

UNIVERSIDAD CARLOS III DE MADRID
ESCUELA POLITÉCNICA SUPERIOR

DEPARTAMENTO DE
TEORÍA DE LA SEÑAL Y COMUNICACIONES



TESIS DOCTORAL

OPTIMIZACIÓN DE ENERGÍA Y EFICIENCIA
DE TRANSMISIÓN CON ANÁLISIS DE
IMPARCIALIDAD EN COMUNICACIONES
INALÁMBRICAS ADAPTATIVAS

JOSÉ JOAQUÍN ESCUDERO GARZÁS
FEBRERO DE 2010

UNIVERSIDAD CARLOS III DE MADRID
ESCUELA POLITÉCNICA SUPERIOR
DEPARTAMENTO DE TEORÍA DE LA SEÑAL Y
COMUNICACIONES



TESIS DOCTORAL

OPTIMIZACIÓN DE ENERGÍA Y EFICIENCIA
DE TRANSMISIÓN CON ANÁLISIS DE
IMPARCIALIDAD EN COMUNICACIONES
INALÁMBRICAS ADAPTATIVAS

Autor:

JOSÉ JOAQUÍN ESCUDERO GARZÁS
INGENIERO DE TELECOMUNICACIÓN

Director:

ANA GARCÍA ARMADA
DOCTOR INGENIERO DE TELECOMUNICACIÓN

LEGANÉS, FEBRERO DE 2010

Tesis Doctoral:
OPTIMIZACIÓN DE ENERGÍA Y EFICIENCIA DE TRANSMISIÓN
CON ANÁLISIS DE IMPARCIALIDAD EN COMUNICACIONES
INALÁMBRICAS ADAPTATIVAS

Autor:
JOSÉ JOAQUÍN ESCUDERO GARZÁS

Director:
ANA GARCÍA ARMADA

El Tribunal nombrado para juzgar la Tesis arriba indicada, compuesto por los siguientes Doctores:

PRESIDENTE: ANTONIO ARTÉS RODRÍGUEZ

SECRETARIO: MARÍA JULIA FERNÁNDEZ-GETINO GARCÍA

VOCAL: LUIS IGNACIO SANTAMARÍA CABALLERO

VOCAL: GONZALO SECO GRANADOS

VOCAL: JAVIER DEL SER LORENTE

Acuerda otorgarle la calificación de:

LEGANÉS, A DE FEBRERO DE 2010

Contenido

Resumen	IX
Abstract	XI
Notación y acrónimos	XIII
1. Introducción	1
1.1. Transmisión adaptativa en redes inalámbricas	2
1.2. Objetivos	3
1.3. Organización y contribuciones	4
1.3.1. Metodología	4
1.3.2. Organización	4
1.3.3. Contribuciones	6
2. Transmisión adaptada al medio: modulación adaptativa, bit loading y radio cognitiva	9
2.1. Modulación adaptativa: conceptos básicos	10
2.1.1. Revisión histórica de los sistemas de modulación adaptativa	11
2.1.2. El sistema de transmisión adaptativa	12
2.1.3. Técnicas básicas de modulación adaptativa	14
2.1.4. Optimización de la eficiencia espectral mediante modulación adaptativa	15
2.1.5. Otras métricas para la evaluación de esquemas de modulación adaptativa	17
2.2. Modulaciones multiportadora: OFDM	17
2.2.1. Fundamentos básicos de OFDM	19
2.2.2. La rejilla tiempo-frecuencia OFDM	21
2.3. Transmisión en sistemas multiportadora punto a punto	22

2.3.1.	Asignación de potencia óptima multiportadora: <i>waterfilling</i>	22
2.3.2.	La aproximación SNR <i>gap</i>	24
2.3.3.	Asignación de bits por subportadora: <i>bit-loading</i>	26
2.4.	Acceso múltiple multiportadora: OFDMA	27
2.4.1.	Diversidad multiusuario	29
2.4.2.	El problema de la asignación de recursos en OFDMA	32
2.5.	Imparcialidad entre los usuarios en OFDMA	33
2.6.	Asignación en OFDMA de los recursos potencia, bits y subportadoras	34
2.6.1.	Algoritmos de asignación conjunta de recursos	36
2.6.2.	Algoritmos subóptimos: asignación en dos pasos de subportadoras, potencia y bits	38
2.7.	Asignación de recursos en OFDMA utilizando agrupaciones de subportadoras	41
2.7.1.	Agrupación de subportadoras: los subcanales	41
2.7.2.	La utilización de los subcanales en la literatura	43
2.8.	Un paso más en la transmisión adaptativa: la radio cognitiva	45
2.8.1.	Redes de nueva generación: acceso oportunista al espectro	46
2.8.2.	Asignación de recursos utilizando radios cognitivas	50
2.9.	Conclusiones	53
3.	Métodos de optimización para problemas con restricciones en comunicaciones inalámbricas	55
3.1.	Optimización de problemas en variable entera	56
3.1.1.	¿Qué es un problema entero binario?	57
3.1.2.	Métodos de optimización entera óptimos	57
3.1.3.	Métodos de optimización entera subóptimos	60
3.1.4.	Bondad de las relajaciones lineal y de Lagrange	61
3.1.5.	Igualdad de las soluciones del problema entero y el problema relajado lagrangiano	63
3.1.6.	Igualdad de las soluciones del problema entero y el problema relajado lineal	64
3.2.	Optimización de problemas en variable continua	64
3.2.1.	Optimización no lineal con restricciones de igualdad: multiplicadores de Lagrange	66
3.2.2.	Optimización no lineal con restricciones de desigualdad: condiciones de Karush-Kuhn-Tucker (KKT)	68

3.2.3. Optimización no lineal con restricciones de desigualdad:	
Teoría de dualidad	70
3.3. Conclusiones	72
4. Modulación adaptativa eficiente en energía para sistemas centralizados	73
4.1. Introducción	74
4.1.1. Esquemas eficientes en energía para redes inalámbricas . .	74
4.1.2. Fiabilidad y comunicaciones eficientes en energía	76
4.1.3. Contribuciones del capítulo	76
4.2. Modelo general del sistema para redes inalámbricas	78
4.2.1. Esquema de transmisión	78
4.2.2. Modelo de pérdidas de los enlaces	79
4.2.3. Modelo de energía	79
4.2.4. Sistema de señalización	80
4.3. Modulación adaptativa basada en intervalos de transmisión variables para redes inalámbricas	80
4.3.1. Esquema de transmisión basado en el número de símbolos del intervalo	82
4.4. Aplicación práctica: optimización de energía en redes de sensores .	84
4.5. Modelo del sistema para WSN	85
4.5.1. Modelo de pérdidas para WSN	86
4.5.2. Justificación frente al estándar 802.15.4	86
4.6. Modulación adaptativa basada en intervalos de transmisión variables para WSN	87
4.6.1. Discusión de la modulación y aproximación SNR <i>gap</i> para modulaciones MQAM	87
4.6.2. Minimización de la energía total de transmisión	88
4.7. Resultados	89
4.8. Conclusiones	94
5. Optimización de la eficiencia de transmisión en sistemas OFDM multiusuario	97
5.1. Introducción	98
5.1.1. El <i>goodput</i> como objetivo de optimización	99
5.1.2. Contribuciones del capítulo	100

5.2. Descripción del sistema: capa física y capa de enlace	101
5.2.1. El nivel físico y el modelo de canal	102
5.2.2. Capa de enlace	105
5.3. Formulación del problema de optimización del <i>goodput</i>	105
5.4. Asignación de subcanales, bits y potencia: método SBPA	108
5.4.1. Cálculo de la matriz de <i>goodput</i> (GMC)	111
5.4.2. El problema de asignación de subcanales (SAP)	112
5.5. Asignación de subcanales, bits y potencia: método subóptimo S-SBPA	113
5.5.1. Cálculo de la matriz de <i>goodput</i> para el método subóptimo	115
5.5.2. El problema de la asignación de subcanales para el método subóptimo	115
5.6. Simulaciones y resultados	121
5.6.1. Canal de variación lenta	122
5.6.2. Canal de variación rápida	125
5.6.3. Evaluación comparativa y de la complejidad	126
5.7. Conclusiones	137
6. Conclusiones y líneas futuras de investigación	139
6.1. Conclusiones generales	139
6.2. Líneas futuras de investigación	142
Anexos	
A. Expresiones matriciales del cálculo diferencial	145
B. Lista de publicaciones	147
Bibliografía	166
Índice de Figuras	169
Índice de Tablas	171

RESUMEN

Las redes inalámbricas se caracterizan por utilizar un medio de transmisión dinámico, en el que las condiciones de transmisión cambian de forma continua. Este cambio continuado lleva a diseñar los sistemas de comunicaciones desde dos puntos de vista.

El más elemental es diseñar nuestro sistema para el caso peor. Aunque sencilla, esta solución lleva aparejado un sobredimensionamiento de los recursos empleados; si consideramos por ejemplo la potencia de transmisión, este sobredimensionamiento resultará, entre las consecuencias más inmediatas, en mayores interferencias y mayor consumo de energía (y por tanto menor duración de las baterías en sistemas así alimentados). Frente al anterior enfoque, la transmisión adaptativa permite adecuar los parámetros de transmisión a las condiciones del canal.

El objetivo de esta tesis es proponer esquemas de transmisión adaptativos para sistemas inalámbricos que, bien mejoran técnicas existentes, como veremos para redes de sensores, o bien resuelven problemas nuevos, como la optimización de la eficiencia de transmisión de una estación base. Contribuimos por tanto al desarrollo de técnicas de transmisión adaptativa en redes inalámbricas para conseguir sistemas eficientes en energía y en tasa de transmisión. Para ello, previamente realizamos un recorrido por las técnicas existentes de transmisión adaptativa (capítulo 2) y presentamos las herramientas matemáticas que se utilizarán para resolver los problemas de optimización que resultan de las técnicas propuestas (capítulo 3).

La transmisión eficiente en energía que proponemos tiene como objetivo optimizar la energía total de transmisión del enlace ascendente para un sistema

centralizado, con comunicaciones de salto único. Se asume que la red es heterogénea, lo que se traduce en que cada nodo tendrá unos requisitos diferentes de calidad de servicio en función de la aplicación que se ejecute en cada nodo. La calidad de servicio vendrá expresada mediante la tasa de error de símbolo y una tasa de transmisión de bit. La optimización propuesta se basa en un esquema de modulación *cross-layer* que asigna adaptativamente la duración de los intervalos de transmisión, y por tanto los bits por símbolo a transmitir, cumpliendo con la calidad de servicio exigida. Este esquema lo particularizamos para redes inalámbricas de sensores.

La optimización de la eficiencia de transmisión (o *goodput*) para el enlace descendente entre una estación base y los usuarios que de ella dependen, la realizaremos mediante la asignación de los recursos disponibles: subcanal (entendido como agrupación de subportadoras), potencia y tasa de bit por símbolo. El modelo del sistema implementa un mecanismo de control de imparcialidad que evita que resulten beneficiados aquellos usuarios que presenten unas condiciones más favorables de transmisión. Para resolver la asignación de recursos, se proponen dos esquemas: un esquema de asignación en dos pasos que reduce la complejidad del problema de asignación inicial, y un esquema subóptimo, orientado a la implementación práctica, que permite una reducción muy considerable de la carga computacional, sin una degradación apreciable del rendimiento respecto al óptimo.

ABSTRACT

Wireless networks are characterized by the use of a dynamic transmission channel, in which transmission conditions are continuously changing. This continuous change asks for designing communication systems from two different viewpoints.

The more basic one is to prepare our system for the worst case; this option, although straightforward, implies an excess of the resources used; for instance, if transmission power is considered, this excess of power results in a high interference and a high energy consumption (and, therefore, provokes a shorter lifetime for the batteries), among the most immediate consequences. Contrary to the previous point of view, adaptive transmission allows to adjust transmission parameters values to channel conditions.

The objective of this thesis is to propose new adaptive transmission schemes for wireless systems. Some of these schemes outperform existing techniques, as for the case of sensor networks, and others solve new problems, for instance the optimization of a base station transmitted goodput. We therefore contribute to the development of adaptive transmission techniques for wireless networks in order to achieve energy-efficient and transmission-efficient systems. To this end, we have previously reviewed the existing adaptive transmission techniques (chapter 2) and the mathematical tools that we apply to the optimization problems resulting from the proposed techniques (chapter 3).

The energy-efficient scheme we propose targets the total transmission energy optimization in the uplink of a centralized system, with single-hop communications. Network heterogeneity is assumed, which translates into different quality

of service requirements depending on the application running at each node. Quality of service is expressed by means of the symbol-error rate and the minimum bit-rate. The presented design is cross-layer-based and adapts time-slot duration to channel conditions, using adaptive modulation to fulfil the required quality of service. This proposal is particularly fitted to wireless sensor networks.

Goodput optimization for the downlink is performed by means of resources allocation, specifically of subchannels, transmit power and bit-rate per symbol. For our purpose, a subchannel is defined as a group of subcarriers. Our system model is provided with a fairness control mechanism that prevents users with favourable channel conditions from obtaining most of the resources. We present a two-step assignment scheme to solve resource allocation that reduces the complexity of the original problem, and a suboptimal implementation-oriented approach that allows a significant reduction of the computational load, without a significant loss of performance.

NOTACIÓN Y ACRÓNIMOS

Dado el uso generalizado del idioma inglés como idioma de trabajo en comunicaciones inalámbricas, algunas expresiones y términos característicos de este ámbito no han sido traducidos, por cuanto su equivalente en castellano pueda llevar a confusión o simplemente no existir, de forma que en esta tesis se ha utilizado directamente la terminología original. Para estos términos y expresiones se han empleado siempre caracteres en cursiva. Por ejemplo, hemos optado por no traducir el término *waterfilling*. Por otro lado, en el caso de ciertos términos y expresiones, aunque sí han sido traducidos, se ha optado por acompañarlas entre paréntesis de la palabra original con la que se denominan en notación anglosajona, también en cursiva, en beneficio de posibles aclaraciones respecto a la terminología usual del ámbito; es el caso por ejemplo de la diversidad multiusuario (*multiuser diversity*). Se ha optado por destacar con caracteres en negrita aquellos conceptos que se definen a lo largo de la tesis y que se consideran de relevancia.

Notación matemática

\gg	Mucho mayor que
\ll	Mucho menor que
$ x $	Módulo de x
$*$	Operador convolución lineal
\otimes	Operador convolución circular
∇f	Gradiente de la función f
\mathbf{x}^T	Traspuesto del vector \mathbf{x}
\mathbf{A}^T	Traspuesta de la matriz \mathbf{A}
$\arg \max_x f(x)$	Valor de la variable x que maximiza la función $f(x)$
$\text{conv}(\mathcal{F})$	Recubrimiento convexo de \mathcal{F}
$\mathbf{E}\{\cdot\}$	Operador valor esperado
$\max_x f(x)$	Valor máximo de $f(x)$ respecto de la variable x
$\min_x f(x)$	Valor mínimo de $f(x)$ respecto de la variable x
\bar{X}	Valor promedio de la variable aleatoria X
\mathbb{R}^m	Conjunto de los números reales de dimensión m
\mathbb{Z}^m	Conjunto de los números enteros de dimensión m
\mathbb{Z}^+	Conjunto de los números enteros positivos

Abreviaturas

μs	microsegundos
cap.	capítulo
s.t.	<i>subject to</i> (sujeto a)
fig.	figura
m	metro
ms	milisegundos
s	segundos

Acrónimos y Siglas

3G	tercera generación
3GPP-LTE	3rd Generation Partnership Project-Long Term Evolution
4G	cuarta generación

ADSL	Asymmetric Digital Subscriber Line
AMC	adaptive modulation and coding
AOR	assignment of resources
ARQ	automatic repeat request
AWGN	additive white Gaussian noise
bps	bits por segundo
BER	bit error rate
BIP	binary integer programming
CDMA	code division multiple access
CSI	channel state information
DAB	Digital Audio Broadcasting
DFT	discrete Fourier transform
DMT	discrete multi-tone
DSA	dynamic spectrum access
DVB-T	Digital Video Broadcasting-Terrestrial
EB	estación base
EGPRS	Enhanced General Packet Radio Service
FFT	fast Fourier transform
FUSC	full usage of subcarriers
FSR	frame success rate
GMC	goodput matrix calculation
GTS	guaranteed time slot
GPRS	General Packet Radio Service
GSC	generalized selection combining
IEEE	Institute of Electrical and Electronics Engineers
IFFT	inverse fast Fourier transform
IP	integer programming
ISI	interferencia entre símbolos

ISM	Industrial, Scientific and Medical (banda de frecuencia)
kbps	kilobits por segundo
KKT	Karush-Kuhn-Tucker
MA	margin adaptive
MAC	medium access control
MIMO	multiple-input multiple-output
MIP	mixed integer programming
MQAM	M quadrature amplitude modulation
MPSK	M phase-shift keying
MRC	maximal ratio combining
MS-GSC	minimum-selection generalized selection combining
OFDM	orthogonal frequency division multiplexing
OFDMA	orthogonal frequency division multiple access
PAN	personal area network
PUSC	partial usage of subcarriers
PHY	physical
RLP	rounded linear programming
SAP	subchannel allocation problem
SBPA	subchannel, bit and power allocation
SDR	software defined radio
SC	selective combining
SER	symbol error rate
SINR	signal-to-interference-plus-noise ratio
SNR	signal-to-noise ratio
S-SBPA	suboptimal-subchannel, bit and power allocation
TDMA	time division multiple access
TDMA-F	time division multiple access-fijo
TDMA-V	time division multiple access-variable

UMTS	Universal Mobile Telecommunications System
VDSL	Very high bit-rate Digital Subscriber Line
WiMAX	Worldwide Interoperability for Microwave Access
WSN	wireless sensor network

CAPÍTULO 1

INTRODUCCIÓN

En los últimos años hemos asistido a la integración de los servicios de telecomunicación en la sociedad, dando lugar a lo que conocemos actualmente como Sociedad de la Información. Por ejemplo, podemos mantener una conversación mediante la telefonía móvil mientras damos un paseo, o podemos hacerlo desde el salón de casa mediante las redes convencionales de telefonía fija. Asimismo disfrutamos de un gran número de servicios para el envío de todo tipo de información (correo electrónico, mensajes SMS y MMS, difusión de señal de televisión para terminales DVB o mediante *streaming* en Internet, o el acceso a información vía web, por citar algunos ejemplos), los cuales pretendemos que estén disponibles en cualquier lugar y en el mismo instante en que lo deseamos. Y éste es precisamente el punto clave hoy día: hemos hecho inherentes a las comunicaciones las características de ubicuidad e inmediatez.

Si sólo dispusiéramos de medios de transmisión guiados, tales como el cable coaxial y la fibra óptica, nuestras aspiraciones a estar siempre conectados (el conocido *always on* en terminología anglosajona) se vendrían abajo, dada la evidente (por costosa) imposibilidad de extender cables por dondequiera que pueda encontrarse un usuario con intención de utilizar un servicio de telecomunicación. Es fácil deducir por tanto la gran importancia para las telecomunicaciones actuales y futuras que tienen los medios no guiados o inalámbricos cuando el objetivo es proporcionar servicios de telecomunicación con la mayor cobertura y calidad

posibles. Así, mediante la utilización del espectro radioeléctrico como medio de transmisión, los operadores de telecomunicaciones pueden realizar el despliegue de redes en plazos sustancialmente más cortos y con menores costes que en el caso de redes guiadas.

No obstante, los canales inalámbricos presentan una serie de desventajas que hay que tener en cuenta y que representan retos muy atractivos desde el punto de vista de los sistemas de comunicaciones. En primer lugar, el espectro radioeléctrico disponible para telecomunicaciones es escaso, ya que el número de bandas de frecuencia adecuadas para transmisiones vía radio es bastante limitado. Por tanto, es muy importante utilizar el espectro de la manera más eficiente posible, lo cual se puede conseguir mediante transmisión y modulación adaptativas. En segundo lugar, el medio radioléctrico es muy hostil para las comunicaciones, puesto que coexisten no sólo las transmisiones correspondientes a diferentes usuarios y sistemas, sino también ruidos provenientes de todo tipo de fuentes, como máquinas industriales o radiaciones solares en el caso de los satélites. Y en tercer lugar, es extremadamente dinámico en cuanto a las condiciones de transmisión para los usuarios, comparado con los medios guiados, de manera que, aunque se contemplen los dos factores anteriores, se han de diseñar soluciones que respondan adecuadamente a los cambios en el canal.

A la vista de lo anterior, esta tesis plantea propuestas para ciertos entornos de comunicaciones inalámbricas con el objetivo de, por un lado, optimizar los recursos disponibles en cada caso y, por otro lado, cumplir con unos compromisos de calidad prefijados.

1.1. Transmisión adaptativa en redes inalámbricas

La creciente demanda por parte de los usuarios de servicios de telecomunicación de banda ancha ha motivado la aparición de estándares para redes inalámbricas, cada uno de ellos enfocado a un área de cobertura de un tamaño determinado; por ejemplo, el estándar 802.16 está dirigido a comunicaciones inalámbricas de área extensa, llegando a cubrir hasta decenas de km. Estas redes, sumadas a la madurez de las redes móviles de tercera generación UMTS, han convertido el entorno de transmisión para comunicaciones inalámbricas en un reto aún mayor, ya que cada uno de los estándares impone sus propias especificaciones en cuanto a la transmisión.

La transmisión adaptativa es una técnica muy adecuada para hacer frente a entornos radioeléctricos complejos, ya que se caracteriza por establecer dinámicamente los valores de los parámetros de transmisión, como la potencia o los bits por símbolo, en función de las circunstancias del medio radioeléctrico en el instante de realizar la transmisión. Como veremos en el capítulo 5, es una técnica muy versátil que permite su aplicación tanto a sistemas de banda ancha, como las redes basadas en OFDMA, como a sistemas con menores necesidades de transmisión, como las redes de sensores.

Intimamente ligada a la transmisión adaptativa, la asignación de los recursos radioeléctricos permite dedicar a cada comunicación del sistema la cantidad de recursos adecuados. Los recursos radio considerados habitualmente son el ancho de banda y la potencia de transmisión, pero también pueden establecerse como criterio para la asignación otros recursos como los bits por símbolo transmitidos o el tiempo destinado a la transmisión. El objetivo último de asignar los recursos radio es, obviamente, hacerlo de la manera más beneficiosa posible, y es aquí donde existe un amplio espectro de posibilidades. Entre ellas figuran la maximización de la capacidad, de la duración de la red o de la tasa de transmisión; la minimización de la interferencia, la potencia de transmisión o la energía consumida, y un largo etcétera.

1.2. Objetivos

A lo largo de los capítulos de esta tesis vamos a estudiar las posibilidades de aplicar las técnicas de transmisión adaptativa a nuevos problemas planteados en el contexto de los sistemas de comunicaciones inalámbricas.

En redes con limitaciones de energía, uno de los objetivos prioritarios, si no el que más, es la minimización de la misma. La mayor parte de estudios realizados en este ámbito sólo tiene en cuenta una de las capas de la torre OSI para conseguir este objetivo, y aquellos que toman en consideración varias capas normalmente no están enfocados a redes centralizadas. Nuestra propuesta de minimización de energía contempla de manera conjunta parámetros de capa física y capa de enlace para dinámicamente adaptar la transmisión mediante modulación adaptativa. A este efecto, establecemos un modelo de minimización de energía para redes inalámbricas centralizadas en general, que luego es particularizado a las redes de sensores. Para realizar la modulación adaptativa, hacemos uso de la aproximación

SNR *gap*.

Para redes centralizadas de aplicaciones orientadas a comunicaciones libres de errores, existe la motivación de optimizar la asignación de recursos basándose en el concepto de *goodput*, que es una medida de los bits transmitidos sin errores. Este tipo de aplicaciones ha sido objeto de muy pocos estudios en entornos de comunicaciones muy extendidos hoy día, como las redes basadas en OFDMA, por lo que proponemos un método de baja complejidad que asigna los recursos radio elegidos (subcanal, potencia y bits por símbolo) para este tipo de redes.

1.3. Organización y contribuciones

1.3.1. Metodología

En esta tesis, para el diseño de cada una de las soluciones aportadas, hemos realizado en primer término un modelado del sistema bajo estudio y una formulación adecuada del problema a resolver. Seguidamente, se proponen uno o varios métodos teóricos de resolución, y finalmente las propuestas son evaluadas mediante simulaciones.

Dado que las limitaciones en la capacidad de procesamiento de los dispositivos y equipos pueden suponer una restricción a la implementación práctica de sistemas de comunicaciones, en los distintos estudios realizados hemos tenido muy presente el grado de complejidad que implican. En esta tesis asociaremos la complejidad a la carga computacional que implica la ejecución de los algoritmos [Bertsimas y Weismantel, 2005], de forma que será alta si el tiempo de ejecución de los algoritmos es exponencial y baja si es polinomial.

Hacemos notar que los diseños propuestos involucran más de una capa de la torre OSI, concretamente las capas física y de enlace. Esta metodología de diseño proporciona mejores resultados a nivel global del sistema que el diseño capa a capa, dado que se optimizan funciones basadas en parámetros de ambas capas, en lugar de optimizar una capa de manera aislada.

1.3.2. Organización

La tesis está dividida en tres partes diferenciadas: los capítulos 2 y 3, el capítulo 4, y el capítulo 5, con la organización que se indica en la figura 1.1. La primera parte, correspondiente como se ha dicho a los capítulos 2 y 3, presenta las herramientas

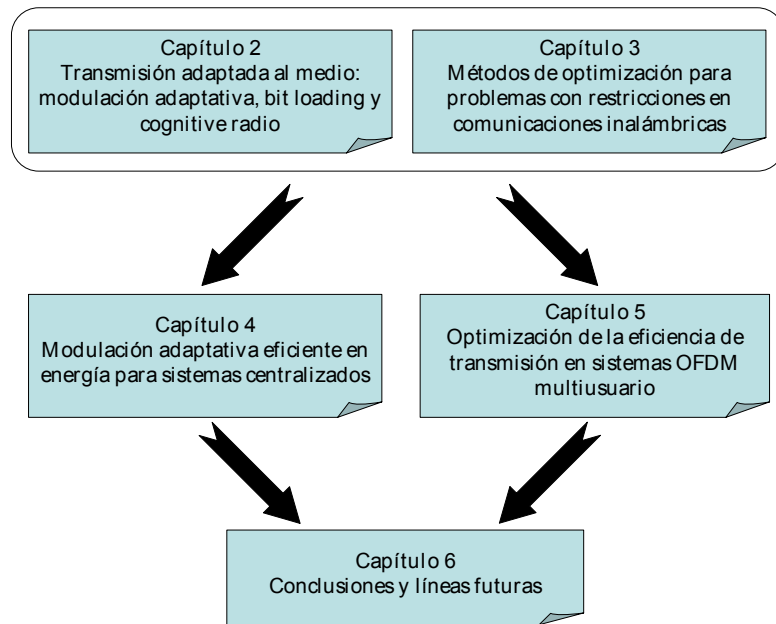


Figura 1.1. Relación entre los capítulos

disponibles para desarrollar y resolver los métodos que se proponen en la tesis. El capítulo 2, que es la base fundamental de los capítulos 4 y 5, está dedicado a la transmisión adaptativa y contiene asimismo una revisión de las diferentes técnicas adaptativas para enlaces punto a punto y punto-multipunto (o multiusuario). El capítulo 3 describe las herramientas matemáticas de las que nos serviremos en los dos siguientes capítulos para resolver problemas de optimización, tanto en variable entera como en variable continua.

El capítulo 4 trata la optimización de la energía en sistemas inalámbricos, la cual resolvemos mediante un esquema adaptativo que calcula la duración del intervalo de transmisión para cada usuario de forma que se cumplan unos requisitos de calidad en cuanto a tasa binaria de transmisión y tasa de error de símbolo. Este esquema es de especial interés en redes de sensores, donde el consumo de energía es vital para la disponibilidad de la red, dado que un uso inadecuado de la energía ocasiona un rápido agotamiento de las baterías y por tanto la indisponibilidad de la red, cuando menos de manera parcial.

En el capítulo 5 presentamos un esquema para la optimización de la eficiencia de transmisión (o *goodput*) de una estación base en sistemas inalámbricos, orientado a aplicaciones con restricciones en la tasa de información entregada sin errores. El método propuesto da lugar a algoritmos óptimos y subóptimos, donde la complejidad adquiere una alta relevancia a la hora de la implementación en sistemas

prácticos.

Finalmente, dedicamos el capítulo 6 a la extracción de las conclusiones generales y al establecimiento de líneas futuras de investigación.

Se han añadido dos anexos: en el anexo A se proporcionan las expresiones matriciales de cálculo diferencial necesarias para el desarrollo del capítulo 2, y en el anexo B se enumeran, desglosadas por capítulos, las publicaciones derivadas de las diferentes contribuciones que constituyen esta tesis.

1.3.3. Contribuciones

A grandes rasgos, describimos a continuación las contribuciones de la tesis, de las que se proporciona un mayor detalle en la introducción de los correspondientes capítulos, donde quedan contextualizadas.

Por una parte, se contribuye con un estudio de las diferentes técnicas de transmisión adaptativa y su utilización en diferentes ámbitos de los sistemas de comunicaciones, y un análisis de los métodos matemáticos para resolver problemas de optimización en variable entera y variable continua y sus diversas aplicaciones a los sistemas inalámbricos; este análisis se completa con el teorema 3.1 y el lema 3.1.

Esta tesis presenta una formalización de la minimización de la energía en sistemas inalámbricos, donde aporta, en primer lugar, un modelo adecuado del sistema y una formulación del problema de optimización de la energía. En segundo lugar, se expone un esquema de transmisión adaptativa basado en la duración variable de los intervalos de transmisión de cada nodo y en la aproximación SNR *gap*, que permite relacionar los parámetros tasa de error de símbolo (*symbol error rate*, SER) y tasa binaria que condicionan los requisitos mínimos de los nodos. Particularizado para redes de sensores, vemos que este esquema permite ahorros de energía significativos frente a la utilización de técnicas convencionales de acceso múltiple TDMA.

Por último, la tesis contribuye con un método para optimizar la eficiencia de transmisión, entendida como el número de bits de información correctamente transmitidos por unidad de tiempo, en el enlace descendente de sistemas de comunicaciones centralizados. El problema de optimización se formula como un problema de asignación de recursos para conseguir la mayor eficiencia de transmisión posible. Dado que el problema está formulado en variable entera, el método propuesto reduce la complejidad significativamente respecto a la utilización de

algoritmos de optimización entera estándar. Una vez aplicado este método, se propone un algoritmo subóptimo que nuevamente reduce la complejidad y consigue que no exista una degradación significativa de las prestaciones con respecto al óptimo.

CAPÍTULO 2

TRANSMISIÓN ADAPTADA AL MEDIO: MODULACIÓN ADAPTATIVA, BIT LOADING Y RADIO COGNITIVA

La característica fundamental de las comunicaciones inalámbricas es la utilización del medio radioeléctrico como canal de transmisión. Este canal inalámbrico aparenta cambiar de manera aleatoria en el tiempo debido, entre otras causas, a la adición de ruido e interferencias sobre la señal transmitida y a los efectos de reflexión y atenuación que dicha señal pueda experimentar en su trayectoria hasta el receptor. La utilización de modulaciones fijas en este tipo de canales conlleva, por ejemplo, que la tasa de error de bit (*bit error rate*, BER) varíe según lo haga la calidad del canal de transmisión, e incluso sea inferior a la BER objetivo que se debe proporcionar. Por el contrario, la adaptación de la modulación a las condiciones del canal permitiría mantener una BER constante independientemente de las buenas o malas condiciones del canal de transmisión. Esta y otras ventajas se derivan de la utilización de la transmisión adaptativa, a la cual dedicaremos el presente capítulo en el contexto de sistemas inalámbricos punto a punto y multiusuario. Dado que las técnicas a emplear varían considerablemente según nos encontremos en sistemas de portadora única o no, vamos a diferenciar dos bloques principales en el capítulo: uno dedicado a las modulaciones adaptativas propia-

mente dichas, y otro dedicado a las técnicas de transmisión adaptativa en sistemas multiportadora.

2.1. Modulación adaptativa: conceptos básicos

La modulación adaptativa, también denominada en algunas referencias como transmisión adaptativa, consiste en adecuar la modulación a emplear por el transmisor a las condiciones del canal variante en el tiempo. La finalidad es hacer la transmisión de manera robusta y espectralmente eficiente, mediante la variación de los parámetros asociados a la transmisión, bien de manera individual, bien de forma conjunta. Dichos parámetros son la potencia de transmisión, la tasa de transmisión, el tamaño de la constelación de la modulación utilizada, la codificación de canal y la tasa de error de bit. Para ajustar convenientemente estos parámetros, es básico realizar una estimación fiable del canal y hacer llegar dicha estimación al transmisor, para que adopte el esquema más conveniente a las condiciones del canal, asignando a los parámetros anteriores los valores adecuados.

Frente a los esquemas de modulación fija, que son diseñados típicamente para hacer frente al caso peor, la modulación adaptativa puede conseguir, por ejemplo, reducir la BER, incrementar la tasa media de transmisión de bit o reducir la potencia necesaria de transmisión; para ello, si las condiciones del canal son favorables, se podrán enviar más bits por símbolo o reducir la potencia, y si el canal se ha degradado, se podrá reducir la tasa de bits por símbolo o incrementar la potencia de transmisión.

En este capítulo, nos centraremos en la modulación adaptativa basada en la variación de los parámetros potencia, nivel de la constelación, tasa de transmisión y probabilidad de error. Por tanto, la adaptación utilizando codificación de canal, aunque es una potente herramienta en este tipo de sistemas (véase, por ejemplo [Song et al., 2004] y [Caire y Kumar, 2007]), no será objeto de estudio en esta tesis.

Existe una serie de factores que determinan la idoneidad o no de usar esquemas de modulación adaptativa, pero hay tres que resultan básicos. En primer lugar, el receptor ha de realizar la estimación del canal, y esta estimación tiene que ser lo más fiable posible. En el caso óptimo, el receptor tendrá un conocimiento perfecto del canal. En segundo lugar, esta información ha de hacerse llegar al transmisor a través de lo que se denomina el canal de realimentación o de *feedback*. Dicho

canal debe ser, al menos en teoría, un canal libre de errores que no altere la estimación enviada. El problema está en que no siempre es posible establecer este canal de *feedback*, lo que hace inviable la implementación de la modulación adaptativa en algunos sistemas. En último lugar, la efectividad de la adaptación está sujeta a la rapidez con que cambia el canal: si el tiempo empleado para hacer llegar la estimación de canal al transmisor y realizar la adaptación misma es superior al tiempo que tarda el canal en cambiar, evidentemente la adaptación de la transmisión proporcionará un pobre rendimiento.

2.1.1. Revisión histórica de los sistemas de modulación adaptativa

A pesar de que la modulación adaptativa es conocida desde hace varias décadas [Hayes, 1968], su aplicación a las comunicaciones inalámbricas no adquiere cierta relevancia hasta mediados de la década de 1990, cuando empiezan a proponerse esquemas de transmisión adaptativa para comunicaciones móviles (ver, por ejemplo, [Webb y Steele, 1995] y [Kamio et al., 1995]). Es, sin embargo, a partir de los inicios del presente siglo cuando se hace patente el interés por la modulación adaptativa, de forma paralela al auge de las comunicaciones inalámbricas y en especial de las comunicaciones móviles de tercera generación (3G), de forma que queda incorporada a estándares como 1xEV-DO [Bhushan et al., 2006], GPRS-136 [Nanda et al., 2000] y EGPRS [Furuskar et al., 1999], o empleada en redes de acceso de área local (familia de estándares IEEE.802.11x) y de área extensa (familia de estándares IEEE 802.16). En esta subsección nos centraremos en la modulación adaptativa en general, dejando para la sección 2.2 la transmisión adaptativa en sistemas multiportadora.

Entre la literatura al respecto algunos trabajos esenciales recogen un conjunto de referencias lo suficientemente completo en determinadas temáticas. En [Svensson, 2007] se ilustran las ventajas de la modulación adaptativa frente a las modulaciones fijas cuando se emplea modulación QAM multinivel, para canal con desvanecimiento plano, tanto para el caso de conocimiento perfecto del canal como para predicción de canal, y proporciona una selección de contribuciones en esquemas de modulación adaptativa hasta 2004. En el cap.9 de [Goldsmith, 2005], que contiene una visión de conjunto de la modulación adaptativa, se citan referencias significativas que tratan la aplicación de esquemas adaptativos en comunicaciones móviles.

La utilización de combinación por diversidad en recepción (*diversity combi-*

ning) es una técnica que permite mejorar las prestaciones de los sistemas de transmisión adaptativos, de forma que se aborda el diseño conjunto del sistema en transmisión y recepción, como se muestra en [Alouini y Goldsmith, 1999]. En dicha referencia, se analiza la capacidad obtenida en canales con desvanecimiento Rayleigh con esquemas de diversidad en recepción MRC (*maximal ratio combining*) y SC (*selection combining*), empleando diferentes técnicas de modulación adaptativa. El interés por el estudio conjunto de ambas técnicas, adaptativas y combinación por diversidad, ha resurgido recientemente entre la comunidad científica que estudia los sistemas inalámbricos adaptativos, lo que ha dado lugar a novedosas propuestas. Los trabajos de [Yang et al., 2007] y [Ko et al., 2007] siguen la línea de [Alouini y Goldsmith, 1999]. El primero propone la combinación de modulación adaptativa con un nuevo esquema de combinación por diversidad denominado *minimum-selection generalized selection combining* (MS-GSC): mientras GSC combina un número fijo de las mejores señales recibidas, MS-GSC está orientado a sistemas eficientes en potencia. El segundo propone una variación de MRC para sistemas con adaptación de la transmisión. En otra línea, en [Nam et al., 2008] los autores presentan un sistema de conmutación de antenas con modulación adaptativa, para receptores multiantena, que aprovecha la correlación temporal entre los canales observados por las diferentes antenas para disminuir tanto la información requerida para la estimación de canal como la frecuencia de conmutación entre antenas. De esta forma, se reduce considerablemente la complejidad de procesamiento respecto a otros esquemas previos de conmutación de antenas en recepción, y mediante la modulación adaptativa por antena se dota al sistema de un cierto grado de libertad en la selección de antena.

2.1.2. El sistema de transmisión adaptativa

Para llevar a cabo la adaptación de la modulación es necesario incorporar nuevos elementos respecto a un sistema con modulación fija, algunos de los cuales ya han sido introducidos en la sección 2.1. La figura 2.1 muestra un diagrama con los bloques más importantes que han de figurar en todo esquema de modulación adaptativa, y que trataremos brevemente a continuación. Una vez realizada la transmisión a través del canal inalámbrico, se realiza en el receptor la estimación del canal, y estos estimadores se hacen llegar al transmisor a través del canal de *feedback*. Cuando el transmisor dispone de esta información, puede adaptar la transmisión de forma consecuente. No obstante, si el tiempo de procesamiento

de la señal (estimación, adaptación de la transmisión) más el retardo asociado al canal de *feedback* es elevado, los estimadores quedan desactualizados, por lo que la adaptación no será la adecuada. En este caso, y en general cuando el canal no es perfectamente conocido, el sistema puede incluir un bloque para la predicción del canal, que a partir de los valores de los estimadores pasados y actuales, conforma un modelo de canal y predice los parámetros a emplear para la transmisión adaptada [Falahati et al., 2004].

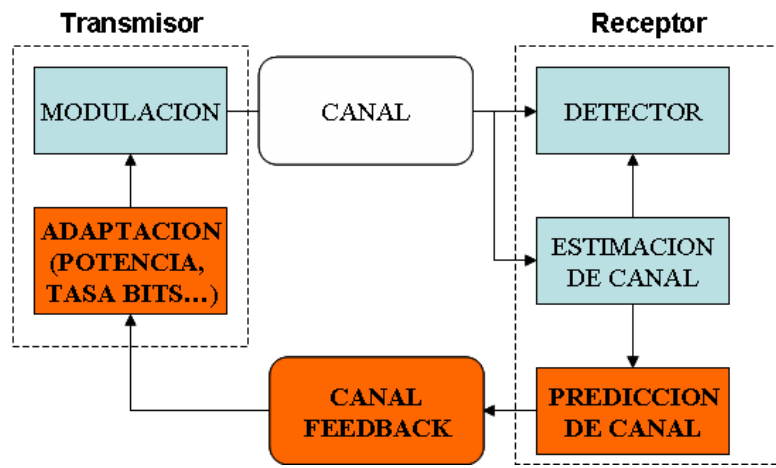


Figura 2.1. Diagrama de bloques del sistema de transmisión adaptativa

En nuestro sistema, asumiremos que el canal de transmisión inalámbrico es modelado como un canal discreto en el tiempo con desvanecimiento plano. El canal tiene una ganancia de potencia $h[i] = |g[i]|^2$ en el instante i , siendo $g[i]$ una variable aleatoria estacionaria y variante en el tiempo con una función de densidad de probabilidad $f(h)$, y $n[i]$ representa el ruido aditivo blanco gaussiano (AWGN, Additive White Gaussian Noise) con densidad espectral de potencia $N_0/2$. La potencia media de la señal transmitida viene representada por \bar{P} , B es el ancho de banda de la señal, y denotaremos por SNR (*signal-to-noise ratio*) a la relación señal a ruido. Si transmitimos con potencia constante \bar{P} , la SNR instantánea es $\bar{\gamma}[i] = \bar{P}h[i]/N_0B$, mientras que denotaremos por $\gamma[i]$ a la SNR instantánea obtenida cuando se transmite con una potencia cualquiera distinta de \bar{P} . Como la variable $g[i]$ es estacionaria, la función de densidad de probabilidad de $\bar{\gamma}[i]$ es independiente de i y se puede denotar por $f(\gamma)$.

