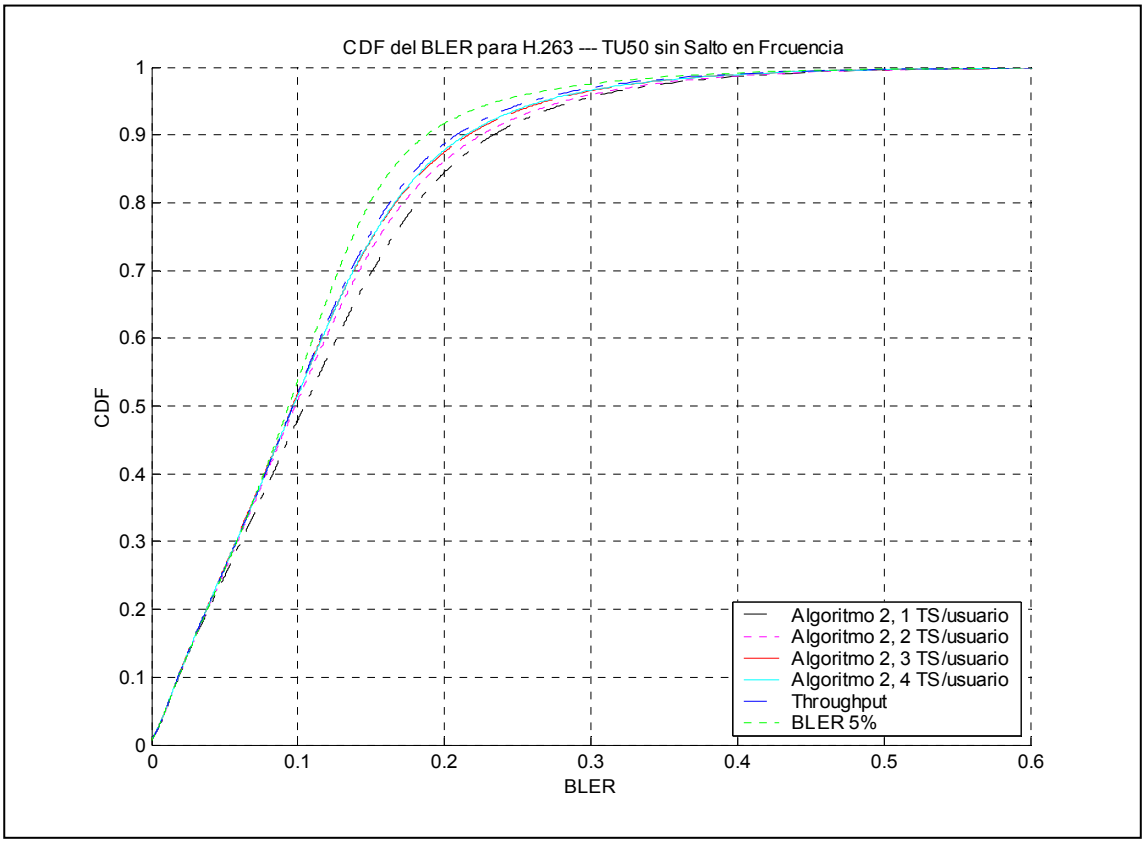


Figura 7.17.- CDF del BLER para el segundo algoritmo propuesto en transmisión de vídeo

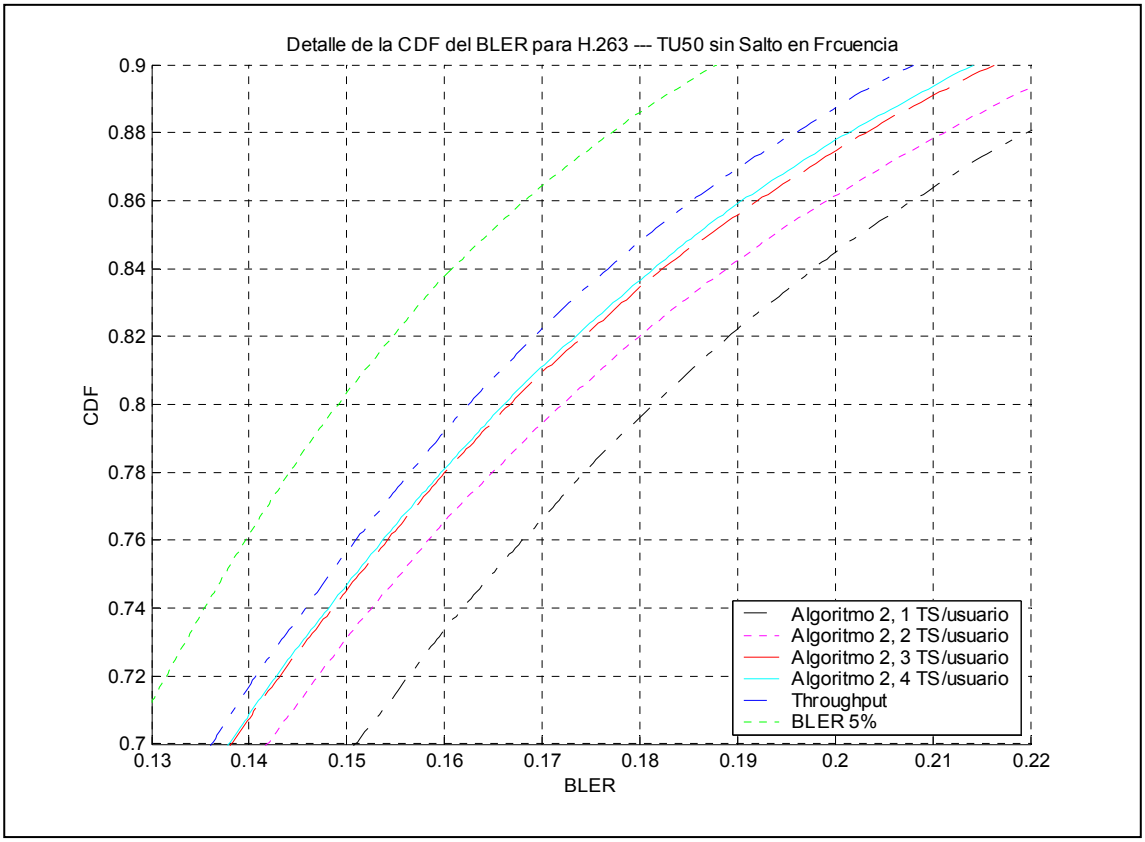


Figura 7.18.- Detalle de la CDF del BLER para el segundo algoritmo propuesto en transmisión de vídeo

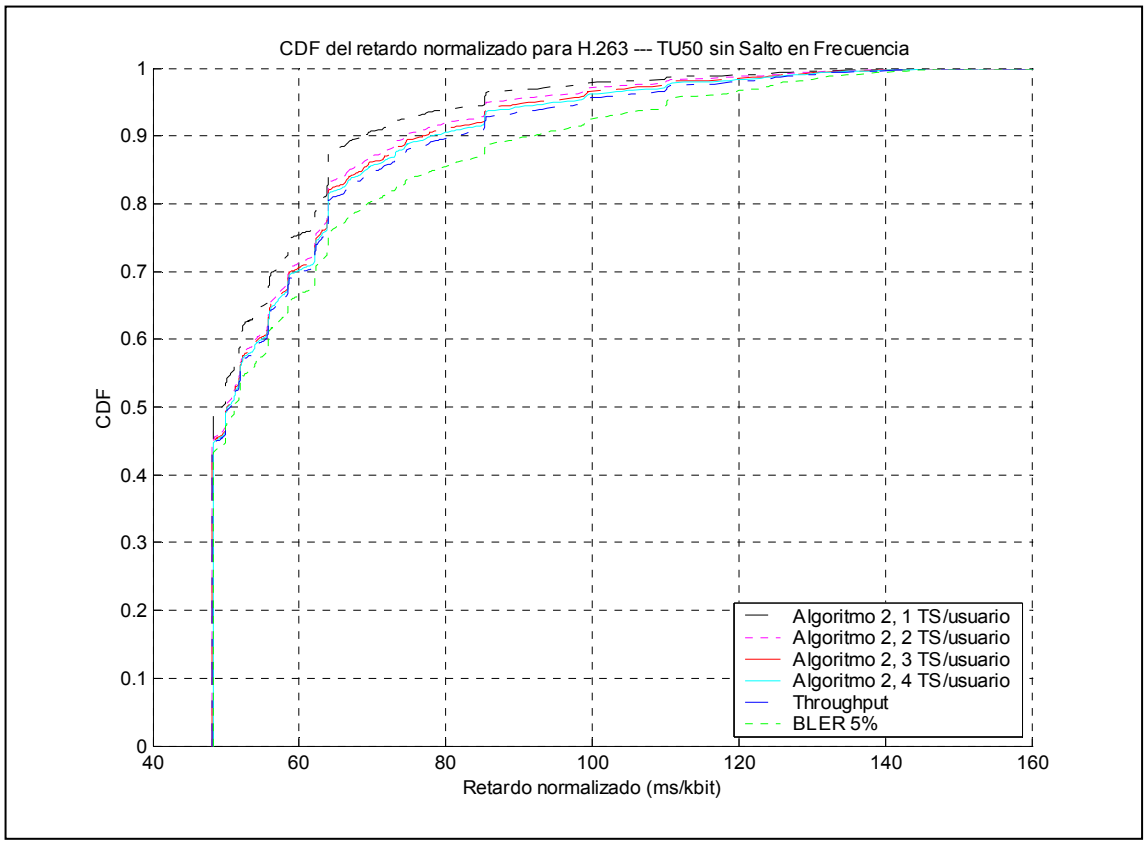


Figura 7.19.- CDF del retardo normalizado para el segundo algoritmo propuesto en transmisión de vídeo

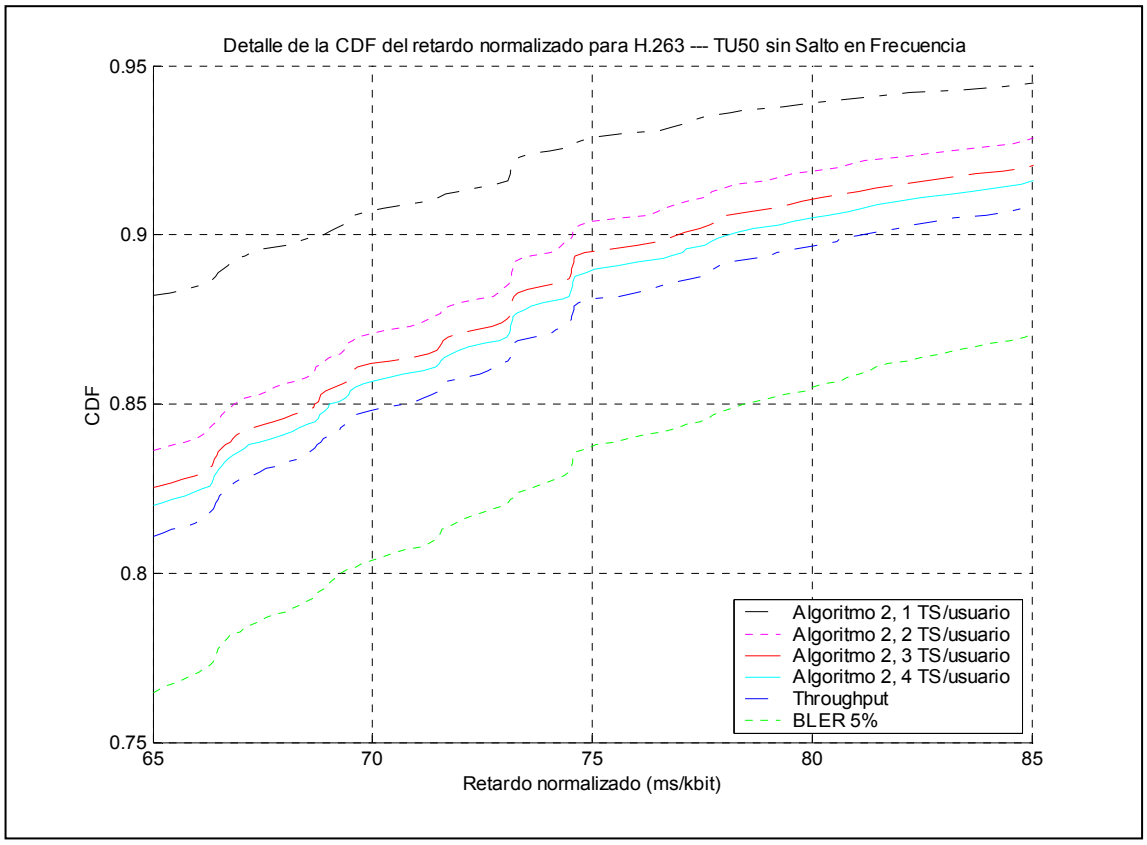


Figura 7.20.- Detalle de la CDF del retardo normalizado para el segundo algoritmo propuesto en vídeo