En general, los parámetros más comunes a través de los cuales se adapta la modulación son la potencia instantánea de transmisión $P(\gamma[i])$, la tasa de

transmisión de bits por segundo $R(\gamma[i])$ y la codificación de canal. Dado que en esta tesis se abordarán esquemas adaptativos sin codificación, nos centraremos en los dos primeros parámetros. Para una potencia $P(\gamma[i])$, la SNR instantánea vendrá expresada como $\gamma[i] = P(\gamma[i])h[i]/N_0B$; por tanto, se deduce que $P(\gamma[i]) = \frac{\gamma[i]}{\gamma[i]} \bar{P}$. Para una modulación de nivel M , la tasa de transmisión se calcula como $R(\gamma[i]) = B \log_2 M(\gamma[i])$ (bps), siendo $M(\gamma[i])$ el nivel de la modulación en el instante i y $m(\gamma[i]) = \log_2 M(\gamma[i])$ (bits/símbolo) el número de bits por símbolo transmitido. La eficiencia espectral de dicha modulación la calculamos como $R(\gamma[i])/B = \log_2 M(\gamma[i])$ (bps/Hz). En adelante, una vez queda clara la dependencia temporal de los parámetros, omitiremos la referencia temporal de γ con respecto al índice i .

Para nuestro sistema, haremos uso de modulación multinivel (MQAM y MPSK) con codificación Gray de los bits. En este caso, para canal con ruido AWGN, es muy habitual aproximar la tasa de error BER mediante la expresión [Chung y Goldsmith, 2001]

$$\text{BER}(\gamma) \approx c_1 \exp \left[\frac{-c_2 \gamma \frac{P(\gamma)}{P}}{2c_3 m(\gamma) - c_4} \right], \quad (2.1)$$

siendo c_1 , c_2 , c_3 y c_4 constantes positivas fijas cuyo valor es función del tipo de modulación empleada.

2.1.3. Técnicas básicas de modulación adaptativa

Las técnicas de modulación adaptativa más sencillas se basan en la adaptación con respecto a un único parámetro. Revisaremos por tanto, en primer lugar, la adaptación empleando sólo la tasa de transmisión como parámetro variable y la modulación adaptativa variando la potencia de transmisión.

- Adaptación mediante la tasa binaria de transmisión. La tasa de transmisión $R(\gamma)$ se adapta con respecto a la ganancia del canal h , o equivalentemente la SNR γ , para cumplir con una tasa de error BER objetivo. Para conseguirlo, bien cambiamos la modulación utilizada (por ejemplo, pasando de 8PSK a 8QAM) o bien variamos el nivel de la modulación $m(\gamma)$.
- Adaptación mediante variación de la potencia de transmisión. El objetivo es, habitualmente, mantener una BER fija o lo que es lo mismo, una SNR constante, que denotaremos por γ_0 . Sabemos que

$$P(\gamma) = \frac{\gamma_0 \bar{P}}{\gamma}, \quad (2.2)$$

y si queremos transmitir empleando toda la potencia media \bar{P} , habrá de cumplirse que

$$\int P(\gamma) p(\gamma) d\gamma = \bar{P}. \quad (2.3)$$

Combinando las ecuaciones 2.2 y 2.3, tenemos que $\gamma_0 = 1/E[1/\gamma]$, luego si el valor de γ es tal que $\gamma_0 > 1/E[1/\gamma]$, no se podrá cumplir con la tasa de error objetivo.

2.1.4. Optimización de la eficiencia espectral mediante modulación adaptativa

Una de las principales motivaciones para emplear modulación adaptativa es su idoneidad para utilizar el espectro de la manera más eficiente posible, lo que ha dado lugar a técnicas de modulación adaptativa basadas en la optimización de la eficiencia espectral $R(\gamma)/B$. En esta sección adoptaremos este enfoque, tomando como base el análisis realizado al respecto en [Goldsmith, 2005].

Estas técnicas, al contrario que las descritas en (2.1.3), no se basan en variar un único parámetro, sino que precisan de la adaptación conjunta de la tasa de bits por símbolo $m(\gamma)$ y la potencia $P(\gamma)$, manteniendo una BER dada. En el caso más general, existirán dos restricciones: primero, la potencia media de transmisión no puede ser mayor que \bar{P} ; segundo, imponemos una tasa de error de bit que ha de cumplirse. Formulando matemáticamente, el problema de optimización a resolver es

$$\begin{aligned} & \underset{m(\gamma), P(\gamma)}{\text{máx}} \quad \frac{R(\gamma)}{B} \\ & \text{s.t.} \quad \int_0^{\infty} P(\gamma) f(\gamma) d\gamma \leq \bar{P} \\ & \quad \text{BER}(\gamma) = \text{BER}_0, \end{aligned} \quad (\text{P-MA})$$

donde BER_0 es la tasa de error de bit requerida en la transmisión. Acerca de este problema, hay que hacer a continuación dos consideraciones:

- La adaptación se puede formular de manera continua o discreta. En el caso continuo, el conjunto de constelaciones de señales no está restringido, lo que implica que se admiten valores de M no enteros, y la eficiencia espectral viene dada por

$$\frac{R(\gamma)}{B} = \int_0^{\infty} m(\gamma) f(\gamma) d\gamma \quad (2.4)$$

Para adaptación discreta, sólo se dispone de un número discreto de constelaciones N , de manera que $M = \{1, \dots, N\}$ y el rango de valores de γ estará dividido en N regiones, a cada una de las cuales le corresponderá un valor del número de bits por símbolo m . Así, se empleará la constelación i -ésima, con m_i bits por símbolo, si $\gamma_i \leq \gamma \leq \gamma_{i+1}$. La eficiencia espectral sigue ahora la expresión

$$\frac{R(\gamma)}{B} = \sum_{k=0}^{N-1} m_k(\gamma) \int_{\gamma_{k+1}}^{\gamma_k} f(\gamma) d\gamma. \quad (2.5)$$

- La restricción en la BER asociada a la modulación se puede plantear en términos de BER instantánea o de BER promedio. En el primer caso, el valor BER_0 ha de mantenerse constante para todo valor de γ , es decir, $BER(\gamma) = BER_0, \forall \gamma$; en otras palabras, ha de cumplirse para cada símbolo transmitido. En el segundo caso, el valor de la BER se obtiene promediando la tasa de error de bit en el tiempo y se emplea habitualmente como métrica de rendimiento en canales con desvanecimiento plano. Definimos la probabilidad de error promedio \overline{BER} como

$$\overline{BER} = \frac{E[\text{número de bits erróneos por transmisión}]}{E[\text{número de bits por transmisión}]}, \quad (2.6)$$

quedando la restricción relativa a probabilidad de error en la forma $\overline{BER} = BER_0$. Esta BER promedio, en función de si la adaptación se formula de forma continua o discreta responde, respectivamente, a las siguientes expresiones:

$$\overline{BER} = \frac{\int_0^\infty BER(\gamma) m(\gamma) f(\gamma) d\gamma}{\int_0^\infty m(\gamma) f(\gamma) d\gamma} \quad (2.7)$$

$$\overline{BER} = \frac{\sum_{k=0}^{N-1} m_k(\gamma) \int_{\gamma_{k+1}}^{\gamma_k} BER(\gamma) f(\gamma) d\gamma}{\sum_{k=0}^{N-1} m_k(\gamma) \int_{\gamma_{k+1}}^{\gamma_k} f(\gamma) d\gamma}. \quad (2.8)$$

El problema de maximizar la eficiencia espectral (P-MA) ha sido resuelto para modulaciones multinivel en [Chung y Goldsmith, 2001] para los tipos de adaptación y de BER descritos en los párrafos previos, aplicando el método de los multiplicadores de Lagrange (ver sección 3.2.1). Los resultados obtenidos ponen de manifiesto que las diferencias en la eficiencia espectral son despreciables entre los cuatro casos anteriores, considerando adaptación continua o discreta y BER instantánea o promedio. En este mismo trabajo, se propone también optimizar la eficiencia espectral pero disminuyendo el número de grados de libertad; en concreto, se analizan los casos de potencia de transmisión constante ($P(\gamma) = P$) y

tasa de bits por símbolo constante ($m(\gamma) = m$). En ambos casos, si bien la complejidad asociada a la adaptación de la transmisión es menor, existe una pérdida no despreciable respecto a realizar adaptación conjunta de potencia y tasa de bits por símbolo.

2.1.5. Otras métricas para la evaluación de esquemas de modulación adaptativa

Existen otras alternativas a la evaluación de los esquemas de modulación adaptativa en términos de la eficiencia espectral o tasa de transmisión. Para ciertas aplicaciones, la restricción crítica no es obtener mayor o menor tasa de transmisión según las condiciones del canal sean más o menos favorables, sino que el retardo máximo de la transmisión está limitado, siendo la tasa de transmisión fija. Este es el caso, por ejemplo, de las aplicaciones multimedia en tiempo real. Esta situación se aborda en [Kamath y Goeckel, 2004], donde utilizando adaptación del nivel de la modulación QAM y con limitación por retardo, se minimiza la probabilidad de corte, habitualmente conocida en el ámbito de comunicaciones como *outage probability*. Relacionado con el trabajo anterior, en [Conti et al., 2007] los autores proponen la misma clase de adaptación para tratar las variaciones en la calidad de servicio, basada en la probabilidad de que la tasa de error de bit supere un determinado umbral (lo que en la literatura especializada se conoce como *bit-error outage probability*). A diferencia de la mayor parte de propuestas de transmisión adaptativa que consideran únicamente los desvanecimientos rápidos, este esquema contempla las variaciones lentas del canal, como pueden ser las debidas a la existencia de zonas de sombra de la señal.

2.2. Modulaciones multiportadora: OFDM

La idea básica de las modulaciones multiportadora es dividir un canal de banda ancha en canales de banda estrecha ortogonales, a los que en la literatura al respecto se les denomina generalmente **subportadoras**, de forma que el flujo de bits a transmitir se reparte entre dichos canales de banda estrecha [Bingham, 1990]. Como consecuencia de la ortogonalidad entre las subportadoras, el canal de banda ancha es equivalente a múltiples canales paralelos de banda estrecha. El número de subportadoras se elige de manera que experimenten desvanecimiento plano en frecuencia, lo que se consigue si el ancho de banda de cada subportadora es lo

suficientemente pequeño con respecto al ancho de banda de coherencia del canal.

La principal ventaja de las modulaciones multiportadora es que son muy robustas frente a la interferencia entre símbolos (*intersymbolic interference*, ISI) que aparece en las comunicaciones de banda ancha con multitrayecto. En sistemas inalámbricos con dispersión de retardo apreciable, mientras los esquemas multiportadora eliminan la ISI de manera sencilla, para los esquemas monoportadora es necesaria una ecualización adicional que requiere un filtrado con un elevado número de etapas, y por tanto ecualizadores bastante complejos. Además, es difícil conseguir que los coeficientes asociados a las etapas del ecualizador sean exactos cuando el canal cambia rápidamente. Así, se ha generalizado la incorporación de las modulaciones multiportadora a estándares de comunicaciones, tanto inalámbricas (entre otros, los estándares IEEE 802.11a/g) como a través de medios guiados (por ejemplo, el estándar ADSL).

La multiplexación por división en frecuencias ortogonales OFDM (*orthogonal frequency division multiplexing*) es una técnica de modulación mutiportadora que ha sido adoptada como esquema de modulación para sistemas de comunicaciones inalámbricas en todo tipo de escenarios, debido fundamentalmente a su robustez frente al desvanecimiento multitrayecto para comunicaciones de banda ancha. Desde el punto de vista de la movilidad del usuario, está presente en estándares tanto de redes fijas (por ejemplo IEEE 802.11a/g y IEEE 802.16-2004) como de redes móviles (IEEE 802.16e-2005 y 3GPP-LTE). Desde el punto de vista de las aplicaciones, la difusión de televisión digital terrestre (DVB-T) y audio digital (DAB) también la contemplan como la modulación a utilizar.

Aunque los fundamentos teóricos de OFDM se fraguan en la década de 1960 [Chang, 1966], la complejidad de su implementación mediante bancos de osciladores y moduladores hacía que su aplicación práctica no resultara viable. La reformulación teórica de OFDM basada en la transformada discreta de Fourier (*discrete Fourier transform*, DFT) [Weinstein y Ebert, 1971] permitía reducir significativamente la complejidad de los moduladores/demoduladores OFDM. No obstante, hubo que esperar a la aparición de dispositivos económicos que proporcionasen una realización sencilla de la DFT requerida para abordar el diseño de moduladores para su producción a gran escala. Es a partir de este momento, al inicio de la década de 1990, cuando el interés por OFDM se generaliza dando lugar a multitud de propuestas que cristalizan, entre otros, en la inclusión de OFDM como esquema de modulación de los estándares mencionados en el anterior párrafo.

Para hacer referencia a esta modulación multiportadora, existen dos términos. Se suele emplear DMT (*Discrete MultiTone*) para sistemas que utilizan medios de transmisión guiados, como cable coaxial o par de cobre, y OFDM para sistemas inalámbricos que usan el medio radioeléctrico. La razón de esta diferenciación obedece a que en sus inicios las modulaciones multiportadora por medios guiados utilizaban modulación adaptativa en cada subportadora, y en OFDM todas las subportadoras estaban moduladas de igual manera. No obstante, hoy día, como se verá a lo largo de este capítulo, se ha generalizado el uso de la modulación adaptativa en OFDM de tal manera que esta diferenciación tiene sentido únicamente en términos históricos.

2.2.1. Fundamentos básicos de OFDM

Como ya hemos adelantado, OFDM consiste en transformar un canal de ancho de banda B en N subportadoras paralelas de ancho de banda B/N , cada uno de los cuales lleva asociada una portadora de frecuencia f_k , $k = 0, \dots, N - 1$. La separación entre portadoras consecutivas es $\Delta f = B/N$, de manera que, si f_0 es la frecuencia más baja de las portadoras consideradas, el resto de valores de las portadoras se obtienen mediante la expresión $f_k = f_0 + k\frac{B}{N}$.

Existen, según mencionamos en la introducción de esta sección, dos enfoques teóricos posibles para OFDM: en tiempo continuo y en tiempo discreto. Cada uno de ellos da lugar a una implementación completamente distinta, como veremos a continuación.

En el caso continuo, el modulador responde al esquema representado en la figura 2.2, donde X es una secuencia de N símbolos complejos y X_0, \dots, X_{N-1} son los símbolos complejos resultantes de aplicar la secuencia X a un conversor serie-paralelo. El símbolo X_k modula a la portadora f_k en la rama k , siendo posteriormente filtrado y resultando la señal $s_k(t)$. La señal OFDM es la suma de las contribuciones de las N ramas del modulador: $s(t) = \sum_{k=0}^{N-1} s_k(t)$.

La formulación discreta permite simplificar de manera sustancial el *hardware* necesario para la implementación del modulador/demodulador OFDM respecto al caso en tiempo continuo y además eliminar la ISI. Se puede demostrar matemáticamente que el resultado obtenido en el caso discreto, siguiendo el esquema de la figura 2.3, equivale exactamente al que se obtiene en el caso continuo de acuerdo al esquema de la figura 2.2. Para el caso discreto, en el transmisor se realiza la transformada inversa discreta de Fourier (IFFT) de los símbolos $X[0], \dots, X[N-1]$

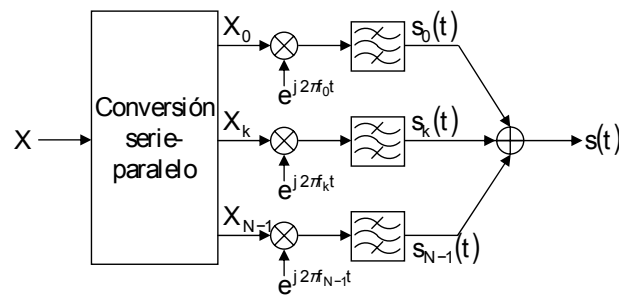


Figura 2.2. Modulador OFDM para tiempo continuo

que conforman la secuencia X , obteniéndose los N coeficientes de la secuencia $x[n] = x[0], \dots, x[N-1]$, que constituirán el símbolo OFDM y al que viene asociado un período de símbolo T_s . Estos coeficientes son transmitidos de forma secuencial, una vez que se ha añadido el prefijo cíclico, del que hablaremos en el siguiente párrafo. En el receptor, una vez discretizada la señal, se elimina el prefijo cíclico y los símbolos a demodular $Y[0], \dots, Y[N-1]$ se obtienen mediante la transformada discreta de Fourier (FFT) de las N muestras de la señal recibida $y[0], \dots, y[N-1]$.

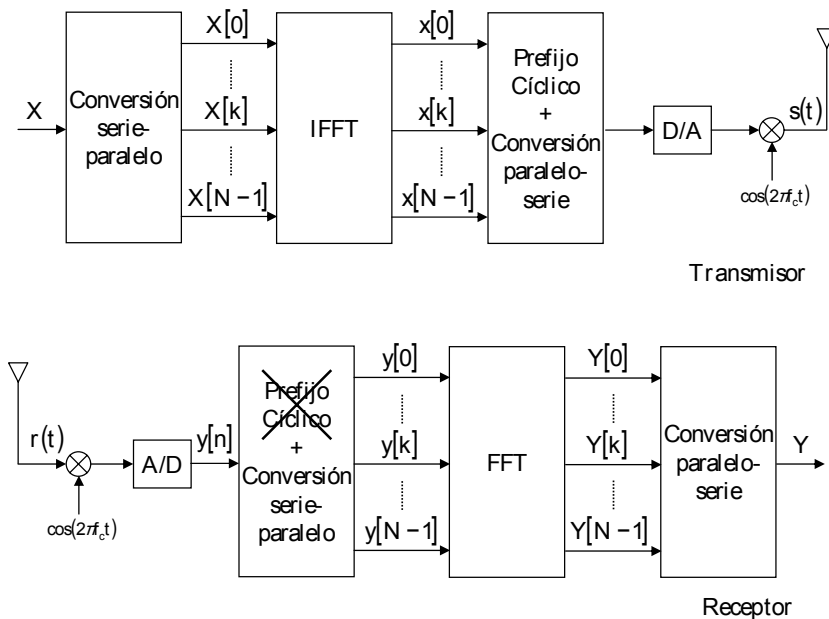


Figura 2.3. Modulador/demodulador OFDM para tiempo discreto

La adición del prefijo cíclico tiene por objeto evitar la interferencia entre símbolos OFDM sucesivos en canales dispersivos en el tiempo. Estos canales se caracterizan mediante la dispersión máxima de retardo T_m , que es la variación máxima en el retardo que pueden sufrir las señales entre el transmisor y el receptor. Para

combatir la ISI, se procede a extender el símbolo OFDM en al menos $N_g = T_m/T_s$ muestras, las cuales son una copia de las últimas N_g muestras del propio símbolo (ver figura 2.4). Por tanto, la señal OFDM $s(t)$ se formará a partir de las muestras $x[N-1-N_g], \dots, x[N-1], x[0], \dots, x[N-1]$, que constituyen la secuencia $\tilde{x}[n]$. Si modelamos el canal como un canal discreto en el tiempo y de respuesta impulsiva $h[n] = h[0], \dots, h[N_g]$, la salida del canal, si se transmite la secuencia $x[n]$, es $y[n] = x[n] * h[n]$, de manera que el símbolo $Y[k]$ es una combinación lineal de los símbolos $X[i], i = 1, \dots, N-1$, y por tanto se pierde la ortogonalidad. Ahora bien, si en vez de efectuar una convolución lineal, conseguimos que $y[n]$ sea el resultado de la convolución circular, es decir $y[n] = x[n] \otimes h[n]$, entonces $Y[k] = X[k]H[k]$, de forma que no existirá ISI y bastará con invertir el canal para obtener el símbolo original $X[k]$. Y esto se consigue si, en vez de transmitir la secuencia original $x[n]$, se transmite $\tilde{x}[n]$, ya que es inmediato comprobar, mediante propiedades de la DFT, que

$$\tilde{x}[n] * h[n] = x[n] \otimes h[n]. \quad (2.9)$$

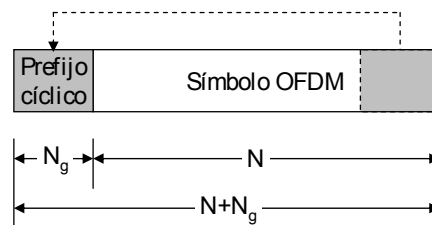


Figura 2.4. Símbolo OFDM de N muestras con prefijo cíclico de N_g muestras

2.2.2. La rejilla tiempo-frecuencia OFDM

Para el propósito de esta tesis, la forma más conveniente de interpretar un sistema OFDM es hacer uso de la rejilla bidimensional tiempo-frecuencia que se muestra en la figura 2.5. Mediante esta rejilla se disponen en el eje de tiempos los símbolos OFDM y el eje frecuencial representará los N canales paralelos obtenidos, cada uno de ellos asociado a una subportadora de frecuencia $f_k, k = 0, \dots, N-1$, y que experimentará una respuesta que será función del ancho de banda de coherencia del canal.

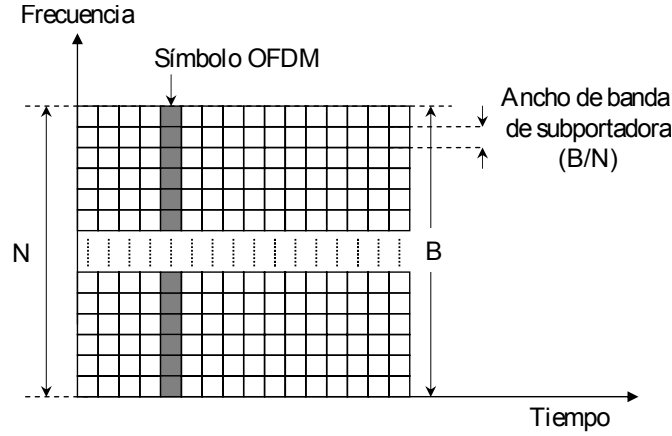


Figura 2.5. Rejilla tiempo-frecuencia de OFDM

2.3. Transmisión en sistemas multiportadora punto a punto

Una vez sentadas las bases de la modulación OFDM en (2.2.1), en esta sección discutiremos la forma de transmitir los bits a través de las subportadoras ortogonales que forman parte del sistema OFDM punto a punto, o lo que es lo mismo, para un único usuario. Para ello, utilizaremos el modelo de canal anticipado anteriormente: el canal de banda ancha, selectivo en frecuencia y al que se le añade un ruido AWGN en el receptor, es equivalente a N canales AWGN paralelos no selectivos en frecuencia de banda estrecha (figura 2.6). En la subportadora i , el símbolo X_i es transmitido con una energía \mathcal{E}_i , siendo H_i la muestra de la respuesta impulsiva en frecuencia de la subportadora, W_i la muestra de ruido AWGN añadido a la subportadora con una densidad espectral de potencia n_i , e Y_i el símbolo recibido en dicha subportadora.

2.3.1. Asignación de potencia óptima multiportadora: *waterfilling*

Para un conjunto de N canales AWGN paralelos, el conocer la capacidad C que se puede obtener nos dirá cuánta información (cuántos bits) podemos transmitir a través de dichos canales, lo que lograremos maximizando la información mutua entre la entrada y la salida del conjunto de las subportadoras. Dado que la energía disponible para transmisión no es infinita, supondremos que está limitada a un valor máximo \mathcal{E}_{max} . Siendo \mathcal{E}_i la energía transmitida en la subportadora i , la capacidad obtenida, en bits por símbolo transmitido, es [Shannon, 1948]

$$c_i = \log_2 \left(1 + \frac{\mathcal{E}_i |H_i|^2}{n_i} \right). \quad (2.10)$$

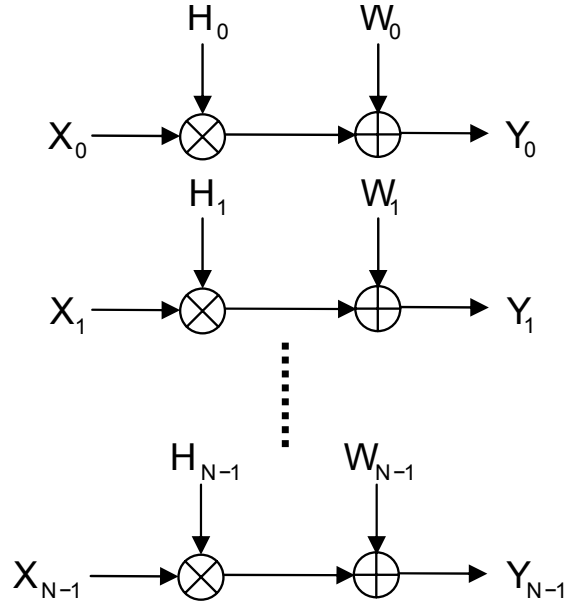


Figura 2.6. Modelo de canales paralelos

Por tanto, para conocer la capacidad hay que resolver la maximización de $C = \sum_{i=0}^{N-1} c_i$ con la restricción de energía, es decir

$$\begin{aligned} \max_{\mathcal{E}_i} C &= \sum_{i=0}^{N-1} c_i = \sum_{i=0}^{N-1} \log_2 \left(1 + \frac{\mathcal{E}_i |H_i|^2}{n_i} \right) = \sum_{i=0}^{N-1} \log_2 (1 + SNR_i) \\ \text{s.t.} \quad &\sum_{i=0}^{N-1} \mathcal{E}_i \leq \mathcal{E}_{max} \\ &\mathcal{E}_i \geq 0, \quad n = 1, \dots, N, \end{aligned}$$

donde $SNR_i = \frac{\mathcal{E}_i |H_i|^2}{n_i}$ representa la SNR en la subportadora i . Este problema se resuelve mediante el método de los multiplicadores de Lagrange (3.2.1), de forma que la distribución de la potencia entre las N subportadoras responde a la expresión

$$\mathcal{E}_i = \begin{cases} 0 & \text{si } \frac{n_i}{|H_i|^2} \geq \mu \\ \mu - \frac{n_i}{|H_i|^2} & \text{si } \frac{n_i}{|H_i|^2} < \mu. \end{cases}$$

Este resultado es la conocida asignación de energía de *waterfilling* (rellenado de agua) [Cover y Thomas, 1991], así denominado porque cada una de las subportadoras es “rellenada” en energía hasta el nivel de la constante μ en función de su SNR, como se muestra en la figura 2.7.

Desde un punto de vista práctico, resulta interesante analizar la capacidad $C = \sum_{i=0}^{N-1} c_i$ del conjunto de los N subportadoras en términos de potencia. Ahora,

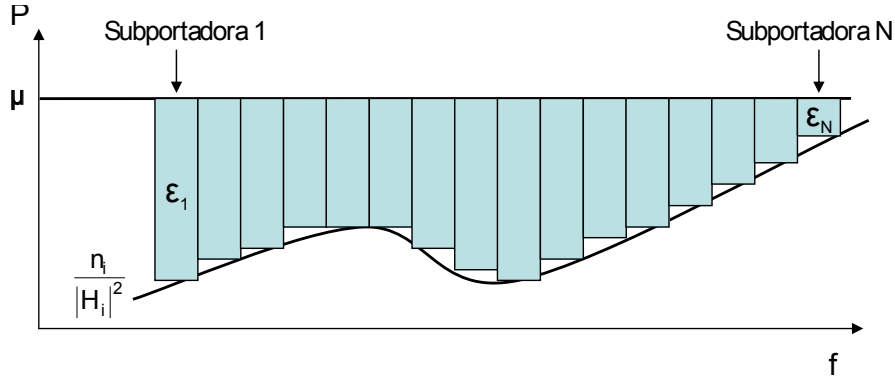


Figura 2.7. Asignación de potencia *waterfilling*

en la subportadora i se transmitirá una potencia de símbolo $P_i \geq 0$, de manera que la capacidad conseguida será $c_i = \log_2 \left(1 + \frac{P_i |H_i|^2}{N_i} \right)$, siendo N_i la potencia del ruido asociado a la subportadora i y teniendo en cuenta que el transmisor dispone de una potencia total máxima P_{max} . Resolviendo análogamente al problema formulado en términos de energía, reemplazando la restricción $\sum_{i=0}^{N-1} \mathcal{E}_i \leq \mathcal{E}_{max}$ por $\sum_{i=0}^{N-1} P_i \leq P_{max}$, la distribución de la potencia transmitida entre las subportadoras siguiendo la política de *waterfilling* es

$$P_i = \begin{cases} 0 & \text{si } \frac{N_i}{|H_i|^2} \geq \nu \\ \nu - \frac{N_i}{|H_i|^2} & \text{si } \frac{N_i}{|H_i|^2} < \nu, \end{cases}$$

donde ν representa el nivel máximo de potencia y juega el mismo papel que μ para la asignación de energías.

2.3.2. La aproximación SNR *gap*

La capacidad en cada una de las subportadoras AWGN, dada por (2.10), corresponde a la tasa máxima de transmisión de bits por símbolo alcanzable libre de errores, siempre que se utilice el código de canal que permite tal tasa de error arbitrariamente pequeña. No obstante, en la práctica tales códigos no pueden ser utilizados, por lo que en vez de hablar de capacidad de canal, hablaremos de tasa de transmisión de bits por símbolo (*throughput* en notación anglosajona), que representaremos mediante b y que determina el tamaño M de la constelación de la modulación utilizada mediante $M = 2^b$.

En sistemas realizables, en función de la modulación y la codificación empleados, la tasa de bits b se puede calcular mediante la **aproximación SNR *gap***. Denotando el SNR *gap* mediante Γ , con $\Gamma \geq 1$, es posible obtener el valor de b

según la expresión [Cioffi et al., 1995]

$$b = \log_2 \left(1 + \frac{SNR}{\Gamma} \right). \quad (2.11)$$

El SNR *gap*, como se observa en 2.11, representa la penalización en términos de SNR que experimenta el receptor, lo que hace que el número de bits por símbolo b_i que se pueden transmitir en la subportadora i se vea reducido respecto a la capacidad c_i , en tanto un esquema de codificación y de modulación determinados son empleados. A este respecto, cuanto más robusto sea el código, menor será el valor de Γ_i en la subportadora.

Además de ser función del esquema de codificación y modulación empleado, Γ depende de la probabilidad de error de bit deseada. Expresiones del SNR *gap* son conocidas para las modulaciones MQAM [Cioffi et al., 1995] y MPSK [García-Armada, 2006] como una función de la tasa de error de símbolo (*symbol error rate*, SER). En general, el valor del SNR *gap* depende de la SER que nos marquemos como objetivo, de manera que $\Gamma = \Gamma(SER)$. Por simplicidad, una vez establecido el vínculo entre Γ y la SER, omitiremos en adelante la dependencia de Γ con respecto a la SER en las distintas formulaciones.

Considerando el SNR *gap*, podemos reformular la maximización en términos de la tasa de bits total Λ , y la distribución de potencia de *waterfilling* será la solución del problema

$$\begin{aligned} \text{máx } \Lambda &= \sum_{i=0}^{N-1} b_i = \sum_{i=0}^{N-1} \log_2 \left(1 + \frac{SNR_i}{\Gamma_i} \right) \\ \text{s.t. } &\sum_{i=0}^{N-1} \mathcal{E}_i \leq \mathcal{E}_{max} \\ &\mathcal{E}_i \geq 0, \quad n = 1, \dots, N, \end{aligned}$$

que procediendo de forma análoga a (2.3.1) resulta

$$\mathcal{E}_i = \begin{cases} 0 & \text{si } \frac{n_i}{|H_i|^2} \Gamma_i \geq \mu \\ \mu - \frac{n_i}{|H_i|^2} \Gamma_i & \text{si } \frac{n_i}{|H_i|^2} \Gamma_i < \mu. \end{cases}$$

Es inmediato, por otro lado, obtener la solución de *waterfilling* para sistemas con restricciones de potencia, obteniéndose

$$P_i = \begin{cases} 0 & \text{si } \frac{N_i}{|H_i|^2} \Gamma_i \geq \nu \\ \nu - \frac{N_i}{|H_i|^2} \Gamma_i & \text{si } \frac{N_i}{|H_i|^2} \Gamma_i < \nu. \end{cases}$$

Observando los resultados obtenidos para el caso del *waterfilling* y para el SNR *gap*, podemos deducir que las asignaciones de energía y de potencia son equivalentes, de manera que a partir de este punto hablaremos indistintamente de potencia y energía.

2.3.3. Asignación de bits por subportadora: *bit-loading*

La distribución de potencia mediante *waterfilling* proporciona una asignación de la misma por subportadora que, aparte de maximizar la tasa de bits, tiene la ventaja de ser única ya que el problema resuelto es convexo [Boyd y Vandenberghe, 2004]. No obstante, desde el punto de vista de su aplicación práctica, en general la solución de *waterfilling* proporcionará valores no enteros de b_i , lo que supone granularidad infinita en el tamaño de la constelación. Esta solución no es realizable ya que los niveles de las constelaciones digitales son valores enteros (o no enteros pero sí discretos si se emplean esquemas de codificación de canal), y por tanto de granularidad finita. La alternativa, conocida en la literatura como *bit-loading*, es optimizar la tasa de bits transmitidos, restringiendo el valor de b_i a valores enteros o discretos, de manera que la distribución de la potencia entre las subportadoras se efectúa de acuerdo a los valores de b_i , en lugar de obtener directamente la potencia P_i que ha de asignarse a cada subportadora.

Existen multitud de algoritmos en la literatura reciente que realizan *bit-loading* en sistemas multiportadora. No obstante, de entre ellos destacan tres propuestas clave que han dado lugar a la mayoría de algoritmos de *bit-loading* para sistemas multiportadora con un usuario. El algoritmo de Hughes-Hartogs [Hughes-Hartogs, 1989] es la primera propuesta aparecida en la literatura para realizar asignación óptima de bits con valores enteros. Este algoritmo se basa en añadir en cada paso una unidad de la tasa de bits, y asignar dicha unidad a aquella subportadora que necesite el menor incremento de energía para transmitirla. No obstante, su alto coste computacional y su lenta convergencia llevan a Chow [Chow, 1993] a formular un algoritmo que mejora el de Hughes-Hartogs. Este algoritmo se basa en redondear la solución óptima de *waterfilling*, por lo que resulta subóptimo, pero la solución es muy cercana a la del algoritmo de Hughes-Hartogs y el tiempo de computación es considerablemente menor. Por último, el algoritmo de Levin-Campello [Campello, 1999] proporciona la solución óptima reduciendo significativamente la complejidad respecto al algoritmo de [Hughes-Hartogs, 1989]. A este tipo de algoritmos, basados en asignar un bit en cada paso del algoritmo

con el menor coste de potencia o energía posible, se les conoce como algoritmos tipo *greedy*.

Desde el punto de vista teórico, el *bit-loading* en su forma más básica se puede enfocar de dos maneras. En el *bit-loading* de tasa adaptativa, al que en lo sucesivo nos referiremos como RA (*rate adaptive*), el objetivo es maximizar la tasa total de bit $\Lambda = \sum_{i=0}^{N-1} b_i$ manteniendo constante la potencia transmitida, lo que conduce a resolver

$$\begin{aligned} \text{máx } \Lambda &= \sum_{i=0}^{N-1} b_i = \sum_{i=0}^{N-1} \log_2 \left(1 + \frac{SNR_i}{\Gamma_i} \right) \\ \text{s.t. } \sum_{i=0}^{N-1} P_i &= P_{max} \\ P_i &\geq 0, \quad n = 1, \dots, N. \end{aligned} \quad (2.12)$$

El segundo planteamiento básico, conocido habitualmente en la literatura como MA (*margin adaptive*), consiste en minimizar la potencia total transmitida, manteniendo la tasa de bit Λ constante, es decir

$$\begin{aligned} \text{mín } P &= \sum_{i=0}^{N-1} P_i \\ \text{s.t. } \sum_{i=0}^{N-1} b_i &= \Lambda \\ P_i &\geq 0, \quad n = 1, \dots, N. \end{aligned} \quad (2.13)$$

Esta segunda formulación se puede replantear de manera inmediata en términos de energía, dado que para cada subportadora i , la potencia transmitida se puede expresar como $\mathcal{E}_i = P_i T_s$, siendo \mathcal{E}_i la energía de símbolo y T_s el período de símbolo.

2.4. Acceso múltiple multiportadora: OFDMA

En la sección anterior hemos tratado la transmisión en sistemas multiportadora para el caso de un único usuario. No obstante, el mayor interés desde el punto de vista de sistemas reales radica en la utilización de esquemas de acceso basados en OFDM en entornos multiusuario. Una posibilidad es emplear acceso múltiple en el tiempo OFDM-TDMA (*time division multiple access*) de forma que los usuarios utilicen todo el espectro OFDM durante un cierto intervalo de tiempo. El caso de OFDM-FDMA (*frequency division multiple access*) es análogo, sólo que ahora

un usuario utiliza la misma sub-banda del espectro de forma permanente. Ambos tipos de acceso adolecen de rigidez en la asignación del recurso compartido, las frecuencias en FDMA y los instantes de transmisión para TDMA. Además, TDMA tiene el inconveniente de que si el número de usuarios es elevado, el retardo en acceder al intervalo de tiempo asignado resultará excesivo. OFDMA (*orthogonal frequency-division multiple access*) es una técnica de acceso múltiple que asigna a cada usuario un conjunto de subportadoras ortogonales, y cada una de las subportadoras es utilizada única y exclusivamente por un usuario durante un intervalo (*slot*) de tiempo. Por tanto, no existe solapamiento de los usuarios en el dominio de la frecuencia, gracias a la ortogonalidad entre las subportadoras OFDM. Las figuras 2.8(a)–2.8(c) ilustran gráficamente la diferencia entre los tres modos de acceso descritos.

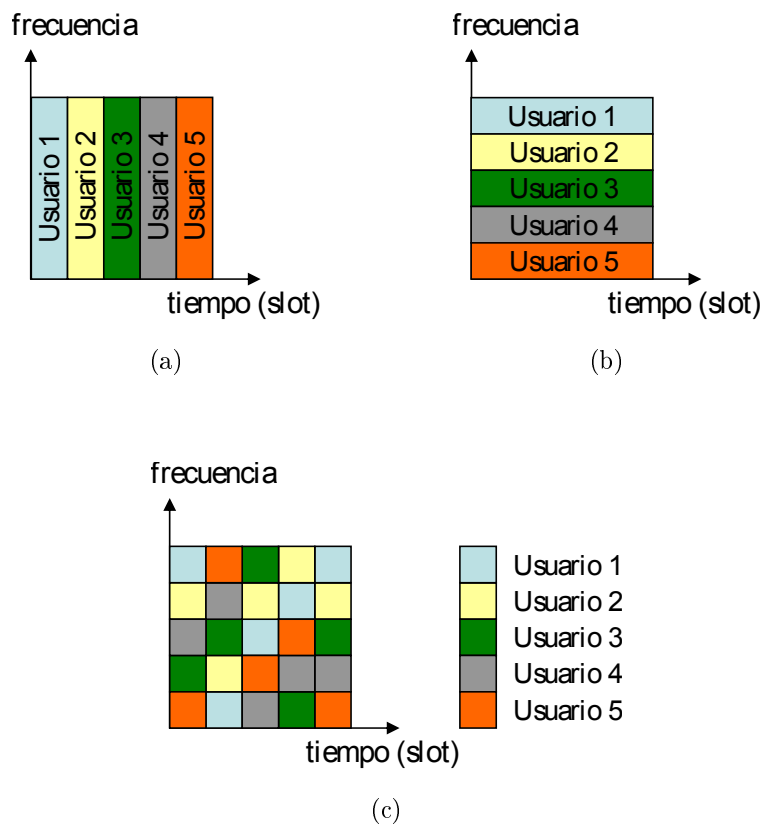


Figura 2.8. Canalización para: (a) TDMA; (b) FDMA; (c) OFDMA

Desde el punto de vista de teoría de la información, OFDMA es el esquema de acceso múltiple óptimo para el enlace descendente de sistemas multiportadora multiusuario con decodificación independiente [Li y Liu, 2005]. En efecto,

para optimizar la capacidad resultante de todas las subportadoras asignadas a los diferentes usuarios con restricción de la potencia de transmisión de la estación base, resulta que la estrategia óptima es que las subportadoras se asignen exclusivamente a un usuario, es decir, empleando OFDMA. Es más, este resultado es extensivo a la utilización de cualquier esquema de modulación adaptativa cuya tasa de transmisión de bits sea una función convexa de SNR o de SINR (relación señal a interferencia más ruido).

2.4.1. Diversidad multiusuario

Una de las características más notorias de OFDMA es que aprovecha la **diversidad multiusuario**, concepto que tiene su origen en el trabajo de [Knopp y Humblet, 1995], donde se muestra que la capacidad suma del canal ascendente, modelado como canal de acceso múltiple gaussiano con desvanecimiento plano de frecuencia, se consigue si transmite únicamente el usuario con mejor canal en cada instante.

Los esquemas de diversidad multiusuario se fundamentan en la naturaleza aleatoria del medio radioeléctrico: si las condiciones de transmisión resultan poco favorables para un usuario (ganancia del canal baja) en un canal e instante determinados, es probable que, simultáneamente, otros usuarios observen una calidad del canal elevada (la ganancia es alta) y por tanto sí puedan aprovechar dicho canal para transmitir. La diversidad multiusuario en OFDMA se basa en asignar cada una de las N subportadoras a aquel usuario que mejor lo explote (figura 2.9), es decir, que mayor tasa de transmisión obtenga, con el consiguiente beneficio adicional en el rendimiento global del sistema. Cuantitativamente, por tanto, existirá una diferencia entre la tasa binaria obtenida con diversidad multiusuario y la obtenida sin diversidad, conocida como ganancia por diversidad multiusuario, y que denotaremos por g_{dmu} . Tomando como base el análisis realizado en [Li y Hui, 2005], mostramos a continuación que se obtiene dicha ganancia si se realiza una asignación de las subportadoras disponibles a los usuarios basada en diversidad multiusuario.

Consideremos un sistema que emplea OFDMA para asignar N subportadoras OFDM a K usuarios, siendo usualmente $N \gg K$. Tomado una de las subportadoras aleatoriamente, por ejemplo n , denotaremos por SNR_i la relación señal a ruido asociada a la subportadora n de cada usuario i , $i = 1, \dots, K$, empleando un esquema de modulación y codificación dados. Todas las SNR_i se modelan co-

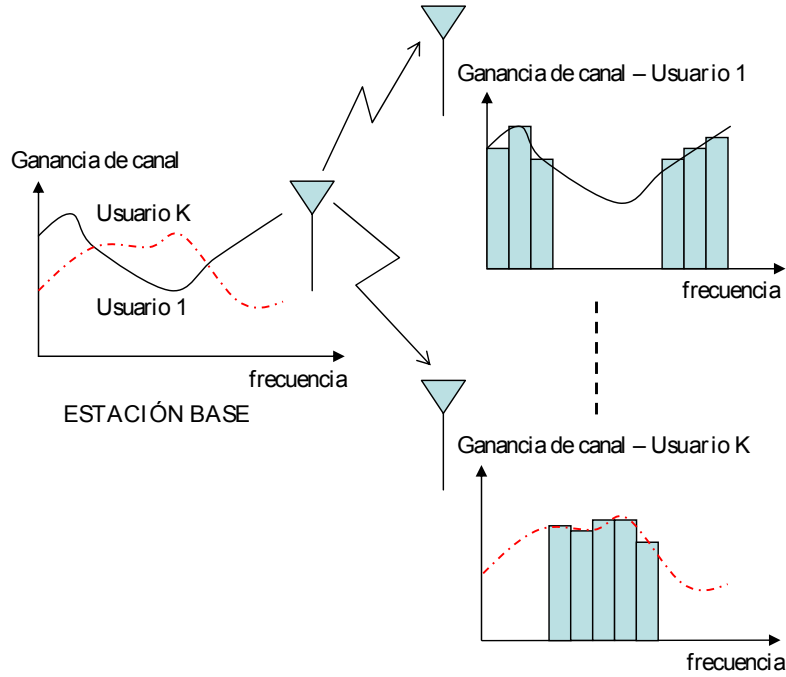


Figura 2.9. Asignación de subportadoras para sistemas multiusuario OFDMA con diversidad multiusuario

mo una variable aleatoria y siguen las mismas funciones de distribución $F(x)$ y densidad de probabilidad $f(x)$. Entre todos los usuarios, existirá un usuario k^* que obtendrá la mayor SNR, de manera que

$$k^* = \arg \max_{i=1, \dots, K} SNR_i,$$

y denotaremos por SNR_{max} al valor máximo de dicha SNR, es decir, $SNR_{max} = SNR_{k^*}$.

La ganancia por diversidad multiusuario la vamos a cuantificar a través de la tasa binaria obtenida esperada. Por tanto, compararemos la tasa binaria esperada en ambos casos: con asignación inteligente, que explota la diversidad multiusuario, y asignación aleatoria. Si asignamos la subportadora n al usuario k^* de entre los K posibles, obtendremos en dicho subportadora una tasa binaria que denotaremos por b_{max}^K y que calculamos mediante la ecuación 2.11

$$b_{max}^K = \log_2(1 + SNR_{max}). \quad (2.14)$$

Por simplicidad, expresaremos mediante $r(x)$ la tasa de transmisión $\log_2(1 + x)$, de forma que $b_{max}^K = r(SNR_{max})$, y la tasa binaria esperada la calculamos mediante

$$\mathbf{E} \{b_{max}^K\} = \mathbf{E} \{r(SNR_{max})\} = \int_0^\infty r(x) f_{SNR_{max}}(x) dx. \quad (2.15)$$

Necesitamos por tanto conocer $f_{SNR_{max}}(x)$, que sabemos que podemos calcular como $f_{SNR_{max}}(x) = \frac{dF_{SNR_{max}}(x)}{dx}$, lo cual su vez implica conocer $F_{SNR_{max}}(x)$. Como $F(x)$ es conocida, podemos evaluar $F_{SNR_{max}}(x)$ a partir de la misma mediante la siguiente relación:

$$\begin{aligned} F_{SNR_{max}}(x) &= \Pr(SNR_{max} < x) = \Pr(SNR_1 < x) \dots \Pr(SNR_K < x) \\ &= F(x) \dots F(x) \\ &= (F(x))^K. \end{aligned} \quad (2.16)$$

Ahora, es inmediato obtener $f_{SNR_{max}}(x)$:

$$f_{SNR_{max}}(x) = \frac{dF_{SNR_{max}}(x)}{dx} = K(F(x))^{K-1} f(x). \quad (2.17)$$

Aplicando integración por partes para resolver (2.15) queda:

$$\begin{aligned} \mathbf{E}\{b_{max}^K\} &= r(x) F_{SNR_{max}}(x) \Big|_0^\infty - \int_0^\infty r'(x) F_{SNR_{max}}(x) dx \\ &= r(\infty) F_{SNR_{max}}(\infty) - r(0) F_{SNR_{max}}(0) - \int_0^\infty r'(x) (F(x))^K dx. \end{aligned} \quad (2.18)$$

$$(2.19)$$

Nótese que el primer sumando de la ecuación 2.19 es exactamente la capacidad C asociada a la subportadora n y que el segundo término es 0, de manera que la tasa esperada será la capacidad menos un cierto término que será función de la distribución de probabilidad de la relación señal a ruido $F(x)$:

$$\mathbf{E}\{b_{max}^K\} = C - \int_0^\infty r'(x) (F(x))^K dx. \quad (2.20)$$

Realicemos ahora el análisis para el caso en que la subportadora n es asignada aleatoriamente a uno de los K usuarios. El valor esperado de la tasa binaria se calcula análogamente al caso anterior, de manera que

$$\mathbf{E}\{b_{cualquiera}\} = \mathbf{E}\{r(SNR_i)\} = \int_0^\infty r(x) f_{SNR_i}(x) dx \quad (2.21)$$

$$= r(x) F(x) \Big|_0^\infty - \int_0^\infty r'(x) F_{SNR_i}(x) dx \quad (2.22)$$

$$= C - \int_0^\infty r'(x) F(x) dx. \quad (2.23)$$

siendo $b_{cualquiera}$ la tasa binaria obtenida por uno de los K usuarios, elegido aleatoriamente. Así, la ganancia por diversidad multiusuario respecto la asignación aleatoria viene dada por la diferencia entre $\mathbf{E}\{b_{max}^K\}$ y $\mathbf{E}\{b_{cualquiera}\}$:

$$g_{dmu} = \mathbf{E}\{b_{max}^K\} - \mathbf{E}\{b_{cualquiera}\} = \int_0^\infty r'(x) \left(F(x) - (F(x))^K \right) dx. \quad (2.24)$$

Como $r(x)$ es una función monótona creciente, $r'(x) \geq 0 \forall x$; y como $0 \leq F(x) \leq 1$, resulta que $F(x) - (F(x))^K \geq 0$. Por tanto, $g_{dmu} \geq 0$, quedando probado que existirá, en el caso peor, un resultado igual al caso aleatorio. Es más, la ganancia esperada será mayor cuanto mayor sea el número de usuarios K , como vemos a continuación. Si comparamos los casos de i e $i - 1$ usuarios en el caso genérico en que $1 \leq i \leq K$ empleando asignación inteligente, la diferencia entre las tasas binarias esperadas es

$$\mathbf{E} \{b_{max}^i\} - \mathbf{E} \{b_{max}^{i-1}\} = \int_0^\infty r'(x) \left[(F(x))^i - (F(x))^{i-1} \right] dx. \quad (2.25)$$

Suponiendo $r'(x) > 0$, como $(F(x))^i - (F(x))^{i-1} = (F(x))^{i-1} (1 - F(x)) > 0$, entonces $\mathbf{E} \{b_{max}^i\} > \mathbf{E} \{b_{max}^{i-1}\}$. Es inmediato, siguiendo el mismo procedimiento, que

$$\mathbf{E} \{b_{max}^K\} > \dots > \mathbf{E} \{b_{max}^i\} > \dots > \mathbf{E} \{b_{max}^1\} = \mathbf{E} \{b_{cualquiera}\}. \quad (2.26)$$

Dada la naturaleza aleatoria e independiente de las ganancias de los canales asociados a los usuarios, el resultado anterior se puede interpretar de la siguiente forma: cuantos más usuarios estén presentes en el sistema, mayor será la probabilidad de que exista un usuario tal que, si se le asigna el canal considerado, el rendimiento global del sistema mejor.

2.4.2. El problema de la asignación de recursos en OFDMA

En redes basadas en OFDM con transmisión centralizada, existen limitaciones en la disponibilidad de los recursos, como por ejemplo el ancho de banda de transmisión y la potencia total disponible. Este es el caso, por ejemplo, del enlace descendente entre una estación base (EB) y los usuarios asociados a la misma. Bajo estas condiciones, es primordial realizar la asignación de los recursos de forma que se extraiga el máximo beneficio posible al emplearlos, especialmente de cara a las altas tasas binarias especificadas en los estándares que utilizan OFDM. Para ello, es básico considerar que el entorno de transmisión va a ser cambiante en el tiempo. Una asignación fija de los recursos, por ejemplo FDMA aplicando *waterfilling* sobre las subportadoras de cada usuario, conducirá entonces a situaciones que, no sólo que no serán óptimas, sino que las prestaciones del sistema (tasa binaria transmitida, por ejemplo) quedarán posiblemente muy degradadas, ya que el esquema de transmisión fue diseñado para unas condiciones específicas. Por el contrario, la asignación dinámica de recursos permite asignar los valores a

los parámetros de transmisión adaptativamente, de manera que sus valores sean los adecuados en cada instante para optimizar el rendimiento del sistema.

La utilización conjunta de diversidad multiusuario y modulación adaptativa constituye una poderosa herramienta para optimizar el rendimiento en sistemas con acceso OFDMA variantes en el tiempo. La diversidad multiusuario resulta esencial como criterio para la asignación de recursos, en concreto de las frecuencias de transmisión (o subportadoras), con el que maximizar el rendimiento global del sistema, bien en cuanto a la tasa de transmisión o capacidad, bien minimizando la potencia total transmitida. La ganancia por diversidad se obtendrá mediante algoritmos de asignación de las subportadoras a los usuarios, siempre que la modulación adaptativa garantice la asignación de los valores adecuados a los parámetros de transmisión asociados a cada una de las subportadoras, de manera que la adaptación a las condiciones del canal se lleva a cabo no sólo a nivel de símbolo OFDM sino a nivel de subportadora, obteniéndose así el máximo nivel de granularidad posible. En la sección 2.6 revisaremos los artículos pioneros en este área y analizaremos los trabajos más recientes relacionados con la asignación de recursos en sistemas OFDM multiusuario.

2.5. Imparcialidad entre los usuarios en OFDMA

La utilización de diversidad multiusuario, como se ha visto en (2.4.1), introduce una ganancia en términos de tasa de transmisión respecto a la asignación aleatoria. No obstante, si los recursos se asignan siempre a los usuarios que mejor canal perciben, cabe la posibilidad de que los usuarios que observen malos canales durante cierto tiempo consigan tasas de transmisión muy pobres, o no transmitir en absoluto, resultando que los usuarios con buenos canales se vean beneficiados en el número de oportunidades para transmitir. Por lo tanto, si bien la asignación de recursos considerando diversidad multiusuario conlleva optimizar el rendimiento del sistema de manera global, puede asimismo conducir a desfavorecer a ciertos usuarios. En sistemas multiusuario en general se busca que los usuarios reciban aproximadamente el mismo servicio, es decir, que exista **imparcialidad** (*fairness*) en la asignación de los recursos, por lo que la optimización del sistema atendiendo sólo a la diversidad multiusuario puede no ser la mejor solución, desde esta perspectiva.

Una solución subóptima es establecer el criterio de imparcialidad proporcional

(*proportional fairness*) [Kelly et al., 1998], de manera que, en vez de asignar los recursos a los usuarios que obtienen la mayor tasa instantánea de bit, se favorece a aquellos que en media hayan obtenido una tasa de transmisión baja; en otras palabras, el objetivo es que la tasa de transmisión se encuentre razonablemente balanceada entre los usuarios. Expresado en términos de tasa de bits transmitidos, sean dos usuarios que consiguen, en una subportadora determinada, unas tasas R_1 y R_2 respectivamente, y que hasta el instante de la asignación de la subportadora han transmitido unas tasas medias \bar{R}_1 y \bar{R}_2 . Si se realiza la asignación considerando imparcialidad proporcional, transmitirá aquel usuario con mayor relación entre la tasa que obtiene de la subportadora y la obtenida hasta el momento, es decir, mayor R_i/\bar{R}_i . Así, si hasta ese instante el usuario 1 ha experimentado una tasa de bit muy pobre, tal que $\bar{R}_1 \ll \bar{R}_2$, la subportadora resultará muy probablemente asignada al usuario 1.

2.6. Asignación en OFDMA de los recursos potencia, bits y subportadoras

Ya hemos comentado anteriormente en (2.4.2) que en redes OFDMA una cuestión fundamental es resolver la asignación de recursos para obtener el mejor rendimiento posible del sistema. Si hacemos uso de modulación adaptativa, los recursos disponibles considerados habitualmente son:

- las subportadoras, que serán asignadas a los usuarios presentes en el sistema individualmente. Alternativamente, las subportadoras también pueden ser asignadas agrupadas, tal y como se recoge en la sección 2.7 y en el capítulo 5, lo cual reduce significativamente la complejidad asociada a esta asignación.
- El esquema de modulación y codificación a utilizar, seleccionable a nivel de subportadora. Así, se determina el número de bits por símbolo que se transmitirán en cada una de las subportadoras.
- La potencia de transmisión por subportadora, que vendrá asociada al esquema de modulación escogido.

Por tanto, la asignación óptima de los recursos (AOR) implica considerar conjuntamente cómo asignar las subportadoras a los usuarios, y cuánta potencia y

cuántos bits transmitir en cada subportadora. Esta solución óptima es evidentemente de una elevada complejidad: implica una búsqueda exhaustiva sobre los niveles de potencia y los posibles esquemas de modulación para cada posible reparto de las subportadoras entre los usuarios. Matemáticamente, el problema de la asignación óptima de recursos en el enlace descendente de sistemas OFDMA lo podemos formular como

$$\begin{aligned} & \text{máx (mín) } U \\ & \text{s.t. } \sum_{k=1}^K \sum_{n=1}^N p_{kn} \leq P, \end{aligned} \quad (\text{AOR})$$

donde U es la función a optimizar, p_{kn} es la potencia asociada a la subportadora k para el usuario n , K es el número total de subportadoras, N representa el número de usuarios, y P la potencia máxima disponible para transmisión. La función U , conocida como función de utilidad, dependerá del criterio adoptado para medir el rendimiento de nuestro sistema, y en general será una función de los parámetros de transmisión p_{kn} y m_{kn} , que designan, respectivamente, la potencia y el nivel de la modulación (o bits por símbolo) por subportadora y por usuario. En la formulación, es habitual incluir un esquema de asignación de las subportadoras a los usuarios mediante una variable binaria, que valdrá 1 si la subportadora es asignada al usuario y 0 en caso contrario.

La alta complejidad del problema AOR se deriva de que es un problema de programación entera con restricciones y para obtener la solución óptima el tiempo de cómputo necesario es de orden exponencial en el número de variables del problema, y no polinómico como sería deseable. Por tanto, de cara a sistemas de comunicaciones reales, la adopción de algoritmos óptimos puede resultar inviable desde el punto de vista de la complejidad computacional. Esto explica los recientes esfuerzos dedicados a resolver este problema desde dos perspectivas. Por un lado, proponiendo métodos óptimos que reduzcan las capacidades de cómputo asociadas a resolver el problema. Por otro, a través de algoritmos subóptimos de baja complejidad cuyas prestaciones sean cercanas a la solución óptima.

Desde el punto de vista del criterio de optimización, entre los distintos autores que han propuesto algoritmos para resolver el problema AOR predominan dos enfoques: maximizar la tasa global de transmisión, en cuyo caso $U = \sum_{i=1}^K b_i$, y minimizar la potencia empleada en la transmisión, siendo ahora $U = \sum_{k=1}^K \sum_{n=1}^N p_{kn}$. En el primer caso hablamos del problema RA, y en el segundo del problema MA,

ambos ya definidos en la sección 2.3.3. No obstante, resulta más interesante diferenciar entre algoritmos que resuelven la asignación de recursos de manera conjunta (es decir, en un único paso) y aquellos que lo resuelven en dos pasos, asignando normalmente primero las subportadoras a los usuarios para a continuación llevar a cabo el *bit-loading* por subportadora. Mientras los primeros proporcionan una solución óptima del problema, a costa de una alta complejidad computacional, los algoritmos en dos pasos, si bien facilitan una solución subóptima, reducen significativamente la complejidad del problema y son más adecuados para implementaciones prácticas. A continuación, revisamos la literatura existente sobre la asignación de recursos en sistemas OFDMA, centrándonos en los relativos al enlace descendente.

2.6.1. Algoritmos de asignación conjunta de recursos

El trabajo de [Wong et al., 1999] se puede considerar pionero al abordar la problemática de asignar recursos, en particular subportadoras, potencia y bits por símbolo por subportadora, en sistemas OFDMA. Los autores plantean la minimización de la potencia total de transmisión, entendida como la suma de las potencias transmitidas por los usuarios en las correspondientes subportadoras, para unas tasas de transmisión dadas y constantes para cada usuario. El problema de optimización resultante es un problema combinatorio, puesto que se asume que los bits por símbolo por subportadora toman valores discretos y que cada subportadora es asignada a un único usuario, y su resolución es muy exigente computacionalmente. La relajación de los valores de los bits a valores continuos permite desarrollar un algoritmo que resuelve de forma conjunta el problema de asignación, aunque proporcionando una solución subóptima. Este algoritmo tiene varios problemas de índole práctica, por lo que los autores proponen un algoritmo en dos pasos, que denominan MAO, que en una primera fase asigna las subportadoras a los usuarios, y luego encuentra las asignaciones de bits y potencia para cada uno de las subportadoras empleando un algoritmo de *bit-loading* de usuario único para las subportadoras asignadas a cada usuario.

En [Song y Li, 2005a] se analiza, desde un punto de vista teórico y para el caso en que la tasa de bits es una función continua, la asignación de potencia y subportadoras en el enlace descendente de la EB a los usuarios. El problema se formula con funciones de utilidad en sentido general, y se deducen las condiciones necesarias y suficientes para que la asignación sea óptima. Este marco teórico general

tiene su continuación en [Song y Li, 2005b], con la propuesta de diferentes algoritmos para la asignación dinámica de subportadoras y la asignación adaptativa de potencia.

En [Mohanram y Bhashyam, 2005], se diseña un algoritmo para la asignación conjunta de subportadoras y potencia que resuelve el problema de MA, a la vez que implementa imparcialidad proporcional entre los usuarios. El reparto de la potencia entre las subportadoras de un usuario se realiza mediante *waterfilling*, lo cual influye en la determinación de las subportadoras asignadas al usuario.

La propuesta de [Han et al., 2005] introduce la Teoría de Juegos Cooperativa como herramienta para la optimización de la asignación de recursos en sistemas OFDMA, de forma que los usuarios acuerdan entre ellos la distribución de las subportadoras disponibles. En cuanto a la formulación matemática del problema, la función de utilidad propuesta a maximizar está basada en la tasa binaria y garantiza la imparcialidad proporcional entre los usuarios, asegurándose también una tasa binaria mínima objetivo por usuario. Un primer algoritmo resuelve el problema de asignación conjunta de potencia y subportadoras para el caso de dos usuarios. Después, los usuarios son agrupados en parejas, denominadas coaliciones, y a través del algoritmo anterior negocian el intercambio de subportadoras para mejorar su utilidad. Sucesivas reagrupaciones de los usuarios en coaliciones conducen a la convergencia del algoritmo. Un punto clave en este trabajo es la formación de las coaliciones, que se puede resolver a través de un algoritmo convencional como puede ser el algoritmo Húngaro [Kuhn, 1955]. No obstante, este algoritmo es considerablemente más complejo y necesita cierto control centralizado en la EB, mientras el propuesto por los autores puede ser implementado de forma distribuida.

Otra forma de resolver el problema AOR es mediante técnicas de dualidad, como se muestra en [Seong et al., 2006]. Los autores proponen resolver la asignación de recursos para los problemas MA y RA, donde las funciones de utilidad incluyen coeficientes de ponderación para cada usuario. Dada la alta complejidad asociada a la solución óptima de ambos problemas, que es exponencial en el número de subportadoras N , los autores proponen sendos algoritmos subóptimos basados en la resolución del problema dual, consiguiendo así una complejidad asociada lineal en N . Se muestra que la diferencia entre las soluciones óptima y subóptima es apreciable para valores de N realistas en aplicaciones prácticas.

En [YuMing et al., 2006], se propone un algoritmo subóptimo para asignar re-

cursos en el problema RA con requisitos de tasa fija de transmisión por usuario, que combina la asignación simultánea de subportadora y de potencia. Inicialmente, las subportadoras son asignadas a los usuarios mediante un algoritmo *greedy* convencional, que favorece la asignación de subportadoras a los usuarios con menor ganancia de canal promedio para minimizar la potencia total de transmisión. Posteriormente, las subportadoras se van reasignando entre los usuarios, con el objetivo de disminuir la potencia total empleada: si la subportadora n asignada al usuario k se asigna al usuario l , dicho cambio se hace efectivo si la disminución de potencia asociada (teniendo en cuenta que para cada usuario la potencia sigue una asignación de tipo *waterfilling*) es mayor que el incremento, es decir, $\Delta P_k > \Delta P_l$.

En [Baamrani et al., 2006], se propone un algoritmo de baja complejidad para la optimización MA con restricción de tasa mínima de transmisión por usuario, tomando como base para la asignación inicial de subportadoras el algoritmo de minimización de energía para un usuario de [Hoo et al., 1998]. El algoritmo se basa en asignar subportadoras a los usuarios ordenados según el siguiente criterio: el usuario más prioritario es aquel cuya suma de ganancias sobre todas las subportadoras disponibles es la más baja.

A partir del algoritmo de asignación de subportadoras propuesto en [Dardari, 2004] para OFDM punto a punto, los autores de [Ermolova y Makarevitch, 2007] extienden su uso a redes multiusuario para minimizar la BER promedio del sistema. El esquema propone, en lugar de asignar la misma energía de bit en cada subportadora, asignar la energía por bit transmitido en cada subportadora de forma inversamente proporcional a la ganancia de potencia de la subportadora.

En [Feiten et al., 2008], los autores proponen un algoritmo heurístico que realiza la asignación conjunta de las subportadoras y tasa binaria por subportadora (la cual es un valor continuo y no discreto), para minimizar la potencia total, sujeto a que los usuarios tienen que recibir una tasa binaria dada.

2.6.2. Algoritmos subóptimos: asignación en dos pasos de subportadoras, potencia y bits

La propuesta de [Rhee y Cioffi, 2000] es realizar la asignación de recursos en el enlace descendente de un sistema que proporciona servicios de tasa binaria variable (*variable bit rate*, VBR), por lo que el objetivo es maximizar la tasa binaria más baja de entre todos los usuarios, es decir, una estrategia máx.mín. Dicha estrategia

es una manera de implementar la imparcialidad, puesto que se pretende así que todos los usuarios obtengan tasas similares. El algoritmo subóptimo distribuye la potencia disponible P por igual entre las N subportadoras, por lo que en cada subportadora se transmite con una potencia P/N , y en una segunda fase se asignan las subportadoras siguiendo el criterio de que cada subportadora se asigne al usuario de mayor SNR. En la línea del anterior, el modelo de asignación de recursos planteado en [Shen et al., 2005] implementa imparcialidad proporcional entre los usuarios para maximizar la capacidad suma, definida ésta como la suma de las capacidades obtenidas por todos los usuarios. Dada la elevada complejidad que supone resolver el problema propuesto, los autores proponen un algoritmo subóptimo en dos pasos que realiza, en primer lugar, la asignación de las subportadoras basándose en el algoritmo de [Rhee y Cioffi, 2000]; el segundo paso consiste en realizar la asignación de potencia, resolviendo el problema de optimización resultante al emplear la asignación de subportadoras obtenida en el paso primero.

En [Yin y Liu, 2000], se propone un algoritmo subóptimo para la optimización MA con restricciones de tasa binaria mínima por usuario. El primer paso del algoritmo consiste en asignar la potencia y el número de subportadoras necesarias para cada usuario, supuesto que los usuarios verán la misma ganancia en todas las subportadoras, y que se calcula promediando la ganancia en los mismos. Una vez resuelta la asignación de potencia y el número de subportadoras necesarias, la asignación de las subportadoras se reduce al problema clásico de asignación de N recursos a N individuos, resoluble con el algoritmo Húngaro. Relacionado con el anterior, el algoritmo presentado en [Kivanc et al., 2003] resuelve el problema RA con restricciones de tasa mínima por usuario, con un primer paso de asignación análogo al de [Yin y Liu, 2000], pero basado en la SNR del usuario promediado entre todas las subportadoras. Para la fase de asignación de las subportadoras en sí, se plantean dos algoritmos diferentes. El primero de ellos se basa en reformular el problema como una versión del conocido problema combinatorio de partición de conjuntos [Bertsimas y Weismantel, 2005]. El segundo, menos exigente computacionalmente, se basa en asignar las subportadoras a los usuarios con mejor ganancia en cada subportadora.

En [Ergen et al., 2003], los autores proponen un algoritmo iterativo alternativo a las hasta ese momento habituales aproximaciones de primero asignar subportadoras, basadas en programación lineal o algoritmo Húngaro, y luego realizar el *bit-loading* mediante algoritmos tipo *greedy*, para resolver la asignación de recursos

para del problema MA. En una primera fase, se realiza conjuntamente una asignación de subportadoras y bits, tratando de encontrar la mejor subportadora para un usuario y asignarle el mayor número posible de bits, con un consumo mínimo de potencia. En una segunda fase, se realiza un reajuste para reducir la potencia transmitida, en dos dimensiones: intercambiando los bits por símbolo entre las subportadoras de un mismo usuario, e intercambiando las subportadoras entre pares de usuarios. Con este algoritmo, los autores consiguen mejorar el control de potencia y la imparcialidad, respecto a propuestas tipo *greedy*.

La compartición de las subportadoras entre los usuarios conduce a interferencias entre aquellos que empleen la misma subportadora. Este tipo de asignación es estudiada en [Jang y Lee, 2003], donde el objetivo es optimizar la tasa binaria total mediante asignación de los recursos subportadoras y potencia. Para resolver este problema, los autores proponen un algoritmo en dos pasos: asignación de subportadoras y asignación de potencia. Para el primer paso, se demuestra que la tasa binaria en una subportadora dada se maximiza sólo si la subportadora se asigna al usuario de mayor ganancia en esa subportadora. Una vez asignadas las subportadoras de esta manera, resulta que la asignación de potencia óptima es aplicar la política de *waterfilling* entre las subportadoras de cada usuario. Para este sistema, los autores muestran que, si se reparte la potencia por igual entre las subportadoras, la diferencia en términos de tasa binaria respecto a aplicar *waterfilling* es marginal.

En [Kim et al., 2006] se aborda la asignación de recursos tanto para optimización MA como RA, mediante la formulación de los respectivos problemas en variable entera. El algoritmo propuesto se basa en la observación de que, tanto para MA como para RA, la solución óptima del problema tiende a asignar, para un usuario dado, la misma cantidad de bits para la mayoría de las subportadoras asignadas; a este valor predominante del número de bits se le denomina constelación dominante [Park et al., 2004]. Bajo la hipótesis de la constelación dominante, la asignación de subportadoras del problema MA, que se realiza en primer lugar, tiene la forma del problema de transporte, cuya resolución es conocida mediante programación lineal [Bertsimas y Weismantel, 2005]. El *bit-loading* se realiza en segundo término mediante un algoritmo tipo *greedy*. Para resolver la optimización RA, los autores se basan en la optimización MA realizada, ya que demuestran la relación matemática entre ambas formulaciones. Similar aproximación se realiza en [Yu et al., 2006] añadiendo restricciones de imparcialidad proporcional de ma-

nera análoga al planteamiento de [Shen et al., 2005]. El algoritmo propuesto alivia la carga computacional respecto a otras propuestas ya que para la asignación de subportadoras la potencia se reparte por igual entre todas ellas.

La casuística particular de la asignación de recursos en redes que proporcionan servicios *multicast*, consistentes en la difusión de información a grupos de usuarios concretos, es analizada en [Changho y Jeonghoon, 2008] desde el punto de vista de la capacidad. En las redes convencionales que proporcionan estos servicios, la capacidad está limitada por la ganancia del usuario que peor canal experimente, ya que en caso contrario dicho usuario no podrá decodificar la información que reciba, y en redes multiportadora esto deberá ser aplicado a cada una de las N subportadoras. La formulación del problema MA propuesto evita esta limitación y plantea un algoritmo óptimo para su resolución. Dada su elevada complejidad, proponen un algoritmo subóptimo en dos pasos, con un criterio similar en la asignación de subportadoras al descrito en [Wong et al., 1999], y una posterior asignación de bits y potencia basada en el algoritmo de Levin-Campello.

Recientemente, los autores de [Mitran et al., 2008] han abordado la asignación de subportadoras, tasa de bits y potencia para los usuarios secundarios (no licenciados) en redes cognitivas, donde este tipo de usuarios hacen un uso oportunista del espectro cuando los usuarios primarios (con licencia) lo permiten, bien por ausencia de transmisión o bien por nivel tolerable de la interferencia creada por los usuarios secundarios. En primer lugar, se plantea el problema de optimización para maximizar la suma de las tasas binarias, que se obtiene añadiendo al problema MA la restricción de la potencia máxima que la EB puede transmitir hacia un usuario secundario sin causar interferencia sobre un usuario primario. Los autores proponen, para resolver el problema de asignación de recursos, dos algoritmos heurísticos que permitirían adaptar dichos recursos con suficiente rapidez para que los cambios no afectasen a los usuarios primarios.

2.7. Asignación de recursos en OFDMA utilizando agrupaciones de subportadoras

2.7.1. Agrupación de subportadoras: los subcanales

Las redes de acceso de banda ancha basadas en OFDM se caracterizan por hacer uso de un elevado número de subportadoras; por ejemplo, en los sistemas basados

en el estándar IEEE 802.16 se contempla la posibilidad de utilizar hasta 4096. Por tanto, el problema de asignación o reparto óptimo de las subportadoras se convierte a menudo en un problema intratable desde el punto de vista computacional para algoritmos en tiempo real. Una alternativa es la utilización de **subcanales**, definidos como una agrupación de subportadoras OFDM [802, 2004]. De esta manera, la asignación de ancho de banda a los usuarios se realizará por bloques de subportadoras en vez de asignar cada una de las mismas individualmente, simplificándose así de forma significativa el esquema de asignación de frecuencias.

La figura 2.10 muestra la estructura de la rejilla espacio-frecuencia cuando se utiliza agrupación en OFDMA. Las N portadoras se pueden agrupar en subcanales, los cuales generalmente van a tener el mismo número de subportadoras N/J , siendo J el número de subportadoras por subcanal. Es posible también la agrupación de varios subcanales en ranuras o *slots*, que se definen como agrupaciones de subcanales durante uno o más símbolos OFDM. A estos *slots* se refieren diferentes autores también con los términos *chunk* (ver, por ejemplo [Pfetschinger et al., 2007]) o *bin* (por ejemplo en [Wan et al., 2007]).

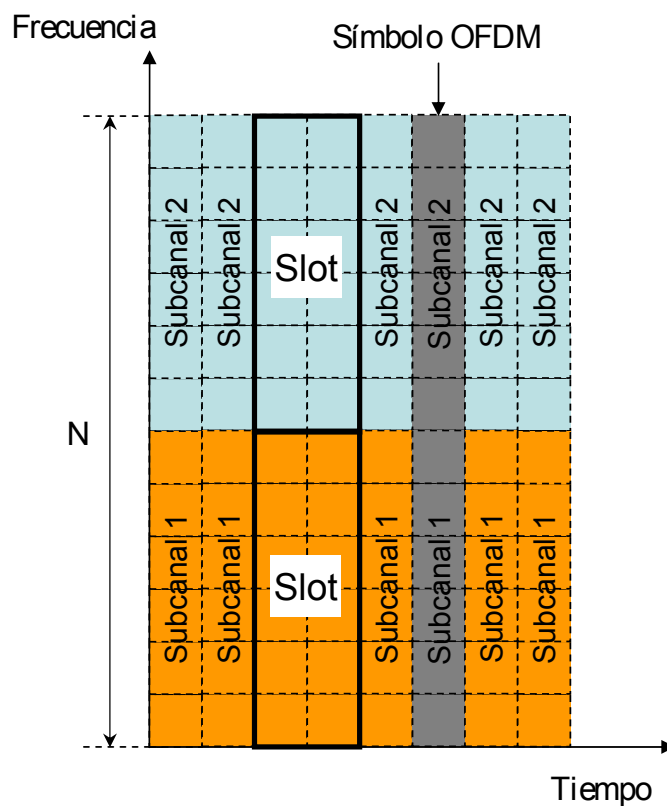


Figura 2.10. Estructura de la rejilla tiempo-frecuencia utilizando agrupaciones de subportadoras e intervalos de tiempo

Las subportadoras que constituyen un subcanal OFDM pueden ser adyacentes o estar distribuidas a lo largo del ancho de banda utilizado, dependiendo de lo que se denomina el modo de permutación de subportadoras. Un modo de permutación distribuido proporciona una mejor diversidad por frecuencia, mientras que la distribución de subportadoras adyacentes permite explotar mejor la diversidad por multiusuario. En el estándar IEEE 802.16e [802, 2006b] se definen, para el enlace descendente, tres modos o esquemas de permutación:

- Uso completo de todas las subportadoras (*full usage of subcarriers*, FUSC), que emplea todas las subportadoras de datos para crear subcanales de 48 subportadoras, distribuidas regularmente sobre el ancho de banda utilizado.
- Uso parcial del conjunto de las subportadoras (*partial usage of subcarriers*, PUSC), análogo al modo FUSC, con la salvedad de que, una vez distribuidas las subportadoras en 6 grupos, pueden asignarse todos o sólo algunos de los 6 grupos para la transmisión.
- Modulación y codificación adaptativa de banda (*band adaptive modulation and coding*, AMC). En este modo, las subportadoras que constituyen un subcanal son adyacentes. Con esta distribución de la subportadoras, se pierde la diversidad por frecuencia, pero permite una explotación de la diversidad por multiusuario mucho más sencilla. Al igual que en el modo FUSC, cada subcanal está formado por 48 subportadoras.

2.7.2. La utilización de los subcanales en la literatura

Uno de los primeros trabajos que hace uso de la asignación de subcanales es la propuesta de [Kwon et al., 2005], donde se expone el diseño de un simulador para redes OFDMA, considerando la asignación de subcanales utilizando el modo AMC, con el objetivo de maximizar la tasa binaria.

Si se emplean los subcanales de forma conjunta con las ranuras de tiempo de la rejilla tiempo-frecuencia OFDM, donde el par subportadora/ranura forma un *slot*, la asignación de recursos en sistemas OFDMA queda simplificada, pasando de dimensionalidad 3 (tiempo, frecuencia y potencia) a dimensionalidad 2 (*slot* y potencia). Este planteamiento es seguido en [Hosein, 2006], donde se propone resolver la asignación de recursos mediante funciones de utilidad genéricas basadas en la tasa de bit. En este caso, los recursos que se asignan son los *slots* que

componen una trama (varios símbolos OFDM), asumiendo que todos los *slots* transmiten con la misma potencia y tienen la misma SINR.

En [Lengoumbi et al., 2007] se compara la eficiencia espectral que se obtiene en el enlace descendente mediante algoritmos de asignación óptima de subportadoras con los modos FUSC y AMC. Si bien existe una ligera pérdida en cuanto a la eficiencia espectral, la complejidad computacional es dos órdenes de magnitud inferior al utilizar la asignación de subcanales.

El trabajo de [Pfleetschinger et al., 2007] está dedicado a un esquema de adaptación de enlace, con desvanecimiento selectivo en frecuencia, que hace uso de los subcanales en la primera fase del proceso completo para realizar modulación adaptativa. A su vez, los subcanales son agrupados durante varios símbolos OFDM, dando lugar a unidad mínima de asignación que es denominada *chunk*. Relacionado con el anterior, en [Sternad et al., 2007] se propone el diseño de un sistema de transmisión OFDMA adaptativo con diferentes estructuras de los subcanales (aquí llamados *bin*) y se discute la utilización de diferentes códigos a nivel de subcanal.

Para realizar el traspaso de conversaciones en entornos móviles, el enfoque usual adoptado es reservar subportadoras específicas para tal fin. En redes WiMAX, los autores de [Qin et al., 2007] plantean la posibilidad de hacer reserva de potencia en vez de reservar ancho de banda. Para ello, proponen un algoritmo de asignación conjunto de subcanales y potencia que tiene como finalidad minimizar el consumo total de potencia, utilizando el modo AMC para la distribución de subportadoras entre los subcanales.

La diversidad en tiempo y frecuencia asociada a los subcanales, aunque sin considerar diversidad multiusuario, es objeto del análisis de [Hosein y Lee, 2007], proponiendo una métrica para cuantificar la diversidad obtenida según la composición del subcanal, en términos de las subportadoras y los símbolos OFDM pertenecientes a una trama. Basado en dicha métrica, se formula el problema para optimizar la asignación de subportadoras y símbolos OFDM a cada subcanal. Dada la complejidad que plantea la resolución del problema, se proponen algunas soluciones heurísticas. Este trabajo tiene su continuación en [Einhaus et al., 2008], que evalúa el impacto de diferentes patrones subportadora/símbolo para formar un subcanal en términos de SNR y tasa de error de paquetes.

En [Brah et al., 2009] se propone un esquema de asignación de potencia y subcanales para el enlace descendente, asumiendo que el conocimiento de la in-

formación del estado del canal (*channel state information*, CSI) no es perfecto: en lugar de conocer la CSI instantánea, el transmisor conoce la función de distribución de probabilidad del canal. Mediante este esquema, se maximiza la tasa binaria ergódica, considerando además imparcialidad a largo plazo.

2.8. Un paso más en la transmisión adaptativa: la radio cognitiva

A lo largo de este capítulo hemos hecho un recorrido por las diferentes técnicas empleadas para modulación adaptativa, desde la más sencilla (modulación adaptativa para un enlace punto a punto) hasta los complejos algoritmos para transmisión multiportadora multiusuario (OFDMA). Una característica común a todas las técnicas anteriores es que es necesario recabar la información del entorno inalámbrico para posteriormente adecuar la transmisión a las circunstancias observadas.

Los escenarios que hasta ahora hemos contemplado son lo suficientemente complejos, sobre todo en el caso de redes OFDMA, como para considerar que, por un lado, la obtención de información fiable del entorno y, por otro, el posterior procesamiento de dicha información para asignar los recursos convenientemente resultan ser procesos de una alta dificultad. Si esto no fuera suficiente, las redes de nueva generación suponen un incremento de la dificultad, dado que está previsto el uso oportunista del espectro incluso en aquellas frecuencias hasta ahora asignadas en exclusiva a determinados servicios, como las frecuencias de los canales de televisión [FCC, 2005b]. Un nuevo tipo de dispositivos radio, las radios cognitivas, están siendo desarrolladas para hacer frente a estos retos [IEE, 2009]. Este tipo de radios son una evolución de la radio definida por software (*software defined radio*, SDR), que es un dispositivo radio cuyos parámetros de transmisión pueden ser alterados mediante un cambio *software*, sin alteración del *hardware* de la misma.

La **radio cognitiva** es un concepto acuñado por Mitolla III [Mitola III y Maguire Jr, 1999], que identifica un sistema de radio capaz de recoger las características del entorno, aprender de la información recogida en el pasado, tomar decisiones inteligentes y adaptar sus transmisiones consecuentemente. Una definición formal ampliamente aceptada entre la comunidad científica es la proporcionada por Haykin [Haykin, 2005], que define la radio cognitiva como “un sistema de comunicación inalámbrico inteligente que es consciente de su entorno circundante”. Si nos basamos en la SDR, una radio cognitiva se puede

interpretar como una SDR dotada de un “cerebro” que no se limita a reaccionar frente a los cambios externos, sino que “piensa” acerca de la acción a tomar. Podemos, por tanto, diferenciar dos características clave asociadas a la radio cognitiva: la capacidad cognitiva y la reconfigurabilidad.

La capacidad cognitiva la podemos considerar constituida, esencialmente, por 3 tareas relacionadas (figura 2.11):

1. Análisis del entorno radio, que lleva a cabo la detección del espectro para determinar los huecos espectrales tiempo-frecuencia de transmisión y la estimación de la interferencia.
2. Caracterización del canal, consistente en estimar la CSI y la capacidad del canal.
3. Gestión del espectro y de potencia, a partir de la información suministrada por las dos tareas anteriores.

La reconfigurabilidad dota a la radio cognitiva de la capacidad de cambiar los valores de los parámetros de transmisión para adaptar la señal transmitida al nuevo entorno. Una SDR soportará generalmente esta funcionalidad.