Al analizar los valores de retardo normalizado obtenidos con el segundo de los algoritmos propuestos (tabla 7.16) se aprecia que los resultados ofrecidos mejoran a los de los dos algoritmos de referencia (un 14% al de throughput y un 22.6% al de BLER del 5% en el 95% de los casos), consecuencia directa de un mayor uso de codificadores capaces de transmitir una mayor cantidad de bits por unidad de tiempo o, de forma equivalente, una misma cantidad de datos en un periodo de tiempo menor. Puesto que el porcentaje de uso de estos codificadores aumenta al reducir el número de time-slots por usuario con el que se diseña el algoritmo, la reducción del retardo normalizado también tiene lugar en este sentido, según se ilustra en las figuras 7.19 y 7.20.

Además, la diferencia significativa en el uso de CS-4 del algoritmo diseñado con un time-slot por usuario respecto a los otros diseños del algoritmo y al algoritmo basado en throughput (tabla 7.15), explica que éste sea capaz de transmitir casi un 5% más de tramas de vídeo con el retardo normalizado mínimo posible que cualquiera de los otros algoritmos (figura 7.19), retardo normalizado que se experimenta en aquellas tramas de vídeo cuya transmisión se realiza empleando únicamente el esquema de codificación CS-4. Para los valores observados en la figura 7.19, este porcentaje se traduce en una mejora relativa del 11%.

Este descenso del retardo normalizado global es la causa de que el porcentaje total de tramas de vídeo que se transmiten sin retardo aumente al aplicar este segundo algoritmo propuesto, tal y como muestra la tabla 7.18. Se observa que este algoritmo es capaz de llegar a transmitir hasta un 4.8% más de tramas de vídeo dentro del plazo de tiempo preestablecido que el algoritmo basado en throughput, lo cual supone una mejora relativa del 7%, bastante apreciable. En comparación con otros algoritmos más conservadores, como el basado en un BLER del 5%, el valor de esta diferencia es aun más importante (del 12%).

El precio a pagar por las mejoras obtenidas en los resultados de retardo es una peor calidad de las tramas de vídeo que se reciben sin retardo, compromiso observado en los análisis anteriores y que se vuelve a observar una vez más en la tabla 7.19. En este parámetro, el algoritmo basado en BLER es el que demuestra ser capaz de ofrecer los mejores resultados. No obstante, conviene recordar las observaciones realizadas en el apartado 7.2 en lo referente a cómo repercute este hecho sobre la calidad final de servicio experimentada por el usuario.

Algoritmo	**Porcentaje de tramas de vídeo transmitidas**	
	Sin retardo	**Con retardo**
Algoritmo 2, 1 TS/u	73.92 %	26.08 %
Algoritmo 2, 2 TS/u	70.67 %	29.33 %
Algoritmo 2, 3 TS/u	70.06 %	29.94 %
Algoritmo 2, 4 TS/u	69.72 %	30.28 %
Throughput	69.16 %	30.84 %
BLER 5%	65.97 %	34.03 %

Tabla 7.18.- Porcentaje de tramas de vídeo transmitidas con y sin retardo para el segundo algoritmo propuesto

Requisito	Alg. 2, 1 TS/u	Alg. 2, 2 TS/u	Alg. 2, 3 TS/u	Alg. 2, 4 TS/u	Throughput	BLER 5%
BLER ≤ 1%	62.53 %	63.89 %	64.40 %	64.42 %	64.67 %	65.81 %
BLER ≤ 2%	62.55 %	63.91 %	64.41 %	64.44 %	64.70 %	65.84 %
BLER ≤ 5%	62.48 %	64.25 %	64.76 %	64.78 %	65.04 %	66.19 %
BER ≤ 10^{-3}	65.07 %	65.82 %	66.25 %	66.21 %	66.41 %	67.34 %
BER ≤ 10^{-4}	57.01 %	57.86 %	58.26 %	58.23 %	58.41 %	59.27 %
BER ≤ 10^{-5}	55.50 %	56.33 %	56.72 %	56.69 %	56.87 %	57.73 %
BER ≤ 10^{-6}	55.36 %	56.18 %	56.57 %	56.55 %	56.73 %	57.58 %

Tabla 7.19.- De las tramas de vídeo transmitidas sin retardo, porcentaje que cumple un determinado requisito de calidad

Requisito	Alg. 2, 1 TS/u	Alg. 2, 2 TS/u	Alg. 2, 3 TS/u	Alg. 2, 4 TS/u	Throughput	BLER 5%
BLER ≤ 1%	46.22 %	45.15 %	45.11 %	44.91 %	44.73 %	43.41 %
BLER ≤ 2%	46.24 %	45.17 %	45.13 %	44.93 %	44.75 %	43.43 %
BLER ≤ 5%	46.45 %	45.40 %	45.37 %	45.16 %	44.98 %	43.66 %
BER ≤ 10^{-3}	48.10 %	46.52 %	46.41 %	46.16 %	45.93 %	44.42 %
BER ≤ 10^{-4}	42.14 %	40.89 %	40.81 %	40.60 %	40.40 %	39.10 %
BER ≤ 10^{-5}	41.02 %	39.81 %	39.74 %	39.52 %	39.33 %	38.08 %
BER ≤ 10^{-6}	40.92 %	39.70 %	39.63 %	39.42 %	39.23 %	37.99 %

Tabla 7.20.- Porcentaje de tramas de vídeo transmitidas sin retardo y cumple un determinado requisito de calidad

A pesar de ello, la mejora de las prestaciones de retardo ofrecidas por parte de este algoritmo se traducen en un mayor porcentaje de tramas de vídeo que satisfacen los diversos requisitos de calidad considerados y, por tanto, en una mayor proporción de tramas capaces de contribuir a la mejora de la calidad final experimentada por el usuario (tabla 7.20), siendo la mejora entorno al 3% respecto al algoritmo basado en un BLER del 5% para los diversos parámetros de la tabla 7.20.

Todas estas observaciones demuestran que el segundo de los algoritmos que se proponen en este trabajo es capaz de ofrecer unas prestaciones de retardo claramente superiores a las de los dos algoritmos tradicionales tomados como referencia, hecho que pone de manifiesto su eficacia en la consecución del objetivo inicial y su capacidad para contribuir a la mejora de la calidad de servicio en una aplicación en tiempo real.

Antes de concluir el presente apartado es necesario comentar algunos aspectos relativos al número de time-slots por usuario utilizado para las simulaciones y al que se estimó en el capítulo sexto como conveniente para realizar la transmisión de vídeo, así como la forma en que esto repercute sobre los resultados obtenidos. En primer lugar, los beneficios logrados por este algoritmo son consecuencia del alto grado de utilización de los esquemas de codificación con mayor velocidad binaria (y del acierto de su elección), hecho que según se comentó al principio de este apartado se debe a que el número de time-slots por usuario considerado al diseñar el algoritmo (uno) es inferior al requerido por la secuencia de vídeo para una transmisión sin retardo de un porcentaje de tramas de vídeo suficientemente elevado (estimado en tres en el capítulo anterior). Este hecho implica que si el algoritmo se diseña considerando tres time-slots por usuario, entonces el porcentaje de utilización de CS-4 se reduciría, con lo que el grado de mejora ya no sería tan significativo como el apreciado. No obstante, no ha de ser necesariamente así. La comparación adecuada con el algoritmo que busca la maximización del throughput de este diseño que se realiza tomando tres time-slots por usuario, sería aquella en la que ambos algoritmos fueran simulados sobre un sistema en el que cada usuario realmente dispone de tres time-slots. Si el algoritmo diseñado y simulado con un time-slot por usuario ofrece mejores resultados que el algoritmo basado en throughput (en throughput y en estadísticas de retardo), y lo sigue haciendo incluso el diseño de tres time-slots por usuario sobre un sistema en el que cada usuario tan sólo dispone de uno, al comparar el algoritmo diseñado y simulado con tres time-slots por usuario cabría esperar una mejora aun más acentuada respecto al algoritmo basado en throughput (empleando también tres time-slots por usuario), que podría llegar a traducirse en una reducción del número de time-slots por usuario necesarios para ofrecer unas prestaciones de retardo similares.

7.5.- SUMARIO

En este capítulo se han presentado los resultados numéricos de relevancia obtenidos al simular los diversos algoritmos estudiados en el presente trabajo y se han analizado de forma comparativa, determinando el grado de mejora o degradación que sobre cada parámetro de interés produce cada uno de los algoritmos estudiados. Tras una comparación del algoritmo basado en throughput y el algoritmo basado en BLER, algoritmos habituales en la literatura cuyos resultados se han tomado como referencia, se han presentado los resultados obtenidos para los dos algoritmos propuestos en el presente trabajo y que constituyen la principal aportación. Son varias las conclusiones de interés obtenidas del análisis realizado sobre dichos resultados, las cuales se recogen en el siguiente y último capítulo del presente trabajo, en el que se sugieren, además, diversas investigaciones como posible vía de continuación de las aquí realizadas.

Capítulo 8

Conclusiones y Trabajo Futuro

Este octavo capítulo con que se da por concluido el presente trabajo, dedicado a estudiar la aplicación de la técnica de gestión de recursos radio *Link Adaptation* sobre la transmisión de datos sensibles al retardo, mostrará las observaciones y conclusiones más reseñables derivadas del análisis de los algoritmos que aquí se proponen, principal aportación de este trabajo, y de la comparación realizada con otros dos algoritmos ya tradicionales en la literatura.

El primero de los algoritmos tomados como referencia (uno de los inicialmente utilizados con *Link Adaptation*) es aquel cuyo criterio de selección se basa en escoger en todo momento aquel esquema de codificación de canal capaz de proporcionar la mayor cantidad de bits correctamente transmitidos por unidad de tiempo (throughput). El segundo de ellos, más conservador, pretende que el codificador se escoja de forma tal que el grado de protección contra errores resultante permita mantener la tasa de error global del sistema por debajo de un determinado límite, el cual se establece a partir de la máxima cantidad de errores tolerable por la aplicación concreta. En el caso que aquí se considera, dicha tasa de errores se ha expresado en términos de la proporción de bloques RLC recibidos con error o BLER (*BLock Error Rate*). La comparación de los resultados ofrecidos por ambos algoritmos muestra que el primero de ellos alcanza el mayor throughput global y la peor tasa de errores, aspecto en el cual se ve superado por el segundo. No obstante, existe un compromiso entre throughput y BLER a nivel de sistema, pues el menor BLER conseguido por este último se ve acompañado por una reducción del throughput respecto al primero.