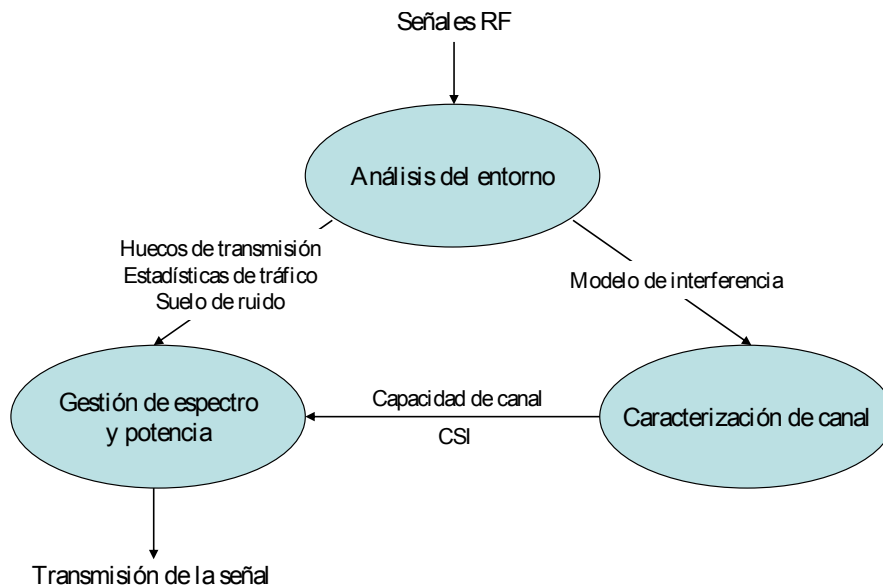


Figura 2.11. Tareas correspondientes a la capacidad cognitiva

2.8.1. Redes de nueva generación: acceso oportunista al espectro

Uno de los problemas inherentes a las comunicaciones inalámbricas es que la banda de frecuencias utilizables es finita, lo que deriva en que el espectro radioeléctrico

se haya considerado históricamente como un recurso escaso. No obstante, en redes que tienen asignadas unas determinadas frecuencias, como la difusión de televisión o las redes de telefonía móvil, el problema es más bien de eficiencia en la utilización del espectro, ya que el porcentaje de tiempo que los usuarios hacen uso de dichas frecuencias es habitualmente bastante bajo [Chiang et al., 2007].

En estas redes, se diferencian los usuarios en función de los derechos adquiridos para la utilización de las frecuencias asignadas a la red en cuestión:

- Los usuarios primarios (también llamados licenciados) son aquellos que están autorizados a utilizar las frecuencias de la red mediante la contratación de servicios, y que sólo emplean dichas frecuencias para sus comunicaciones. En este caso encontramos, por ejemplo, los usuarios de una red de telefonía móvil 3G.
- Los usuarios secundarios (o no-licenciados) que no pagan por el uso de esas frecuencias, de manera que harán un uso oportunista de las mismas en la medida que lo permitan las transmisiones de los usuarios primarios. Los usuarios secundarios se han de comprometer a no interferir las comunicaciones de los usuarios primarios.

La diferenciación anterior nos lleva a considerar dos casos bien distintos en cuanto al espectro. Si el espectro es libre, esta diferenciación entre usuarios no tiene mayor relevancia, por cuanto todos los usuarios accederán al espectro de manera oportunista. Si el espectro está licenciado, existen dos políticas de transmisión para los usuarios secundarios:

1. Solapamiento espectral (*spectrum underlay*): los usuarios secundarios pueden compartir con los primarios las mismas bandas de frecuencia simultáneamente, a condición de que el nivel de la interferencia provocada sobre estos últimos se encuentre por debajo de un determinado umbral.
2. Sin solapamiento espectral (*spectrum overlay*): los usuarios secundarios sólo pueden transmitir en aquellas bandas que se encuentren libres en un determinado instante de tiempo; a estos períodos de tiempo durante los que una banda de frecuencia no está ocupada por un usuario se les denomina **huecos espectrales** (figura 2.12).

Resumiendo lo anterior, los usuarios secundarios no podrán causar ni colisiones en el canal ni interferencia perjudicial para los usuarios primarios.

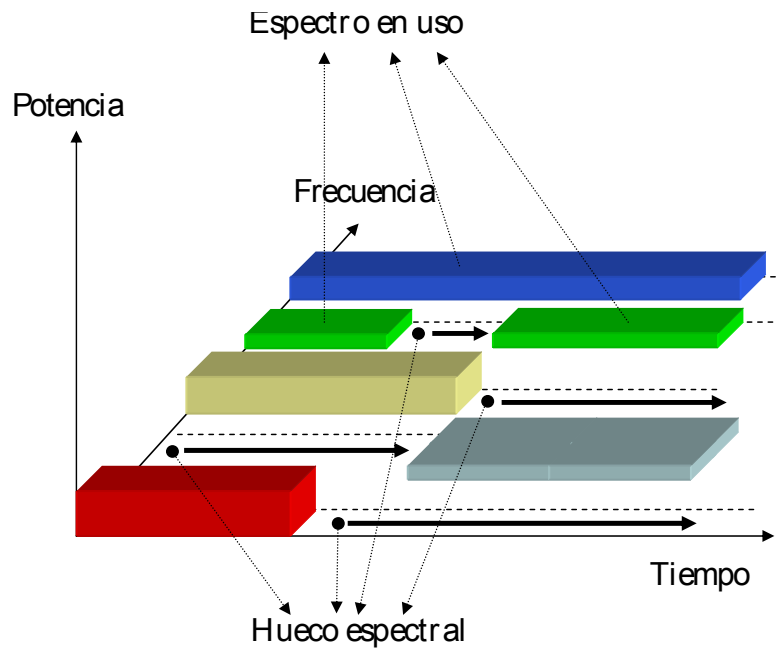


Figura 2.12. Huecos espectrales

Acceso al espectro

En redes cognitivas, el acceso o asignación del espectro añade nuevas características respecto a las redes convencionales. Por un lado, el espectro de transmisión no consiste en una banda concreta, sino que está formado por los espectros asignados a distintos servicios (bandas de TV, bandas de telefonía móvil). Además, coexisten dos tipos de usuarios, licenciados y no-licenciados, con diferentes privilegios a la hora de acceder al espectro. Si comparamos esta situación con una red basada en OFDMA, donde la dificultad se encuentra fundamentalmente en asignar las J subportadoras disponibles entre los K usuarios, en una red de radios cognitivas con uso oportunista del espectro no es posible conocer a priori ni siquiera las bandas de frecuencia disponibles para la transmisión, tanto en espectro licenciado como no licenciado. Eso implica que la asignación de los canales de transmisión para los usuarios estará íntimamente ligada a la detección del espectro disponible.

Dada la naturaleza de estas redes, el acceso al espectro es dinámico, y nos referiremos a él por sus siglas en inglés DSA (*dynamic spectrum access*). El DSA en redes cognitivas se caracteriza por ser un proceso en tres etapas, como se detalla en la figura 2.13 [Akyildiz et al., 2006]:

1. Detección del espectro, enfocada en la detección de los huecos de transmisión
2. Análisis espectral, para la caracterización de las bandas libres, fundamen-

talmente en términos de capacidad.

3. Decisión de espectro a usar, en base a la información proporcionada por las etapas anteriores.

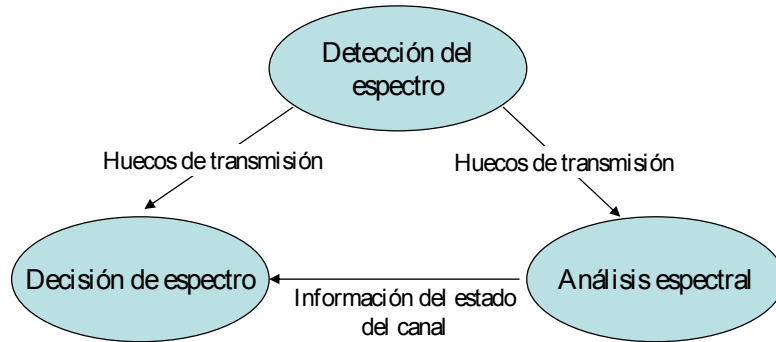


Figura 2.13. Acceso al espectro en redes cognitivas

El estudio conjunto de las tres etapas del DSA descritas ha dado lugar a un número importante de protocolos de acceso al medio MAC (*medium access control*), entre los que cabe destacar las propuestas de [Vuran y Akyildiz, 2007] y [Chou et al., 2007].

Un concepto directamente relacionado con el DSA es la compartición dinámica del espectro (*dynamic spectrum sharing*), consistente en realizar el reparto del ancho de banda disponible. Al contrario que en el problema de asignación de recursos discreto, por ejemplo para OFDM, donde hay una asignación discreta de las subportadoras, ahora el espectro se modela como una variable continua cuyo valor es el ancho de banda asignado al usuario. Entre otros factores, hay que considerar la interferencia originada, tanto para usuarios primarios como para secundarios, tratado por ejemplo en [Atakan y Akan, 2007]. En este tipo de redes, la compartición del espectro no se restringe a los usuarios, sino que también los proveedores de servicio que operan en una misma área pueden actuar de igual manera para atender a sus usuarios en cada instante de tiempo (ver, por ejemplo, [Acharya y Yates, 2007]).

Sin embargo, una de las características más importantes de estas radios es que su capacidad cognitiva les permite tomar decisiones inteligentes respecto a la utilización del espectro, basándose en las estimaciones realizadas a partir de la detección de espectro. De esta forma, el DSA se puede enfocar desde una perspectiva completamente diferente a los habituales esquemas basados en protocolos de acceso al medio MAC centralizados. En este contexto, la Teoría de Juegos es una

herramienta ampliamente utilizada, ya que permite modelar el comportamiento egoísta de los usuarios en estos casos. Una buena introducción a este tema es [Ji y Liu, 2007], así como las referencias allí contenidas.

2.8.2. Asignación de recursos utilizando radios cognitivas

Las expectativas creadas por los radios cognitivos han motivado un crecimiento en el número de estudios de la aplicación de estos radios en diferentes áreas de las comunicaciones inalámbricas. Así, la actividad investigadora es intensa en aspectos considerados clave en el nuevo entorno de las redes cognitivas, tales como la detección de espectro o el análisis de interferencia.

La detección del espectro afronta en redes con radios cognitivas un paradigma completamente nuevo caracterizado por:

- la detección de la señal transmitida tanto por usuarios primarios como por secundarios;
- la detección de las señales transmitidas en múltiples bandas de frecuencia, tantas como espectros licenciados y no licenciados se pretenda utilizar.

Esto conlleva la necesidad de esquemas de detección de espectro distintos a los tradicionales, como por ejemplo las propuestas de [Wild y Ramchandran, 2005] y [Ganesan y Ye, 2007].

En cuanto a la determinación de la interferencia, el reto principal es caracterizarla a través de la denominada temperatura de interferencia [FCC, 2005a], representada por T_L , la cual generará una potencia de ruido térmico equivalente en el receptor $P_L = KBT_L$. Una vez determinado el valor de P_L , hay que garantizar que este valor está limitado a un umbral P_{lim} , para que no interfiera sobre el usuario primario, es decir $P_L < P_{lim}$; esta aproximación sirve de base para el estudio de la potencia de interferencia en [Bater et al., 2007].

Otras áreas, como la diversidad cooperativa y las técnicas CDMA (Code Division Multiple Access) y UWB (UltraWide Band), también se han visto beneficiadas con la aparición de este tipo de radios; véanse por ejemplo [Simeone et al., 2007], [Qu et al., 2008] y [Sahin et al., 2007], respectivamente.

En nuestro caso, nos centraremos en la utilización de radios cognitivas para asignación de recursos, contemplando su aplicación a la gestión de potencia y la modulación adaptativa. Si en otro tipo de redes la asignación de recursos tiene gran importancia, más lo tiene, si cabe, en redes cognitivas, donde la presencia o

ausencia de usuarios primarios implica situaciones completamente nuevas respecto a las redes tradicionales. La siguiente situación puede ilustrar estos nuevos retos. Supongamos un escenario de espectro licenciado sin solapamiento, donde no hay canales disponibles para la transmisión, y algunos de ellos se encuentran ocupados por usuarios secundarios. Si en estas condiciones aparece un usuario primario que quiere acceder a la red, algún usuario secundario habrá de ser desalojado del canal licenciado que ocupa. El secundario desalojado sufrirá una suerte de “traspaso” de la comunicación a otro canal, muy probablemente perteneciente a otra banda diferente. Si, bajo las mismas condiciones, suponemos que puede existir solapamiento en las transmisiones, siempre por debajo de un determinado umbral de interferencia, el usuario secundario puede que no deba abandonar el canal que ocupa, pero los usuarios secundarios habrán de adaptar sus transmisiones para no interferir sobre el usuario primario que acaba de acceder a la red.

Asignación de potencia de transmisión

En las redes inalámbricas basadas en el despliegue de estaciones base o nodos de acceso para cubrir áreas de servicio, el control de la potencia transmitida se puede realizar de forma centralizada, en el caso de que la EB realice las asignaciones de potencia tanto para la EB como para los usuarios, o de forma distribuida, de manera que son los usuarios quienes ejecutan el control de la potencia. En cualquier caso, en redes con radios cognitivas la dificultad es mucho mayor, puesto que el control de potencia debe realizarse sin tener conocimiento ni de la topología de la red ni de la naturaleza de la misma.

Si además el acceso al espectro es con solapamiento entre los usuarios primarios y los secundarios, la asignación de potencia presenta una restricción adicional: los usuarios secundarios tienen limitada su potencia de transmisión para no causar ninguna degradación sobre los usuarios primarios. Por tanto, la asignación de potencia ha de hacerse teniendo presente la restricción impuesta por la temperatura de interferencia. En [Bansal et al., 2007], la asignación de potencia se enfoca a maximizar la capacidad del enlace descendente en una red multiusuario basada en OFDM, con interferencia; este mismo objetivo se persigue en [Wang et al., 2007b], para redes basadas en CDMA. En [Setoodeh y Haykin, 2009], se analiza la asignación de potencia mediante el conocido algoritmo de *waterfilling* iterativo en redes cognitivas.

No obstante, no sólo han de ser tenidas en cuenta las transmisiones de los

usuarios primarios, sino también las de los secundarios. En este sentido, la asignación de potencia puede resultar más efectiva si se realiza conjuntamente con el control de admisión de usuarios secundarios, como muestran los autores de [Qian et al., 2007] y de [Zhang et al., 2007].

Como ya mencionamos en la sección anterior, la Teoría de Juegos resulta ser una herramienta matemática muy adecuada para las redes de radios cognitivas, y resulta especialmente útil para la asignación de potencia, dado que cada usuario, primario o secundario, se comportará egoístamente y tratará de transmitir la mayor potencia posible para obtener el máximo beneficio¹, en forma de tasa binaria u otra métrica de rendimiento adecuada. Por ejemplo, el problema de asignar potencia a los usuarios secundarios mediante Teoría de Juegos es estudiado en [Wang et al., 2007a] y [Clemens y Rose, 2005], para espectro licenciado y no licenciado respectivamente.

Modulación adaptativa

Las radios cognitivas permiten no sólo variar la potencia de transmisión en función del entorno, sino adoptar esquemas de transmisión adaptativos más complejos, como la modulación adaptativa. Por ejemplo, en [Gao et al., 2007], el objetivo es maximizar el tiempo de vida de una red de sensores basada en radios cognitivas para transmisión monoportadora. Esta maximización se logra con la adecuada elección del tamaño de la constelación QAM, cuya expresión del valor óptimo es obtenida analíticamente en dicho trabajo.

En redes con radios cognitivas, la utilización de OFDM y OFDMA resulta muy ventajosa gracias a la capacidad cognitiva y la reconfigurabilidad proporcionadas por estas radios. En [Budiarjo et al., 2006] se propone la combinación de la técnica de *bit loading*, usando diversos algoritmos como el de Chow [Chow, 1993], con el enventanado de la señal OFDM para reducir la interferencia sobre las bandas ocupadas en un sistema con espectro licenciado. Con un objetivo similar, los autores de [Qin y Leung, 2007] proponen un algoritmo de asignación de subportadoras y potencia por subportadora para optimizar la interferencia creada por los usuarios secundarios sobre un usuario primario en una red OFDM.

Otros sistemas multiportadora pueden asimismo aprovechar las características de las radios cognitivas. En [Qu et al., 2008], los autores definen un marco para el diseño de sistemas basados en radios cognitivas con modulación CDMA

¹Siempre que el objetivo no sea minimizar la energía

multiportadora, con dos bloques bien diferenciados. En primer lugar, para obtener información del canal como la disponibilidad de subportadoras y ganancia de canal, se especifica el procedimiento de detección de espectro y estimación de parámetros, realizado de forma distribuida por los usuarios. En segundo término, los recursos (subportadoras, y la potencia y bits por símbolo por subportadora) son asignados de manera que existe protección para los usuarios presentes en el sistema frente a la potencial interferencia creada por los usuarios que intentan acceder al mismo.

2.9. Conclusiones

En este capítulo se proporciona una revisión de las diferentes técnicas empleadas para adaptar la transmisión al medio radioeléctrico: modulación adaptativa, *bit-loading* y radio cognitiva. La utilización adecuada de estas técnicas permite cubrir los escenarios de única portadora y multiportadora, así como comunicaciones punto a punto y entornos multiusuario. Hemos destacado la importancia que tiene para el rendimiento de sistemas OFDMA la asignación de los recursos radio disponibles, y las ventajas que se derivan de emplear la asignación de agrupaciones de subportadoras, a las que hemos denominado subcanales. En sistemas que emplean acceso múltiple OFDMA, de especial interés en redes inalámbricas, hemos visto que la explotación de la diversidad multiusuario aporta una ganancia en el rendimiento del sistema respecto a si dicha diversidad no es explotada.

CAPÍTULO 3

MÉTODOS DE OPTIMIZACIÓN PARA PROBLEMAS CON RESTRICCIONES EN COMUNICACIONES INALÁMBRICAS

El propósito de este capítulo es proporcionar una revisión de las herramientas y métodos matemáticos para la resolución de problemas de optimización sujetos a un determinado conjunto de restricciones, que permitirán resolver los problemas formulados en los siguientes capítulos. La primera parte, constituida por la sección 3.1, presenta los métodos de optimización en variable entera a utilizar en el capítulo 5. En la segunda parte del capítulo, la sección 3.2, expondremos la resolución de problemas de variable continua que será empleada en el capítulo 4.

Como advertirá el lector, la sección 3.1 está formulada en términos de problemas de maximización, mientras que la sección 3.2 trata problemas de minimización, en concordancia con su aplicación a los problemas de los capítulos 5 y 4, respectivamente. La formulación de ambas teorías, de optimización entera y continua, en los términos contrarios es inmediata y no ha sido desarrollada en beneficio de la claridad expositiva y brevedad del capítulo.

3.1. Optimización de problemas en variable entera

La **programación entera** (en notación anglosajona *integer programming*, IP) presenta características muy diferentes respecto a la optimización en problemas continuos, como por ejemplo la no diferenciabilidad de las funciones, que hacen que los métodos y algoritmos sean completamente distintos y que en general demanden una mayor complejidad en su resolución. Una tendencia creciente entre la comunidad científica en materia de comunicaciones es formular ciertos problemas de optimización en términos de problemas de optimización enteros, fundamentalmente para la asignación de recursos no divisibles, como los canales de transmisión en sistemas multiportadora (ver, por ejemplo, [Lee y Leung, 2006], [Fu et al., 2008] y [Li y Liu, 2006]), lo cual hace pensar en su idoneidad para resolver problemas generales de asignación de recursos en entornos inalámbricos multiusuario OFDM, según vimos en el capítulo 2.

Esta sección está dedicada a describir cómo realizar la optimización de problemas enteros en general, entendiendo por problemas enteros aquellos en los que las variables del problema se restringen a valores enteros. Este enfoque generalista en los problemas enteros a tratar hace que no sean aquí objeto de estudio métodos más específicos como el algoritmo Húngaro [Kuhn, 1955], ya que, aunque empleado con cierta asiduidad para resolver problemas en comunicaciones (ver, por ejemplo, [Choi y Bahk, 2008]), su uso queda limitado a resolver de manera óptima el problema de asignación de N recursos a N usuarios. Dada la dificultad que supone en general el resolver los problemas IP, trataremos no sólo los métodos de resolución óptimos sino también los subóptimos; esto nos lleva a poder resolver los problemas de forma más sencilla, aunque a costa de obtener una solución aproximada del problema en lugar de la solución exacta.

El contenido de esta sección se basa en diferentes capítulos de [Bertsimas y Weismantel, 2005] y [Wolsey, 1998]; por consiguiente, remitimos al lector a dichas referencias para más detalles o ampliación de los aspectos aquí tratados. Empezaremos definiendo en la sección 3.1.1 los problemas enteros, y dedicaremos la sección 3.1.2 a los principales métodos de optimización entera y sus características. En la sección 3.1.3 y siguientes, trataremos diferentes métodos para resolver de forma subóptima problemas IP, la relación entre las soluciones obtenidas mediante estos métodos, y veremos bajo qué circunstancias la solución subóptima del problema es igual a la solución óptima.

3.1.1. ¿Qué es un problema entero binario?

Supongamos que tenemos el problema de optimización siguiente, que consiste en maximizar la función de coste o función objetivo z con unas ciertas restricciones:

$$\begin{aligned} \text{máx } z &= \mathbf{c}^T \mathbf{x} \\ \text{s.t. } \mathbf{Ax} + \mathbf{By} &= \mathbf{b} \\ \mathbf{x} &\geq 0, \mathbf{y} \geq 0 \end{aligned}$$

donde las matrices $\mathbf{A} \in \mathbb{Z}^{m \times n}$ y $\mathbf{B} \in \mathbb{Z}^{m \times k}$, los vectores $\mathbf{b} \in \mathbb{Z}^m$ y $\mathbf{c} \in \mathbb{Z}^n$ son vectores columna de dimensiones m y n , respectivamente, y $\mathbf{x} \in \mathbb{Z}^n$ e $\mathbf{y} \in \mathbb{Z}^k$ son vectores columna de n y k variables, respectivamente. A este problema, en el que todas las variables son enteras, es decir, cumplen lo que se conoce como restricción de integralidad, le denominaremos **problema entero** (*integer programming*, IP). Si alguna de las variables no es entera, por ejemplo \mathbf{y} , el problema se denomina **problema entero mixto** (*mixed integer programming*, MIP), quedando formulado como

$$\begin{aligned} \text{máx } z &= \mathbf{c}^T \mathbf{x} \\ \text{s.t. } \mathbf{Ax} + \mathbf{By} &= \mathbf{b} \\ \mathbf{x} &\geq 0, \mathbf{y} \geq 0, \mathbf{y} \in \mathbb{R}. \end{aligned}$$

Si las variables están restringidas a tomar los valores binarios 0–1, lo llamaremos **problema binario** (*binary integer programming* BIP), formulado como

$$\begin{aligned} \text{máx } z &= \mathbf{c}^T \mathbf{x} \\ \text{s.t. } \mathbf{Ax} + \mathbf{By} &= \mathbf{b} \\ \mathbf{x} &\in \{0, 1\}, \mathbf{y} \in \{0, 1\}. \end{aligned}$$

3.1.2. Métodos de optimización entera óptimos

El principal problema que afrontan los métodos para resolver problemas IP es que, a diferencia de la programación lineal (*linear programming*, que aborda la optimización de problemas lineales, y de la programación continua en general, no se trata con regiones de soluciones convexas o susceptibles de convertir en convexas, sino con conjuntos de puntos que, en general, no proporcionan óptimos locales que sean óptimos globales, lo que implica aproximaciones diferentes para calcular la solución óptima. De los métodos que se proponen en la literatura, tratamos a continuación los dos más empleados, *branch-and-bound* y *cutting plane*.

- *Método Branch-and-bound.* El método *branch-and-bound* es la forma más usual de resolver problemas de optimización enteros empleando enumeración implícita. Se basa en una exploración del tipo “divide y vencerás” del conjunto de soluciones posibles, de forma que se divide el conjunto de soluciones en subconjuntos. Si \mathcal{S} es el conjunto de soluciones del problema entero

$$\begin{aligned} \text{máx } z &= \mathbf{c}^T \mathbf{x} \\ \text{s.t. } \mathbf{x} &\in \mathcal{S} \end{aligned}$$

podemos partir \mathcal{S} en subconjuntos $\mathcal{S}_1, \dots, \mathcal{S}_k$ y resolver por separado cada uno de los k subproblemas resultantes, que denotaremos por \mathcal{P}_i

$$\begin{aligned} \text{máx } z &= \mathbf{c}^T \mathbf{x} \\ \text{s.t. } \mathbf{x} &\in \mathcal{S}_i, \quad i = 1, \dots, k. \end{aligned}$$

A continuación, una vez resueltos los k subproblemas, se elige la mejor de las soluciones. Ahora bien, como cada \mathcal{P}_i puede ser de la misma complejidad que el original, estos son susceptibles a su vez de ser resueltos dividiéndolos en nuevos subproblemas hasta que su resolución sea lo suficientemente sencilla, lo que representa la parte *branching* del método, proporcionando un árbol de subproblemas.

Lo habitual, sin embargo, es que la resolución exacta de los subproblemas no sea sencilla. Resulta en este caso más eficiente, para cada \mathcal{P}_i , calcular una cota superior $b(\mathcal{S}_i)$ del valor óptimo de la función objetivo del subproblema mediante un cierto algoritmo (por ejemplo, relajando la restricción entera), esto es $b(\mathcal{S}_i) \geq \text{máx}_{\mathbf{x} \in \mathcal{S}_i} \mathbf{c}^T \mathbf{x}$. Esta es la clave del método: mientras obtener una solución exacta puede ser muy complejo, obtener una cota superior puede resultar muy sencillo. Esto permite mantener una cota U del valor óptimo de z , correspondiente al valor de la mejor solución encontrada hasta el momento. Así, se podrán entonces descartar aquellos subproblemas \mathcal{P}_i para los que se haya obtenido una cota superior $b(\mathcal{S}_i) \leq U$.

A la hora de emplear los métodos *branch-and-bound*, existen dos tipos de problemas a los que se aplican principalmente. Como adelantábamos en la introducción de esta sección, proporciona buenos resultados en la minimización de la potencia total de transmisión en sistemas multiusuario multiporadora (por ejemplo, OFDM y OFDMA) mediante asignación de canales o

subportadoras ([Zhang y Letaief, 2006], [Feiten et al., 2008]). Por otro lado, el *bit-loading* (tratado en 2.3.3) se puede resolver formulándolo como un problema de optimización entera, donde los bits por portadora a asignar m_k se consideran variables enteras que pueden tomar valores de un conjunto finito $\mathcal{M} = \{0, 2, \dots, K\}$, siendo 2^K el nivel de la constelación empleada. Entre los trabajos que adoptan este enfoque para resolver el *bit loading* destacamos [Mao y Wang, 2008], [Madan et al., 2006] y [Zhang y Letaief, 2004].

- Algoritmos *cutting plane*. El método *cutting plane* consiste en refinar de manera iterativa el conjunto de soluciones de un problema entero \mathcal{P} mediante la adición de nuevas restricciones lineales, a las que se denomina *cut*. Este método funciona de la siguiente forma. Se resuelve el problema original relajado, que designamos mediante $\tilde{\mathcal{P}}$, eliminando la restricción entera. Si la solución obtenida \mathbf{x}^* es entera, ésta es la solución del problema; en caso contrario, se añade una nueva restricción lineal de la forma $\mathbf{A}\mathbf{x} \leq \mathbf{b}$ que permita separar la solución no entera del conjunto de soluciones enteras, y se resuelve el nuevo problema resultante, procediendo así hasta obtener una solución entera. La clave de este método, y lo que da lugar a la gran mayoría de los diferentes algoritmos, es cómo obtener las restricciones lineales o *cuts* que van separando la solución entera óptima del recubrimiento convexo del conjunto de soluciones posibles. Formalmente, dado el problema entero \mathcal{P}

$$\begin{aligned} \text{máx } z &= \mathbf{c}^T \mathbf{x} \\ \text{s.t. } \mathbf{x} &\in \mathcal{S} \\ \mathbf{x} &\in \mathbb{Z}^n, \end{aligned}$$

en primer lugar se resuelve el problema relajado $\tilde{\mathcal{P}}$

$$\begin{aligned} \text{máx } z &= \mathbf{c}^T \mathbf{x} \\ \text{s.t. } \mathbf{x} &\in \mathcal{S}, \end{aligned}$$

que proporciona una solución \mathbf{x}_0^* ; si $\mathbf{x}_0^* \in \mathbb{Z}^n$, ésta es la solución del problema entero. En caso contrario, añadimos una nueva restricción lineal. Sea $\mathcal{F} = \{\mathbf{A}_q \mathbf{x} \leq \mathbf{b}_q, q = 1, \dots, Q\}$ el conjunto de Q restricciones lineales para

realizar la separación; ahora el problema a resolver es $\tilde{\mathcal{P}}_1$

$$\begin{aligned} \text{máx } z &= \mathbf{c}^T \mathbf{x} \\ \text{s.t. } \mathbf{A}_1 \mathbf{x} &\leq \mathbf{b}_1 \\ \mathbf{x} &\in \mathcal{S} \end{aligned}$$

cuya solución es \mathbf{x}_1^* ; si $\mathbf{x}_1^* \in \mathbb{Z}^n$, ésta es la solución óptima y si no, continuamos añadiendo restricciones de \mathcal{F} hasta obtener la solución entera óptima.

Por ejemplo, en [Liu et al., 2008], se desarrolla un algoritmo *cutting plane* empleado en la resolución conjunta de enrutamiento y las asignaciones de potencia y canales para redes *ad hoc* que utilizan antenas MIMO.

Una variación de los métodos *branch-and-bound* son los métodos *branch-and-cut*, combinación de los dos métodos anteriormente descritos y que consisten, básicamente, en obtener las cotas superiores de los subproblemas mediante *cuts* en lugar de mediante relajación lineal en un método *branch-and-bound*. En [Althaus et al., 2003] se propone un algoritmo *branch-and-cut* para la optimización de la potencia total consumida en redes *ad hoc*; para optimizar la potencia, el problema se formula a través de la asignación de enlaces entre los diferentes nodos que componen la red.

La cuestión inmediata que surge ahora, una vez vistos ambos métodos, es cuándo emplear *branch-and-bound* y cuándo *cutting plane*. La respuesta está, fundamentalmente, en la estructura del problema a resolver: si se comprende cómo es el problema será relativamente sencillo encontrar *cuts* que permitan solucionar el problema de forma óptima y garanticen la convergencia en pocas iteraciones. Por otro lado, el método *branch-and-bound* es más sencillo de aplicar y es computacionalmente menos exigente, ya que, al contrario que los métodos *cutting plane* que añaden nuevas restricciones al problema, se basan en la resolución de problemas más sencillos que el original.

3.1.3. Métodos de optimización entera subóptimos

A menudo, la resolución de problemas IP de manera óptima representa una complejidad excesiva, de forma que, si podemos encontrar un método más sencillo que nos proporcione una solución subóptima z' lo suficientemente buena, optaremos por el método subóptimo. Por tanto, se trata de encontrar cotas superiores (o inferiores, si el problema de optimización es de minimización) lo más ajustadas

posibles a la solución óptima z^* . La aproximación más importante es mediante **relajación**, que consiste en reemplazar un problema IP difícil de resolver por otro más sencillo cuyo valor óptimo de la función objetivo será mayor (menor, si es minimización) que z^* . La primera y más inmediata relajación es obvia: convertir el problema IP en un problema en variable real, relajando la restricción de integralidad de forma que ahora $\mathbf{x} \in \mathbb{R}^n$; así, en vez de resolver el problema entero

$$\begin{aligned} \text{máx } z &= \mathbf{c}^T \mathbf{x} \\ \text{s.t. } \mathbf{Ax} &\leq \mathbf{b} \\ \mathbf{Dx} &\leq \mathbf{d} \\ \mathbf{x} &\in \mathbb{Z}^n \end{aligned} \tag{P}$$

se resuelve el problema

$$\begin{aligned} \text{máx } z &= \mathbf{c}^T \mathbf{x} \\ \text{s.t. } \mathbf{Ax} &\leq \mathbf{b} \\ \mathbf{Dx} &\leq \mathbf{d} \\ \mathbf{x} &\in \mathbb{R}^n. \end{aligned} \tag{PL}$$

No obstante, incluso así, el problema relajado puede resultar difícil de resolver. En este caso, podemos hacer uso de la dualidad y transformar el problema en su dual, relajando la restricción que más complica el problema; hablamos entonces de **relajación de Lagrange**, por cuanto se resuelve el problema dual, en lugar del problema primario (P). Si, por ejemplo, la condición que complica la resolución de (P) es $\mathbf{Ax} \leq \mathbf{b}$, podemos reformular el problema construyendo el problema dual relajado

$$\begin{aligned} \text{máx } z_D(\lambda) &= \mathbf{c}^T \mathbf{x} - \lambda(\mathbf{Ax} - \mathbf{b}) \\ \text{s.t. } \mathbf{Dx} &\leq \mathbf{d}, \end{aligned} \tag{PD}$$

es decir, añadimos a la función de coste z la restricción de interés como el término de penalización $\lambda(\mathbf{Ax} - \mathbf{b})$, donde $\lambda \geq 0$ es la variable dual y representa el “precio” asociado a la restricción que relajamos.

3.1.4. Bondad de las relajaciones lineal y de Lagrange

La resolución de problemas subóptimos resulta interesante, lógicamente, si el valor de z que proporcionan, z' , es una buena aproximación del valor óptimo asociado

al problema original z^* . De hecho, si conseguimos que ambos valores coincidan, estaremos en el caso ideal, ya que hemos resuelto el problema óptimo con mucho menos esfuerzo. Desafortunadamente, el caso más habitual es que esto no suceda, como indican los siguientes teoremas y proposiciones. Haremos notar que se incluyen las demostraciones del teorema 3.1 y del lema 3.1, dado que en la bibliografía consultada sólo figuran sus equivalentes para problemas enteros de minimización.

Teorema 3.1 (Dualidad débil). *Si el problema (P) tiene solución óptima que proporciona el valor óptimo z^* , y $z_D^*(\lambda)$ es el valor óptimo asociado al problema dual (PD), entonces $z^* \leq z_D^*(\lambda)$.*

Demostración. Sea \mathbf{x}^* una solución del problema (P); entonces, ha de cumplirse que $\mathbf{Ax}^* - \mathbf{b} \leq \mathbf{0}$; además, como $\mathbf{x}^* \in \{\mathbf{x} \in \mathbb{Z}^n | \mathbf{Dx} \leq \mathbf{d}\}$:

$$z^* = \mathbf{c}^T \mathbf{x}^* \leq \mathbf{c}^T \mathbf{x}^* + \lambda (\mathbf{b} - \mathbf{Ax}^*),$$

y dado que $z_D(\lambda) = \max \mathbf{c}^T \mathbf{x} + \lambda (\mathbf{b} - \mathbf{Ax})$, obtenemos $\mathbf{c}^T \mathbf{x}^* + \lambda (\mathbf{b} - \mathbf{Ax}^*) \leq z_D^*(\lambda)$. Por tanto $z^* \leq z_D^*(\lambda)$. \square

Este teorema nos dice que la solución del problema dual (PD) supondrá una cota superior del problema original o primario (P). Para obtener la mejor cota posible z_D^* en relación a z^* , no habrá más que resolver el problema

$$\begin{aligned} z_D^* &= \min z_D(\lambda) && \text{(Pd)} \\ &\text{s.t. } \lambda \geq 0. \end{aligned}$$

De forma análoga, se relacionan el valor asociado al problema dual y el obtenido mediante la relajación lineal. Para ello, haremos uso del siguiente teorema [Bertsimas y Weismantel, 2005], que caracteriza el dual lagrangiano (Pd) como un problema lineal.

Teorema 3.2. *Sea $\mathcal{F} = \{\mathbf{x} \in \mathbb{Z}^m | \mathbf{Dx} \leq \mathbf{d}\}$ y $\text{conv}(\mathcal{F}) = \{\mathbf{x} | \mathbf{Dx} \leq \mathbf{d}, \mathbf{x} \geq \mathbf{0}\}$ el recubrimiento convexo de \mathcal{F} . El valor óptimo z_D^* del problema (Pd) es igual a la solución del siguiente problema:*

$$\begin{aligned} \max z_L &= \mathbf{c}^T \mathbf{x} \\ \text{s.t. } \mathbf{Ax} &\leq \mathbf{b} && \text{(Peq)} \\ \mathbf{x} &\in \text{conv}(\mathcal{F}). \end{aligned}$$

De este teorema, podemos extraer el siguiente lema.

Lema 3.1. *El valor óptimo asociado al problema lineal z_L^* es una cota superior del valor óptimo del dual z_D^* : $z_L^* \geq z_D^*$.*

Demostración. Dado que $\text{conv}(\mathcal{F}) \subseteq \{\mathbf{x} \in \mathbb{Z}^m \mid \mathbf{D}\mathbf{x} \leq \mathbf{d}\}$, el problema lineal (PL) es una relajación del problema (Peq), y por tanto, proporciona una cota superior a la solución de (Peq). Como el valor óptimo del dual z_D^* viene dado por la resolución de (Peq) (teorema 3.2), de manera inmediata tenemos que $z_L^* \geq z_D^*$. \square

Estamos ahora en condiciones de al menos establecer una relación entre el valor óptimo del problema original (P) y las cotas obtenidas mediante la relajación de Lagrange y la de la restricción entera. Considerando el lema 3.1 y el teorema 3.1, obtenemos que

$$z_L^* \geq z_D^* \geq z^*. \quad (3.1)$$

Ahora bien, ¿existen condiciones bajo las cuales las cotas corresponden al valor exacto de la solución? En este caso, sería posible calcular la solución óptima mediante cualquiera de los métodos subóptimos expuestos, lo que tiene indudables ventajas en términos de complejidad. En las siguientes secciones, veremos cuáles son estas condiciones para las dos desigualdades de (3.1).

3.1.5. Igualdad de las soluciones del problema entero y el problema relajado lagrangiano

El óptimo correspondiente a (P) se podrá obtener a través del problema relajado lagrangiano aplicando el siguiente corolario [Bertsimas y Weismantel, 2005], que se basa en la geometría de las restricciones del problema.

Corolario 3.1. $z_D^* = z^* \Leftrightarrow \text{conv}(\mathcal{F} \cap \{\mathbf{x} \mid \mathbf{A}\mathbf{x} \leq \mathbf{b}\}) = \text{conv}(\mathcal{F}) \cap \{\mathbf{x} \mid \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$.

Existe asimismo un conjunto de condiciones que garantizan que la resolución del problema entero y la del problema dual proporcionan el mismo valor de la función objetivo [Wolsey, 1998], contempladas en la siguiente proposición.

Proposición 3.1. *Sea $z_D(\mathbf{x}^*(\lambda))$ el valor óptimo del problema dual. Para $\lambda \geq 0$, si se cumple que*

1. $\mathbf{D}\mathbf{x}^*(\lambda) \leq \mathbf{d}$
2. $(\mathbf{D}\mathbf{x}^*(\lambda))_i = \mathbf{d}_i$ para $\lambda_i > 0$,

entonces el óptimo del problema entero es $z_D^ = z_D(\mathbf{x}^*(\lambda)) = z(\mathbf{x}^*(\lambda)) = z^*$.*

3.1.6. Igualdad de las soluciones del problema entero y el problema relajado lineal

En el mejor de los casos, podremos resolver el problema entero simplemente relajando la restricción de variable entera. Para ello, empleamos el siguiente corolario, similar al corolario 3.1 [Bertsimas y Weismantel, 2005].

Corolario 3.2. $z_L^* = z_D^* \Leftrightarrow \text{conv}(\mathcal{F}) = \{\mathbf{x} \mid \mathbf{D}\mathbf{x} \leq \mathbf{d}\}$.

Por tanto, si se cumplen las condiciones de los corolarios 3.1 (o la proposición 3.1) y 3.2, tendremos que $z^* = z_L^* = z_D^*$. Ahora bien, probar las igualdades de los recubrimientos convexos de ambos corolarios puede resultar bastante complicado, especialmente si el número de restricciones del problema es elevado. Una alternativa ampliamente utilizada es recurrir a la teoría de matrices unimodulares, dado que es posible resolver el problema entero mediante el problema relajado si la matriz de la restricción es totalmente unimodular.

Definición 3.1 (Matriz totalmente unimodular). *Una matriz \mathbf{A} es totalmente unimodular si cada submatriz cuadrada de \mathbf{A} tiene determinante de valor ± 1 ó 0 .*

Probar que una matriz es totalmente unimodular es sencillo [Wolsey, 1998], por lo que el siguiente teorema permite calcular la solución óptima del problema entero resolviendo el problema lineal relajado [Wolsey, 1998].

Teorema 3.3. *El problema lineal*

$$\begin{aligned} & \text{máx } \mathbf{c}^T \mathbf{x} \\ & \text{s.t. } \mathbf{A}\mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \in \mathbb{R}^n \end{aligned}$$

tiene una solución óptima entera \mathbf{x}^ si y sólo si la matriz \mathbf{A} es totalmente unimodular.*

3.2. Optimización de problemas en variable continua

Existe una extensa bibliografía que trata la optimización de problemas de variable continua con restricciones (ver, por ejemplo, [Bertsekas, 1999],

[Bazaraa et al., 2006] y [Luenberger, 2003]) donde encontrar una completa variedad de métodos y algoritmos sobre esta materia. El objeto de esta sección no es realizar un estudio exhaustivo de dichos métodos, para lo cual remitimos al lector a las mencionadas referencias, sino hacer una breve discusión de las diferentes formas de resolver la optimización en variable continua y los métodos de optimización más utilizados en comunicaciones inalámbricas. Nos centraremos por tanto en la teoría de los multiplicadores de Lagrange (subsección 3.2.1), los métodos de optimización mediante las condiciones de Karush-Kuhn-Tucker (subsección 3.2.2) y la teoría de dualidad (subsección 3.2.3); estas dos últimas clases de métodos tienen precisamente como base la teoría de Lagrange de (3.2.1).