En lo que respecta al comportamiento en cuanto a retardo, aspecto de especial relevancia para el tipo de servicio aquí considerado, se concluye que el criterio de maximización del throughput es el que muestra los mejores resultados de los dos: obtiene el menor retardo normalizado (hasta un 17% mejor) y el mayor porcentaje de tramas transmitidas sin retardo (con una mejora relativa del 11%). La aportación del algoritmo basado en BLER en este aspecto consiste en una mejor calidad de las tramas transmitidas sin retardo, las cuales llegan a satisfacer diferentes objetivos de calidad en un mayor porcentaje de ocasiones que las tramas transmitidas sin retardo empleando el algoritmo basado en throughput, aun constituyendo éstas últimas un porcentaje más alto. Se observa, por tanto, un compromiso entre cantidad de tramas sin retardo y calidad de las tramas transmitidas sin retardo.

Se ha señalado la posibilidad de que esta mejora alcanzada por el algoritmo basado en BLER, en lo que a calidad se refiere, no se traduzca finalmente en una mejor calidad de servicio final para el usuario, al menos bajo las condiciones de simulación en las que se ha realizado el estudio, pues el porcentaje de tramas con retardo al aplicar este algoritmo, más conservador, no sólo es mayor sino que, además, la cantidad de información que en ellas se pierde también lo es respecto al algoritmo basado en throughput. Se planteó la posibilidad de que la aplicación de este algoritmo a servicios en tiempo real no mejore la calidad de servicio final que experimenta el usuario (que es lo que en definitiva se persigue con este algoritmo) sino que, por el contrario, origine un deterioro por el excesivo retardo en que se incurre, el cual no se ve compensado por la mejor calidad de las tramas sin retardo, con lo que la cantidad final de información correcta en el receptor resulta ser menor y, por tanto, la calidad de servicio final, peor. No obstante, este hecho no implica que el algoritmo no pueda ser aplicado a otro tipo de servicios de tipo *streaming*, con requisitos de retardo menos exigentes, en los cuales sí puede lograr una mejora de la calidad [104].

Al analizar los resultados ofrecidos por el algoritmo basado en BLER, se apreció que era incapaz de alcanzar la tasa de errores que se proponía como objetivo porque en su diseño no se tenía en cuenta el hecho de que la probabilidad de utilización del esquema de codificación CS-4 resultaba ser muy elevada, incluso en aquellos casos en que la calidad del canal se podía considerar más bien pobre. Esta elevada utilización de CS-4 era la causante de que el algoritmo no lograra su objetivo, pues se demostró que forzando un uso más frecuente de codificadores con mayor grado de robustez, el BLER obtenido disminuía de manera significativa. Por lo tanto, la definición del algoritmo que realmente logre el objetivo ha de implementar un mecanismo de control que limite de forma deliberada el uso de los codificadores menos robustos. La implementación de este mecanismo se hace necesaria porque la modificación del mapa de decisión del algoritmo no resulta suficiente por sí sola, pues es incapaz de paliar el uso excesivo del esquema de codificación CS-4.

Una posible implementación de este mecanismo de control consistiría en fijar un margen de histéresis $\pm\Delta$ entorno al valor de BLER buscado, del 5% por ejemplo. El procedimiento consistiría en realizar medidas, en tiempo real, del BLER obtenido en el receptor. Si en algún momento el BLER alcanza o supera el valor $5\% + \Delta$, entonces se inhibiría el empleo de CS-4 hasta que las mediciones del BLER ofrecieran un valor igual a $5\% - \Delta$ o inferior, momento a partir del cual se volvería a permitir el empleo de CS-4. Durante este periodo de tiempo en el que el empleo de CS-4 no está permitido, en aquellas ocasiones en las que el mapa de decisión indicara la utilización de CS-4 se procedería a utilizar, en su lugar, CS-3. Si el empleo de CS-3 resultara insuficiente para obtener el BLER buscado, se inhibiría también el uso de CS-3 e incluso el de CS-2 si fuera preciso, situación que podría llegar a darse en aquellos casos en los que la calidad del canal empeorara de forma considerable. Una vez alcanzado el margen $5\% - \Delta$ o un valor inferior, se volvería a permitir el uso de todos los esquemas de codificación. El método, por tanto, consistiría en restringir progresivamente el uso de los codificadores menos robustos en función de la calidad real que se experimenta en el enlace radio.

Este mecanismo debería ser capaz, en principio, de mantener estable el valor del BLER alrededor del margen deseado. No obstante, de esta forma no se define por completo el algoritmo *Link Adaptation* que debería ser realmente capaz de conseguir el BLER marcado, ya que sería necesario analizar en este caso el sentido que tendría la aplicación del mapa de decisión correspondiente al BLER que se busca. En efecto, si el BLER deseado se logra por la aplicación de este mecanismo de control, sería indiferente la utilización de un mapa de decisión u otro. Una posibilidad sería aplicar, de forma conjunta, el mecanismo de control que pretende preservar un determinado BLER y el mapa de decisión correspondiente a ese mismo BLER. Sin embargo, sería interesante analizar la aplicación de este mecanismo junto con el mapa de decisión del algoritmo basado en throughput ya que, en principio, el mecanismo de control se encargaría de mantener estable el BLER en el valor deseado mientras que el mapa de decisión sería el responsable de lograr el máximo throughput posible para el objetivo de BLER buscado. La implementación de estas ideas y su análisis completo así como el de otras posibles soluciones alternativas requiere un análisis más detenido que se deja propuesto como trabajo futuro.

A la luz de los resultados relativos al primero de los algoritmos propuestos en el presente trabajo, es posible afirmar que no sólo consigue establecer un compromiso entre el throughput global y el BLER global a nivel de sistema, objetivo inicial para el cual fue diseñado sino que, además, permite que el grado de compromiso establecido entre ambos factores pueda ser modificado de forma sencilla (con tan sólo cambiar el valor de un parámetro de diseño) en función de cuál de los dos se considere de mayor relevancia respecto al otro. Debido a que a nivel de sistema existe un compromiso no sólo entre el throughput y el BLER, sino también entre éste y el retardo de transmisión, la variación del compromiso establecido entre los dos primeros implica también, como pudo observarse, la del compromiso mantenido entre tasa de errores y retardo global a nivel de sistema.

Cuando la variación del parámetro de diseño mencionado suponía la obtención de un algoritmo más conservador en cuanto a tasa de errores global del sistema, se observaba una disminución del throughput y un deterioro del retardo de transmisión, el cual desembocaba en un aumento de la cantidad de tramas transmitidas con retardo y en una mayor cantidad de información descartada en el receptor. El beneficio que se obtiene a cambio es una mejora en la calidad de las tramas transmitidas sin retardo. La variación de este parámetro en sentido inverso originaba el comportamiento opuesto. El algoritmo supone, por tanto, una forma sencilla de ajustar el grado de compromiso que existe entre los diversos parámetros de un sistema de comunicaciones móviles.

Este hecho podría ser aprovechado para conseguir un algoritmo dinámico en el cual el valor del parámetro k variara en tiempo real en función de algún parámetro relevante del sistema como, por ejemplo, la carga de usuarios o el retardo global. Este principio de funcionamiento habilitaría un medio sencillo para lograr una adaptación dinámica a las condiciones particulares del sistema en cada momento.

A pesar de la consecución del objetivo marcado, este algoritmo no es más que el principio que sirve de motivación para la búsqueda de otros algoritmos similares, los cuales pueden ser fácilmente derivados a partir del planteamiento inicial. Según se vio en el diseño de este algoritmo, existía una función matemática que se aplicaba sobre el BLER a fin de conseguir el comportamiento deseado. La empleada para el algoritmo mostró ser suficiente para la consecución del objetivo marcado. No obstante, no existe motivo alguno para pensar que es la que ha de ofrecer el mejor compromiso. Cabría la posibilidad de que aplicando otras funciones diferentes sobre el BLER en el diseño del algoritmo se obtuvieran mejores prestaciones de throughput para una misma tasa de errores o viceversa, razón por la cual la búsqueda de otras funciones y el estudio de los resultados ofrecidos con su aplicación constituye una posible vía de continuación de las investigaciones llevadas a cabo en este aspecto.

Frente a los algoritmos basados en throughput y BLER, cuyos criterios estaban basados, respectivamente, en la mejora de la cantidad de bits correctos transmitidos por unidad de tiempo y de la tasa de bloques RLC recibidos con error, el segundo de los algoritmos que se proponen en el presente trabajo establece un criterio de selección del esquema de codificación de canal en el que interviene, no sólo el throughput, sino también el retardo (reflejado en el concepto de retardo de trama, empleado en el diseño del algoritmo), el cual constituye un factor clave para el tipo de servicio considerado en el presente trabajo.

A diferencia de estos dos algoritmos, que mostraban una clara predilección por el empleo de codificadores más robustos o con mayor velocidad binaria en función del objetivo perseguido, este segundo algoritmo no es propenso a fomentar una utilización más frecuente de un determinado grupo de esquemas de codificación de canal u otro.

Respecto a los algoritmos tomados como referencia, se distinguía por mostrar un comportamiento dinámico en el cual el empleo más frecuente de unos codificadores u otros no era permanente sino que venía determinado por el tamaño y la duración de la trama de vídeo a transmitir en cada situación o, más exactamente, por la tasa binaria que requería para su transmisión sin retardo así como por la cantidad de recursos por usuario de los que se disponía. Esta capacidad de adaptación en tiempo de ejecución a las características particulares del flujo de información en cada momento, rasgo que no caracterizaba a ninguno de los otros algoritmos, sea probablemente la causa de que los resultados ofrecidos sean superiores en aquellos aspectos que se pretendían mejorar.

Por una parte, el throughput que proporciona este algoritmo supera el ofrecido por el propio algoritmo cuyo objetivo consiste en maximizar el throughput. La mejora llega a ser del 8.3% en el 95% de los casos, la cual se debe a que el algoritmo propuesto selecciona el codificador de canal óptimo en un mayor porcentaje de ocasiones que el algoritmo basado en throughput. De forma paralela, esta mejora del throughput se ve acompañada por una mejoría apreciable de las estadísticas de retardo. Respecto al algoritmo basado en throughput, el retardo normalizado global del sistema desciende (un 14% respecto al algoritmo basado en throughput y un 22.6% respecto al basado en un BLER del 5%), se incrementa en un 4.8% el porcentaje de tramas transmitidas sin retardo (mejora relativa del 7%) y en un 5% el de las que se transmiten con el mínimo retardo normalizado posible (mejora relativa del 11%). Incluso cuando el algoritmo se diseña sobreestimando los recursos radio disponibles, los resultados ofrecidos siguen siendo (ligeramente) superiores.

Todas estas observaciones demuestran que el segundo de los algoritmos que se proponen en este trabajo es capaz de ofrecer unas prestaciones de retardo claramente superiores a las de los dos algoritmos tradicionales tomados como referencia, hecho que pone de manifiesto su eficacia en la consecución del objetivo inicial y su capacidad para contribuir a la mejora de la calidad de servicio en una aplicación en tiempo real. Estos resultados llevan a plantear la posibilidad de que la búsqueda de una solución dinámica, como la que se propone con el segundo algoritmo, constituya la elección más adecuada para responder a los exigentes requisitos de retardo que caracterizan a los servicios de tiempo real, lo cual sirve de motivación para la investigación de otras alternativas capaces de adaptar su comportamiento al del flujo de datos transmitido. En este sentido, *Link Adaptation* todavía tiene mucho que decir.

Anexo A

Codificación de vídeo

A.1.- OBJETIVO DEL ANEXO

En el capítulo 4 se trató superficialmente la codificación de vídeo H.263, dando una explicación de los tipos de imágenes que se generan en el proceso de codificación. La finalidad de este anexo es ofrecer una descripción general y sencilla del proceso de codificación H.263 para mostrar la forma en que se codifica cada tipo de imagen.

A.2.- CODIFICACIÓN DE VÍDEO H.263

Un codificador de vídeo H.263 tiene dos modos básicos de funcionamiento [71]: modo *Intra* y modo *Inter*. Las imágenes codificadas en el primer modo se denominan imágenes I, mientras que aquellas codificadas en el segundo modo se denominan imágenes P [71]. Las imágenes de tipo I son aquellas en las que se codifica toda la información de vídeo contenida en ellas, de forma que pueden ser reconstruidas por completo de forma independiente; se dice en este caso que la imagen original se codifica sin referencia a ninguna otra. Por el contrario, las imágenes de tipo P son aquellas que se codifican haciendo referencia a la imagen anterior. Cuando el codificador funciona en modo *Inter*, tan sólo se codifica la diferencia entre la imagen actual y su predecesora. Las zonas de la imagen que no varían respecto a la imagen anterior no son codificadas, ya que el receptor posee esa información.

La recomendación ITU-T H.263 contempla diversos formatos de imagen. Todos ellos se definen como múltiplos o submúltiplos de un formato intermedio común o CIF (*Common Intermediate Format*), cuyo tamaño es de 352 píxeles (sentido horizontal) por 288 líneas (sentido vertical). La tabla A.1 muestra estos formatos de imagen.

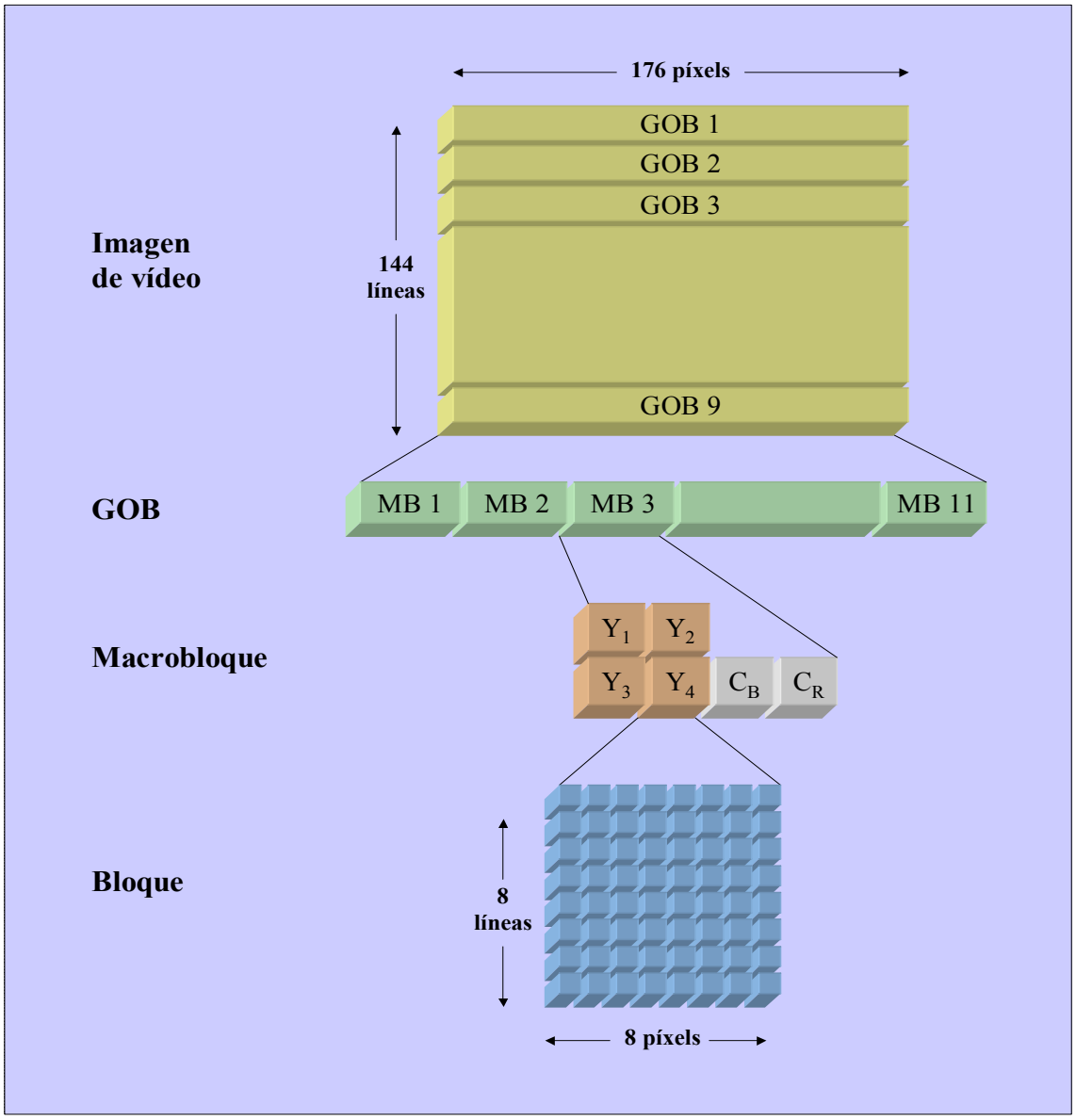
Formato de imagen	Número de píxeles (d_x)	Número de líneas (d_y)
Sub-QCIF	128	96
QCIF	176	144
CIF	352	288
4CIF	704	576
16CIF	1408	1152

Sub-QCIF: Sub-cuarto de CIF **QCIF:** Cuarto de CIF
4CIF: Cuatro veces CIF **16CIF:** 16 veces CIF

Tabla A.1.- Formatos de imagen normalizados de la recomendación H.263 [71]

Una imagen de vídeo se divide en varias entidades de menores dimensiones. La figura A.1 muestra, a modo de ejemplo, la estructura para el formato de imagen más ampliamente utilizado [78], el formato QCIF. La imagen se divide en varios grupos de bloques (GOB, *Group of Blocks*), cada uno de los cuales está compuesto, a su vez, por varios macrobloques (MB). El número de GOBs y de MBs en cada GOB vendrá dado por el formato de imagen. Se puede observar en la tabla A.1 que tanto el número de píxeles como de líneas de cada formato es múltiplo entero de 16. Ello se debe a que un MB siempre tiene, para todos los formatos de imagen, un tamaño de 16 píxeles por 16 líneas. Cada uno de los píxeles que conforman una imagen queda perfectamente definido por la combinación de tres colores básicos aditivos: el rojo (R, *Red*), el verde (G, *Green*) y el azul (B, *Blue*). Con las componentes RGB se obtiene la información de

luminancia (Y) y crominancia (C_B, C_R) para cada píxel[64]. Habitualmente, la información de luminancia se muestrea con la misma resolución que la imagen, mientras que la información de crominancia se muestrea a la mitad de resolución [69]. Por tanto, para cada macrobloque (16 píxeles por 16 líneas) se obtendrá una matriz Y de 16 × 16 valores, una matriz C_B de 8 × 8 valores y una matriz C_R también de 8 × 8 valores. Cada matriz de 8 × 8 valores recibe el nombre de bloque. La codificación H.263 se realiza a nivel de bloque, es decir, cada imagen se codifica bloque a bloque empleando para ello un codificador como el de la figura A.2 donde se muestra, de forma muy simplificada, la estructura general de un codificador de vídeo H.263 básico.

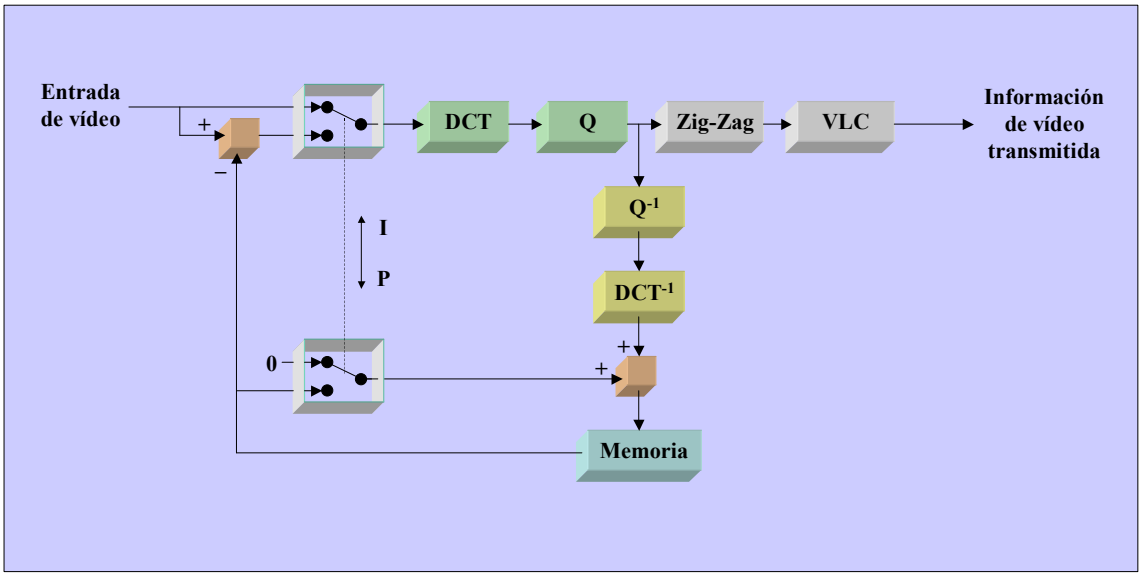


MB: Macrobloque **GOB:** *Group of Blocks* (Grupo de Bloques)
Y: Matriz de luminancia **C_B, C_R:** Matrices de crominancia

Figura A.1.- Estructura del formato de imagen QCIF [74]

[64] La relación es la siguiente [79]:

$$\begin{bmatrix} Y \\ C_B \\ C_R \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & 0.500 \\ 0.500 & -0.419 & -0.081 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$



DCT: *Discrete Cosine Transform* (Transformada de Coseno Discreto) **DCT^{-1}:** Transformada DCT inversa
Q: Cuantificador **Q^{-1}:** Proceso inverso a la cuantificación Q
VLC: *Variable Length Code* (Código de Longitud Variable)

Figura A.2.- Diagrama de bloques simplificado de un codificador H.263 [71]

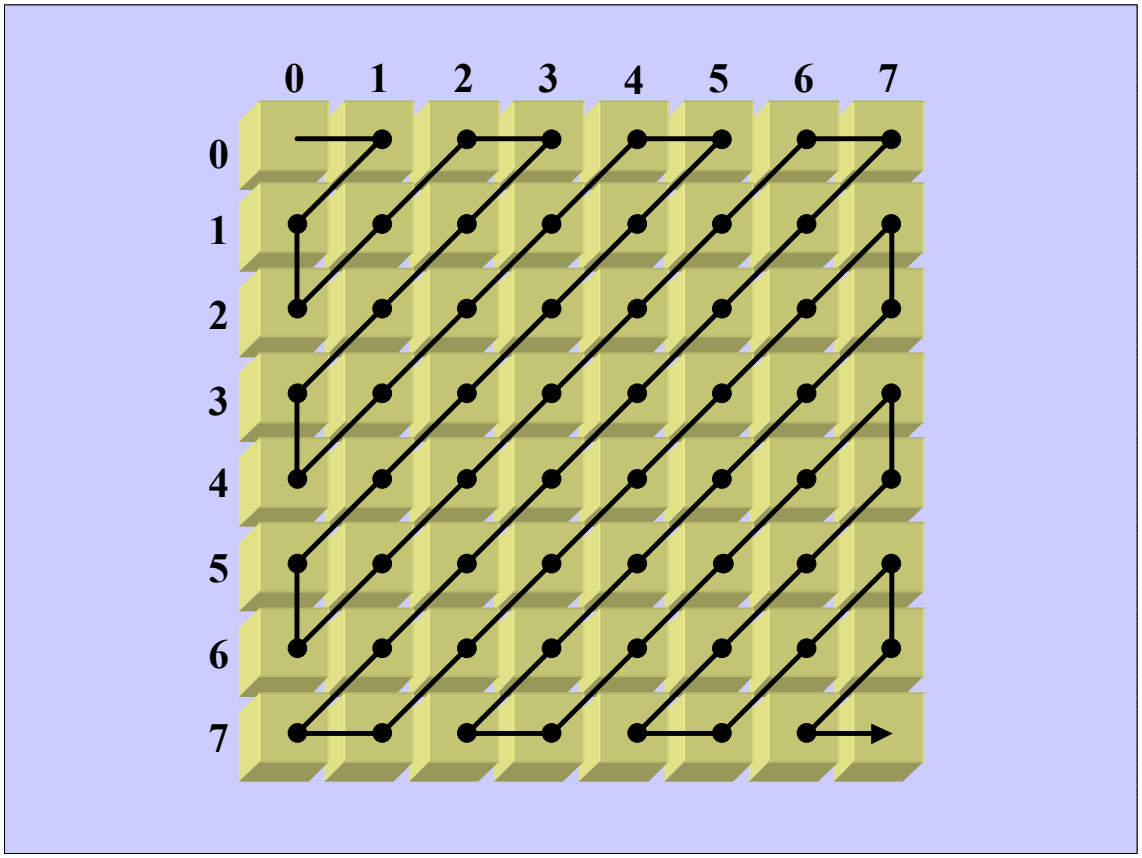


Figura A.3.- Reordenación en zig-zag de los coeficientes de las matrices 8 × 8 [72, 74]