A la hora de abordar la optimización de problemas de variable continua, es esencial diferenciar entre los distintos tipos de problemas:

- *Optimización lineal*: generalmente conocida como **programación lineal**, un problema de optimización lineal es aquel cuya función objetivo a optimizar y todas las restricciones son lineales. Aun siendo un caso particular de la programación no lineal, las condiciones de linealidad permiten métodos específicos especialmente eficientes.
- *Optimización convexa*: este tipo de optimización trata problemas en los que tanto las restricciones como la función objetivo son convexas, lo que da lugar, al igual que en el caso de optimización lineal, a un paradigma que permite desarrollar algoritmos muy eficientes. Su aplicación al ámbito de las telecomunicaciones es cada vez más relevante, como se refleja en [Luo y Yu, 2006], donde los autores hacen una revisión de los conceptos básicos de optimización convexa y de los avances más significativos en la aplicación de métodos convexos a problemas de comunicaciones y procesado de señal (ver referencias citadas en el mencionado trabajo).
- *Optimización no lineal*: también conocida como **programación no lineal**, constituye el tipo más genérico de problemas. Tiene por objeto resolver problemas de optimización donde bien las funciones objetivo, bien las restricciones, o bien ambas, son no lineales.

Como se verá en el capítulo 4, nuestro interés se centrará en problemas no lineales, por lo que dedicaremos el resto de esta sección a los principales métodos de resolución de este tipo de problemas.

3.2.1. Optimización no lineal con restricciones de igualdad: multiplicadores de Lagrange

Mediante los multiplicadores de Lagrange es posible obtener la solución óptima en problemas de optimización cuyas restricciones se establecen en términos de igualdad. De forma general, el método de los multiplicadores de Lagrange proporciona óptimos locales de la función objetivo, como veremos a continuación. No obstante, en determinadas circunstancias, como indica el teorema 3.6, existe una condición necesaria y suficiente para que el óptimo obtenido sea un óptimo global.

En la literatura, podemos encontrar numerosos trabajos en los que el método de Lagrange es empleado para resolver problemas de optimización. Más concretamente, en entornos OFDM multiusuario y a modo de muestra representativa, podemos citar las referencias siguientes. En [Jang y Lee, 2003], mediante multiplicadores de Lagrange la potencia disponible es distribuida, de forma que se maximiza la tasa de transmisión total, habiendo previamente asignado cada portadora al usuario que mejor ganancia de canal obtiene en dicha portadora. Mientras, en [Wong et al., 1999] y en [Hosein, 2006] se aplica a la asignación de recursos radio: en el primer caso, para la asignación de potencia, portadora y bits por símbolo; en el segundo, para asignar ancho de banda, considerando funciones de utilidad basadas en la tasa de transmisión de bits. Y más recientemente, los autores de [Lau et al., 2008] realizan el análisis asintótico del compromiso entre la ganancia de *goodput*¹ y la ganancia por diversidad de paquete, bajo condiciones no perfectas de conocimiento del canal de transmisión, vía multiplicadores de Lagrange.

Para formular el método de Lagrange y enunciar los teoremas asociados, partimos del problema con restricciones en igualdad siguiente

$$\begin{aligned} & \text{mín } f(\mathbf{x}) \\ & \text{s.t. } g(\mathbf{x}) = \mathbf{b} \\ & \mathbf{x} \in \mathbb{R}^n, \mathbf{b} \in \mathbb{R}^m; \end{aligned} \tag{Q}$$

se define la **función de Lagrange** o **lagrangiano** como

$$L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) - \boldsymbol{\lambda}^T [g(\mathbf{x}) - \mathbf{b}], \tag{3.2}$$

y el vector m -dimensional $\boldsymbol{\lambda}$ recibe el nombre de vector de los **multiplicadores de Lagrange**. Definimos también el vector gradiente de la función de Lagrange

¹Este concepto será ampliamente tratado en el capítulo 5

como

$$\nabla L(\mathbf{x}, \boldsymbol{\lambda}) = \begin{pmatrix} \frac{\partial L(\mathbf{x}, \boldsymbol{\lambda})}{\partial x_1} \\ \dots \\ \frac{\partial L(\mathbf{x}, \boldsymbol{\lambda})}{\partial x_n} \\ \frac{\partial L(\mathbf{x}, \boldsymbol{\lambda})}{\partial \lambda_1} \\ \dots \\ \frac{\partial L(\mathbf{x}, \boldsymbol{\lambda})}{\partial \lambda_m} \end{pmatrix} = \begin{pmatrix} \nabla_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}) \\ \nabla_{\boldsymbol{\lambda}} L(\mathbf{x}, \boldsymbol{\lambda}) \end{pmatrix} = \begin{pmatrix} \nabla_{\mathbf{x}} f(\mathbf{x}) - \sum_{i=1}^m \lambda_i \nabla g_i(\mathbf{x}) \\ -[g(\mathbf{x}) - \mathbf{b}] \end{pmatrix}.$$

Una vez definida la función de Lagrange y su gradiente, podemos calcular los puntos candidatos a óptimos locales, llamados puntos críticos, y establecer la condición necesaria de óptimo local (teorema 3.4).

Definición 3.2 (Punto crítico). *El punto $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$ es un punto crítico de la función de Lagrange 3.2 si $\nabla L(\mathbf{x}^*, \boldsymbol{\lambda}^*) = 0$.*

Teorema 3.4 (Condición necesaria de óptimo local). *Si \mathbf{x}^* es un óptimo local del problema (Q), existe un vector de multiplicadores $\boldsymbol{\lambda}^*$ tal que $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$ es un punto crítico de la función de Lagrange.*

Por tanto los puntos críticos obtenidos mediante la condición necesaria son los candidatos a puntos óptimos. Ahora bien, es necesaria una condición suficiente que determine los verdaderos puntos óptimos.

Teorema 3.5 (Condición suficiente de óptimo local). *Dado el problema (Q), donde las funciones f y g son diferenciables con continuidad dos veces; si $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$ es un punto crítico de la función de Lagrange, \mathbf{x}^* es un máximo local si la forma cuadrática asociada a la matriz hessiana de la función de Lagrange respecto a \mathbf{x} , $H_{\mathbf{x}}L(\mathbf{x}^*, \boldsymbol{\lambda}^*)$, es definida negativa (positiva para un mínimo local) (ver Apéndice A) respecto a los vectores del plano tangente a la superficie restricción $g(\mathbf{x}) = b$, es decir, respecto a $\mathbf{h} \in \mathbb{R}^n$ tales que $\mathbf{h}^T Jg(\mathbf{x}^*) = 0$, siendo $Jg(\mathbf{x})$ la matriz jacobiana de la función vectorial g (ver Apéndice A).*

El inconveniente de las condiciones necesaria y suficiente anteriores (teoremas 3.4 y 3.5 respectivamente) es que se restringen a óptimos locales, con lo que la obtención del óptimo global estaría aún por determinar. En ciertas circunstancias podemos encontrar una única condición necesaria y suficiente para calcular el óptimo de manera global sin que sea necesario acudir a los óptimos locales. Dicha condición se concreta en el teorema siguiente.

Teorema 3.6 (Condición necesaria y suficiente de solución óptima global). *En problemas convexos con restricciones de igualdad, \mathbf{x}^* es la solución óptima global si y sólo si existe $\boldsymbol{\lambda}^*$ tal que $\nabla L(\mathbf{x}^*, \boldsymbol{\lambda}^*) = \mathbf{0}$.*

3.2.2. Optimización no lineal con restricciones de desigualdad: condiciones de Karush-Kuhn-Tucker (KKT)

Las condiciones de Karush-Kuhn-Tucker ([Karush, 1939] y [Kuhn y Tucker, 1951]), o condiciones KKT, establecen condiciones necesarias de optimalidad para una solución en problemas no lineales, siempre que se cumplan ciertas condiciones de regularidad, y se pueden interpretar como una extensión del método de los multiplicadores de Lagrange a problemas con restricciones en desigualdades.

Un ejemplo clásico de aplicación de las condiciones KKT a problemas en comunicaciones inalámbricas es la distribución de la potencia total disponible entre N canales independientes gaussianos y paralelos mediante *waterfilling* (2.3.1). Más recientemente, su aplicación ha permitido resolver problemas de muy diversa naturaleza en sistemas OFDM, desde la asignación óptima de recursos, como potencia y subportadoras, hasta la minimización de la potencia transmitida. Entre estos estudios, podemos citar los dos siguientes a modo ilustrativo. En [Gault et al., 2007], se aplican las condiciones KKT a un algoritmo para la asignación de portadoras y potencia que minimiza la potencia total transmitida por una estación base; en este caso, el problema de optimización se formula en base a la capacidad ergódica. Y en [Li y Liu, 2007], los autores se sirven de las condiciones KKT para estudiar en qué circunstancias OFDMA es el acceso múltiple óptimo, considerando que la suma de las tasas de transmisión de los usuarios es la función objetivo a optimizar.

En general, este método permite resolver problemas tanto con restricciones en igualdades como en desigualdades de la forma

$$\begin{aligned} & \text{mín } f(\mathbf{x}) \\ & \text{s.t. } g_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, r \\ & \quad h_i(\mathbf{x}) = 0, \quad i = 1, \dots, q \end{aligned} \tag{P1}$$

En las condiciones KKT tiene especial relevancia si las restricciones sobre las funciones g_i se cumplen en términos de igualdad o desigualdad. La diferenciación entre ambas situaciones se concreta en la siguiente definición.

Definición 3.3 (Restricciones activas e inactivas). *Se dice que una restricción $g(\mathbf{x}^*) \leq \mathbf{0}$ está activa o saturada en el punto \mathbf{x}^* si se cumple que $g(\mathbf{x}^*) = \mathbf{0}$. En el caso de estricta desigualdad, la restricción estará inactiva o no saturada. El conjunto $\mathcal{A}(\mathbf{x}^*)$ representa los índices de las restricciones activas en \mathbf{x}^* .*

Si una restricción está inactiva en un punto \mathbf{x}^* , esto implica que el recurso no está agotado a pesar de estar limitado en cantidad. Por lo tanto, esta restricción no tendría relevancia a la hora de determinar el óptimo en \mathbf{x}^* . Por el contrario, si la restricción se encuentra activa en \mathbf{x}^* , el recurso se encuentra agotado y la restricción sí es relevante para determinar el óptimo en \mathbf{x}^* .

En general, el problema (P1) tendrá un conjunto de soluciones factibles \mathcal{F} , entre las que se encontrará la solución óptima \mathbf{x}^* . No obstante, aquellos puntos de \mathcal{F} que verifiquen una cierta propiedad, cumplirán las condiciones necesarias de optimalidad. Esta propiedad es la hipótesis de cualificación de las restricciones, que se define a continuación.

Definición 3.4 (Hipótesis de cualificación de las restricciones). *[Guerrero-Casas, 1994] Se dice que un punto \mathbf{x}^* del conjunto de soluciones factibles de (P1) verifica la hipótesis de cualificación de las restricciones si se cumple que $\mathcal{T}_L(\mathbf{x}^*, g, h) = \mathcal{T}(\mathbf{x}^*, \mathcal{F})$, siendo \mathcal{F} el conjunto de soluciones factibles del problema (P1), $\mathcal{T}(\mathbf{x}^*, \mathcal{F})$ el cono tangente a \mathbf{x}^* y $\mathcal{T}_L = \{\mathbf{s} \in \mathbb{R}^n : \mathbf{s}^T \nabla g_i(\mathbf{x}^*) \leq 0 \forall i \in \mathcal{A}(\mathbf{x}^*), \mathbf{s}^T \nabla h(\mathbf{x}^*) = 0\}$.*

El problema es que la hipótesis de cualificación de restricciones es, en general, difícil de comprobar. No obstante, existen una serie de condiciones que son suficientes para que se cumpla esta hipótesis y que resultan más sencillas de probar. La más utilizada es la **condición de Slater** [Boyd y Vandenberghe, 2004]: si todas las funciones de las restricciones h_j y g_i son convexas, y el conjunto de soluciones posibles \mathcal{F} tiene un interior no vacío, todo punto factible $\mathbf{x} \in \mathcal{F}$ verifica la hipótesis de cualificación de las restricciones. Teniendo en cuenta esto y las definiciones previas de esta sección, las condiciones de Karush-Kuhn-Tucker permiten comprobar si el punto \mathbf{x}^* es el óptimo global del problema (P1) anterior.

Teorema 3.7 (Condiciones necesarias de Karush-Kuhn-Tucker). *Si \mathbf{x}^* es un óptimo del problema (P1) y se satisface la hipótesis de cualificación de las*

restricciones, entonces

$$\nabla f(\mathbf{x}^*) + \sum_{j=1}^q \lambda_j \nabla h_j(\mathbf{x}^*) + \sum_{i=1}^r \mu_i \nabla g_i(\mathbf{x}^*) = 0 \quad (3.3)$$

$$h_j(\mathbf{x}^*) = 0, \quad \forall j = 1, \dots, q \quad g_i(\mathbf{x}^*) \leq 0, \quad \forall i = 1, \dots, r \quad (3.4)$$

$$\lambda_j \geq 0, \quad \forall j = 1, \dots, q \quad \mu_i \geq 0, \quad \forall i = 1, \dots, r \quad (3.5)$$

$$\mu_i g_i(\mathbf{x}^*) = 0, \quad \forall i = 1, \dots, r. \quad (3.6)$$

La condición (3.3), conocida como condición de estacionariedad, se puede poner de forma compacta como $\nabla_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = 0$, siendo

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f(\mathbf{x}) + \sum_{j=1}^q \lambda_j h_j + \sum_{i=1}^r \mu_i g_i \quad (3.7)$$

la función de Lagrange correspondiente al problema (P1). Esta condición tiene una interpretación geométrica inmediata: el gradiente de la función de coste ha de expresarse como una combinación lineal de los gradientes de todas las restricciones del problema.

La condición (3.6), también conocida como condición de inactividad complementaria (*complementary slackness*), refleja el comportamiento de los multiplicadores de Lagrange asociados a las restricciones en desigualdad: mientras se cumpla la restricción con desigualdad estricta ($g_i(\mathbf{x}) < 0$), la restricción del multiplicador de Lagrange correspondiente $\mu_i \geq 0$ ha de estar activa ($\mu_i = 0$); por el contrario, si la restricción está activa ($g_i(\mathbf{x}) = 0$), ha de cumplirse que $\mu_i \geq 0$.

3.2.3. Optimización no lineal con restricciones de desigualdad: Teoría de dualidad

La teoría de dualidad es una herramienta matemática que, utilizando como base la teoría de los multiplicadores de Lagrange, permite resolver el problema original, denominado **problema primal**, a través de otro problema equivalente, que se denomina **problema dual**. La idea esencial es relajar el problema primal transfiriendo las restricciones a la función objetivo en forma de una suma ponderada, lo cual permite afrontar la resolución de problemas complejos de forma más sencilla. El estudio de sistemas de comunicaciones en particular se ha visto muy beneficiado del uso de la dualidad, como vemos en los siguientes ejemplos. En [Yu y Lui, 2006] y en [Luo y Zhang, 2008] se proponen diferentes métodos para la gestión dinámica del espectro para sistemas multiportadora. En [Chiang y Bell, 2004], mediante

dualidad se desarrollan algoritmos que llevan a cabo un control equilibrado entre la tasa de transmisión de bits que se puede obtener de un enlace y la tasa de transmisión que se debe proporcionar a cada usuario. Finalmente, los autores de [Brah et al., 2009] proponen estrategias de asignación de potencia y agrupaciones de subportadoras en el enlace descendente de sistemas WiMAX, asumiendo que el conocimiento del canal de transmisión no es perfecto.

Sea el siguiente problema de partida, que consideraremos como el problema primal

$$\begin{aligned} & \text{mín } f(\mathbf{x}) \\ & \text{s.t. } g_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, r \\ & \quad h_i(\mathbf{x}) = 0, \quad i = 1, \dots, q \end{aligned} \quad (\text{P-I})$$

se define la función objetivo dual asociada a (P-I) como el mínimo valor de la función de Lagrange sobre \mathbf{x} , es decir:

$$g(\boldsymbol{\lambda}, \boldsymbol{\mu}) = \inf_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \inf_{\mathbf{x}} \left(f(\mathbf{x}) + \sum_{j=1}^q \lambda_j h_j + \sum_{i=1}^r \mu_i g_i \right), \quad (3.8)$$

donde la variable \mathbf{x} se denomina variable primal y los multiplicadores de Lagrange $\boldsymbol{\lambda}$ y $\boldsymbol{\mu}$ asociados a las restricciones son las variables duales. Dado que se cumple siempre que $f(\mathbf{x}) \geq g(\boldsymbol{\lambda}, \boldsymbol{\mu})$, para cualquier \mathbf{x} y $(\boldsymbol{\lambda}, \boldsymbol{\mu})$ que sean soluciones factibles, obtendremos la mejor cota inferior posible de la función objetivo primal f si se maximiza la función objetivo dual g , esto es, resolviendo el problema dual, dado por

$$\begin{aligned} & \text{máx}_{\boldsymbol{\lambda}, \boldsymbol{\mu}} g(\boldsymbol{\lambda}, \boldsymbol{\mu}) \\ & \text{s.t. } \boldsymbol{\lambda} \geq \mathbf{0}, \end{aligned} \quad (\text{P-II})$$

que resulta ser un problema de optimización convexa incluso si el problema primal (P-I) no lo fuera, dado que la función g es siempre cóncava. La resolución del problema dual nos proporciona entonces una cota g^* del valor óptimo de la función objetivo primal f^* , de manera que $f^* \geq g^*$, y hablaremos en este caso de dualidad débil. A la diferencia entre ambos valores de le denomina **separación de dualidad**, que será siempre no negativa: $f^* - g^* \geq 0$. Ahora bien, si la separación de dualidad es nula, es decir si $f^* = g^*$, el problema primal (P-I) se resolverá de manera equivalente a través de la resolución del problema dual (P-II). En este caso, hablaremos de dualidad fuerte, y esto sucede cuando el problema primal

es convexo y se cumple la hipótesis de cualificación de las restricciones, definida en la sección 3.2.2 [Boyd y Vandenberghe, 2004]. De hecho, es posible establecer condiciones necesarias y suficientes para obtener la solución del primal a través del dual, haciendo uso de las condiciones KKT (3.3)–(3.6), de acuerdo al siguiente teorema [Boyd y Vandenberghe, 2004].

Teorema 3.8. *Sea el problema (P-I) que satisface la hipótesis de cualificación de las restricciones; \mathbf{x}^* es óptimo y $f^* = f(\mathbf{x}^*) = g^* = g(\mathbf{x}^*, (\boldsymbol{\lambda}, \boldsymbol{\mu})) \Leftrightarrow$ existen $(\boldsymbol{\lambda}, \boldsymbol{\mu})$ que satisfacen las condiciones KKT.*

3.3. Conclusiones

La optimización de sistemas de comunicaciones inalámbricas presenta a menudo limitaciones en los valores que pueden tomar las variables que intervienen. Dicha optimización se puede llevar a cabo formulando un problema de optimización con restricciones, el cual se puede resolver mediante los métodos matemáticos expuestos en este capítulo. En función de si las variables toman valores enteros o reales, haremos uso, respectivamente, de la optimización en variable entera o en variable continua. A lo largo del capítulo hemos dado asimismo diferentes ejemplos de la utilización de los métodos presentados en aplicaciones tan diversas como la asignación de canales radioeléctricos, *bit-loading* o la asignación de potencia de transmisión.

CAPÍTULO 4

MODULACIÓN ADAPTATIVA EFICIENTE EN ENERGÍA PARA SISTEMAS CENTRALIZADOS

La autonomía de los dispositivos y equipos es un factor crítico en redes de comunicaciones inalámbricas, y está directamente relacionada con el consumo de energía. En este capítulo trataremos la optimización de la energía de transmisión en redes inalámbricas centralizadas, y en particular en redes de sensores inalámbricas, en las que cada usuario envía la información a un nodo central y puede requerir una tasa binaria de transmisión y una fiabilidad distintas, de acuerdo a su aplicación particular. Para ello, se propone un esquema entre capas (*cross layer*) cuyo objetivo es minimizar la energía total consumida por el sistema, con restricciones de tasa binaria fijada y tasa de error de símbolo (*symbol error rate*, SER).

El diseño del esquema involucra la capa física y la capa de enlace. En la capa física, el nodo central que recibe la información realiza la estimación de canal y la selección de la modulación a utilizar para cada usuario. En la capa de enlace, el protocolo de acceso al medio (*medium access control*, MAC) asigna a cada usuario el tiempo necesario para la transmisión, siendo este tiempo un número entero de veces el período de símbolo. El nivel de la modulación y el número de períodos de símbolo correspondientes a cada usuario serán tales que se cumplan las tasas binarias y SER requeridas y la asignación resulte eficiente en energía. La clave para relacionar la tasa binaria de bit, la energía de transmisión y la SER, es la

aproximación SNR *gap* introducida en la sección 2.3.2).

4.1. Introducción

Para el despliegue de redes inalámbricas en general son ampliamente aceptados como aspectos más relevantes la fiabilidad, la escalabilidad y, especialmente, la duración de las baterías. Por tanto, los esquemas de comunicación eficientes en energía con fiabilidad han adquirido una gran importancia para conseguir que los dispositivos en este tipo de redes estén activos el mayor tiempo posible. En definitiva, el objetivo buscado es minimizar el consumo de energía cumpliendo con unas garantías de servicio. Ahora bien, el problema es complejo de tratar, dado que para diseñar un sistema de comunicaciones de bajo consumo de energía hay que tener en cuenta varios factores, como por ejemplo el diseño del amplificador, el esquema de codificación y modulación o el protocolo de enrutamiento, con la complicación adicional de que estos factores están relacionados entre sí.

Una solución inmediata para conseguir eficiencia en energía es hacer uso de intervalos de transmisión muy largos. No obstante, esto impone, lógicamente, unos tiempos entre transmisiones sucesivas muy elevados que aplicaciones con restricciones de retardo no se pueden permitir. Esto lleva a esquemas que alcanzan un compromiso entre el retardo y la eficiencia en energía, como los propuestos en [Schurgers et al., 2001] y en [Yu et al., 2004]. Otra aproximación diferente examina las comunicaciones de salto único en redes de sensores usando acceso múltiple por división en el tiempo TDMA, proporcionando algoritmos óptimos y subóptimos para minimizar la energía de transmisión para una capacidad dada [Yao y Giannakis, 2005]. De esta manera, se obtienen ganancias teóricas de energía al comparar con la capacidad ideal obtenida empleando sólo TDMA. Otras aproximaciones, para redes multisalto, han sido desarrolladas en [Haapola et al., 2005] y [Shelby et al., 2005].

4.1.1. Esquemas eficientes en energía para redes inalámbricas

Las aproximaciones más habituales para optimizar el consumo de energía son los esquemas de capa única, donde el objetivo está orientado a minimizar la energía con parámetros correspondientes a una sola capa. Para la capa física, una opción es utilizar modulación adaptativa y transmitir con el nivel de la modulación óptimo que minimiza el consumo de energía [Schurgers et al., 2001]. Dicho valor óptimo se

puede obtener de diferentes maneras, como por ejemplo a partir del tiempo óptimo del modo activo del dispositivo [Cui et al., 2005a], o bien haciendo un seguimiento continuado del estado de las baterías [Yang y Wang, 2008].

En la línea de los anteriores, Verdú considera en [Verdu, 2002] el compromiso entre eficiencia espectral y energía por bit de información como la medida clave de la capacidad de canal en sistemas limitados en potencia de banda ancha. En este contexto, propone esquemas que se caracterizan por su baja eficiencia espectral pero a la vez la energía por bit resultante está muy próxima al valor mínimo necesario para comunicaciones fiables.

Una de las herramientas más empleadas en la capa de enlace es el diseño de protocolos MAC eficientes en energía. Estos protocolos se centran fundamentalmente en tres aspectos: evitar las colisiones, ya que las retransmisiones de paquetes son muy costosas energéticamente hablando; mantener los transmisores y receptores en el modo de espera (*standby*) siempre que sea posible; y adaptar la potencia de transmisión al alcance del enlace, en vez de transmitir a la máxima potencia. Una selección bastante completa de estos protocolos, orientados al control de potencia, se puede encontrar en [Kumar et al., 2006]. Otra forma de minimizar la energía en la capa de enlace es a través de esquemas de planificación de la transmisión (*scheduling*). Por ejemplo, en [Prabhakar et al., 2001] se propone un algoritmo para minimizar la energía de transmisión variando el tiempo de transmisión asociado al paquete para aplicaciones con restricciones de retardo.

Finalmente, y sólo por completitud puesto que en este capítulo nos centraremos en redes de un único salto centralizadas, haremos mención a los protocolos de enrutamiento, que pueden diseñarse para minimizar la energía en redes multisalto (ver, por ejemplo, [Wieselthier et al., 2000]).

No obstante, recientemente se ha optado por esquemas más eficientes, aunque más complejos, que basan el diseño del esquema en la consideración conjunta de parámetros pertenecientes a varias capas, por ejemplo la potencia del nivel físico y el enrutamiento del nivel de red. Este tipo de enfoque, conocido como diseño *cross-layer* en la literatura, optimiza el rendimiento del sistema considerando conjuntamente parámetros de varias capas, mientras que la optimización capa a capa, aunque simplifica el diseño, no tiene por qué resultar siempre en un rendimiento óptimo global del sistema. Distintos enfoques son por tanto posibles, en función de las capas diseñadas conjuntamente. Por ejemplo, en [Cui et al., 2005b] se proponen, para redes de sensores, dos tipos de estrategias de *cross-layer*: una, basada en

el diseño conjunto de la capa MAC y el enrutamiento; otra, considerando conjuntamente el enrutamiento, la planificación de los intervalos de transmisión y la adaptación de la transmisión (bits por símbolo a transmitir). En [Pollin et al., 2008] se propone considerar la energía utilizada para la transmisión (capa PHY), la planificación de las transmisiones (capa de enlace) y el consumo de los componentes de radiofrecuencia.

4.1.2. Fiabilidad y comunicaciones eficientes en energía

Junto con la eficiencia en energía, la fiabilidad en canales inalámbricos de baja potencia es un objetivo a cumplir, especialmente en redes heterogéneas donde los requisitos de cada nodo de la red pueden diferir de acuerdo al servicio o aplicación implementado. En nuestro caso, vamos a considerar que la fiabilidad se traduce en la conservación de la tasa binaria objetivo R_n y una SER objetivo, que denotaremos por SER_n , asociados al usuario n . Como luego veremos, ambos parámetros se pueden relacionar a través de la aproximación SNR *gap*. Considerando la fiabilidad así especificada, lo que proponemos es una reformulación del problema MA (2.13), que conduce a una minimización del consumo de energía.

4.1.3. Contribuciones del capítulo

En este capítulo, proponemos un esquema de modulación adaptativa eficiente en energía con diseño *cross-layer* considerando la capa física y la capa de enlace, para el enlace ascendente de redes centralizadas. El esquema es formulado en primer lugar para redes inalámbricas en general, y en la segunda parte del capítulo es particularizado para uno de los casos donde más interés suscita la optimización de la energía, las redes inalámbricas de sensores. Las principales características de la propuesta, que se concretarán en las siguientes secciones, son las siguientes:

- Es un esquema de modulación adaptativa, basado en el diseño adaptativo de la duración de los intervalos de transmisión, siendo la duración de los mismos un número entero de períodos de símbolo. Dicho período de símbolo permanece invariante para todos los nodos.
- La utilización de la aproximación SNR *gap* de la sección 2.3.2, a través de la cual relacionamos de una manera sencilla los dos parámetros que garantizan la fiabilidad de las comunicaciones, SER y la tasa de transmisión.

- Minimiza la energía total del sistema \mathcal{E}_T , que es la suma de las energías de transmisión de los N nodos de la red. Matemáticamente, siendo \mathcal{E}_n la energía transmitida por el nodo n -ésimo: $\mathcal{E}_T = \sum_{n=1}^N \mathcal{E}_n$.

El esquema, *grosso modo*, sigue el siguiente procedimiento. En primer lugar se hace una estimación del canal entre cada nodo y el nodo central, modelando las pérdidas de trayecto y el desvanecimiento de pequeña escala. A continuación, el protocolo de acceso al medio hace uso de esta información, junto con las restricciones de fiabilidad impuestas (tasa de transmisión y tasa de error de símbolo) para diseñar la duración de los intervalos de transmisión de cada uno de los nodos, de forma que la trama TDMA de transmisión estará dividida en intervalos de transmisión diferentes. Finalmente, la consideración conjunta de la duración de los intervalos y la tasa de transmisión requerida determinará el nivel adecuado de la modulación para esa trama. Para mostrar el rendimiento del esquema, comparemos la energía necesaria para nuestro esquema de modulación adaptativa con el esquema tradicional de duración fija de los intervalos TDMA, en el entorno de redes de sensores inalámbricas.

En el contexto de las redes de sensores inalámbricas, se discute en este capítulo la modulación multinivel a utilizar para la adaptación de la tasa binaria. De entre las opciones barajadas, la que mejor compromiso ofrece entre ancho de banda y energía consumida resulta ser MQAM, para el rango de distancias de nuestro interés. El esquema adaptativo propuesto lo comparamos con esquemas de transmisión convencionales TDMA, tanto para el escenario correspondiente al estándar IEEE 802.16.4 como para Bluetooth.

El resto del capítulo está organizado como sigue. En la sección 4.2 establecemos el modelo del sistema para redes inalámbricas en general, consistente en el esquema de transmisión y el modelo de pérdidas. Dedicamos la sección 4.3 a la formulación del problema de acuerdo al modelo propuesto. Esto constituye la parte genérica del capítulo. En la sección 4.4 afrontamos la problemática concreta de la optimización de la energía en redes de sensores y discutimos las diferentes soluciones adoptadas hasta el momento. Las siguientes secciones están orientadas a las redes de sensores: en (4.5) se elabora el modelo para estas redes y en (4.6) adaptamos nuestro esquema para las redes de sensores y se resuelve el problema mediante el método de los multiplicadores de Lagrange. Finalmente, las secciones 4.7 y 4.8 las dedicamos a presentar los resultados obtenidos y a la extracción de conclusiones, respectivamente.

4.2. Modelo general del sistema para redes inalámbricas

Nos centraremos en el enlace ascendente de redes inalámbricas, donde un conjunto de nodos envían información a un nodo central haciendo uso de un ancho de banda B , como se muestra en la figura 4.1. Los nodos de la red pueden, individualmente, tener diferentes requisitos de transmisión en términos de tasa binaria y SER, R_n y SER_n respectivamente para $n = 1, \dots, N$, en función de la aplicación que cada uno de ellos ejecute.

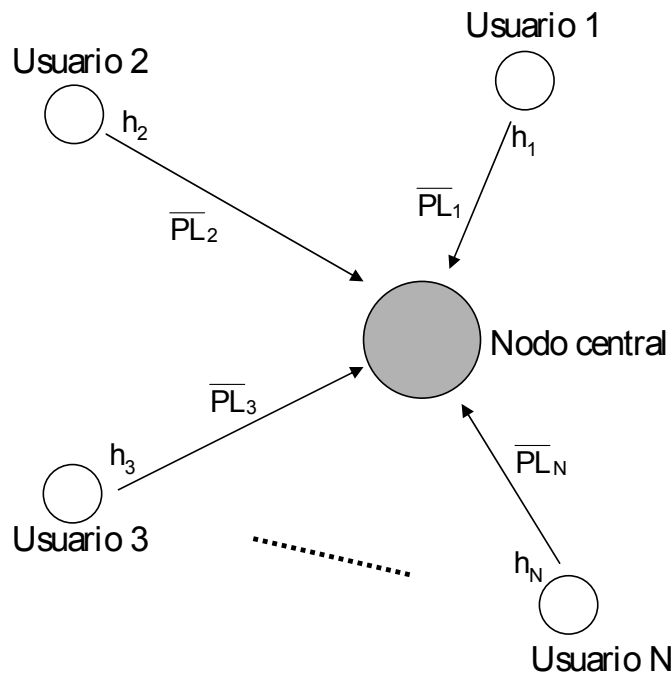


Figura 4.1. Modelo del sistema

4.2.1. Esquema de transmisión

Asumimos que la transmisión de los N nodos se realiza en una trama de duración T dividida en N intervalos de tiempo. Estos intervalos representan la duración de la transmisión asociada al nodo n , la cual denotaremos por T_n . Al contrario que en los esquemas de transmisión tradicionales TDMA, vamos a considerar que el valor de T_n es variable, de manera que puede ser adaptado para cada trama TDMA, como se muestra en la figura 4.2. Según lo anterior, tenemos que:

$$\sum_{n=1}^N T_n = T. \quad (4.1)$$

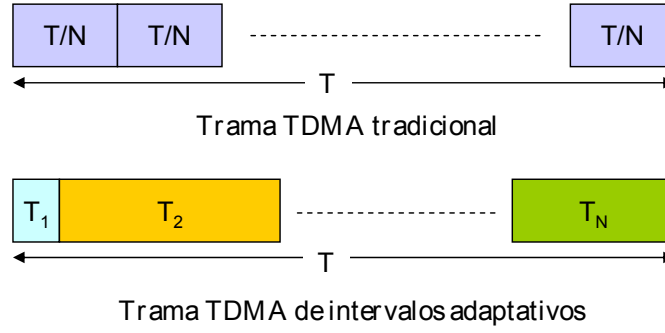


Figura 4.2. Tramas TDMA con longitud de los intervalos de transmisión fija y variable

4.2.2. Modelo de pérdidas de los enlaces

Definimos un modelo de pérdidas para estimar la energía recibida que tendrá dos componentes: las pérdidas de trayecto y el desvanecimiento de pequeña escala. Para las pérdidas de trayecto, vamos a considerar que cada uno de los nodos experimenta en su transmisión hacia el nodo central unas pérdidas medias de propagación \overline{PL}_n (dB). Asociado al valor anterior, definimos $\alpha_n = 10^{\overline{PL}_n/10}$. El desvanecimiento de pequeña escala lo modelamos mediante una distribución Rayleigh, de forma que el coeficiente asociado al canal del nodo n es una variable aleatoria gaussiana compleja de valor h_n , cuyo módulo es una variable aleatoria Rayleigh. Para ambas pérdidas, asumiremos que el canal es invariante durante la duración de la transmisión de la trama.

Las pérdidas totales asociadas al enlace entre el nodo n y el nodo central las denotaremos por g_n . Teniendo en cuenta las dos contribuciones anteriores, la expresión de g_n vendrá dada por

$$g_n = \frac{\alpha_n}{|h_n|^2}. \quad (4.2)$$

4.2.3. Modelo de energía

La energía total transmitida la denotaremos mediante \mathcal{E}_T y la energía transmitida por cada nodo mediante \mathcal{E}_n . La energía transmitida por símbolo está representada por \mathcal{E}_s , que asumimos será la misma para todos los nodos de la red. \mathcal{E}_n la podemos calcular basándonos en la duración de cada intervalo de tiempo, considerando que se transmitirán $\frac{T_n}{T_s}$ símbolos en el intervalo n , como

$$\mathcal{E}_n = \frac{T_n}{T_s} \mathcal{E}_s. \quad (4.3)$$

La energía recibida por símbolo en el nodo central desde el nodo n la denotaremos mediante \mathcal{E}_{s-rx}^n . Considerando los factores del modelo de pérdidas de (4.2.2), \mathcal{E}_{s-rx}^n se calcula como

$$\mathcal{E}_{s-rx}^n = \frac{\mathcal{E}_s}{g_n}. \quad (4.4)$$

Recordemos que la relación señal a ruido, que denotaremos mediante γ , se puede expresar a partir de la energía recibida por símbolo como $\gamma = \frac{\mathcal{E}_s}{N_0}$. Por tanto, la SNR en el nodo central correspondiente a la transmisión del nodo n , representada por γ_n , se puede formular a partir de \mathcal{E}_{s-rx}^n como

$$\gamma_n = \frac{\mathcal{E}_{s-rx}^n}{N_0}, \quad (4.5)$$

siendo N_0 la densidad espectral de potencia de ruido AWGN. Así, es inmediato a partir de la expresión anterior conocer \mathcal{E}_{s-rx}^n en función de la relación señal a ruido correspondiente:

$$\mathcal{E}_{s-rx}^n = \gamma_n N_0. \quad (4.6)$$

4.2.4. Sistema de señalización

Queda claro que alguna entidad en la red debe, en primer lugar, calcular los valores de los intervalos T_n y coordinar las asignaciones de los mismos a los usuarios, así como estimar los parámetros de canal necesarios para realizar la modulación adaptativa. Por tanto, algún tipo de sistema de señalización debe ser implementado para informar a los nodos acerca de estas asignaciones. Dado que la topología de la red es centralizada, lo razonable es que estas tareas, incluida la estimación de los parámetros h_n , sean realizadas por el nodo central, dado que se le suponen limitaciones de energía mucho menores que a los nodos y mayor capacidad de procesamiento para llevar a cabo las diferentes tareas. El diseño de este sistema está fuera del objeto de este capítulo, y asumiremos que su correcto funcionamiento está garantizado. Esta información, que ha de ser trasladada a los usuarios en el enlace descendente, en general no supondrá un tráfico añadido significativo que pueda afectar el rendimiento del sistema [Johansson, 2004].

4.3. Modulación adaptativa basada en intervalos de transmisión variables para redes inalámbricas

Para conseguir que el sistema sea eficiente en energía, resolvemos el problema de optimización de la energía total del sistema, la cual calculamos como la suma de las

energías individuales de los nodos del sistema. Por tanto, el problema resultante es el siguiente:

$$\text{mín } \mathcal{E}_T = \sum_{n=1}^N \mathcal{E}_n. \quad (4.7)$$

La aproximación SNR *gap*, como se vió en (2.3.2), permite relacionar γ , la tasa de bit y la SER para una modulación y esquema de codificación dados, y facilita la implementación de algoritmos de *bit-loading* para modulación adaptativa. Recordamos que para obtener una tasa binaria de r bits por símbolo, para una modulación multinivel, recurrimos a la expresión

$$r = \log_2 \left(1 + \frac{\gamma}{\Gamma} \right) = \log_2 M, \quad (4.8)$$

donde M representa el nivel de la modulación, γ representa la SNR y Γ el valor del SNR *gap*, el cual depende de la SER que nos marquemos como objetivo concreto, según vimos en la sección 2.3.2.

Dos consideraciones han de hacerse desde el punto de vista de las aplicaciones prácticas empleando modulación adaptativa. En primer lugar los valores de r y de M tienen que ser discretos. En este sentido, consideraremos que ΔM especifica el incremento entre valores consecutivos de M , por lo que los resultados obtenidos dependerán de este parámetro. En segundo lugar, la forma usual de especificar la tasa binaria es en bits por segundo (bps), por lo que por conveniencia utilizaremos en nuestra formulación la tasa binaria en estas unidades, la cual denotaremos mediante R . Para el ancho de banda B utilizado en el sistema, tenemos que, para una trama de duración T , la tasa binaria resultante es:

$$R = B \log_2 \left(1 + \frac{\gamma}{\Gamma} \right) = \frac{1}{T_s} \log_2 \left(1 + \frac{\gamma}{\Gamma} \right) = \frac{r}{T_s}. \quad (4.9)$$

Por tanto, la aproximación SNR *gap* nos es útil para obtener la expresión de γ basada en la tasa binaria R y la tasa de error de símbolo SER. Para obtener la tasa binaria de transmisión asociada al nodo n , que transmite durante un tiempo T_n , modificamos convenientemente la expresión anterior, resultando

$$R_n = B \frac{T_n}{T} \log_2 \left(1 + \frac{\gamma_n}{\Gamma_n} \right) = \frac{1}{T_s} \frac{T_n}{T} \log_2 \left(1 + \frac{\gamma_n}{\Gamma_n} \right), \quad (4.10)$$

A partir de (4.10) podemos obtener la expresión de γ_n en función de R_n y T_n :

$$\gamma_n = \Gamma_n \left(2^{\frac{R_n T_n}{B T_s}} - 1 \right) \quad (4.11)$$

y combinando (4.6) y (4.11), es inmediato obtener la expresión de la energía recibida por símbolo \mathcal{E}_{s-rx}^n

$$\mathcal{E}_{s-rx}^n = \Gamma_n \left(2^{\frac{R_n T}{B T_n}} - 1 \right) N_0. \quad (4.12)$$

Para obtener la expresión de \mathcal{E}_n en función de T_n , partimos de la ecuación (4.3), donde una vez insertado el valor de \mathcal{E}_{s-rx}^n obtenido a partir de (4.4), tenemos

$$\mathcal{E}_n = \frac{T_n}{T_s} g_n \mathcal{E}_{s-rx}^n; \quad (4.13)$$

y sustituyendo en la ecuación (4.12), tenemos

$$\mathcal{E}_n = \frac{T_n}{T_s} g_n N_0 \Gamma_n \left(2^{\frac{R_n T}{B T_n}} - 1 \right). \quad (4.14)$$

Ahora, podemos formular la minimización de la energía en función de los intervalos T_n añadiendo la restricción (4.1) al problema (4.7) y considerando la ecuación (4.14):

$$\begin{aligned} \min_{T_n} \mathcal{E}_T &= \sum_{n=1}^N \mathcal{E}_n = \sum_{n=1}^N g_n \frac{T_n}{T_s} \Gamma \left(2^{\frac{R_n T}{B T_n}} - 1 \right) N_0 \\ \text{s.t.} \quad &\sum_{n=1}^N T_n = T \\ &R_n = D_n \quad \forall n = 1, \dots, N, \end{aligned} \quad (4.15)$$

donde D_n representa la tasa binaria en bps fijada para el nodo n , y R_n representa la tasa binaria que se obtiene mediante modulación adaptativa. La resolución del problema anterior garantizará, para cada sensor n , los valores especificados de R_n y $SE R_n$, adaptando la modulación y la duración del tiempo de transmisión T_n a los cambios en el valor de γ producidos por la variación del canal entre dos tramas consecutivas.