Sobre cada bloque 8 × 8 se aplica una transformada de coseno discreto (DCT, *Discrete Cosine Transform*). Esta transformada es preferible a otro tipo de transformadas (como, por ejemplo, la transformada de Fourier) por proporcionar una mayor cantidad de coeficientes prácticamente iguales a cero, lo que permite representar las matrices 8 × 8 con una menor cantidad de coeficientes [69]. Tras aplicar la transformada DCT, los valores de las matrices 8 × 8 son cuantificados[65], lo que hace que los coeficientes prácticamente nulos se cuantifiquen como cero. Estas matrices 8 × 8 se transforman en matrices 64 × 1 siguiendo el orden de zig-zag mostrado en la figura A.3. Lo que se consigue con ello es que el primer coeficiente corresponda al que representa el nivel de continua del bloque 8 × 8 y el resto de coeficientes queden ordenados de menor a mayor frecuencia. Aquellos coeficientes cuantificados como cero serán (idealmente) consecutivos [74]. Para conseguir un mayor nivel de compresión, estos coeficientes son codificados mediante códigos de longitud variable (VLC, *Variable Length Code*) antes de ser transmitidos.

En la figura A.2 se puede apreciar que los procesos de cuantificación y de transformada DCT se invierten. Si la imagen fue codificada en modo *Intra* (imagen I), entonces es fácil observar que el proceso inverso devuelve la imagen "original", que queda almacenada en la memoria interna del codificador. Si la siguiente imagen se codifica en modo *Inter* (imagen P) entonces, observando la figura A.2, se ve que sólo se codifican y transmiten aquellos macrobloques que difieren respecto a la imagen anterior. Al invertir el proceso de cuantificación y transformada DCT se obtiene la diferencia codificada y transmitida. Incorporando esta diferencia a la memoria de imagen (que contenía la imagen anterior) se obtiene, como resultado, que la memoria pasa a almacenar la imagen actual completa, quedando disponible para ser utilizada como referencia para la siguiente imagen P.

Aunque en la figura A.2 no se muestre por simplicidad, existe una técnica importante denominada *estimación y compensación del movimiento* que se aplica cuando el codificador funciona en modo *Inter* y que permite conseguir mayores niveles de compresión. Una vez determinados los macrobloques que varían respecto a la imagen anterior, se comparan con la imagen almacenada en la memoria con la intención de encontrar coincidencias. Si se encuentra que un macrobloque a codificar coincide con una región de 16 píxeles por 16 líneas de la imagen anteriormente codificada, entonces dicho macrobloque no se codifica. En su lugar se calcula un vector de movimiento (determinado por dos componentes) que indica el desplazamiento de dicho macrobloque respecto a la imagen anterior. Estos vectores de movimiento se codifican aplicando también los códigos VLC antes de ser transmitidos hacia el receptor. La codificación de las dos componentes del vector de movimiento supone un ahorro significativo respecto a la codificación de los valores de las seis matrices 8×8 asociadas a un macrobloque, consiguiendo un nivel de compresión mucho mayor.

Antes de concluir el presente anexo, conviene aclarar que el procedimiento de codificación descrito no es particular y exclusivo del método de codificación H.263 sino que, por el contrario, corresponde a la línea general (salvo las particularidades de cada caso) seguida por la mayoría de técnicas de codificación de vídeo actuales [69].

[65] En el apartado 4.2.2 se dijo que la codificación H.263 es irreversible o con pérdidas, ya que la imagen decodificada no es idéntica, píxel a píxel, a la original. La cuantificación es la principal responsable de ello [69].

Anexo B

Número Necesario de Canales por Usuario

B.1.- OBJETIVO DEL ANEXO

La finalidad de este anexo es demostrar, de una forma más detallada a como se realizó en el apartado 5.3.4 del capítulo 5, que la transmisión en condiciones aceptables de la secuencia de vídeo empleada en las simulaciones requiere el uso de más de un canal o time-slot por usuario. Hay que tener presente que el número de canales que realmente requiere un usuario para poder recibir la secuencia de vídeo con una calidad aceptable depende de muchos factores como, por ejemplo, la eficacia de las técnicas de gestión de recursos radio utilizadas. Algunas técnicas ofrecen, con la misma cantidad de recursos por usuario, un mejor rendimiento que otras, lo cual es equivalente a decir pueden obtener las mismas prestaciones con una cantidad de recursos menor. Por este motivo, los resultados aquí mostrados deben ser considerados como una indicación y no como una conclusión acerca del número de time-slot que realmente precisa cada usuario para que la calidad de la transmisión de vídeo sea aceptable. No obstante, sí son suficientes para demostrar que empleando un único time-slot por usuario no es viable una transmisión de vídeo en condiciones satisfactorias.

B.2.- NÚMERO DE CANALES POR USUARIO

Cada trama de vídeo generada por el modelo de vídeo tiene un tamaño concreto así como una determinada duración o tiempo al cabo del cual se genera una nueva trama. Para que una trama de vídeo se transmita sin retardo es necesario que la velocidad binaria de transmisión sea igual o mayor a la velocidad binaria requerida por dicha trama, que no es más que el cociente entre su tamaño y su duración. De esta forma, cada trama de vídeo llevará asociado un valor de tasa binaria requerida. La figura B.1 muestra la CDF de la tasa binaria requerida por la secuencia de vídeo que se ha empleado para parametrizar el modelo de vídeo descrito en el capítulo quinto, la cual ha sido codificada para que su velocidad media sea de 16 kbps.

Suponiendo, por ejemplo, que se utiliza CS-2 (13.4 kbps/time-slot) como esquema de codificación de canal durante toda la transmisión, entonces en dicha figura se puede apreciar que el 100% – 85.58% = 14.42% y el 100% – 98.65% = 1.35% de las tramas de vídeo requiere más de dos time-slots (26.8 kbps) y tres time-slots (40.2 kbps) por usuario, respectivamente, para poder ser transmitidas sin retardo. Es decir, para poder transmitir la secuencia de vídeo (98.65% de las tramas) sin retardo con CS-2 se requieren tres time-slots por usuario.

Extendiendo el análisis de la secuencia de vídeo para los cuatro esquemas de codificación se obtienen los resultados mostrados en la tabla B.1, donde se detalla el porcentaje de tramas que se podría transmitir sin retardo (antes de que se genere la siguiente) según el esquema de codificación empleado (supuestamente utilizado durante la transmisión completa de la secuencia) y el número de canales por usuario.

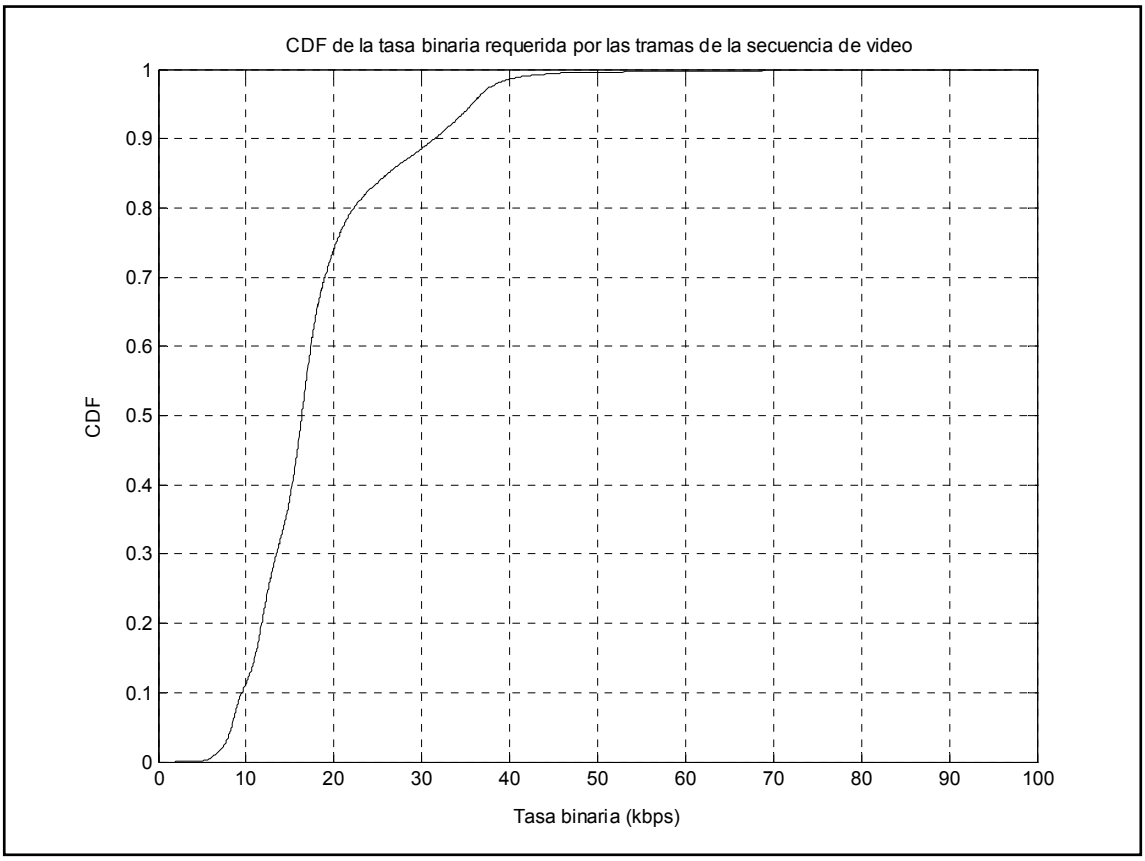


Figura B.1.- CDF de la tasa binaria requerida por las tramas de la secuencia de vídeo

Esquema de codificación	Número de time-slots	Tasa binaria disponible (kbps)	Porcentaje de tramas sin retardo
CS-1 (9.05 kbps/time-slot)	1	9.05	8.26%
	2	18.10	65.30%
	3	27.15	85.20%
	4	36.20	95.69%
CS-2 (13.4 kbps/time-slot)	1	13.4	29.67%
	2	26.8	85.58%
	3	40.2	98.65%
	4	53.6	99.66%
CS-3 (15.6 kbps/time-slot)	1	15.6	42.49%
	2	31.2	89.73%
	3	46.8	99.53%
	4	62.4	99.72%
CS-4 (21.4 kbps/time-slot)	1	21.4	78.19%
	2	42.8	99.17%
	3	64.2	99.75%
	4	85.6	100.0%

Tabla B.1.- Número de time-slots requeridos por usuario

Si establecemos como criterio que la calidad de la transmisión de vídeo será aceptable si como mínimo el 95% de las tramas se recibe sin retardo, entonces se requerirán 4, 3, 3 ó 2 time-slots por usuario según se emplee CS-1, CS-2, CS-3 o CS-4, respectivamente, para transmitir la secuencia de vídeo. Esto significa que en el mejor de los casos (que correspondería a una calidad del canal tal que permitiese utilizar CS-4 durante toda la transmisión) se requeriría un mínimo de dos canales por usuario para transmitir sin retardo un porcentaje de tramas aceptable ya que, suponiendo posible la utilización continua de CS-4, la asignación de un canal a cada usuario permitiría la transmisión sin retardo de menos del 80% de las tramas.

Por tanto, a partir del análisis que se acaba de realizar en este anexo es posible afirmar la necesidad de más de un time-slot por usuario para transmitir la secuencia de vídeo con una calidad aceptable.

Anexo C

Demostración

C.1.- OBJETIVO DEL ANEXO

La finalidad de este anexo es demostrar que para aquellas tramas de vídeo que, teniendo un tamaño y una duración diferentes pero requiriendo una misma velocidad binaria para su transmisión sin retardo, los umbrales de decisión obtenidos a partir de las expresiones (6.6) y (6.7) del capítulo sexto son los mismos, aun cuando el retardo de trama que estas tramas proporcionan según la expresión (6.5) sea distinto.

C.2.- DEMOSTRACIÓN

En el capítulo sexto se definió el retardo de trama según la expresión

$$Retardo\ de\ trama = \frac{T}{R_{CS-i}} - D \qquad \text{(C.1)}$$

donde T representa el tamaño de la trama de vídeo, D su duración o tiempo que transcurre hasta que se genera la siguiente trama de vídeo y $R_{CS\text{-}i}$ la velocidad binaria del esquema de codificación CS-*i*.

Para los fines de este anexo, se aplicará la expresión (C.1) a dos conjuntos diferentes de tramas de vídeo. El primero, mostrado en la tabla C.1, agrupa varias tramas con diferentes tamaños T y duraciones D pero que tienen en común una misma tasa binaria T/D requerida para su transmisión sin retardo de 5 kbps. Esta velocidad binaria corresponde a una de las más bajas observadas en la secuencia de vídeo empleada en las simulaciones y se ha escogido porque la transmisión de una trama con esta velocidad binaria empleando cualquier esquema de codificación (incluso CS-1) da lugar a un retardo de trama negativo según la expresión (C.1). El segundo grupo, mostrado en la tabla C.2, también agrupa varias tramas con diferentes tamaños y duraciones pero con una velocidad binaria común de 30 kbps. Esta otra tasa binaria corresponde a una de las más altas observadas en la secuencia de vídeo que se ha utilizado en las simulaciones y la transmisión de una trama con esta velocidad binaria empleando cualquier esquema de codificación (incluso CS-4) proporciona un retardo de trama positivo según la expresión (C.1). Observando las tablas C.1 y C.2 se deduce que si el tamaño y la duración de varias tramas es diferente, también lo es el retardo de trama que se obtiene según la expresión (C.1), aun cuando la tasa binaria requerida por todas ellas para su transmisión sin retardo sea la misma.

***T* (bytes)**	**25**	**50**	**75**	**100**	**125**	**150**
***D* (ms)**	**40**	**80**	**120**	**160**	**200**	**240**
CS-1	– 17.90	– 35.80	– 53.70	– 71.60	– 89.50	– 107.40
CS-2	– 25.07	– 50.15	– 75.22	– 100.30	– 125.37	– 150.45
CS-3	– 27.18	– 54.36	– 81.54	– 108.72	– 135.90	– 163.08
CS-4	– 30.65	– 61.31	– 91.96	– 122.62	– 153.27	– 183.93

Tabla C.1.- Retardo de trama según la expresión (C.1) para tramas de diferentes tamaños y duraciones pero con una tasa binaria común de 5 kbps.

T (bytes)	150	300	450	600	750	900
D (ms)	40	80	120	160	200	240
CS-1	92.60	185.19	277.79	370.39	462.98	555.58
CS-2	49.55	99.10	148.66	198.21	247.76	297.31
CS-3	36.92	73.85	110.77	147.69	184.62	221.54
CS-4	16.07	32.15	48.22	64.30	80.37	96.45

Tabla C.2.- Retardo de trama según la expresión (C.1) para tramas de diferentes tamaños y duraciones pero con una tasa binaria común de 30 kbps.

El segundo algoritmo propuesto en el presente trabajo se definió en el capítulo sexto evaluando la expresión

$$\frac{Throughput}{Retardo\ de\ trama} = \frac{Throughput}{\frac{T}{R_{CS-i}} - D} = \frac{Throughput \cdot R_{CS-i}}{T - D \cdot R_{CS-i}} \qquad si\ Ret.\ de\ trama > 0 \quad \text{(C.2a)}$$

$$Throughput \cdot \left| Retardo\ de\ trama \right| = Throughput \cdot \left(D - \frac{T}{R_{CS-i}} \right) \ si\ Ret.\ de\ trama < 0 \text{ (C.2b)}$$

Según la expresión (C.1), para el conjunto de tramas que requiere una velocidad binaria de 30 kbps, el retardo de trama obtenido con cualquier esquema de codificación siempre será positivo, lo que significa que para este conjunto se aplica, sobre los cuatro esquemas de codificación, la expresión (C.2a). Análogamente, según la expresión (C.1), para el conjunto de tramas que requiere una velocidad binaria de 5 kbps, el retardo de trama obtenido con cualquier esquema de codificación siempre será negativo, lo que significa que para este conjunto se aplica, sobre los cuatro esquemas de codificación, la expresión (C.2b). Bajo esta consideración se ha representado, para todos los esquemas de codificación, la expresión (C.2) en función de la relación CIR tanto para retardos de trama positivos (figura C.1) como para retardos de trama negativos (figura C.2). En estas figuras se observa que para tramas de diferentes tamaños y duraciones pero que requieren la misma velocidad binaria para una transmisión sin retardo, el valor que se obtiene para los umbrales de decisión a partir de las expresiones (6.6) y (6.7), repetidas en las expresiones (C.2a) y (C.2b) respectivamente, es el mismo. El valor del retardo de trama, como se ha visto anteriormente, es diferente para cada par de valores de tamaño y duración, por lo que las curvas toman valores distintos (observar el eje de ordenadas de las curvas); sin embargo, el valor de los umbrales de decisión no varía. Con esto se concluye, por tanto, que para el análisis de las expresiones (6.6) y (6.7) no es necesario considerar el valor del tamaño de la trama (T) y su duración (D), sino que es suficiente considerar únicamente el valor de la tasa binaria (T/D) que dicha trama requiere para ser transmitida sin retardo.

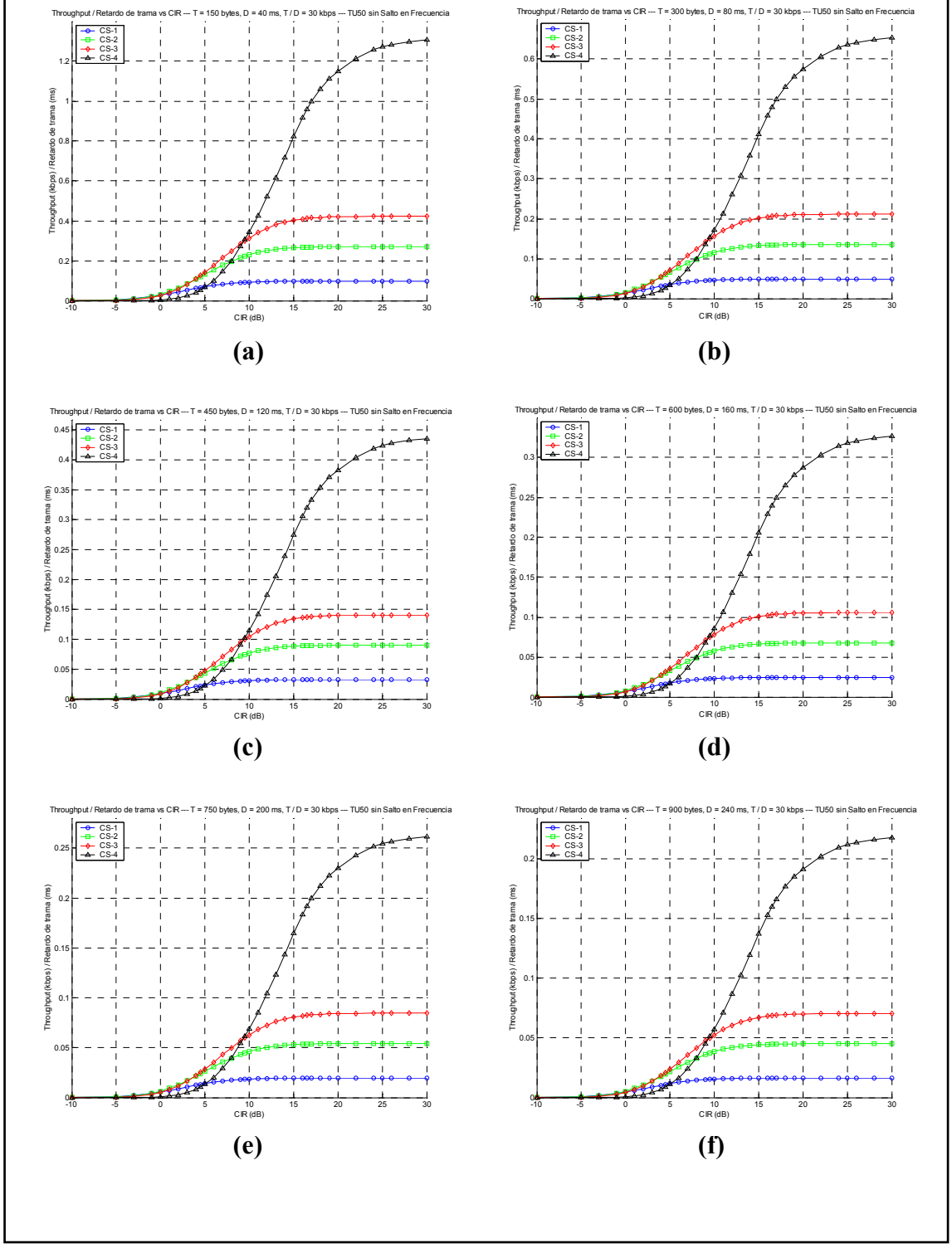


Figura C.1.- Representación de la expresión (C.2a) en función de la relación CIR para diversos tamaños (T) y duraciones (D) de trama pero velocidad binaria (T/D) de 30 kbps: (a) T = 150 bytes, D = 40 ms, (b) T = 300 bytes, D = 80 ms, (c) T = 450 bytes, D = 120 ms, (d) T = 600 bytes, D = 160 ms, (e) T = 750 bytes, D = 200 ms, (f) T = 900 bytes, D = 240 ms.