4.3.1. Esquema de transmisión basado en el número de símbolos del intervalo

Dada las dificultades prácticas que implica considerar períodos de símbolo T_s variables, la adaptación de la duración de los intervalos T_n la haremos precisamente considerando un valor de T_s constante. La trama, de duración T , está compuesta por un número entero de símbolos $K \in \mathbb{Z}^+$, de manera que $T = K T_s$, como se

observa en la figura 4.3. Por tanto, el intervalo de transmisión del n -ésimo nodo tendrá una duración de N_n símbolos, de manera que

$$T_n = N_n T_s, \quad (4.16)$$

siendo N_n un número entero mayor o igual que 0. Insertando (4.16) en (4.10), obtenemos la tasa de bit R_n de cada nodo en función de N_n :

$$R_n = \frac{N_n}{T} \log_2 \left(1 + \frac{\gamma}{\Gamma} \right). \quad (4.17)$$

La restricción de la duración de la trama ha de ser modificada en consecuencia: ahora, tenemos que $\sum_{n=1}^N N_n = \sum_{n=1}^N \frac{T_n}{T_s}$. Podemos así reformular el problema de optimización de la energía en función de los $\{N_n\}_{n=1}^N$ como

$$\begin{aligned} \min_{N_n} \mathcal{E}_T &= \sum_{n=1}^N \mathcal{E}_i = \sum_{n=1}^N g_n N_n \Gamma \left(2^{\frac{R_n T}{N_n}} - 1 \right) N_0 \\ \text{s.t.} \quad &\sum_{n=1}^N N_n = K \\ &N_n \in \mathbb{Z}^+, \quad \forall n \\ &R_n = D_n, \quad \forall n = 1, \dots, N. \end{aligned} \quad (4.18)$$

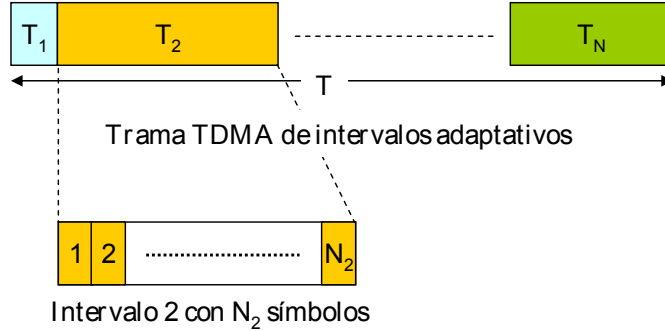


Figura 4.3. Tramas TDMA con longitud de los intervalos de transmisión variable en número de símbolos

A la vista de este esquema de transmisión, podemos realizar las siguientes observaciones:

- La imparcialidad entre los nodos está asegurada, ya que aunque aparentemente estos se dedican a “robarse” tiempo de transmisión unos a otros, cada nodo lo único que hace es mantener sus requisitos de fiabilidad especificados por sus respectivos valores de R_n y $SE R_n$.

- De la ecuación (4.17) se deduce que la tasa binaria R_n requerida se garantiza mediante la modulación adaptativa: según las condiciones del canal cambien entre dos tramas sucesivas, el cambio correspondiente en la relación señal a ruido γ se traslada al número de bits por símbolo a través del término $\log_2(1 + \frac{\gamma_n}{\Gamma})$, y este cambio se compensa mediante el valor adecuado de N_n para mantener R_n constante.

4.4. Aplicación práctica: optimización de energía en redes de sensores

Dentro de las redes inalámbricas, merece especial atención la optimización del consumo de energía en las redes de sensores inalámbricas, identificadas en la literatura mediante el acrónimo WSN (*wireless sensor network*). Estas redes se caracterizan porque sus nodos, los sensores, están la inmensa mayoría de las veces alimentados por batería. Por tanto, es esencial el diseño de protocolos de comunicaciones eficientes en energía, ya que se consigue prolongar la vida útil de las baterías. Esta cuestión es de gran importancia en este tipo de redes ya que su despliegue se hace a menudo en áreas de difícil y/o remoto acceso, lo que dificulta en gran medida el cambio de las baterías, a la vez que implica un largo tiempo de inactividad del sensor cuya batería esté extinguida.

Las redes WSN son susceptibles de aplicarse a diferentes campos tales como áreas de la industria y el comercio (por ejemplo, monitorización y control del entorno), automatización del hogar y edificios inteligentes (seguridad, iluminación, control de temperatura), periféricos de computadores (ratón, impresora), electrónica de consumo, medicina y asistencia sanitaria (monitorización de enfermos, diagnóstico, sensores corporales), y vigilancia y mantenimiento, entre otras [Akyildiz et al., 2002]. Algunos ejemplos relevantes de estas aplicaciones son los trabajos de [Sharp et al., 2005], [Malan et al., 2004], [Werner-Allen et al., 2005] y [Szewczyk et al., 2004].

La disponibilidad de productos comerciales ha fomentado a su vez el número de potenciales aplicaciones de las WSN. En particular, los dispositivos inalámbricos conformes al estandar IEEE 802.15.4 y las especificaciones Zigbee parecen ganar mercado debido a sus características de bajo consumo de potencia y bajo coste, para aplicaciones con necesidades de tasa binaria poco exigentes. Estas características hacen que ambas especificaciones sean muy adecuadas para la mayor parte

de las aplicaciones de WSN en las redes de área personal (*personal area network*, PAN), como muestran, entre otros, los trabajos de [Timmons y Scanlon, 2004] y [Yu et al., 2004].

El problema de la optimización de la energía ha sido tratado bajo distintos puntos de vista en la literatura, para redes WSN donde un nodo central recoge la información suministrada por los sensores que componen la red. Existen estudios orientados a optimizar el consumo de energía a través de la capa PHY, como es el caso de la propuesta de [Cui et al., 2004]. Otros autores optan por basar la minimización de la energía en el protocolo de acceso al medio, como por ejemplo [Yu et al., 2004], [Yao y Giannakis, 2005] y [Shelby et al., 2005]. Al igual que en redes inalámbricas en general, hay una tendencia creciente a realizar la minimización de energía mediante el diseño de protocolos entre capas o *cross-layer*, principalmente a través del diseño conjunto del enrutamiento y la capa de enlace (véase, por ejemplo, [Sichitiu, 2004]). No obstante, en el caso de una red WSN centralizada de un salto es evidente que no existe enrutamiento, por lo que hemos de acudir a diseños conjuntos de capa de enlace y capa física. Además, la capa física proporciona un mayor número de grados de libertad para abordar la minimización del consumo de energía. Por ejemplo, para redes multisalto, en [Kwon et al., 2006] se hace uso de la gestión de la potencia transmitida, mientras que en [Karvonen y Pomalaza-Raez, 2006] se propone la codificación como parámetro de diseño.

4.5. Modelo del sistema para WSN

La configuración de la red es, a grandes rasgos, la misma que la presentada en el modelo genérico de la sección 4.2. Recordamos a continuación brevemente las principales características. Vamos a considerar una red WSN centralizada, donde un nodo central recoge la información de los N sensores de la red. Centraremos el análisis en este sentido ascendente de la comunicación. Los sensores no tienen por qué estar dedicados necesariamente a la misma aplicación o servicio, de forma que cada uno de ellos potencialmente requeriría distintos parámetros de fiabilidad, es decir, el sensor n -ésimo solicitará una tasa binaria R_n y una tasa de error de símbolo $SE R_n$. Análogamente a como se definió para el caso general en (4.2.1), cada sensor transmitirá en el intervalo n -ésimo, que tendrá una duración T_n , de forma que $\sum_{n=1}^N T_n = T$, y la señalización implícita en el sistema estará a cargo

del nodo central.

Cada sensor utilizará una energía en su intervalo de transmisión $\mathcal{E}_n = \frac{T_n}{T_s} \mathcal{E}_s$, según la ecuación (4.3). Hacemos notar que este modelo de consumo de energía podría comprender asimismo el consumo de los circuitos de potencia y las transiciones entre los estados activo y de espera [Shih et al., 2001]. No obstante, en nuestro caso el objetivo es analizar el impacto en la minimización de la energía debido a la transmisión exclusivamente, por lo que no hemos incluido dichas contribuciones en este modelo. Aun así, la extensión de nuestro modelo a un modelo que incluyera dichos consumos resultaría bastante sencillo.

4.5.1. Modelo de pérdidas para WSN

Para redes WSN, vamos a modelar las pérdidas de trayecto \overline{PL} de acuerdo al modelo de propagación descrito en [Janssen et al., 1996], donde la distancia entre el transmisor y el receptor está en el rango de las comunicaciones personales (hasta 10 m), de manera que las pérdidas de trayecto correspondientes al enlace entre el sensor n y el nodo central vienen determinadas por la expresión

$$\overline{PL}_n = S_0 + 10a \log d_n + b(\text{dB}), \quad (4.19)$$

siendo d_n la distancia entre el sensor n y el nodo central, S_0 las pérdidas de trayecto a una distancia de 1 m, y los parámetros a y b corresponden al escenario con visión directa en la banda libre ISM de 2,4 GHz para entornos de interior. La altura de las antenas se asume en este modelo que es de 1 m para el receptor y entre 1 y 3 m para el nodo transmisor.

Por lo que se refiere a las pérdidas de pequeña escala, el modelo es el mismo que el descrito en la sección 4.2.2.

4.5.2. Justificación frente al estándar 802.15.4

El protocolo MAC del estándar IEEE 802.15.4 [802, 2006a] proporciona un modo de operación para sensores que soliciten garantía de servicio, haciendo uso de intervalos de tiempo garantizados (*guaranteed time slot*, GTS) y CDMA ranurado. Sin embargo, la solicitud de esta garantía de servicio puede ser rechazada por el coordinador de la red, por lo que tal garantía de servicio no está completamente asegurada. Aun cuando los GTS sean aceptados para un determinado sensor, la minimización de energía sólo puede ser implementada por medio del control de potencia puesto que el estándar no contempla ningún mecanismo de gestión de la

potencia. En cambio, el esquema adaptativo propuesto en este capítulo optimiza el consumo de energía simplemente ajustando la duración de los intervalos de tiempo de transmisión de cada sensor, manteniendo constante la potencia de transmisión.

4.6. Modulación adaptativa basada en intervalos de transmisión variables para WSN

La implementación de la modulación adaptativa puede hacerse de manera sencilla mediante arquitecturas *hardware* en paralelo, dado que en la práctica el rango de niveles de la modulación es moderado. Otra alternativa es la implementación mediante SDR o radio cognitiva. Esta alternativa es muy razonable para el nodo central, donde no hay especiales restricciones de capacidad de procesamiento ni de energía. Para los sensores, aunque esta implementación pueda parecer *a priori* poco factible, ya existen algunas propuestas que abogan por esta solución, como [Shen et al., 2006] y [Gao et al., 2007].

4.6.1. Discusión de la modulación y aproximación SNR *gap* para modulaciones MQAM

Diferentes modulaciones multinivel pueden ser adoptadas para transmisión adaptativa. MQAM proporciona una SER más baja, comparada con MPSK para el mismo valor de SNR. Desde el punto de vista de la energía consumida, MQAM es considerada la modulación a utilizar en comunicaciones eficientes en energía ([Shih et al., 2001], [Schurgers et al., 2001]). Si nos atenemos a la diferencia en SNR requerida para una determinada tasa binaria, es conocido que, para valores usuales de r comprendidos entre 1 y 6 bits por símbolo, MPSK se comporta mejor para valores bajos de r y MQAM es más conveniente para valores altos [Garcia-Armada, 2006]. En nuestro caso, al basar nuestro esquema en modulación adaptativa y no conocer de antemano los niveles de la modulación, la discusión basada en la diferencia de SNR no nos permite tomar una decisión acerca del tipo de modulación a utilizar.

Por otra parte, la eficiencia espectral (bps/Hz) es un inconveniente para las modulaciones MFSK, ya que puede llegar a ser mucho menor que en MQAM (hasta 8 veces para $M = 16$) [Wang et al., 2001], lo que implica que el ancho de banda necesario sería varias veces el utilizado en MQAM. Lo habitual es que el período de símbolo sea constante, es decir, que el ancho de banda sea constante,

de manera que si queremos la mayor eficiencia espectral posible, optaremos por MQAM.

Además, las redes WSN son sistemas de banda estrecha, que requieren regímenes binarios no elevados, en torno a decenas de kbps. Esto, unido a que el ancho de banda disponible es constante, la elección de la modulación recaerá en aquella de mayor eficiencia espectral, con lo que optaremos por MQAM.

Si consideramos la distancia del enlace, la energía puede ser minimizada empleando tanto MQAM como MFSK [Cui et al., 2005a]; la energía por bit es menor para MQAM para distancias menores de 30 m, que es nuestro rango de interés. Otros trabajos relacionados consideran asimismo la utilización de MQAM para comunicaciones eficientes en energía [Schurgers et al., 2001], [Shih et al., 2001].

Como utilizaremos la aproximación SNR *gap* proporcionada por la ecuación (4.9), de manera que la tasa de bit vendrá dada por $R = \frac{1}{T_s} \log_2 \left(1 + \frac{\gamma}{\Gamma}\right)$, para el caso particular de MQAM la expresión del SNR *gap* Γ es:

$$\Gamma = \frac{1}{3} \left(Q^{-1} \left(\frac{SER}{4} \right) \right)^2, \quad (4.20)$$

donde $Q(x)$ es la función de distribución complementaria de la gaussiana normalizada y se calcula como

$$Q(x) = \int_x^{\infty} \frac{\exp(-u^2/2)}{\sqrt{2\pi}} du. \quad (4.21)$$

4.6.2. Minimización de la energía total de transmisión

En este punto, vamos a particularizar el esquema de modulación adaptativa de la sección (4.3) a las redes WSN. Para minimizar la energía, resolvemos el problema (4.18) tomando en consideración el modelo de red WSN presentando en la sección anterior utilizando la modulación MQAM. El problema resultante es por tanto:

$$\begin{aligned} \min_{N_n} \mathcal{E}_T &= \sum_{n=1}^N g_n N_n \Gamma_n \left(2^{\frac{R_n T}{N_n}} - 1 \right) N_0 \\ \text{s.t.} \quad \sum_{n=1}^N N_n &= K \\ N_n &\in \mathbb{Z}^+, \forall n \\ R_n &= D_n \forall n = 1, \dots, N, \end{aligned} \quad (4.22)$$

donde el valor de Γ es el correspondiente a la modulación MQAM. Dado que las variables N_n toman valores enteros, este problema es de programación entera (IP),

requiriendo una complejidad computacional prohibitiva si se quiere adaptación continua de la transmisión, especialmente para valores de K elevados. Por ello, vamos a obtener una solución subóptima aplicando la relajación de la condición de integralidad según se detalla en la sección 3.1.3, y una vez tenemos el problema en variable real, se aplica el método de los multiplicadores de Lagrange de la sección 3.2.1. Sin pérdida de generalidad, vamos a suponer que la energía de transmisión de cada sensor \mathcal{E}_n está normalizada con respecto a N_0 . La formulación resultante con ambas consideraciones es:

$$\begin{aligned} \min_{N_n} \mathcal{E}_T &= \sum_{n=1}^N g_n N_n \Gamma_n \left(2^{\frac{R_n T}{N_n}} - 1 \right) \\ \text{s.t.} \quad \sum_{n=1}^N N_n &= K \\ N_n &\in \mathbb{R}^+, \quad \forall n \\ R_n &= D_n \quad \forall n = 1, \dots, N. \end{aligned} \tag{4.23}$$

Aplicando el método de los multiplicadores de Lagrange mencionado, el conjunto $\{N_n\}_{n=1}^N$ que optimiza la energía total consumida ha de satisfacer el valor obtenido del multiplicador de Lagrange λ

$$\lambda = \Gamma_n g_n \left(1 + 2^{\frac{R_n T}{N_n}} \left(\frac{R_n}{N_n} \ln 2 - 1 \right) \right), \tag{4.24}$$

donde el valor de λ puede ser calculado mediante búsqueda numérica. A partir de los valores $\{N_n\}_{n=1}^N$, es inmediato obtener la duración de los intervalos.

En la práctica, como apuntamos en secciones anteriores, los valores de N_n han de ser enteros y no reales. Para obtener valores enteros una opción inmediata es redondear los valores de N_n obtenidos al entero más próximo, teniendo en cuenta que dicho redondeo no puede originar que el total de los símbolos de la trama supere el valor de K . Ahora bien, el valor de energía óptimo obtenido para el caso real y el caso redondeado es muy similar, como se muestra en la siguiente sección, por lo que las simulaciones han realizadas sido con variable real.

4.7. Resultados

El sistema descrito se ha simulado teniendo como referencia el estándar IEEE 802.15.4 para efectuar una elección realista de los parámetros de simulación. La banda de transmisión es la banda ISM (2,4 GHz), que está especificada como la

banda primaria para este tipo de redes. El ancho de banda es de 62,5 kHz y el período de símbolo es $T_s = 16\mu s$.

Para que efectivamente el canal resulte invariante durante el tiempo de transmisión de una trama, el tiempo de coherencia del canal T_c debe ser mayor que la duración de una trama. T_c se puede calcular como $T_c = \frac{0,423}{f_m}$ [Rappaport, 2002], siendo $f_m = \frac{vf_c}{c}$, $c = 300000$ m/s, $v = 3$ km/h (velocidad de paseo) y f_c la frecuencia de portadora, esto es, 2,4 GHz. Con estos parámetros obtenemos que $T_c = 63,45$ ms.

La longitud de la trama se ha elegido de acuerdo al valor de T_c y al estándar IEEE 802.15.4, que establece una duración de la trama de entre 15 ms y 250 s. Establecemos 896 símbolos por trama, obteniendo entonces una duración de la misma de 14,336 ms. El resto de parámetros toman los siguientes valores: la tasa binaria total de la red es, conforme al estándar, de 250 kbps y consideramos una $SER = 10^{-3}$, para todos los sensores.

Los resultados presentados en esta sección se refieren al ahorro de energía que supone el utilizar el esquema propuesto, respecto a la alternativa convencional de TDMA con modulación fija, de manera que la ganancia de energía a que nos referimos en las figuras representa dicho ahorro. Distinguiremos dos casos para comparar:

1. TDMA con longitud de los intervalos fija, al que nos referiremos mediante TDMA-F, con modulación 16QAM;
2. TDMA con longitud de los intervalos variable, al que nos referiremos mediante TDMA-V, con modulación 16QAM y 64QAM.

La figura 4.4 muestra a modo de ejemplo, para el caso en que todos los sensores tienen el mismo R_n objetivo con escenario IEEE 802.15.4, que la diferencia de energía entre redondear los valores de N_n y hacer uso de los valores reales de N_n es muy pequeña, lo que nos permite trabajar con los valores reales de N_n para el resto de las simulaciones. En la figura 4.5 representamos la ganancia de energía en función de la heterogeneidad de la red, representada por la desviación entre los usuarios del requisito en régimen binario, en bps; así, el valor 0 de la desviación corresponde a que todos los usuarios requieren el mismo R_n . Se han considerado 16 usuarios, que es el máximo admitido por el estándar, distribuidos uniformemente en torno al nodo central a una distancia entre 1 y 10 m. Notamos que en este caso la ganancia es creciente a medida que aumenta la heterogeneidad de la red.

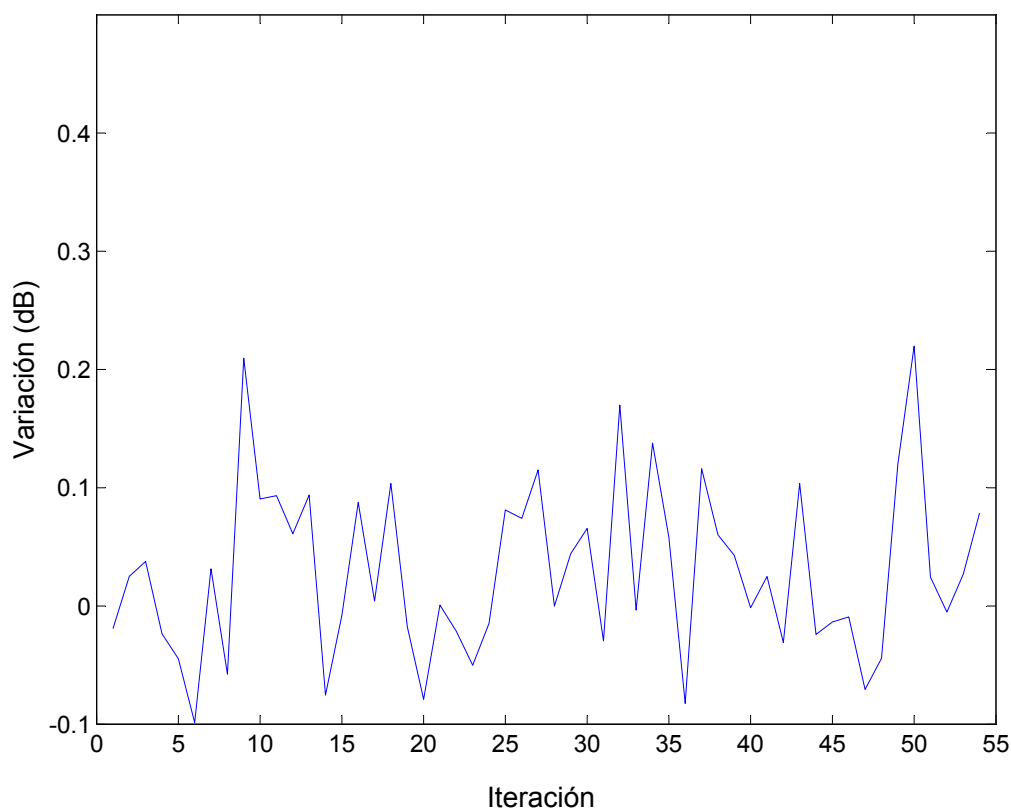


Figura 4.4. Variación de energía (dB) del esquema adaptativo con N_n redondeado respecto al valor real, para escenario IEEE.802.15.4

La figura 4.6 es análoga a la anterior, comparando nuestro esquema adaptativo con TDMA-V para 64QAM. Ahora las ganancias son mayores, cercanas a los 6 dB, y la influencia de la heterogeneidad de la red es baja en el comportamiento del esquema adaptativo.

Otro entorno típico para redes de sensores inalámbricas es el especificado para Bluetooth. En este caso, la duración de la trama está limitada a 3 valores: 0,625, 1,875 y 3,125 ms. A efecto de las simulaciones, hemos elegido una trama de 64 símbolos correspondiente a un valor de 1,024 ms, que se sitúa entre los valores anteriores, donde los 16 sensores están distribuidos de la misma forma que en el escenario IEEE 802.15.4. Vemos que los resultados son similares (figuras 4.7 y 4.8), aunque se observa una menor ganancia y menor variación debida a la heterogeneidad al comparar con TDMA-V con 64QAM (figura 4.8).

Tanto para el escenario 802.16.15.4 como para Bluetooth, se ha observado que los sensores ubicados a la menor distancia posible del nodo central, que es 1 m, usan un número de símbolos bajo, tomando N_n típicamente un valor de entre 3 y 5 símbolos por intervalo de transmisión. Esto obliga a que dichos sensores empleen

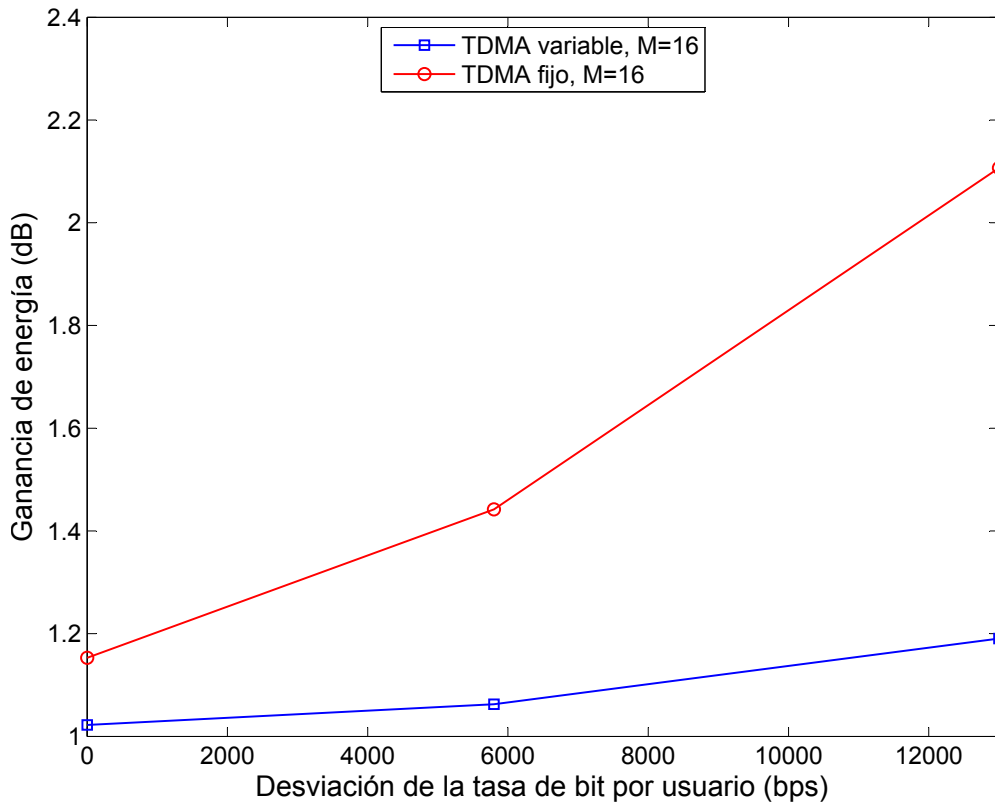


Figura 4.5. Ganancia de energía (dB) del esquema adaptativo respecto a TDMA-F y TDMA-V con modulación 16QAM, para escenario IEEE.802.15.4

un alto nivel de modulación.

En entornos como los descritos, donde las distancias son pequeñas, es probable que exista correlación entre los canales de los diferentes sensores. Para ver el efecto que tendría esto, vamos a considerar que los coeficientes del desvanecimiento de pequeña escala h_n están correlacionados, donde el parámetro σ^2 indica el grado de correlación entre los coeficientes. En la tabla 4.1 se muestran los resultados obtenidos para el escenario IEEE.802.15.4. En principio, cabría esperar una degradación respecto al caso no correlado, lo cual no sucede como se desprende de los resultados. La explicación radica en que los coeficientes de pérdidas de trayecto α_n son mucho mayores que los valores de $|h_n|^2$, por lo que la energía optimizada, a la vista de (4.23), está fuertemente dominada por el efecto de las pérdidas de trayecto.

En las simulaciones, se ha tenido en cuenta la realización de *bit-loading* discreto, como se explica en (4.3). Establecemos $\Delta M = 1$, asegurando siempre un valor de tasa binaria mayor que el R_n objetivo, considerando que cualquier valor entero

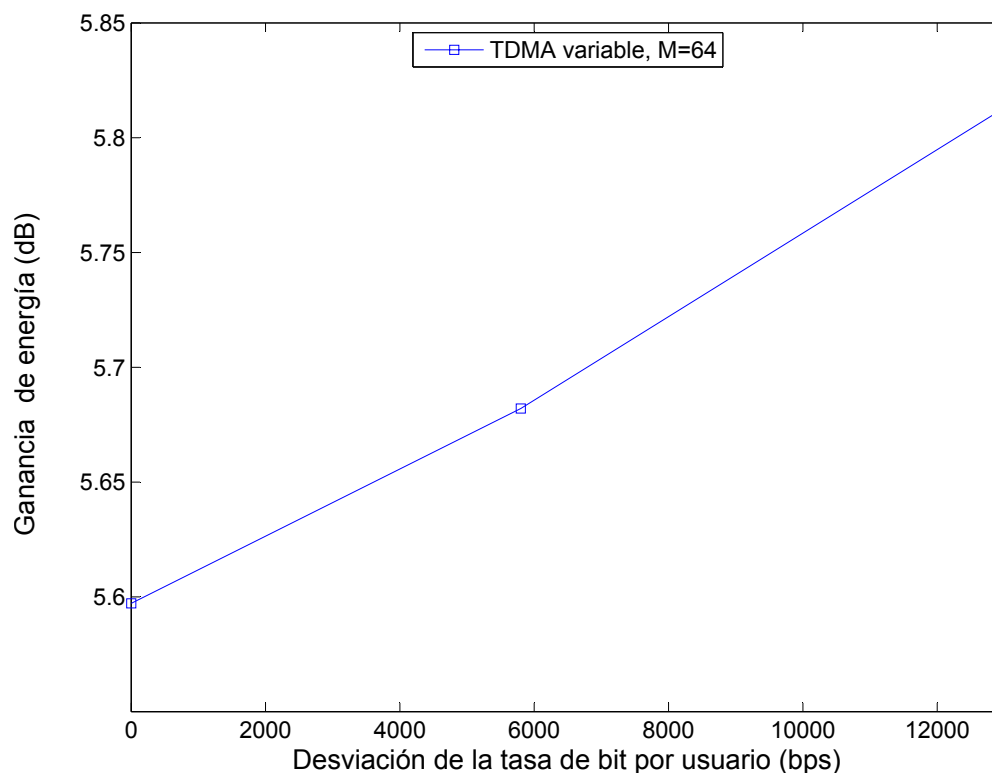


Figura 4.6. Ganancia de energía (dB) del esquema adaptativo respecto a TDMA-V con modulación 64QAM, para escenario IEEE.802.15.4

de M se podría alcanzar usando la codificación apropiada.

Es importante considerar las limitaciones de potencia que están generalmente impuestas a nivel mundial por las agencias regulatorias a las transmisiones de radiofrecuencia, especialmente en transmisiones que involucran bandas de uso libre, como la banda ISM. En el caso de aplicaciones conforme al estándar IEEE 802.15.4, la potencia máxima de transmisión permitida es de 100 mW, pero en la práctica es muy común que el valor de potencia transmitida se encuentre en torno a 1 mW en este tipo de redes. En las simulaciones realizadas utilizando el esquema adaptativo propuesto la potencia necesaria se encuentra siempre por debajo del umbral especificado, y con valores adecuados para su uso práctico, ya que el valor máximo obtenido es ligeramente superior a 1 mW y los valores usuales son de cientos de μW .

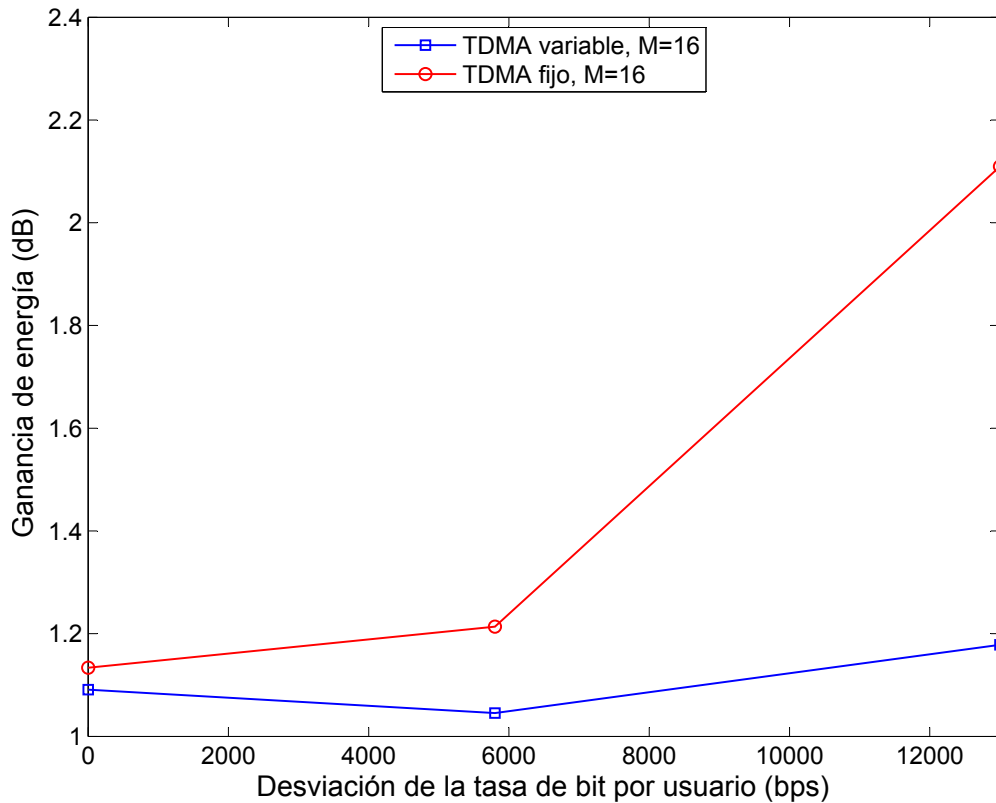


Figura 4.7. Ganancia de energía (dB) del esquema adaptativo respecto a TDMA-F y TDMA-V con modulación 16QAM, para escenario Bluetooth

4.8. Conclusiones

La eficiencia en energía es de vital importancia para la autonomía de los dispositivos y el rendimiento de las redes inalámbricas en general, y crítica para las redes de sensores en particular. En este capítulo hemos desarrollado un esquema basado en modulación adaptativa eficiente en energía y que está diseñado empleando parámetros de capa física y de capa de enlace, formulado como un problema de minimización de la energía total de transmisión de los nodos. La topología de la red elegida es centralizada, por ser ésta la más habitual en redes de sensores de corta distancia.

El rendimiento de la red se aborda garantizando la tasa binaria y la probabilidad de error (SER) requerida por la aplicación que se ejecuta en cada sensor de la red, lo que se consigue adaptando el nivel de la modulación MQAM. A este respecto, aunque hemos desarrollado nuestra propuesta usando MQAM, la modulación MPSK también tendría aquí cabida. La elección de MQAM está sobre todo

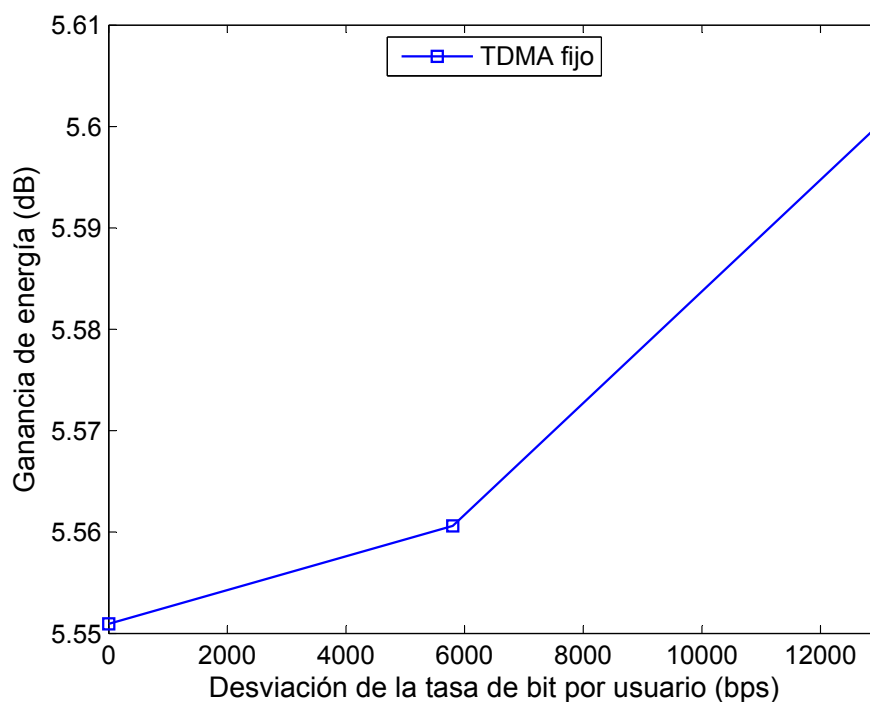


Figura 4.8. Ganancia de energía (dB) del esquema adaptativo respecto a TDMA-V con modulación 64QAM, para escenario Bluetooth

motivada por su mayor eficiencia en energía en el entorno concreto estudiado y la gran diferencia existente en eficiencia espectral respecto a MPSK. Sin embargo, MPSK puede ser de interés bajo otros enfoques; por ejemplo, para aplicaciones con limitaciones en la potencia de pico, ya que aunque pueda requerir mayor consumo de energía, la relación entre potencia de pico y potencia media (PAR, *peak to average power ratio*) es menor. Esta posibilidad queda por explorar.

Hemos mostrado que efectivamente el esquema adaptativo propuesto supone un ahorro de energía significativo respecto a TDMA, tanto si el tamaño de los intervalos de transmisión es fijo como si es variable. También se ha comprobado que el efecto de la correlación entre los canales, por lo que se refiere al desvanecimiento de pequeña escala, afecta poco al rendimiento del esquema, puesto que predomina en el consumo de energía el efecto del desvanecimiento de trayecto. Finalmente, la potencia necesaria al aplicar este esquema está entre los valores usuales para aplicaciones en redes de sensores, y muy por debajo del límite especificado por el estándar IEEE 802.15.4.

Otro punto importante a tratar es la movilidad en este tipo de redes: el modelo

Tabla 4.1. Ganancia de energía con respecto a TDMA-F y TDMA-V, con correlación entre canales σ^2 , para escenario IEEE.802.16.4

M	Desviación (kbps)	σ^2	G_{TDMA-F} (dB)	G_{TDMA-V} (dB)
16	0.0	1	1,15	1,02
16	0.0	20	1,24	1,06
16	5.8	1	1,44	1,07
16	5.8	20	1,34	1,07
16	13	1	2,11	1,19
16	13	20	2,12	1,23
64	0.0	1	–	5,6
64	0.0	20	–	5,62
64	5.8	1	–	5,68
64	5.8	20	–	5,7
64	13	1	–	5,81
64	13	20	–	5,93

de canal que hemos considerado asume que los sensores son estáticos o restringidos a un movimiento muy lento, como pueden ser sensores corporales o dispositivos para el hogar. La extensión del trabajo de este capítulo a entornos con movilidad puede resultar adecuado con la inclusión del modelo de canal apropiado. Este escenario puede resultar de gran interés para un amplio rango de aplicaciones como la transmisión de imágenes de vigilancia y sensores para control de seguridad, siempre que la velocidad no sea elevada.

CAPÍTULO 5

OPTIMIZACIÓN DE LA EFICIENCIA DE TRANSMISIÓN EN SISTEMAS OFDM MULTIUSUARIO

Para ciertas aplicaciones no resulta suficiente que la tasa de error del nivel físico sea baja, sino que es necesario que todos los bits se entreguen correctamente. La eficiencia de transmisión, o *goodput*, es un criterio muy adecuado para estas aplicaciones, puesto que permite cuantificar los bits entregados sin errores en la capa de enlace.

En este capítulo trataremos la optimización de sistemas OFDM multiusuario desde la perspectiva del *goodput* del enlace descendente entre la estación base y los usuarios. Para ello, se resuelve el problema de asignación de los recursos subcanal, tasa de bit por símbolo y potencia, tomando como referencia lo expuesto en el capítulo 2. Ahora bien, dado que optimizamos el *goodput* asociado a la EB, existe el riesgo de favorecer a unos usuarios frente a otros al asignar los recursos disponibles, es decir, de que no haya imparcialidad entre los usuarios. Para controlar el reparto de los recursos se introduce en el sistema un mecanismo que permite ajustar el grado de imparcialidad deseado del sistema a través de un único parámetro, de forma que se podrá o no favorecer a los usuarios de acuerdo a un cierto patrón. Todo ello se realiza en el marco de la matemática entera binaria, donde algunas

de las variables de asignación de los recursos toman los valores 0 ó 1, de manera que haremos uso de las herramientas presentadas en el capítulo 3.

5.1. Introducción

OFDM es una técnica bien conocida y ampliamente utilizada actualmente en los sistemas de comunicaciones inalámbricos, principalmente por su robustez frente al desvanecimiento selectivo en frecuencia y al multitrayecto, y se prevé que juegue asimismo un papel esencial en las comunicaciones inalámbricas 4G. Por tanto, resultará de vital importancia optimizar el rendimiento de estos sistemas por medio de una buena gestión de los recursos radio disponibles. En OFDM, como ya vimos en la sección 2.3.3, la técnica más extendida es el *bit-loading*: la potencia y el número de bits por símbolo (o lo que es lo mismo, el nivel de la modulación) que lleva cada subportadora OFDM se determinan con el objetivo de optimizar la transmisión, bien minimizando una función de la potencia de transmisión o bien maximizando una función de utilidad formulada en términos de tasa de información.

Si hablamos de entornos multiusuario, la utilización de la asignación de potencia y bits resulta muy ventajosa en sistemas con acceso OFDMA: si procedemos a asignar a un usuario una subportadora o un grupo de las mismas, su señal se puede adaptar a las características de su canal de comunicación particular mediante la asignación de la potencia y el número de bits adecuados. Es más, dado que los usuarios, casi con toda certeza, percibirán unas condiciones del canal diferentes, podremos hacer uso de la diversidad multiusuario, que introdujimos en la sección 2.4.1.