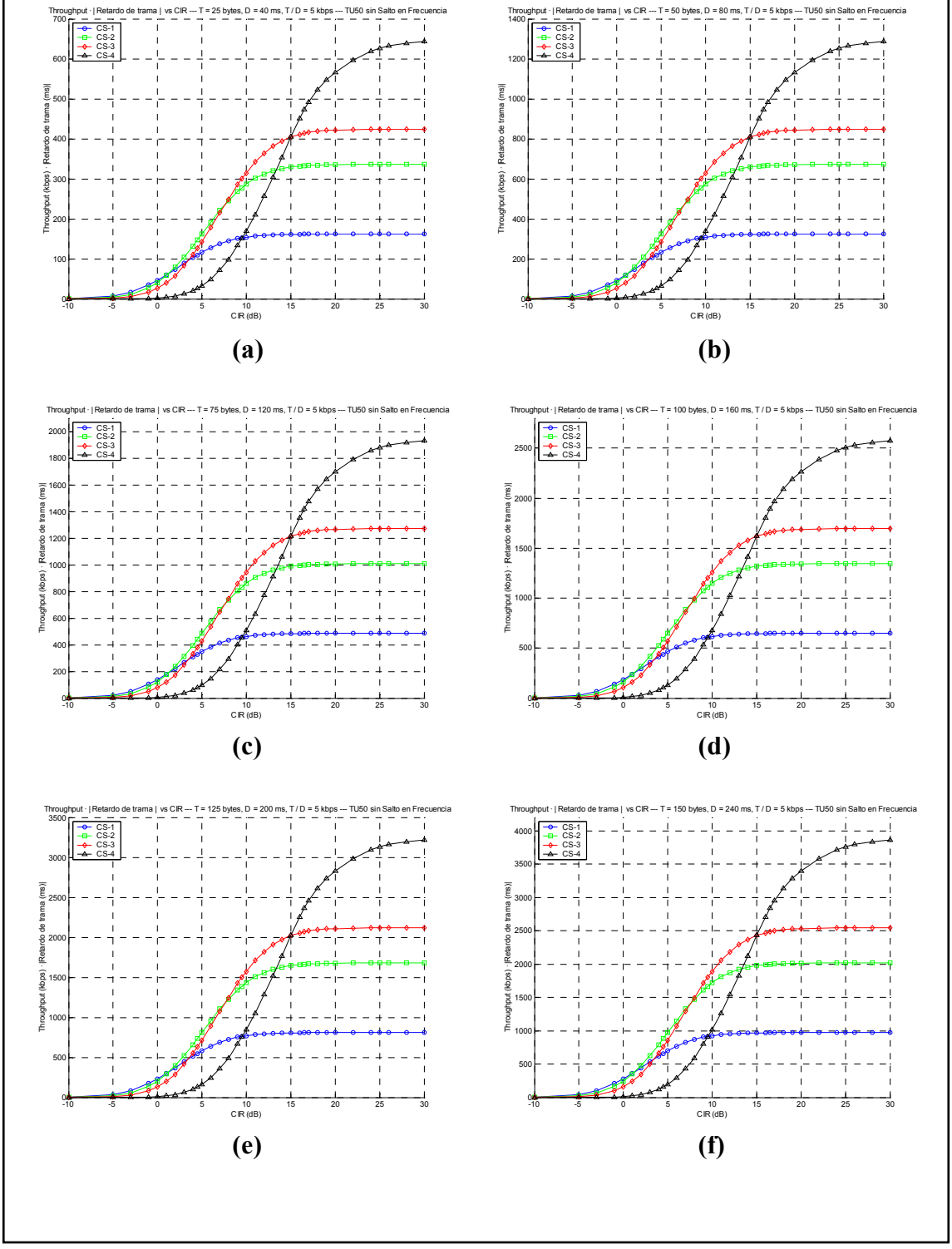


Figura C.2.- Representación de la expresión (C.2b) en función de la relación CIR para diversos tamaños (T) y duraciones (D) de trama pero velocidad binaria (T/D) de 5 kbps: (a) T = 25 bytes, D = 40 ms, (b) T = 50 bytes, D = 80 ms, (c) T = 75 bytes, D = 120 ms, (d) T = 100 bytes, D = 160 ms, (e) T = 125 bytes, D = 200 ms, (f) T = 150 bytes, D = 240 ms.

Bibliografía

[1] A. Segura, E. M. Tébar,"Introducción a las Comunicaciones Móviles", Universidad Miguel Hernández, 2002.

[2] A. Segura, E. M. Tébar, "Tecnología de las Comunicaciones Móviles", Universidad Miguel Hernández, 2003.

[3] A. Arregui, "Sistemas de Telefonía Móvil Digital GSM", "Evolución de GSM como Enlace a los Sistemas 3G", Servicio de Publicaciones de la Universidad Politécnica de Valencia, 1998, SPUPV-98.2071, pp. 1 – 25.

[4] ETSI, "(GSM 03.02 version 7.1.0 Release 1998) Digital Cellular Telecommunications System (Phase 2+); Network Architecture".

[5] T. Li, "An overview of General Packet Radio Service", 31 Marzo 2001. URL: "http://www.nd.edu/~tli/gprs_edge/dnet_proj.doc".

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

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

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

[9] J. Gozálvez, "Performance Analysis of Link Adaptation for Circuit and Packet Switched TDMA Mobile Communications Systems", Tesis doctoral, Department of Electronic and Electrical Engineering, University of Strathclyde, 2002, Capítulo 3, pp. 52 – 71.

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

[11] ETSI, "(GSM 05.01 version 8.4.0 Release 1999) Digital Cellular Telecommunications System (Phase 2+); Physical Layer on the Radio Path; General Description".

[12] ETSI, "(GSM 05.02 version 8.5.1 Release 1999) Digital Cellular Telecommunications System (Phase 2+); Multiplexing and Multiple Access on the Radio Path".

[13] ETSI, "(GSM 05.03 version 8.5.1 Release 1999) Digital Cellular Telecommunications System (Phase 2+); Channel Coding".

[14] ETSI, "(GSM 05.04 version 8.1.2 Release 1999) Digital Cellular Telecommunications System (Phase 2+); Modulation".

[15] ETSI, "(GSM 05.05 version 8.5.1 Release 1999) Digital Cellular Telecommunications System (Phase 2+); Radio Transmission and Reception".

[16] J. Zander, "Trends in Resource Management Future Wireless Networks", IEEE Wireless Communications and Networking Conference, Septiembre 2000, vol. 1, pp. 159 – 163.

[17] T. Ojanperä, R. Prasad, "An Overview of Third-Generation Wireless Personal Communications: A European Perspective", IEEE Personal Communications Magazine, Diciembre 1998, vol. 5, nº 6, pp. 59 – 65.

[18] M. Flament, F. Gessler, F. Lagergren, O. Queseth, R. Stridh, M. Unbedaun, J. Wu, J. Zander, "An Approach to 4th Generation Wireless Infraestructures–Scenarios and Key Research Issues", IEEE Vehicular Technology Conference, Mayo 1999, vol. 2, pp. 1742 – 1746.

[19] A. Acampora, "Wireless ATM: A Perspective on Issues and Prospects", IEEE Personal Communications Magazine, Agosto 1996, vol. 3, nº 4, pp. 8 – 17.

[20] A. Furüskar, "Can 3G Services Be Offered in Existing Spectrum?", Techinal Licenciate, Radio Communications Systems Laboratory, Department of Signals, Sensors and Systems, Royal Institute of Technology, 2001.

[21] J. Zander, "Radio Resource Management in Future Wireless Networks: Requierements and Limitations", IEEE Communications Magazine, Agosto 1997, vol. 35, nº 8, pp. 30 – 36.

[22] J. M. H. Rábanos, "Comunicaciones móviles", Editorial Centro de Estudios Ramón Areces, S.A., Madrid, 1997.

[23] J. M. H. Rábanos, "Transmisión por radio", Tercera edición, Editorial Centro de Estudios Ramón Areces, S.A., Madrid, 1998.

[24] M. Andersin, "Power Control and Admission Control in Cellular Radio Systems", Tesis Doctoral, Radio Communications Systems Laboratory, Royal Institute of Technology, 1996.

[25] J. Gozálvez, "Performance Analysis of Link Adaptation for Circuit and Packet Switched TDMA Mobile Communications Systems", Tesis doctoral, Department of Electronic and Electrical Engineering, University of Strathclyde, 2002, Capítulo 2, pp. 24 – 51.

[26] D. Todinca, P. Perry, J. Murphy, "Resource Allocation Algorithms for GPRS Network Performance Improvement During Congested Periods", IEI/IEE Postgraduate Symposium on Telecommunications Systems Research, Noviembre 2001.

[27] J. Bada, F. Casadevall, "Sensibilidad en la elección de la disciplina de servicio sobre el sistema GPRS", XII Jornadas de I+D en Telecomunicaciones", Telecom I+D, Noviembre 2002.

[28] H. Oloffson, "Interference Diversity as Means for Increasing Capacity in GSM", Proceedings of the European Personal Mobile Communications Conference, Noviembre 1995.

[29] S. Tabbane, "Handover procedures in ATDMA project", RACE Mobile Workshop, Mayo 1993.

[30] K. David, P. Blanc, T. Kassing, "Handover and Link Adaptation Evaluation of the RACE/ATDMA System Concept for a 3rd Generation Air Interface by Simulation", IEEE Vehicular Technology Conference, Abril – Mayo 1996, vol. 1, pp. 451 – 455.

[31] J. Zander, "Performance of Optimum Transmitter Power Control in Radio Systems", IEEE Transactions on Vehicular Technology, Febrero 1992, vol. 41, nº 1, 57 – 62.

[32] H. Panzer, R. Beck, "Adaptive Resource Allocation in Metropolitan Area Cellular Mobile Radio Systems", IEEE Vehicular Technology Conference, Mayo 1999, pp. 638 – 645.

[33] M. Frodigh, "Performance Bounds for Power Control Supported DCA-Algorithms in Highway Micro Cellular Radio Systems", IEEE Transactions on Vehicular Technology, Mayo 1995, vol. 44, nº2, pp. 238 – 243.

[34] G. J. Foscini, Z. Miljanic, "Distributed Autonomous Wireless Channel Assignment Algorithm with Power Control", IEEE Transactions on Vehicular Technology, Agosto 1995, vol. 44, nº 3, pp. 420 – 429.

[35] R. D. Yates, C.-Y. Huang, "Integrated Power Control and Base Station Assignment", IEEE Transactions on Vehicular Technology, Agosto 1995, vol. 44, nº 3, pp. 638 – 644.

[36] S. V. Hanly, "An Algorithm for Combined Cell-Site Selection and Power Control to Maximize Cellular Spread Spectrum Capacity", IEEE Journal on Selected Areas in Communications, Septiembre 1995, vol. 13, nº 7, pp. 1332 – 1340.

[37] A. Krämling, M. Siebert, M. Lott, M. Weckerle, "Interaction of Power Control and Link Adaptation for Capacity Enhancement and QoS Assistance", IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, Septiembre 2002, vol. 2, pp. 697 – 701.

[38] J. Dunlop, P. Cosimini, G. Graham, E. Le Strat, "Estimation of the Performance of an Adaptive Air Interface in Mobile Radio", Proceedings of the RACE Mobile Telecommunications Workshop, Mayo 1994, pp. 47 – 51.

[39] J. Dunlop, J. Irvine, P. Cosimini, "Estimation of the Performance of Link Adaptation in Mobile Radio", IEEE Vehicular Technology Conference, Julio 1995, vol. 1, pp. 326 – 330.

[40] N. B. Mehta, A. J. Goldsmith, "Performance Analysis of Link Adaptation in Wireless Data Networks", IEEE Global Telecommunications Conference, Noviembre – Diciembre 2000, vol. 3, pp. 1422 – 1426.

[41] J. Gozálvez, J. Dunlop, "GPRS Link Adaptation Switching Thresholds and Intervals", IEEE Electronics Letters, Julio 2000, vol. 36, nº 15, pp. 1311 – 1312.

[42] O. Queseth, F. Gessler, M. Frodigh, "Algorithms for Link Adaptation in GPRS", IEEE Vehicular Technology Conference, Mayo 1999, vol. 2, pp. 943 – 947.

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

[44] J. de Maré, "Link Adaptation in GPRS", Master's Thesis Report 20p., Department of Signals, Sensors and Systems, Royal Institute of Technology, 1998.

[45] K. Balachandran, S. Kadaba, S. Nanda, "Rate Adaptation over Mobile Radio Channels Using Channel Quality Information", IEEE Global Telecommunications Conference, Noviembre 1998.

[46] K. Balachandran, S. Kadaba, S. Nanda, "Channel Quality Estimation and Rate Adaptation for Cellular Mobile Radio", IEEE Journal on Selected Areas in Communications, Julio 1999, vol. 17, nº 7, pp. 1244 – 1256.

[47] M. D. Austin, G. L. Stüber, "In-Service Signal Quality Estimation for TDMA Cellular Systems", IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, Septiembre 1995, vol. 2, pp. 836 – 840.