A la vista de lo anterior, queda claro que la asignación de recursos radio es un factor clave a la hora de optimizar sistemas OFDMA. No obstante, dicha asignación de recursos dista mucho de resultar un problema fácil de resolver, ya que es un problema con relaciones de dependencia entre los recursos disponibles, y cuya solución óptima se puede conseguir mediante búsqueda exhaustiva entre todos los posibles usuarios, portadoras, esquemas de codificación y modulación y niveles de potencia. Para OFDMA, los algoritmos de asignación de bit y potencia siguen uno de los dos siguientes enfoques, ya discutidos en la sección 2.6: un enfoque subóptimo en dos pasos, asignando en primer lugar las portadoras OFDM a los usuarios, y a continuación se asignan la potencia y los bits a cada una

de las portadoras, y la optimización conjunta de asignación de bits, potencia y portadoras. Frente al comportamiento óptimo de los algoritmos de optimización conjunta, los algoritmos en dos pasos contraponen una menor complejidad que los hace más adecuados para implementaciones prácticas, si bien a costa de un comportamiento subóptimo.

En este capítulo, en lugar de considerar la asignación de subportadoras a los usuarios, haremos uso de la agrupación de las mismas en subcanales, tal y como se describió en la sección 2.7.1. Así, los recursos a asignar son los subcanales, la potencia de transmisión y los bits por símbolo por subportadora.

5.1.1. El *goodput* como objetivo de optimización

La demanda de nuevos servicios y aplicaciones por parte de los usuarios requiere de las redes inalámbricas tasas de transmisión y compromisos de calidad de servicio (BER) cada vez más exigentes. En el caso de algunas aplicaciones como la transmisión de voz, la optimización del sistema se lleva a cabo, generalmente, de dos maneras: maximizando la tasa de transmisión para una BER dada, o minimizando la BER para una tasa de transmisión dada. Sin embargo ciertas aplicaciones, como la transmisión de datos, requieren que todos los bits sean transmitidos correctamente. Esto conllevaría, si se formula el problema en términos de BER y tasa de transmisión, el conseguir una BER nula, lo cual no es posible dado que tratamos con sistemas limitados en potencia.

El *goodput* se define como el número de bits de información transmitidos correctamente por unidad de tiempo. Por tanto, para las aplicaciones arriba mencionadas, dado que el interés radica en conseguir transmisiones libres de errores, un objetivo adecuado es maximizar el *goodput*, o una función del mismo. Ahora bien, esto supone tener en cuenta la tasa de bit conseguida por la capa PHY y los esquemas de corrección de errores y mecanismos de retransmisión de trama de la capa de enlace; por lo tanto, el esquema de optimización ha de involucrar información y parámetros pertenecientes a ambas capas, siguiendo por tanto un diseño *cross-layer* similar a los descritos en la sección 4.1.1, para conseguir un compromiso entre tasa de bit y BER.

Las estrategias habituales de optimización para redes OFDM están formuladas bien como minimización de la potencia transmitida, bien como maximización de la tasa de bit, por lo que no se alcanza un compromiso entre BER y tasa de bit (veánse las referencias mencionadas en el capítulo 2). Por otro lado, es probable que los

usuarios estén interesados no sólo en obtener una comunicación libre de errores, sino en obtener una tasa mínima de bits de información libre de errores, es decir un *goodput* mínimo, por lo que el esquema de optimización debe garantizar este requisito. En la literatura, el *goodput* ha sido principalmente estudiado desde una perspectiva analítica (ver, por ejemplo, [Manzoli y Merani, 2006]); no obstante, trabajos como el descrito en [Lau et al., 2008], que estudia asintóticamente la ganancia de *goodput* en sistemas OFDMA, muestran la validez del *goodput* como métrica en la que basar el esquema de optimización.

5.1.2. Contribuciones del capítulo

En este capítulo, resolvemos el problema de optimizar el *goodput* transmitido por una EB en redes OFDMA mediante la asignación de subcanales, potencia y bits. Este planteamiento del problema, donde asignamos subcanales en lugar de subportadoras, hace que el problema sea menos exigente desde el punto de vista computacional para su resolución. Además, se garantiza una calidad de servicio en términos de *goodput* mínimo recibido por el usuario. Para ello, se propone un método para la asignación de recursos en OFDMA, que denominamos SBPA (*subchannel, bit and power allocation*), que constituye la primera aportación del capítulo y que se caracteriza por:

- tener un carácter no centralizado: si bien no totalmente distribuido, sí permite un reparto de las tareas de asignación entre los usuarios y la EB, diferenciando así de la mayoría de esquemas existentes, según vimos en la sección 2.6;
- estar basado en parámetros de capa física y de enlace;
- resuelve la asignación de ancho de banda en términos de subcanales, en lugar de subportadoras.

El problema de asignación de subcanal, bit y potencia, como veremos en la sección 5.3, está formulado en términos de matemática entera (también llamada en la literatura discreta), lo cual quiere decir que la solución óptima se obtendrá a costa de una elevada complejidad computacional. Como alternativa a la resolución óptima, proponemos en segundo lugar una aproximación más adecuada para sistemas implementables y que funcionen en tiempo real, que identificaremos mediante el acrónimo S-SBPA (subóptimo-SBPA). Dicha aproximación, como detallaremos, la podemos resolver de tres maneras:

- Por medio de métodos estándar de optimización entera, computacionalmente costosos.
- Relajando la restricción de integralidad para obtener un problema lineal que proporciona una solución no entera sino real, mucho más sencillo de resolver.
- Mediante el algoritmo aquí propuesto, que denotaremos por RLP (*rounded linear programming*). Este algoritmo proporciona una solución entera que solventa el inconveniente asociado a la solución relajada de no ser realizable para sistemas reales, pues los recursos a asignar no son divisibles: no podemos asignar la mitad de un subcanal a un usuario y a la mitad a otro.

El resto del capítulo está estructurado como sigue. En la sección 5.2 describimos el modelo de capa física y de capa de enlace que emplearemos a lo largo del capítulo. A continuación, en la sección 5.3, formulamos el problema de optimización del *goodput*, y en la sección 5.4 describimos el método de optimización en dos pasos propuesto para la asignación de recursos. La sección 5.5 detalla la aproximación subóptima y el algoritmo RLP. Los resultados obtenidos se presentan en la sección 5.6 y dedicamos la sección 5.7 a la extracción de conclusiones.

5.2. Descripción del sistema: capa física y capa de enlace

El modelo del sistema se basa en la combinación de técnicas de capa PHY, donde la unidad elemental intercambiada entre transmisor y receptor es el bit, y capa de enlace (*data link*, DL), que emplea tramas como unidad elemental, para optimizar el *goodput* transmitido por la EB (figura 5.1).

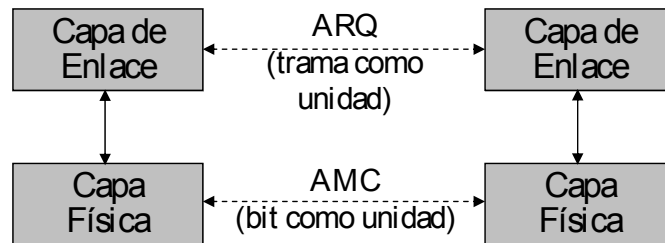


Figura 5.1. Modelo del sistema

En el nivel PHY, el esquema de transmisión entre la EB y cada usuario consta de dos componentes: la modulación adaptativa, que comprende asignación de bits y de potencia con las técnicas presentadas en el capítulo 2, y la asignación de

subcanales a los usuarios. Asimismo, en este nivel se implementa un mecanismo de corrección de errores. En la capa DL, se dispone de un mecanismo de retransmisión de tramas erróneas, en concreto el protocolo *automatic repeat request* (ARQ), para las tramas transmitidas entre la EB y los usuarios. El sistema OFDMA consta de una EB, K usuarios, y las comunicaciones entre la EB y cada usuario constan de un único salto. En este capítulo nuestro interés se centra en el enlace descendente, es decir, el envío de información desde la EB hacia los usuarios. Asumimos también la existencia de un sistema de señalización, necesario para transportar la información empleada en la asignación de los recursos adaptativa.

5.2.1. El nivel físico y el modelo de canal

La asignación de subcanal, número de bits por símbolo y potencia por subportadora se realiza en el nivel PHY. El número total de subportadoras disponibles en el sistema lo denotaremos mediante M . Estas subportadoras las agrupamos uniformemente en subcanales, de manera que en lugar de asignar subportadoras a los usuarios, la EB realizará la asignación de uno o varios subcanales a cada usuario en el enlace descendente. Por tanto, la EB dispone de $N = M/J$ subcanales, siendo J el número de subportadoras por subcanal. Hay que hacer notar que, aunque la asignación del ancho de banda entre los usuarios se hace a nivel de subcanal, la carga de bits y de potencia tiene granularidad de subportadora, de forma que se preservan todas las ventajas inherentes a las técnicas de *bit-loading* a la vez que el proceso total de asignación de recursos queda simplificado.

El reparto de las subportadoras entre los subcanales podría asimismo formar parte del diseño del sistema. En nuestro modelo este reparto está realizado de antemano, con lo que la asignación de las subportadoras a los subcanales es fija y se mantendrá en el tiempo, independientemente de las variaciones del canal. El objetivo último de esta asignación fija es reducir la complejidad global del modelo. Además, la posterior asignación de subcanales al mejor usuario, basada en el criterio del *goodput*, dota al sistema de diversidad multiusuario. Una asignación habitual es que las subportadoras que conforman el subcanal sean consecutivas, empleando el modo AMC del estándar descrito en la sección 2.7.1 (véase, por ejemplo, [Kaneko et al., 2008]), lo que reduce la complejidad del sistema ya que en la práctica las ganancias de varias subportadoras adyacentes se pueden considerar iguales [Hanzo et al., 2000]. Sin embargo, en este caso la diversidad por frecuencia quedaría bastante limitada, sobre todo en el caso de que un usuario utilice un

único subcanal. En nuestro caso, dado que no existe ninguna restricción en cómo disponer las subportadoras en los subcanales, hemos adoptado una distribución aleatoria de las subportadoras entre los subcanales para conseguir un cierto grado de diversidad por frecuencia.

En cuanto al modelo de canal, asumiremos un canal variante con desvanecimiento multitrayecto por bloques, siendo la ganancia de potencia de canal en cada subportadora constante durante el tiempo de transmisión de una trama. Suponemos también que el canal de subida o ascendente, empleado para enviar la información necesaria a la EB desde los usuarios, está libre de errores, por lo que esa información es transmitida idealmente hasta la EB. La potencia total de transmisión de la EB está limitada a P_T y es distribuida entre las M subportadoras. Finalmente, las muestras de ruido corresponden a una variable aleatoria gaussiana compleja de media nula y varianza σ^2 .

Transmisión multiportadora en enlaces punto a punto

Consideremos, en primer lugar, el caso de transmisión multiportadora con un único usuario. Cada subportadora j tendrá asociada, en el dominio de la frecuencia, una ganancia de canal H_j modelada como una variable aleatoria gaussiana compleja, cuyo módulo sigue una distribución Rayleigh. La modulación en cada subportadora j se realiza mediante MQAM (*quadrature amplitude modulation*). Denotamos mediante ϵ_j la tasa de error de bit (BER) asociada a la decisión dura en el receptor sobre los bits codificados correspondientes a una trama. ϵ_j se puede calcular mediante la aproximación dada en [Chung y Goldsmith, 2001] para modulaciones MQAM con codificación Gray, como

$$\epsilon_j = 0,2 \exp\left(-\frac{1,6 |H_j|^2 p_j}{(2^{m_j} - 1)\sigma^2}\right) \quad (5.1)$$

siendo p_j y m_j , respectivamente, la potencia y bits por símbolo asignados a la subportadora j . Como mecanismo corrector de errores utilizamos un código convolucional a nivel de trama de tasa r con decodificación Viterbi dura.

La formulación del *goodput* se basa en la probabilidad de que una trama sea correctamente recibida, esto es, la tasa de éxito de trama (*frame success rate*, FSR), y para cada usuario la transmisión se organiza en tramas de N_f bits de información. En el caso de subportadora única cuya BER asociada es ϵ , la FSR se puede aproximar a una exponencial negativa [Liu et al., 2004] como

$$FSR(\epsilon) = d \exp\left(-\left(c_\tau \epsilon^\tau + c_{\tau-1} \epsilon^{\tau-1} \dots + c_1 \epsilon\right)\right), \quad (5.2)$$

donde $\tau, d, c_1, \dots, c_\tau$ son parámetros que deben ser ajustados para que la FSR aproximada se ajuste lo más posible a la curva FSR real. Para multiportadora con un único usuario, la utilización conjunta de un decodificador Viterbi y un entrelazador ideal aleatorio permite considerar el canal multiportadora como un canal cuya probabilidad de error, que denotaremos por ϵ_{av} , es la BER promediada sobre las J subportadoras empleadas para la transmisión. Esta BER promedio se calcula como

$$\epsilon_{av} = \frac{1}{\sum_{j=1}^J m_j} \sum_{j=1}^J m_j \epsilon_j, \quad (5.3)$$

donde ϵ_j viene dado por (5.1). En este caso, la expresión de la FSR es [Devillers et al., 2008]

$$FSR(\epsilon_{av}) = d \exp \left(-(c_\tau \epsilon_{av}^\tau + c_{\tau-1} \epsilon_{av}^{\tau-1} + \dots + c_1 \epsilon_{av}) \right) \quad (5.4)$$

Transmisión multiportadora multiusuario

Generalizamos ahora el análisis para el caso de multiportadora multiusuario, considerando que se asignarán subcanales y no subportadoras. La ganancia de canal H_{knj} asociada a la subportadora j del subcanal n y asignada al usuario k es, al igual que los H_j del apartado anterior, una variable aleatoria gaussiana compleja, y dichos coeficientes H_{knj} se considerarán invariantes durante al menos un período de trama. Denotamos por p_{knj} la potencia asignada al usuario k en la subportadora j correspondiente al subcanal n . De forma análoga, el usuario k tendrá asignados m_{knj} bits por símbolo en la subportadora j del subcanal n . Definimos el conjunto N_L como las L posibles combinaciones de subcanales, esto es, los elementos de N_L son grupos de subcanales que serán asignados a los usuarios. Por ejemplo, dado un total de $N = 2$ subcanales, dispondremos de $L = \sum_{n=1}^{N=2} \binom{2}{n} = 3$ posibles agrupaciones de los subcanales, y $N_L = \{\{1\}, \{2\}, \{1, 2\}\}$. Dado que la BER promedio de interés es la BER asociada al usuario k cuando se le asigna el conjunto de subcanales $N_l \in N_L$, tenemos que

$$\epsilon_{av}(k, l) = \frac{1}{\sum_{n \in N_l} \sum_{j=1}^J m_{knj}} \sum_{n \in N_l} \sum_{j=1}^J m_{knj} 0,2 \exp \left(-\frac{1,6 |H_{knj}|^2 p_{knj}}{(2^{m_{knj}} - 1) \sigma^2} \right). \quad (5.5)$$

La expresión de la FSR para el caso multiusuario multiportadora, denotada por $FSR(\epsilon_{av}(k, l))$, la obtendremos entonces sencillamente insertando (5.5) en (5.4), quedando entonces

$$FSR(\epsilon_{av}(k, l)) = d \exp \left(-(c_\tau \epsilon_{av}^\tau(k, l) + c_{\tau-1} \epsilon_{av}^{\tau-1}(k, l) + \dots + c_1 \epsilon_{av}(k, l)) \right). \quad (5.6)$$

5.2.2. Capa de enlace

En la capa de enlace, un protocolo de petición automática de repetición (*automatic repeat request*, ARQ) es el encargado de la retransmisión de las tramas incorrectamente recibidas por el usuario. Recordamos que el *goodput* se define como el número de bits de información libres de errores que son transmitidos por unidad de tiempo. En lo que sigue, consideramos que el *goodput* se define para un símbolo OFDM. Las expresiones correspondientes al *goodput* para distintos protocolos ARQ han sido deducidas en [Devillers et al., 2008] para enlaces OFDM punto a punto. Sin pérdida de generalidad, adoptamos en nuestro modelo el protocolo ARQ de repetición selectiva, ya que resulta inmediato modificar la formulación para otros protocolos ARQ. Para nuestro caso, el *goodput* recibido, que denotamos mediante χ_{kl} , al asignar los subcanales N_l al usuario k tiene la expresión

$$\chi_{kl} = r \left(\sum_{n \in N_l} \sum_{j=1}^J m_{knj} \right) FSR(\epsilon_{av}(k, l)), \quad (5.7)$$

siendo r la tasa del código convolucional empleado.

5.3. Formulación del problema de optimización del *goodput*

Nuestro objetivo es optimizar el *goodput* transmitido por la EB de acuerdo al modelo descrito en la sección 5.2. Para ello, primero definiremos la función de utilidad U a maximizar, formulada en términos de *goodput*. Una posibilidad es definir U como la suma de los *goodputs* recibidos por cada usuario k , siendo entonces $U = \sum_{k=1}^K \chi_k$. Sin embargo, esta función de utilidad puede conducir a situaciones de no-imparcialidad, en las que algunos usuarios se podrían ver favorecidos y otros penalizados. Por esta razón introducimos el control de imparcialidad en la formulación del problema mediante un parámetro, α , que tomará valores reales entre 0 y 1. La utilidad de la EB queda entonces formulada como

$$U = \sum_{k=1}^K \frac{1}{\bar{\chi}_k^{1-\alpha}} \chi_k, \quad (5.8)$$

siendo $\bar{\chi}_k$ el *goodput* promedio recibido por el usuario k durante un cierto número de tramas recibidas anteriores y χ_k el *goodput* total recibido por el usuario k en el instante considerado. Es la introducción de $\bar{\chi}_k$ en la función U lo que nos permite implementar una imparcialidad a largo plazo. Mediante el parámetro α se controla

el grado de imparcialidad del sistema: si $\alpha = 0$, los usuarios reciben el *goodput* siguiendo una asignación de imparcialidad proporcional; en el otro extremo, $\alpha = 1$ supone una política de asignación de *goodput* máximo. Así, α se puede ajustar para obtener el grado de imparcialidad deseado para el sistema en su conjunto.

La asignación de los subcanales a los usuarios se hace a través de la variable binaria π_{kl} , la cual indica si un grupo de subcanales N_l es asignado a un usuario o no, es decir

$$\pi_{kl} = \begin{cases} 1 & \text{conjunto de subcanales } N_l \text{ asignado al usuario } k \\ 0 & \text{resto} \end{cases} \quad (5.9)$$

Debemos garantizar que sólo un conjunto de subcanales N_l es asignado a cada usuario en cada trama transmitida, para que los bits correspondientes a los subcanales asignados al usuario formen parte de la misma trama. Por tanto, sólo uno de los valores de π_{kl} es igual a 1 para un usuario k determinado, y se ha de cumplir que $\sum_{l=1}^L \pi_{kl} = 1$ para cada usuario k . Necesitamos también una variable que indique la asignación de cada subcanal individualmente a los usuarios; para ello definimos las variables binarias β_{kn} , cuyos valores determinan si el subcanal n se asigna al usuario k , por lo que tenemos

$$\beta_{kn} = \begin{cases} 1 & \text{subcanal } n \text{ asignado al usuario } k \\ 0 & \text{resto} \end{cases} \quad (5.10)$$

Vemos claramente que un subcanal debe ser asignado a uno y sólo a un usuario, por lo que la igualdad $\sum_{k=1}^K \beta_{kn} = 1$ debe cumplirse para todo subcanal n .

El problema de optimización consiste en maximizar la utilidad de la EB (5.8) encontrando los valores óptimos de asignación de bits por símbolo, nivel de potencia y subcanal, esto es m_{knj} , p_{knj} y π_{kl} , respectivamente. Como p_{knj} y m_{knj} toman valores reales y π_{kl} y β_{kn} son valores enteros, el problema se formula como un problema de programación entero mixto (MIP). En el contexto de nuestro problema, hay que considerar las siguientes restricciones:

- i. La EB dispone de una potencia total de transmisión P_T que se distribuye entre las M subportadoras. Esto se traduce en la siguiente desigualdad:

$$\sum_{k=1}^K \sum_{n=1}^N \sum_{j=1}^J p_{knj} \leq P_T \quad (5.11)$$

- ii. Cada usuario k debe recibir, al menos, un *goodput* mínimo, denotado por χ_{kmin} .

- iii. Para cumplir con la restricción anterior, cada usuario ha de tener asignado al menos un subcanal, de forma que la solución asociada a $N_l = \emptyset$ no está contemplada.

Con estas restricciones, escribimos el problema como

$$\begin{aligned}
 \underset{\pi_{kl}, \beta_{kn}, p_{knj}, m_{knj}}{\text{máx}} \quad & U = \sum_{k=1}^K \frac{1}{\bar{\chi}^{n(1-\alpha)}} \chi_k \\
 \text{s.t.} \quad & \sum_{k=1}^K \sum_{n=1}^N \sum_{j=1}^J p_{knj} \leq P_T \\
 & \sum_{k=1}^K \beta_{kn} = 1, \quad \forall n = 1, \dots, N \\
 & \chi_k \geq \chi_{kmin}, \quad \forall k = 1, \dots, K \\
 & \sum_{l=1}^L \pi_{kl} = 1, \quad \forall k = 1, \dots, K
 \end{aligned} \tag{P1}$$

donde χ_k representa el *goodput* recibido por el usuario k , que se calcula como $\chi_k = \sum_{l=1}^L \chi_{kl} \pi_{kl}$, siendo χ_{kl} el *goodput* obtenido mediante (5.7).

Los problemas MIP, como (P1), van a ser, por lo general, problemas \mathcal{NP} -hard, que son difíciles de resolver mediante métodos de optimización estándar. Una posible solución vendría dada por la aplicación del teorema 3.3, dado que así podemos resolver el problema entero mediante la relajación de la integralidad de manera óptima. Desafortunadamente, en nuestro caso la matriz asociada al problema no es totalmente unimodular, por lo que no es posible aplicar el mencionado teorema. Por lo que se refiere a aplicar teoría de dualidad, utilizando el corolario 3.1 o la proposición 3.1, nos encontramos, si cabe, con mayores dificultades aún al tener que demostrar la igualdad de los recubrimientos de los conjuntos que definen las restricciones en desigualdades.

El análisis de nuestro problema nos lleva a determinar que la condición que más complica su resolución es la restricción de potencia (5.11). En [Chow, 1993] se muestra, a partir de observaciones empíricas, que mientras se utilice el ancho de banda adecuado, el asignar la potencia por igual entre los distintos canales de transmisión implica una pérdida despreciable con respecto a la asignación de potencia óptima. Este hecho ha sido recientemente probado analíticamente en [Yu y Cioffi, 2006]. Siguiendo el planteamiento de [Sternad et al., 2007], entre otros, optamos por relajar la restricción de potencia y asumiremos que la potencia

total P_T es repartida por igual entre los N subcanales, de manera que la restricción de potencia (5.11) se convierte en

$$\sum_{k=1}^K \sum_{j=1}^J p_{knj} \leq \frac{P_T}{N}, \quad (5.12)$$

y el problema resultante es

$$\begin{aligned} \max_{\pi_{kl}, \beta_{kn}, p_{knj}, m_{knj}} \quad & U = \sum_{k=1}^K \frac{1}{\bar{\chi}^{(1-\alpha)}} \chi_k \\ \text{s.t.} \quad & \sum_{k=1}^K \sum_{j=1}^J p_{knj} \leq \frac{P_T}{N}, \quad \forall n = 1 \dots N \\ & \sum_{k=1}^K \beta_{kn} = 1, \quad \forall n = 1, \dots, N \\ & \sum_{l=1}^L \pi_{kl} \chi_{kl} \geq \chi_{kmin}, \quad \forall k = 1, \dots, K \\ & \sum_{l=1}^L \pi_{kl} = 1, \quad \forall k = 1, \dots, K \end{aligned} \quad (\text{P2})$$

Desafortunadamente, incluso con la relajación aplicada, el problema (P2), si bien resoluble mediante algoritmos de optimización estándar, sigue siendo muy exigente desde el punto de vista computacional. En la siguiente sección planteamos un método que reduce sustancialmente la complejidad del problema.

5.4. Asignación de subcanales, bits y potencia: método SBPA

En esta sección, describimos el esquema propuesto SBPA (*subchannel, bit and power allocation*) para asignar los subcanales, bits y potencia correspondientes al problema (P2). El esquema SBPA básicamente consiste en elaborar una matriz de *goodput* por usuario y grupo de subcanales para posteriormente realizar la asignación de subcanales. A continuación haremos una descripción del esquema, y dedicamos las subsecciones 5.4.1 y 5.4.2 a los detalles de formulación de los diferentes pasos del método.

Para resolver nuestro problema de asignación de recursos, se pueden considerar dos tipos de esquemas: centralizados, donde la EB se encarga de realizar todo el proceso de asignación, y distribuidos, donde los usuarios ejecutan los algoritmos

para la asignación de recursos. En ambos casos, los esquemas estarán dotados de un canal de retorno desde los usuarios a la EB. En nuestro caso, descomponemos el proceso total de asignación en dos pasos diferentes, lo cual conduce a un método semidistribuido (figura 5.2), como veremos luego.

Existe un paso previo al método propiamente dicho, durante el cual cada usuario ha de realizar la estimación de canal para conocer la ganancia de canal de las subportadoras, y esta información, representada en la figura por CSI (*channel state information*), sirve para comenzar el proceso de asignación en sí.

Como hemos dicho, el objetivo de la primera fase es proporcionar la matriz de *goodput* por usuario que se empleará en la segunda fase para realizar la asignación de subcanales. Consta el método por tanto de los siguientes pasos:

1. En paralelo, cada usuario k realiza la asignación de bits y potencia por subportadora mediante un algoritmo de *bit-loading* (ver detalles en 5.4.1), y calcula su vector de *goodput* $\mathbf{x}_k = [\chi_{k1} \dots \chi_{kL}]^T$ de dimensión L , donde L es el número de posibles combinaciones de los subcanales. A continuación, los usuarios envían la información necesaria (\mathbf{x}_k y la asignación de bits y potencia) a la EB a través del canal de retorno.
2. La EB ejecuta el algoritmo de asignación de subcanales y envía el resultado de la asignación a los usuarios por el enlace descendente. Acto seguido, la EB completa la transmisión, enviando la información hacia los usuarios organizada de acuerdo a la asignación de subcanales, bits y potencia resultante.

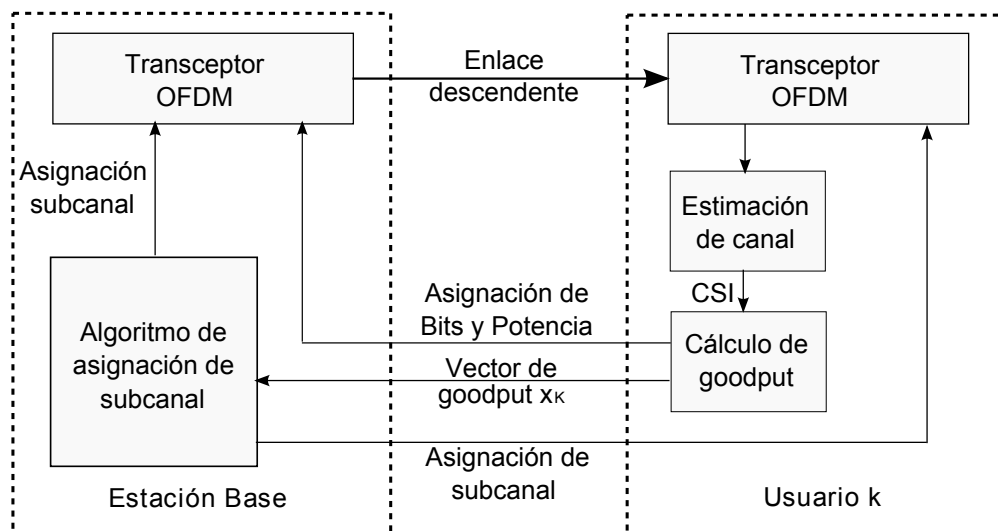


Figura 5.2. Modelo del sistema

En el caso de que hiciésemos uso de un esquema centralizado, el proceso completo de optimización sería realizado por la EB: el canal de retorno transmitiría, desde cada usuario, la ganancia de potencia por subportadora, y el enlace descendente transportaría la asignación de subcanal, bits y potencia a cada uno de los usuarios.

A la luz de la discusión previa, es sencillo comprobar que la información intercambiada, en términos de número de variables, es menor en el esquema semidistribuido propuesto comparado con el esquema centralizado genérico. Tenemos, por cada usuario, M variables de asignación de bit y M variables de asignación de potencia, así como L componentes del vector de *goodput*, luego en total los K usuarios harán llegar a la EB $(2M + L)K$ variables. Si añadimos a esto las N variables correspondientes a la asignación de los subcanales, tenemos un total de $2KM + KL + KN$ variables en total para el esquema semidistribuido. Para el caso centralizado, los usuarios envían los M valores correspondientes a las ganancias de las subportadoras, y han de recibir las M asignaciones de bit y de potencia y las N asignaciones de subcanal, lo que nos lleva a un total de $3KM + KN$ variables. En sistemas reales OFDMA, el número de subportadoras M es mucho mayor que el número de subcanales N [802, 2004], por lo que $L < M$ y en consecuencia el número de variables será menor en el esquema semidistribuido¹, a la vez que menos complejo computacionalmente. Además, el cálculo en paralelo de la asignación de bit y potencia que realiza cada usuario conduce a sistemas más rápidos y robustos que los sistemas centralizados.

Se puede cuestionar si la información del canal de retorno penaliza el rendimiento del sistema, dado que se debe destinar alguna tasa binaria a tal efecto. Los esquemas eficientes para compresión de información del canal de retorno pueden minimizar la tasa dedicada a la información de este canal [Eriksson y Ottosson, 2007], garantizando que las ganancias derivadas de la utilización de modulación adaptativa y por diversidad multiusuario son preservadas.

En lo que resta de esta sección, procedemos a detallar el método en dos pasos SBPA esbozado anteriormente, que permite reducir la complejidad del problema (P2). Mediante el primer paso, que denominamos GMC (*goodput matrix calcula-*

¹Por ejemplo, en el estándar IEEE 802.16e, si se utiliza el modo 512 el número de subportadoras para datos de usuarios es de 384, que con la asignación de 48 subportadoras por subcanal proporciona 8 subcanales. Por tanto, tenemos que $L = 256 < M = 384$. Para el resto de modos considerados de banda ancha, el número de subportadoras es mayor, siendo aún mayor por tanto la diferencia entre M y L

tion), hallamos las asignaciones de bit y potencia y determinamos la matriz de *goodput* \mathbf{X} , de dimensión $K \times L$. GMC nos permite reformular (P2) como un nuevo problema BIP, que se reduce a realizar la asignación de los subcanales. El segundo paso, que denominamos SAP (*subchannel allocation problem*), consiste precisamente en resolver dicho problema.

5.4.1. Cálculo de la matriz de *goodput* (GMC)

Este primer paso consiste en obtener la matriz de *goodput* \mathbf{X} , que estará formada por los vectores de *goodput* como

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1^T \\ \vdots \\ \mathbf{x}_K^T \end{bmatrix}. \quad (5.13)$$

En consecuencia, cada usuario k tiene que calcular el *goodput* χ_{kl} que recibiría si se le asignase el grupo de subcanales l , para todos los valores de l , y forma su vector de *goodput* $\mathbf{x}_k^T = [\chi_{k1} \dots \chi_{kL}]$. Para calcular χ_{kl} , nos basamos en el algoritmo de *bit-loading* descrito en [Devillers et al., 2008], que optimiza el *goodput* sobre un conjunto $\mathcal{J} = \{1, \dots, J\}$ de subportadoras dado, para enlaces OFDM punto a punto. Este algoritmo proporciona las expresiones de asignación de bits y potencia para conseguir la misma BER en todas las subportadoras, de manera que para dos subportadoras $j, j' \in \mathcal{J}$ cualesquiera, con ganancias de canal H_j y $H_{j'}$, respectivamente, ha de cumplirse la relación

$$\frac{2^{m_j}}{|H_j|^2} = \frac{2^{m_{j'}}}{|H_{j'}|^2}, \quad \forall j, j' \in \mathcal{J}, \quad (5.14)$$

y la potencia necesaria para la subportadora j es

$$p_j = \frac{(2^{m_j} - 1) \sigma^2 P_T / N}{B |H_j|^2} \frac{1}{\sum_{i=1}^J \frac{(2^{m_i} - 1) \sigma^2}{B |H_i|^2}} \quad (5.15)$$

Calcular los bits por símbolo en cada subportadora es sencillo una vez se ha asignado el número de bits m_1 a la subportadora con la ganancia de canal más baja. Previamente, por supuesto, las subportadoras han de ser ordenadas en orden ascendente según su ganancia de canal, de forma que $|H_1|^2 < \dots < |H_J|^2$, y así obtener el resto de valores de m_j a partir de m_1 mediante la expresión

$$m_j = m_1 + \log_2 \left(\frac{|H_j|^2}{|H_1|^2} \right), \quad \forall j \in \mathcal{J}. \quad (5.16)$$

Para calcular m_1 , recurrimos a la expresión siguiente, que optimiza el *goodput* sobre las J subportadoras:

$$\begin{aligned} \frac{2^{m_1}}{|H_1|^2} \left(\sum_{i=1}^{J'} \log_2 \left(\frac{|H_i|^2}{|H_1|^2} \right) + J' m_1 \right) (L c_L \epsilon_{av}^L + \dots + c_1 \epsilon_{av}) \ln 2 \frac{B P_T / N}{\sigma^2} \quad (5.17) \\ = \left(J' \frac{2^{m_1}}{|H_1|^2} - \sum_{i=1}^{J'} \frac{1}{|H_i|^2} \right)^2 \end{aligned}$$

Cabe la posibilidad de que la primera subportadora designada para obtener m_1 no sea válida, es decir, que no se pueda usar para transmitir ya que $m_1 \leq 0$. En realidad, la peor subportadora será la primera para la que, comenzando por la de ganancia $|H_1|^2$, se obtenga $m_k > 0$, con lo que el número total de subportadoras utilizadas va a ser en realidad $J' \leq J$.

En nuestro caso, extendemos el algoritmo al escenario OFDMA, donde el *bit-loading* será realizado sobre los subcanales (y por tanto sobre las subportadoras) asignados a cada usuario, de manera que el objetivo es encontrar los bits por símbolo m_{knj} y el nivel de potencia p_{knj} , para todos los usuarios y todas las subportadoras. En este algoritmo, los valores $\{m_{knj}\}$ son valores reales, lo que es un inconveniente para aplicaciones prácticas. Por tanto, una vez obtenida la solución real, para evaluar las prestaciones del método los restringimos a que tomen valores discretos para modulaciones MQAM (por ejemplo, $\Omega = \{0, 2, 4, 6\}$), y obtenemos los valores finales de m_{knj} redondeando al entero menor más próximo de Ω . Esto implica que la solución es subóptima, pero se muestra en [Devillers et al., 2008] que el efecto de este redondeo es prácticamente despreciable. Notar que $m_{knj} = 0$ significará que $p_{knj} = 0$. El objetivo final, que en esta fase es obtener cada χ_{kl} para construir \mathbf{X} , se alcanzará insertando los valores obtenidos de m_{knj} y p_{knj} en (5.5)–(5.7).

5.4.2. El problema de asignación de subcanales (SAP)

El objetivo del segundo paso es asignar los subcanales a los usuarios, esto es, encontrar la matriz óptima $\mathbf{\Pi}$ de dimensión $K \times L$, cuya entrada (k, l) es π_{kl} , y la matriz \mathbf{B} de dimensión $K \times N$, cuya entrada (k, n) es β_{kn} . Para obtener las matrices de asignación $\mathbf{\Pi}$ y \mathbf{B} , formulamos el problema (P3), que se deriva del problema (P2) una vez que \mathbf{X} es conocida. Procedemos de la siguiente manera: en lugar de considerar m_{knj} y p_{knj} como variables, imponemos la restricción $\chi_{kl} \in \mathbf{X}$,

obteniendo así el problema

$$\begin{aligned}
\max_{\pi_{kl}, \beta_{kn}} \quad & U = \sum_{k=1}^K \frac{1}{\bar{\chi}^{(1-\alpha)}} \chi_k \\
\text{s.t.} \quad & \sum_{k=1}^K \beta_{kn} = 1, \quad \forall n = 1, \dots, N \\
& \sum_{l=1}^L \pi_{kl} \chi_{kl} \geq \chi_{kmin}, \quad \chi_{kl} \in \mathbf{x}_k^T, \quad \forall k = 1, \dots, K, \\
& \sum_{l=1}^L \pi_{kl} = 1, \quad \forall k = 1, \dots, K
\end{aligned} \tag{P3}$$

siendo (P3) claramente menos complejo que (P2). En este punto podemos completar la resolución del problema de optimización del *goodput*. La opción inmediata es resolver (P3) por medio de algoritmos de programación entera, según se trató en la sección 3.1.2. Ahora bien, el problema (P3) sigue siendo un problema del tipo \mathcal{NP} -hard, y ninguno de los métodos estándar es capaz de resolver este tipo de problemas con complejidad polinómica, por lo que se opta por un algoritmo tipo *branch-and-bound*, dada su sencillez frente a otros métodos estándar. Por tanto, la correspondiente solución óptima de (P3), que denotamos por \mathbf{B}_{IP} , todavía adolece de una complejidad alta. A fin de reducir la misma, presentamos a continuación una aproximación que realiza la asignación de los subcanales de una forma más sencilla.

5.5. Asignación de subcanales, bits y potencia: método subóptimo S-SBPA

Supongamos ahora que, en lugar de transmitir los bits de todos los subcanales secuencialmente en la misma trama, es posible realizar la transmisión de los bits correspondientes a diferentes subcanales en paralelo. O dicho de otra manera, que es posible que los subcanales sean codificados independientemente. Esto significa que cada transmisor (EB) y receptor (usuario) procesan cada subcanal independientemente, con lo que se requieren tantos codificadores y decodificadores como subcanales en ambos extremos. Esta aproximación planteada la denominamos S-SBPA (subóptimal-SBPA) y proporcionará unos resultados subóptimos respecto al esquema presentado en la sección anterior.

La hipótesis de codificación independiente de los subcanales implica reformular el problema por completo, a partir del *goodput* χ_{kn} recibido por el usuario k cuando se le asigna el subcanal n . Dicho *goodput* lo calcularemos como

$$\chi_{kn} = r \left(\sum_{j=1}^J m_{knj} \right) FSR(\epsilon_{av,kn}). \quad (5.18)$$

El *goodput* total recibido por el usuario k es ahora la suma de *goodput* recibido individualmente por cada subcanal: $\chi_k = \sum_{n=1}^N \beta_{kn} \chi_{kn}$, siendo β_{kn} las variables de asignación de subcanales definidas en (5.10). Modificando adecuadamente las ecuaciones (5.4) y (5.5), obtenemos las expresiones de FSR y BER promedio como

$$FSR(\epsilon_{av}(k, n)) = d \exp(- (c_\tau \epsilon_{av}^\tau(k, n) + c_{\tau-1} \epsilon_{av}^{\tau-1}(k, n) + \dots + c_1 \epsilon_{av}(k, n))) \quad (5.19)$$

$$\epsilon_{av}(k, n) = \frac{1}{\sum_{j=1}^J m_{knj}} \sum_{j=1}^J m_{knj} 0,2 \exp \left(- \frac{1,6 |H_{knj}|^2 p_{knj}}{(2^{m_{knj}} - 1) \sigma^2} \right), \quad (5.20)$$

resultando con ello la función de utilidad U siguiente:

$$U = \sum_{k=1}^K \frac{1}{\bar{\chi}_k^{(1-\alpha)}} \sum_{n=1}^N \beta_{kn} \chi_{kn}. \quad (5.21)$$

Ahora, el problema consiste en maximizar la utilidad de la EB dda por la ecuación (5.21) encontrando los valores óptimos de asignación de subcanal, bit y potencia. Asumiendo de nuevo que la potencia total P_T está repartida por igual entre los N subcanales, el problema (P1) se convierte en

$$\begin{aligned} \max_{\beta_{kn}, p_{knj}, m_{knj}} \quad & U = \sum_{k=1}^K \frac{1}{\bar{\chi}_k^{(1-\alpha)}} \sum_{n=1}^N \chi_{kn} \beta_{kn} \\ \text{s.t.} \quad & \sum_{k=1}^K \sum_{j=1}^J p_{knj} \leq \frac{P_T}{N}, \quad \forall n = 1, \dots, N \\ & \sum_{k=1}^K \beta_{kn} = 1, \quad \forall n = 1, \dots, N \\ & \sum_{n=1}^N \beta_{kn} \chi_{kn} \geq \chi_{kmin}, \quad \forall k = 1, \dots, K. \end{aligned} \quad (P4)$$

Si comparamos (P4) respecto a (P2), advertimos que la transmisión de los bits correspondientes a cada subcanal codificados de forma independiente trae como consecuencia la desaparición de las variables de asignación de los subcanales π_{kl} de la formulación, suponiendo por tanto una mayor sencillez en la resolución del problema. Estudiaremos la solución del problema (P4) de manera similar a (P2), adaptando el método en dos pasos de la sección 5.4.