[48] M. Andersin, N. B. Mandayam, R. D. Yates, "Subspace Based Estimation of the Signal to Interference Ratio for TDMA Cellular Systems", IEEE Vehicular Technology Conference, Abril – Mayo 1996, vol. 2, pp. 1155 - 1159.

[49] M. Turkboylari, G. L. Stüber, "An Efficient Algorithm for Estimating the Signal to Interference Ratio in TDMA Cellular Systems", IEEE Transactions on Communications, Junio 1998, vol. 46, nº 6, pp. 728 – 731.

[50] A. Brandäo, L. Lopes, D. MacLernon, "Quality Assessment for Pre-Detection Diversity Switching", IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, Septiembre 1995, vol. 2, pp. 577 – 581.

[51] J. Pons, J. Dunlop, "In-Service Quality Estimation for Link Adaptation Algorithms, Applied to GSM", IEEE International Conference on Universal Personal Communications, Octubre 1998, vol. 2, pp. 1169 – 1173.

[52] J. Pons, J. Dunlop, "Bit Error Rate Based Link Adaptation for GSM", IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, Septiembre 1998, vol. 3 , pp. 1530 – 1534.

[53] R. Ferrús, F. Casadevall, "Using Time Advance Information to Choose Coding Scheme in GPRS", IEEE Vehicular Technology Conference, Octubre 2001, vol. 2, pp. 562 – 566.

[54] J. Irvine, P. Cosimini, J. Dunlop, "Implementation Considerations for Gross Rate Link Adaptation", IEEE Vehicular Technology Conference, Abril – Mayo 1996, vol. 3, pp. 1766 – 1770.

[55] J. C.-I. Chuang, "Improvement of Data Throughput in Wireless Packet Systems with Link Adaptation and Efficient Frecuency Reuse", IEEE Vehicular Technology Conference, Mayo 1999, vol.1, pp. 821 – 825.

[56] J. Gozálvez, J. Dunlop, "On the Dynamics of Link Adaptation Updating Periods for Packet Switched Systems", Fourth International Symposium on Wireless Multimedia Personal Communications, Septiembre 2001.

[57] O. Corbun, M. Almgren, K. Svanbro, "Capacity and Speech Quality Aspects using Adaptive Multi-Rate (AMR)", IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, Septiembre 1998, vol. 3, pp. 1535 – 1539.

[58] J. Chuang, X. Qiu, J. Whitehead, "Data Throughput Enhancement in Wireless Packet Systems by Improved Link Adaptation with Application to the EDGE System", IEEE Vehicular Technology Conference, Septiembre 1999, vol. 1, pp. 456 – 460.

[59] X. Qiu, J. Chuang, "Link Adaptation in Wireless Data Networks for Throughput Maximization under Retransmissions", IEEE International Conference on Communications, Junio 1999, vol. 2 , pp. 1272 – 1277.

[60] S. Eriksson, A. Furuskär, M. Höök, S. Jäverbring, H. Olofsson, J. Sköld, "Comparison of Link Quality Control Strategies for Packet Data Services in EDGE", IEEE Vehicular Technology Conference, Mayo 1999, vol. 2, pp. 938 – 942.

[61] J. Pons, J. Dunlop, J. Kelliher, "On the Application of Link Adaptation Concepts to GSM", IEEE Vehicular Technology Conference, Mayo 1998, vol. 2, pp. 1219 – 1223.

[62] E. Esteves, P. J. Black, M. I. Gurelli, "Link Adaptation Techniques for High-Speed Packet Data in Third Generation Cellular Systems", European Wireless, Febrero 2002.

[63] T. Holliday, A. Goldsmith, P. Glynn, "Optimal Link Adaptation in Wideband CDMA Systems", IEEE Global Telecommunications Conference, Noviembre 2002, vol. 1, pp. 721 – 726.

[64] J. Irvine, J. Pons, J. Dunlop, "Link Adaptation to Improve Coverage in the TETRA Private Mobile Radio System", IEEE Vehicular Technology Conference, Septiembre 1999, vol. 2, pp. 1091 – 1095.

[65] D. Qiao, S. Choi, K. G. Shin, "Goodput Analysis and Link Adaptation for IEEE 802.11a Wireless LANs", IEEE Transactions on Mobile Computing, vol. 1, nº 4, Octubre – Diciembre 2002, pp. 278 – 292.

[66] J.-F. Huard, A. A. Lazar, "On QoS Mapping in Multimedia Networks", IEEE Computer Software and Applications Conference, Agosto 1997, pp. 312 – 317.

[67] B. G. Marchent, M. Wilson, A. Rouz, "Support of Mobile Multimedia over Radio for a Range of QoS and Traffic Profiles", IEEE Vehicular Technology Conference, Mayo 1999, vol. 2, pp. 1540 – 1544.

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

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

[70] S. N. Fabri, S. Worrall, A. Sadka, A. Kondoz, "Real-Time Video Communications over GPRS", IEE 3G Mobile Communication Technologies, Mayo 2000, Publicación nº 471, pp. 426 – 430.

[71] UIT-T, "(Recomendación H.263, Febrero 1998) Codificación de Vídeo para Comunicación a Baja Velocidad Binaria".

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

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

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

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

[76] M. Krunz, "Bandwidth Allocation Strategies for Transporting Variable-Bit-Rate Video Traffic", IEEE Communications Magazine, Enero 1999, vol. 37, nº 1, pp. 40 – 46.

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

[78] B. Girod, N. Färber, "Feedback-Based Error Control for Mobile Video Transmission", Procedings of the IEEE, Octubre 1999, vol. 87, nº 10, pp. 1707 – 1723.

[79] B. Sklar, "Digital Communications: Fundamentals and Applications", Segunda Edición, Editorial Prentice Hall, Inc., New Jersey, 2001, p. 877.

[80] T. Ren, I. Koutsopoulos, L. Tassiulas, "QoS Provisioning for Real-Time Traffic in Wireless Packet Networks", IEEE Global Telecommunications Conference, Noviembre 2002, vol. 2, pp. 1673 – 1677.

[81] ETSI, "(3GPP TS 22.060 version 5.3.0 Release 5) Digital Cellular Telecommunications System (Phase 2+); Universal Mobile Telecommunications Sysmtem (UMTS); General Packet Radio Service (GPRS); Service Description; Stage 1".

[82] P. Cherriman, L. Hanzo, "Programable H.263-based Wireless Video Transceivers for Interference-limited Environments", IEEE Transactions on Circuits and Systems for Video Technology, Junio 1998, vol. 8, nº 3, pp. 275 – 286.

[83] P. Cherriman, B. L. Yeap, L. Hanzo, "The Performance of H263-Based Video Telephony Over Turbo-Equalized GSM/GPRS", IEEE Transactions on Circuits and Systems for Video Technology, Octubre 2002, vol. 12, nº 10, pp. 909 – 915.

[84] ETSI, "(3GPP TS 22.105 version 6.2.0 Release 5) 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Service Aspects; Services and Service Capabilities".

[85] ETSI, "(3GPP TS 23.107 version 5.2.0 Release 5) 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; QoS Concept and Architecture".

[86] ETSI, "(3GPP TR 26.912 version 4.0.0 Release 4) Universal Mobile Telecommunications System (UMTS); QoS for Speech and Multimedia Codec; Quantitave Performance Evaluation of H.324 Annex C over 3G".

[87] K. K. Lee, S. T. Chanson, "Packet Loss Probability for Real-Time Wireless Communications", IEEE Transactions on Vehicular Technology, Noviembre 2002, vol. 51, nº 6, pp. 1569 – 1575.

[88] J. Gozálvez, "Performance Analysis of Link Adaptation for Circuit and Packet Switched TDMA Mobile Communications Systems", Tesis doctoral, Department of Electronic and Electrical Engineering, University of Strathclyde, 2002, Capítulo 4, pp. 72 - 99.

[89] Communications Networks Class Library (CNCL), Aachen University of Technology. URL: "http://www.comnets. rwth-aachen. de/cnroot_engl.html". También "http://www.comnets.rwth-aachen. de/doc/cncl.html".

[90] A. Arregui, "Benefits and Feasibility of the Partial Loading Approach in Current and Future Cellular Mobile Radio Systems", Tesis doctoral, Department of Electronic and Electrical Engineering, University of Strathclyde, 1998.

[91] K. Feher, "Wireless Digital Communications: Modulation and Spread Spectrum Applications", Editorial Prentice Hall, Inc., New Jersey, 1995.

[92] C. You, K. Chandra, "Time Series Models for Internet Data Traffic", Conference on Local Computer Networks, Octubre 1999, pp. 164-171.

[93] C. Johansson, H. Nyberg, P. Bruin, "Streaming Services in GSM/EDGE – Radio Resource Management Concepts and System Performance", IEEE Vehicular Technology Conference, Octubre 2001, vol. 3, pp. 1765 – 1769.

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

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

[96] O. Lázaro, "Dynamic Radio Resource Management Algorithms and Traffic Models for Emerging Mobile Communication Systems", Tesis doctoral, Department of Electronic and Electrical Engineering, University of Strathclyde, 2002, Capítulo 5, pp. 101 – 134.

[97] F. H. P. Fitzek, M. Reisslein, "MPEG–4 and H.263 Video Traces for Network Performance Evaluation", URL: "http://www-tkn.ee.tu-berlin.de/research/trace/trace.html".

[98] J. Gozálvez, "Performance Analysis of Link Adaptation for Circuit and Packet Switched TDMA Mobile Communications Systems", Tesis doctoral, Department of Electronic and Electrical Engineering, University of Strathclyde, 2002, Capítulo 7, pp. 151 - 209.

[99] J. Gozálvez, J. Dunlop, "On the Effect of Correlation in Multislot Link Layer Analysis for GPRS", IEEE Vehicular Technology Conference, Septiembre 2000, vol. 1, pp. 444 – 450.

[100] J. Gozálvez, J. Dunlop, "Impact of the Interference on the GPRS Multislot Link Level Performance", European Wireless, Febrero 2002.

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

[102] J. Gozálvez, J. Dunlop, "On the Importance of Using Appropiate Link-to-System Level Interfaces for the Study of Link Adaptation", Proceedings of the IST Mobile & Wireless Communications Summit, Junio 2003, pp. 441 – 445.

[103] J. Gozálvez, "Performance Analysis of Link Adaptation for Circuit and Packet Switched TDMA Mobile Communications Systems", Tesis doctoral, Department of Electronic and Electrical Engineering, University of Strathclyde, 2002, Capítulo 6, pp. 121 - 150.

[104] K. K. Leung, P. F. Driessen, K. Chawla, X. Qiu, "Link Adaptation and Power Control for Streaming Services in EGPRS Wireless Networks", IEEE Journal on Selected Areas in Communications, Octubre 2001, vol. 19, nº 10, pp. 2029 – 2039.

[105] K. K. Leung, L.-C. Wang, "Controlling QoS by Integrated Power Control and Link Adaptation in Broadband Wireless Networks", European Transactions on Telecommunications, Julio – Agosto 2000, vol. 11, pp. 383-394.

[106] C. Kodikara, S. N. Fabri, A. M. Kondoz, "Performance Improvement for Real-Time Video Communications by Link Adaptation in E-GPRS Wireless Networks", IEE 3G Mobile Communication Technologies, Mayo 2002, Publicación nº 489, pp. 489 – 494.

[107] W. Luo, K. Balachandran, S. Nanda, K. Chang, "Packet Size Dependent Link Adaptation for Wireless Packet Data", IEEE Global Telecommunications Conference, Noviembre – Diciembre 2000, vol. 1. pp. 53 – 56.

[108] T. Holliday, A. Goldsmith, P. Glynn, "Wireless Link Adaptation Policies: QoS for Deadline Constrained Traffic with Imperfect Channel Estimates", IEEE International Conference on Communications, Abril – Mayo 2002, vol. 5, pp. 3366 – 3371.