5.5.1. Cálculo de la matriz de *goodput* para el método subóptimo

Respecto al caso óptimo, la diferencia radica en que la fila k -ésima de la matriz de *goodput* \mathbf{X} es ahora $\mathbf{x}_k^T = [\chi_{k1} \dots \chi_{kN}]$; por tanto, la dimensionalidad de \mathbf{X} es $K \times N$. Cada usuario calcula \mathbf{x}_k y lo envía a la EB. Procediendo de forma similar a como hicimos en (5.4.1), se determinan las asignaciones de bit y potencia, y cada elemento x_{kn} de la matriz de *goodput* \mathbf{X} se obtiene insertando (5.19)–(5.20) en (5.18), ya que $x_{kn} = \chi_{kn}$.

5.5.2. El problema de la asignación de subcanales para el método subóptimo

Como ya vimos en (5.4.2), después de calcular la matriz de *goodput* el problema se reduce a resolver la asignación de los subcanales a los usuarios, es decir, encontrar los valores de $\beta_{kn} \in \mathbf{B}$ óptimos, siendo \mathbf{B} la matriz de asignación de subcanales de dimensión $K \times N$. Aunque la naturaleza del problema es de nuevo combinatoria,, es considerablemente menos complejo que (P3), al evitar el cálculo de la matriz $\mathbf{\Pi}$. El problema a resolver queda por tanto

$$\begin{aligned} \max_{\beta_{kn}} U &= \sum_{k=1}^K \frac{1}{\bar{\chi}_k^{(1-\alpha)}} \sum_{n=1}^N \chi_{kn} \beta_{kn} \\ \text{s.t.} \quad &\sum_{k=1}^K \beta_{kn} = 1, \quad \forall n = 1, \dots, N \\ &\sum_{n=1}^N \beta_{kn} \chi_{kn} \geq \chi_{kmin}, \quad \forall k = 1, \dots, K \\ &\beta_{kn} \in \{0, 1\}. \end{aligned} \tag{P5} \tag{5.22}$$

donde χ_{kn} se obtienen de la k -ésima fila de \mathbf{X} . Para ello resolver (P5), discutimos a continuación tres propuestas diferentes, dos de ellas estándar y el algoritmo RLP (*rounding linear programming*) que se propone.

Solución entera óptima

Aquí se plantea la misma discusión que en (5.4.2). La matriz de asignación de subcanales \mathbf{B}_{IP} representa la solución de (P5) obtenida directamente mediante un algoritmo tipo *branch-and-bound* y $U(\mathbf{B}_{IP})$ es el valor de la función de utilidad correspondiente. No obstante, esta aproximación óptima implica todavía una alta complejidad.

Solución subóptima: relajación de la restricción de integralidad

Una aproximación subóptima muy extendida en la práctica para resolver problemas BIP es la relajación de la condición de integralidad, para obtener un problema de programación lineal, como vimos en (3.1.3). Más concretamente, referido a (P5) significa que la condición $\beta_{kn} \in [0, 1]$ sustituye a la restricción $\beta_{kn} \in \{0, 1\}$, de manera que el problema a resolver es ahora

$$\begin{aligned}
 \max_{\beta_{kn}} U &= \sum_{k=1}^K \frac{1}{\bar{\chi}_k^{(1-\alpha)}} \sum_{n=1}^N \chi_{kn} \beta_{kn} \\
 \text{s.t.} \quad &\sum_{k=1}^K \beta_{kn} = 1 \quad \forall n = 1, \dots, N \\
 &\sum_{n=1}^N \beta_{kn} \chi_{kn} \geq \chi_{kmin}, \quad \forall k = 1, \dots, K \\
 &\beta_{kn} \in [0, 1].
 \end{aligned} \tag{P5-LP}$$

La solución a este problema, que denotaremos por \mathbf{B}_{LP} , proporciona el valor de la función de utilidad $U(\mathbf{B}_{LP})$, el cual es obviamente una cota superior de $U(\mathbf{B}_{IP})$, según se vió en (3.1.4).

Existe un amplio espectro de algoritmos para resolver problemas LP. Desde un punto de vista teórico y práctico, los algoritmos de punto interior primal-dual resultan ser preferibles a otros métodos [Bertsekas, 1999]. Entre estos, hemos seleccionado el algoritmo de [Mehrotra, 1992] por su probada eficiencia en la práctica. Sin embargo, la solución obtenida no resulta válida para implementaciones prácticas: como β_{kn} puede tomar valores reales, esto implicaría que, tomando como ejemplo el caso de dos usuarios k_1 and k_2 , tal que $\beta_{k_1n} \neq 0$ y $\beta_{k_2n} \neq 0$, ambos deben compartir el mismo subcanal, lo que no es coherente con la formulación original del problema, ya que cada subcanal es asignado a un único usuario. Aun así, esta solución no entera \mathbf{B}_{LP} la utilizaremos para generar una solución entera que supera esta limitación, como describimos a continuación.

Solución subóptima: algoritmo RLP

Las dos soluciones anteriores plantean serios inconvenientes para ser aplicadas en la práctica. De un lado, la complejidad de los algoritmos *branch-and-bound*, que puede hacer inviable su aplicación en sistemas de tiempo real. De otro, la relajación de la condición de integralidad; hemos visto que una aplicación práctica requeriría una señalización no trivial para realizar el reparto de las subportadoras

de los subcanales, por ejemplo mediante FDMA, y en cualquier caso no estaríamos resolviendo el problema planteado de asignar subcanales. Por estas razones, proponemos un algoritmo, al que nos referiremos como RLP (*rounded linear programming*), que posee las siguientes características:

- i. Proporciona una solución entera válida, a la que nos referiremos como \mathbf{B}_{RLP} , tomando como punto de partida la solución del problema relajado \mathbf{B}_{LP} .
- ii. Satisface la restricción de *goodput* mínimo requerido.
- iii. Consigue que la función de utilidad $U(\mathbf{B}_{RLP})$ tome un valor tan próximo como sea posible a $U(\mathbf{B}_{LP})$.
- iv. Hereda la baja complejidad que es inherente a la aproximación LP.

Notar que la solución entera inmediata, el simple redondeo de \mathbf{B}_{LP} al entero más próximo, no aseguraría que se satisficieran siempre (ii) y (iii). Además, podría suponer la no asignación de un subcanal n si todas las variables de asignación asociadas toman valores inferiores a 0.5, es decir, $\beta(k, n) \leq 0,5 \forall n$.

El algoritmo RLP hace que sean posibles implementaciones prácticas, ya que la solución es entera y cumple a la vez con la restricción de *goodput* mínimo. Este algoritmo está basado en la observación de que la solución óptima \mathbf{B}_{IP} y la solución \mathbf{B}_{LP} tienen significativas similitudes en los valores de las variables de asignación de subcanales β_{kn} . Por otro lado, evitamos la complejidad de los algoritmos óptimos BIP al considerar como punto de inicio la solución \mathbf{B}_{LP} .

El resto de esta sección lo dedicamos a explicar los principales pasos del algoritmo, cuyo pseudocódigo se proporciona inmediatamente a continuación. Por lo que se refiere a la notación, tanto en la descripción del algoritmo como en el pseudocódigo nos referiremos a las variables $\beta(k, n)$ correspondientes a \mathbf{B}_{LP} y \mathbf{B}_{RLP} como $\beta_{LP}(k, n)$ y $\beta_{RLP}(k, n)$ respectivamente.

En primer lugar, definimos los conjuntos S_K como el conjunto total de usuarios, y S_N como el conjunto formado por los subcanales disponibles (paso 1). Inicialmente, los subcanales son asignados redondeando \mathbf{B}_{LP} de forma que el subcanal n quede asignado sólo al usuario k que mayor proporción del subcanal haya obtenido (paso 2). Este redondeo puede causar que la restricción de *goodput* mínimo deje de cumplirse para algunos usuarios. A continuación, los subcanales quedan asignados sólo a aquellos usuarios que obtienen el *goodput* mínimo requerido con la asignación de un único subcanal (pasos 3–4). Estos usuarios forman el conjunto

S'_K , y S'_N es el conjunto de subcanales asignados en 3-4, es decir los subcanales asignados a los usuarios $k \in S'_K$ (paso 5). A partir de los anteriores, formamos ahora dos conjuntos más, S''_K y S''_N (paso 6). $S''_K = S_K - S'_K$ representa aquellos usuarios cuyo *goodput* resultante, después de los pasos 3-5, es inferior al mínimo χ_{kmin} . Los subcanales que aún no han sido asignados constituirán el conjunto $S''_N = S_N - S'_N$.

Ahora, los subcanales de S''_N tenemos que asignarlos a los usuarios de S''_K . La clave para esta asignación es asignar cada subcanal n'' de S''_N al usuario $k'' \in S''_K$ que obtenga el mayor *goodput* en n'' , pero sólo teniendo en cuenta la solución \mathbf{B}_{LP} (pasos 7-9): entre los usuarios de S''_K que comparten n'' en la solución \mathbf{B}_{LP} (usuarios k'' tal que $\beta_{LP}(k'', n'') \neq 0$), seleccionamos el usuario k^* que recibe el mayor *goodput* de n'' . Sin embargo, algunos usuarios en este punto pueden aún no haber conseguido el *goodput* mínimo requerido χ_{kmin} , pero tomando como base la asignación obtenida hasta este instante, se procede a refinar la solución \mathbf{B}_{RLP} en dos fases, para lo que previamente se define el conjunto S'''_K conteniendo aquellos usuarios (paso 10): $S'''_K = \{k \in S''_K \mid \chi_k < \chi_{kmin}\}$.

El ajuste grueso de la asignación (pasos 12-23) es realizado si más de un subcanal ha de ser reasignado, es decir, $\|S''_N\| > 1$. Mediante el intercambio de subcanales entre $k \in S'''_K$ y el resto de usuarios de S''_K se intenta alcanzar la tasa mínima $\chi_k \geq \chi_{kmin}$ para todos los usuarios, con un criterio similar al empleado en los pasos 7-9.

Un ajuste final (pasos 24-38) puede ser necesario en el caso de que $S'''_K \neq \emptyset$ todavía. Se realiza un intercambio de subcanales entre pares de usuarios, basado en ordenar los pares subcanal/usuario según el parámetro heurístico v . Este parámetro garantiza una desviación lo menor posible de la solución óptima \mathbf{B}_{LP} . Para calcular v , consideramos de nuevo la solución \mathbf{B}_{LP} . Todos los pares (k, n) tales que $round(\beta_{LP}(k, n)) = 1$ en el paso 2 son tenidos en cuenta para la reasignación de subcanales. Definimos v para estos pares (k, n) como $v(k, n) = \frac{\chi_{kn} - \chi_{kmin}}{\hat{\chi}_k}$, siendo $\hat{\chi}_k$ el *goodput* total obtenido por el usuario k tras el paso 9. La primera opción para un usuario k_1 que no haya conseguido χ_{kmin} es intercambiar su subcanal n_1 con el usuario k de máximo v .

Algoritmo RLP

- 1: Definir $S_K = \{1, \dots, K\}$, $S_N = \{1, \dots, N\}$
- 2: Inicializar \mathbf{B}_{RLP} como $\mathbf{B}_{LP} \forall n$:

$$\beta_{RLP}(k, n) = \begin{cases} 1 & \text{si } k = \arg \max_{k \in S_K} \beta_{LP}(k, n) \\ 0 & \text{resto} \end{cases}$$

- 3: Definir $A = \{(k, n) \mid \beta_{RLP}(k, n) = 1, \chi_{kn} \geq \chi_{kmin}\}$
- 4: Asignar subcanales de acuerdo a:

$$\beta_{RLP}(k, n) = \begin{cases} 1 & \text{si } (k, n) \in A \\ 0 & \text{resto} \end{cases}$$

- 5: Definir los conjuntos S'_K y S'_N :

$$S'_K = \{k \in S_K : \exists n \in S_N \mid (k, n) \in A\}$$

$$S'_N = \{n \in S_N : \exists k \in S_K \mid (k, n) \in A\}.$$

- 6: Definir $S''_K = S_K - S'_K$ and $S''_N = S_N - S'_N$.
- 7: **for** each $n'' \in S''_N$ **do**
- 8: $\beta_{RLP}(k^*, n'') = 1 \Leftrightarrow k^* = \arg \max_{k'' \in S''_K} (\chi_{k''n''})$ and $\beta_{LP}(k^*, n'') \neq 0$
- 9: **end for**
- 10: Identificar usuarios que no reciben χ_{kmin} : $S'''_K = \{k \in S''_K \mid \chi_k < \chi_{kmin}\}$. Identificar subcanales sin asignar: $S'''_N = \{n \in S_N \mid \beta_{RLP}(k, n) = 0, \forall k \in S_K\}$
- 11: **if** $S'''_K \neq \emptyset$ **then**
- 12: **if** $\|S'''_N\| > 1$ **then**
- 13: **for all** $k \in S'''_K$ **do**
- 14: $S_k^n = \{n \in S_N \mid \beta_{RLP}(k, n) = 1\}$
- 15: **while** $\chi_k \leq \chi_{kmin}$ and $S_k^n \neq \emptyset$ **do**
- 16: Tomar $n \in S_k^n$ aleatoriamente
- 17: $n^* = \arg \max_r \chi_{kr}$, $r \in \{n \in S_N \mid \beta_{RLP}(q, n) = 1, q \in S'''_K\}$
- 18: $k^* = \arg_q \{q \in S'''_K \mid \beta_{RLP}(q, n^*) = 1\}$
- 19: Intercambiar los subcanales n y n^* entre k y k^* :

$$\beta_{RLP}(k, n) = 0, \beta_{RLP}(k, n^*) = 1$$

$$\beta_{RLP}(k^*, n^*) = 0, \beta_{RLP}(k^*, n) = 1$$
- 20: **end while**

```

21:   end for
22: end if
23: end if
24: if  $S_K''' \neq \emptyset$  then
25:   for all  $k_1 \in S_K'''$  do
26:     Encontrar todos los pares  $(k, n)$ ,  $k \neq k_1$ , tales que  $\text{round}(\beta_{LP}(k, n)) = 1$ 
27:      $S_{k_1}^n = \{n \in S_N \mid \beta_{RLP}(k_1, n) = 1\}$ 
28:     Calcular el parámetro  $v(k, n)$  y formar  $\mathbf{v}^T = (v_1, \dots, v_j, \dots, v_{N-|S_{k_1}^n|})$ ,
        con  $v_1 \geq \dots \geq v_{N-|S_{k_1}^n|}$ 
29:     Inicializar  $j = 1$  y tomar el primer  $n_1 \in S_{k_1}^n$ 
30:     while  $\chi_{k_1} \leq \chi_{kmin}$  do
31:       Intercambiar  $n_1$  con el correspondiente  $n_j$  de  $v_j(k_j, n_j)$ , entre  $k_1$  y  $k_j$ 
32:       if  $\chi_k \leq \chi_{kmin}$  then
33:         Deshacer el intercambio de subcanales
            $j = j + 1$ 
34:       end if
35:       Tomar el siguiente  $n_1 \in S_{k_1}^n$  si  $j = N - |S_k^n| + 1$ 
36:     end while
37:   end for
38: end if

```

Transmisión secuencial empleando RLP

En la sección anterior, hemos supuesto que los bits correspondientes a diferentes subcanales eran transmitidos en paralelo, con codificación independiente en cada subcanal. Llegados a este punto, cabe plantearse si es posible adaptar el algoritmo RLP a sistemas reales. Esto supone realizar la transmisión de los bits de forma secuencial en una única trama, pero utilizando la asignación obtenida mediante el algoritmo RLP, sin tener una pérdida significativa de prestaciones respecto a la asignación óptima que obteníamos en la sección 5.4. Por tanto, la asignación de bit y potencia obtenida mediante la aproximación S-SBPA (sección 5.5) y la asignación de subcanales obtenida usando el algoritmo RLP, se utilizan para transmitir los bits de información sin tener que codificarlos independientemente para cada subcanal, que era la hipótesis de partida del método S-SBPA. La utilización del algoritmo RLP para la transmisión secuencial de los bits la denotaremos como algoritmo ST-RLP (*sequential transmission-RLP*), y en la sección de resultados

mostramos que mediante ST-RLP se consigue un ahorro importante del tiempo de procesamiento, a la vez que se consigue un muy buen compromiso entre complejidad y rendimiento.

5.6. Simulaciones y resultados

En esta sección se presentan los resultados de las simulaciones realizadas, a fin de evaluar el rendimiento de los algoritmos y su complejidad. Los parámetros del sistema han sido elegidos de acuerdo al modo 512 definido en [802, 2006b], que proporciona 384 subportadoras de datos y cuenta con 48 subportadoras por subcanal ($N = 8$ subcanales), con un ancho de banda de 10937,5 Hz asociado a cada subportadora. Como esquema de corrección de errores empleamos codificación convolucional con polinomios generadores [5, 7] de tasa 1/2, y la modulación adaptativa en cada subportadora es 4/16/64QAM, que son los niveles admitidos en [802, 2006b]. Para estos valores, la expresión resultante de la FSR según la ecuación (5.4) es:

$$FSR(\epsilon_{av}) = 0,987 \exp(15441\epsilon_{av}^3 + 462,8\epsilon_{av}^2 - 5\epsilon_{av})$$

para una duración de trama de 1024 bits.

La densidad espectral de ruido toma un valor de $N_0 = -80$ dBm/Hz y la relación entre la potencia máxima recibida por subcanal y la densidad espectral de ruido es de 40 dBHz, esto es $\frac{P_{max}/N}{N_0} = 40$ dBHz. El período del símbolo OFDM es de 102.9 μ s, como en [802, 2006b], y vamos a considerar que el *goodput* mínimo recibido es 875 kbps, esto es, 90 bits de información por símbolo OFDM. En nuestro escenario de simulación, el *goodput* promedio $\bar{\chi}_k$ es actualizado en cada símbolo OFDM transmitido, de manera que la utilidad $U = \sum_{k=1}^K \frac{1}{\bar{\chi}_k^{(1-\alpha)} \chi_k}$ es actualizada de forma similar. Las simulaciones se han realizado para un período de 100 símbolos OFDM.

Por lo que respecta a la variación del canal, vamos a considerar dos casos extremos. Primero, analizamos el caso de variación lenta del canal, donde los coeficientes de ganancia del canal por subportadora H_{knj} no cambian de símbolo a símbolo. En segundo lugar, analizamos el caso de canal de variación rápida, donde el sistema experimenta desvanecimiento por bloques y los coeficientes H_{knj} varían de símbolo a símbolo. Este segundo caso, aunque poco realista por extremo, nos sirve para comprobar que incluso en estas circunstancias el método propuesto

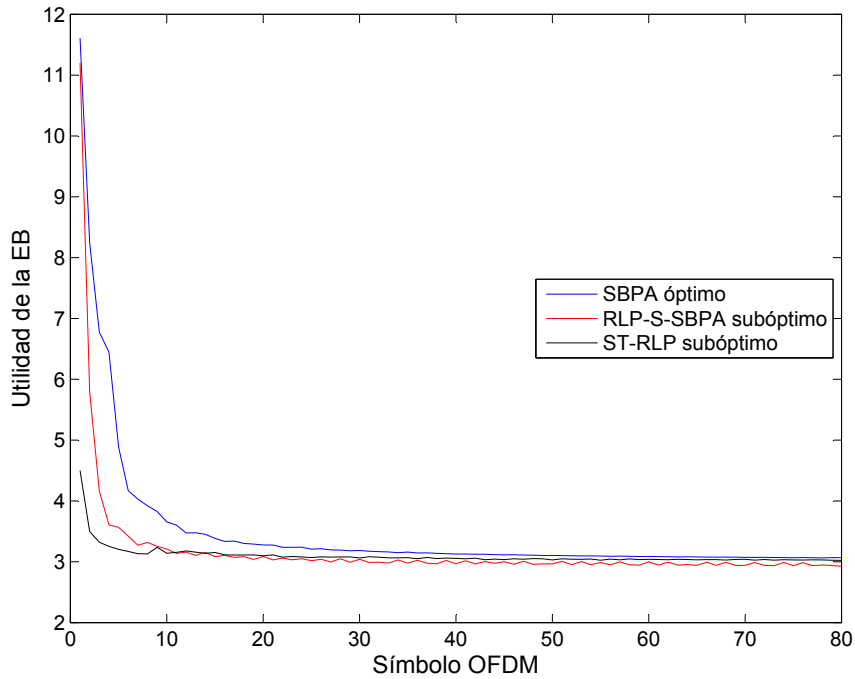


Figura 5.3. Algoritmos óptimo y subóptimos: utilidad obtenida por la estación base con canal de variación lenta, para 3 usuarios con imparcialidad proporcional

se comporta correctamente.

5.6.1. Canal de variación lenta

Para este canal, consideramos que los coeficientes H_{knj} son constantes durante los 100 símbolos de la simulación. Las figuras 5.3–5.5 muestran la utilidad de la EB para los algoritmos ST-RLP (*sequential transmission*-RLP), SBPA (*subchannel, bit and power allocation*) y RLP-S-SBPA (RLP-*suboptimal*-SBPA), para los valores de $\alpha = 0$ (asignación con imparcialidad proporcional), $\alpha = 0,5$ y $\alpha = 1$ (asignación para máximo *goodput*), para el caso de 3 usuarios. Observamos que las pérdidas de los algoritmos subóptimos ST-RLP RLP-S-SBPA son despreciables comparadas con el óptimo SBPA.

Para verificar que se consigue el grado de imparcialidad deseado, utilizamos como métrica el *goodput* recibido acumulado por un usuario k en bits de información por símbolo OFDM, al que denominamos $\chi_{cum}(k)$. Este *goodput* acumulado

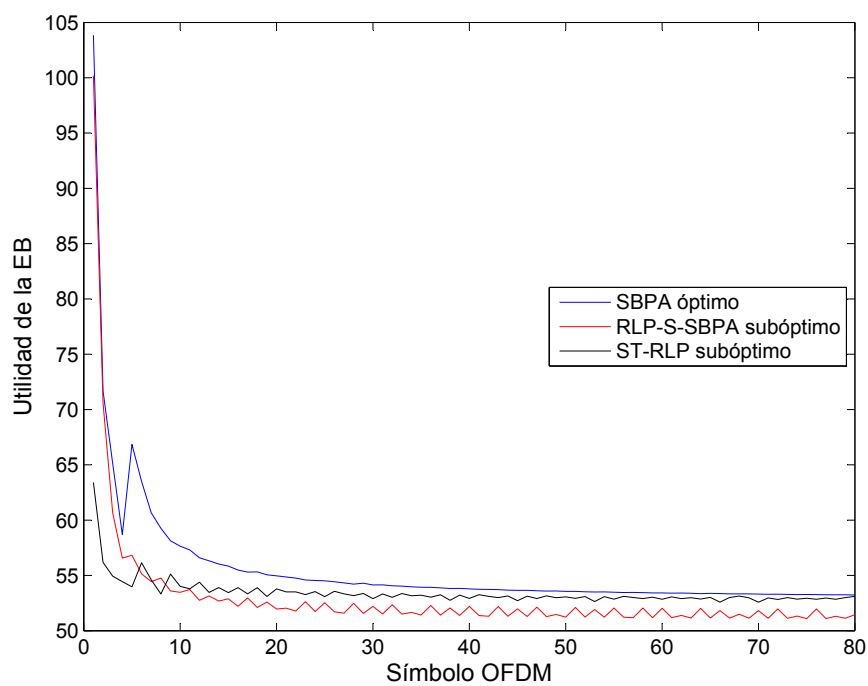


Figura 5.4. Algoritmos óptimo y subóptimos: utilidad obtenida por la estación base con canal de variación lenta, para 3 usuarios con $\alpha = 0,5$

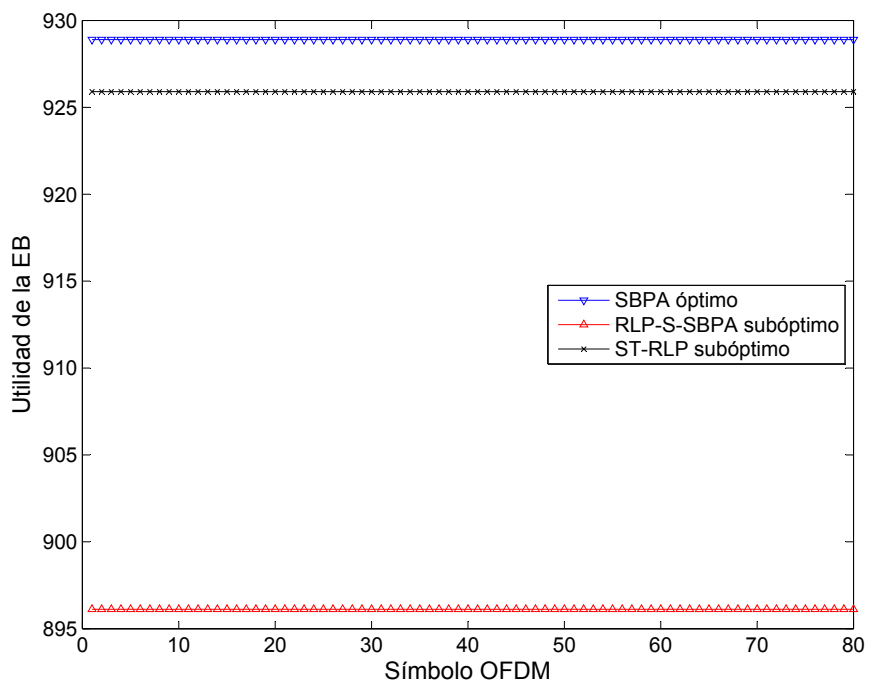


Figura 5.5. Algoritmos óptimo y subóptimos: utilidad obtenida por la estación base con canal de variación lenta, para 3 usuarios con asignación de máximo *goodput*

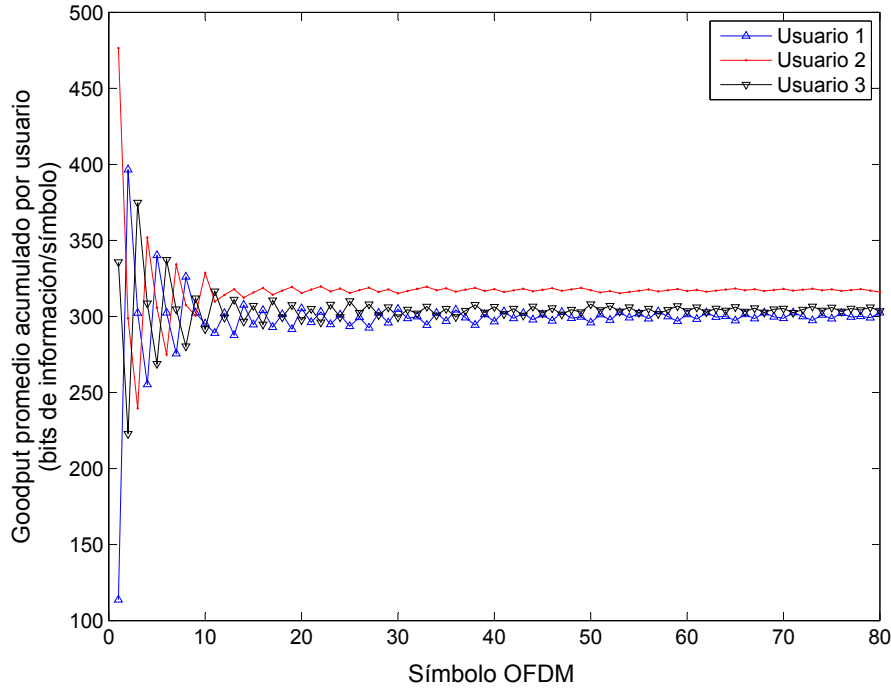


Figura 5.6. Algoritmo ST-RLP, 3 usuarios: goodput acumulado recibido por usuario, con canal de variación lenta, con imparcialidad proporcional

lo calculamos como

$$\chi_{cum}(k) = \frac{1}{F} \sum_{f=1}^F \chi_{kf}, \quad (5.24)$$

siendo F el número de símbolos OFDM transmitidos y χ_{kf} el *goodput* recibido por el usuario k cuando fue transmitido el símbolo f . Las figuras 5.6–5.8 representan el *goodput* acumulado cuando se emplea el algoritmo ST-RLP, mostrando que cuanto mayor es el valor de α , menor es el grado de imparcialidad, como era de esperar. Si variamos el número de usuarios, el sistema se sigue comportando en la forma prevista, como ilustran las figuras 5.9–5.11 para 7 usuarios.

5.6.2. Canal de variación rápida

Ahora, estudiaremos el rendimiento del sistema para un canal de variación rápida. El modelado de esta situación lo haremos considerando que los coeficientes H_{knj} varían aleatoriamente entre símbolos consecutivos. Se muestra en las figuras 5.12–5.14 que, al igual que para canal de variación lenta, la utilidad obtenida mediante ST-RLP es casi-óptima, para 3 usuarios. Debido a la variación aleatoria de

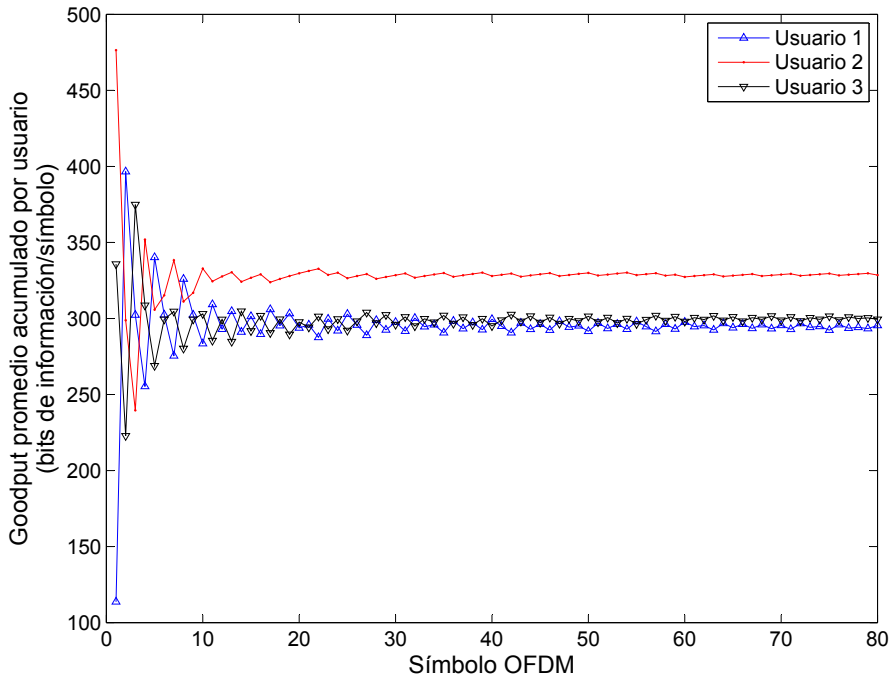


Figura 5.7. Algoritmo ST-RLP, 3 usuarios: goodput acumulado recibido por usuario, con canal de variación lenta, con $\alpha = 0,5$

los coeficientes de canal de las subportadoras, cuyo valor es diferente para cada símbolo OFDM, se espera que, después de la transmisión de un cierto número de símbolos, el efecto del parámetro de imparcialidad α se desvanezca. No obstante, las simulaciones realizadas, para valores de K entre 2 y 7 usuarios, confirman que el algoritmo ST-RLP tiene un comportamiento cercano al óptimo, como se muestra para el caso de 5 usuarios. La figura 5.15 muestra cómo la imparcialidad proporcional se mantiene para el canal de variación rápida, mientras que en el caso de asignación para máximo *goodput* (figura 5.16), el comportamiento egoísta característico con $\alpha = 1$ se conserva para un alto porcentaje de los símbolos OFDM.

5.6.3. Evaluación comparativa y de la complejidad

En la literatura no existen otros esquemas similares que optimicen la asignación de recursos formulada en términos de *goodput*. Por eso, a efectos comparativos, se han considerado otros algoritmos que, aun no teniendo como objetivo maximizar el *goodput* transmitido por la estación base, sí proporcionen buenos resultados al realizar *bit-loading*. De entre estos algoritmos, comparamos el algoritmo ST-RLP

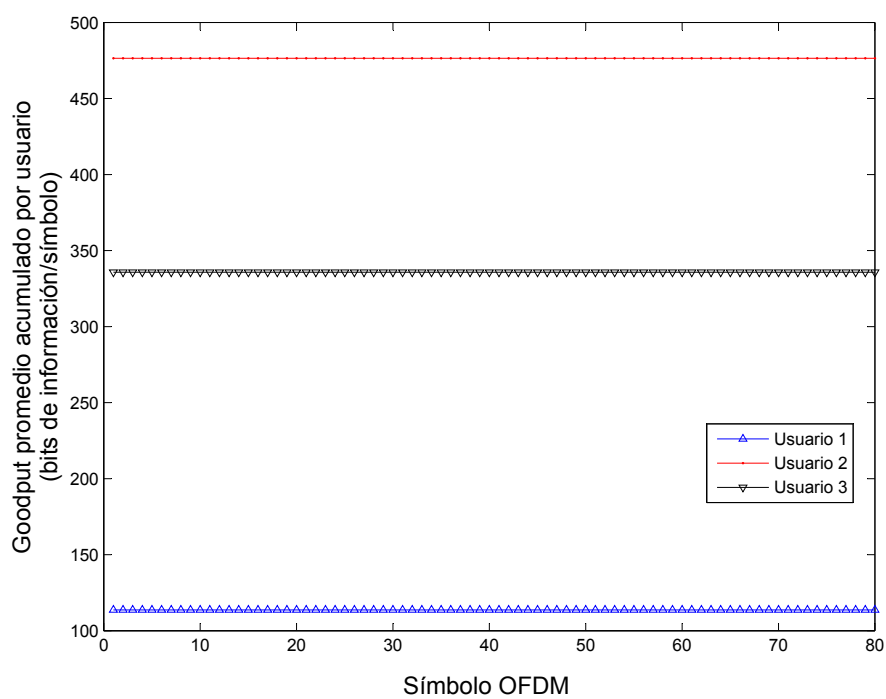


Figura 5.8. Algoritmo ST-RLP, 3 usuarios: goodput acumulado recibido por usuario, con canal de variación lenta, con máximo goodput

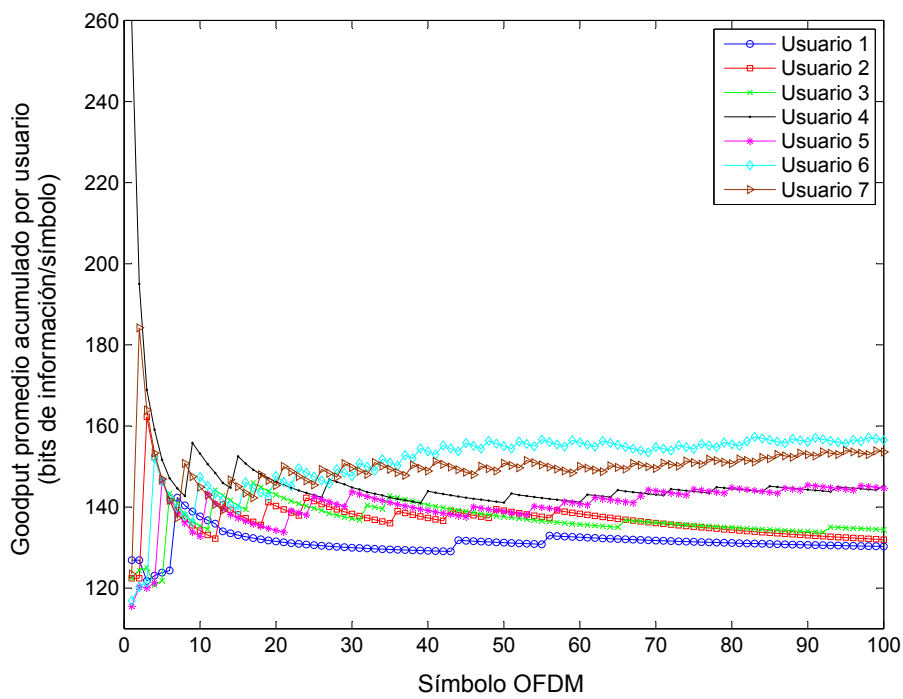


Figura 5.9. Algoritmo ST-RLP, 7 usuarios: goodput acumulado recibido por usuario, con canal de variación lenta, con imparcialidad proporcional

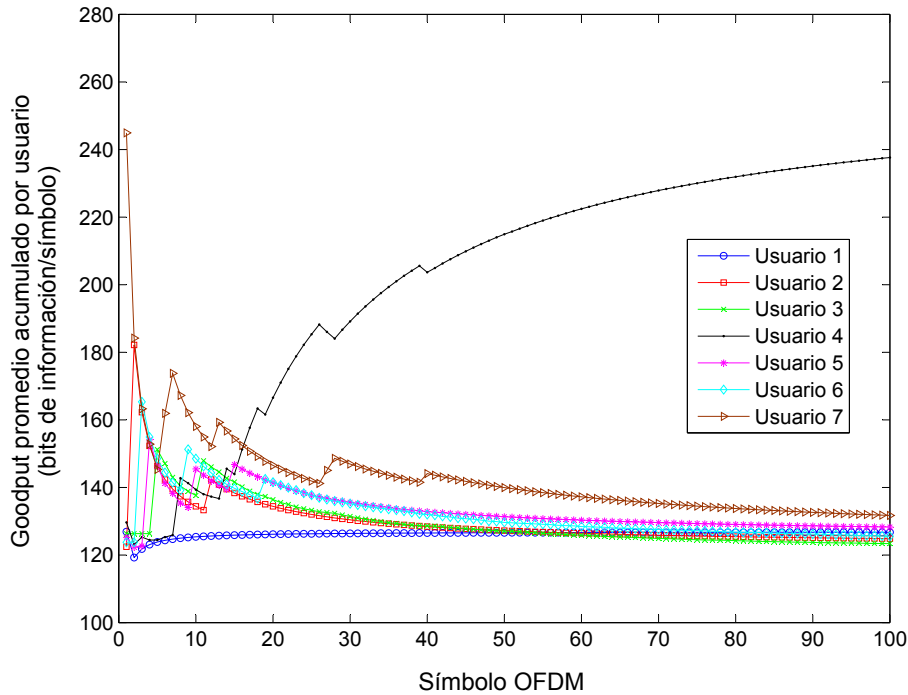


Figura 5.10. Algoritmo ST-RLP, 7 usuarios: goodput acumulado recibido por usuario, con canal de variación lenta, con $\alpha = 0,5$

propuesto con el algoritmo de asignación conjunta de subportadora y potencia de [Mohanram y Bhashyam, 2005], al que denominaremos REF, el cual, aunque no incluye *goodput* en su formulación, ofrece un rendimiento comparable al de algoritmos de probada eficiencia como el desarrollado en [Shen et al., 2005], e incorpora imparcialidad en su formulación. Mostramos como ejemplo el caso de $K = 2$ usuarios en la figura 5.17, para canal rápidamente variante, donde observamos que el *goodput* total transmitido por la EB es significativamente mayor empleando ST-RLP que utilizando el algoritmo REF y que no hay apenas diferencia respecto al óptimo SBPA.

En términos de complejidad, la diferencia en el rendimiento la establecemos en términos de tiempo de cómputo de CPU. Las simulaciones han sido realizadas utilizando un procesador Pentium M 1.6 GHz. Los resultados, mostrados en la figura 5.18, indican una complejidad mucho más elevada en el caso del algoritmo REF, incluso superior a la complejidad asociada al caso óptimo resuelto con el método propuesto SBPA, de forma que alcanza hasta cientos de veces a favor del algoritmo ST-RLP. En media, los valores obtenidos son 0.45 s para el algoritmo ST-RLP, 30.81 s para el algoritmo óptimo SBPA y 345.7 s para el algoritmo REF

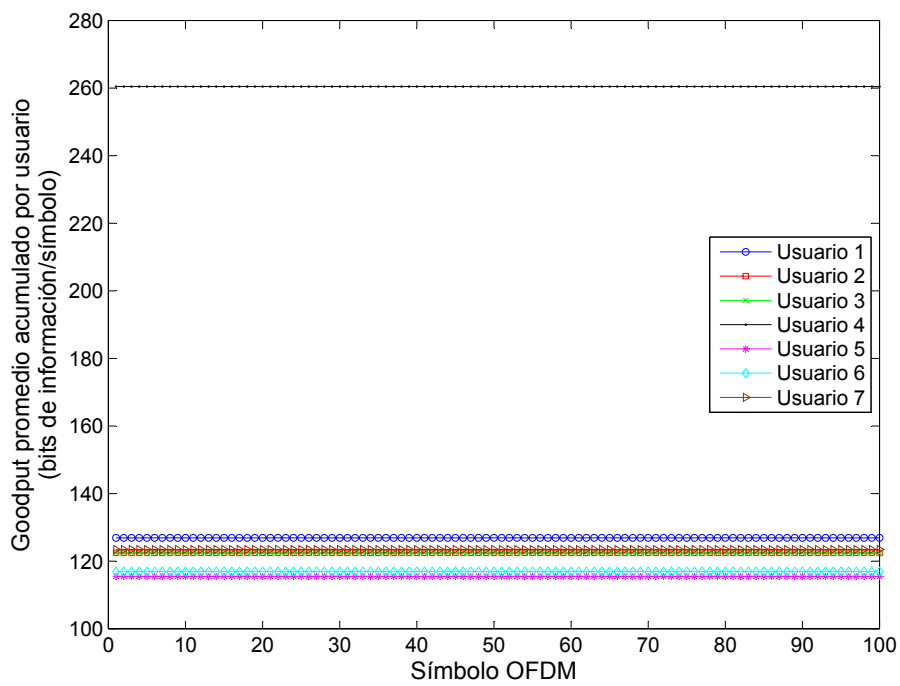


Figura 5.11. Algoritmo ST-RLP, 7 usuarios: goodput acumulado recibido por usuario, con canal de variación lenta, con máximo goodput

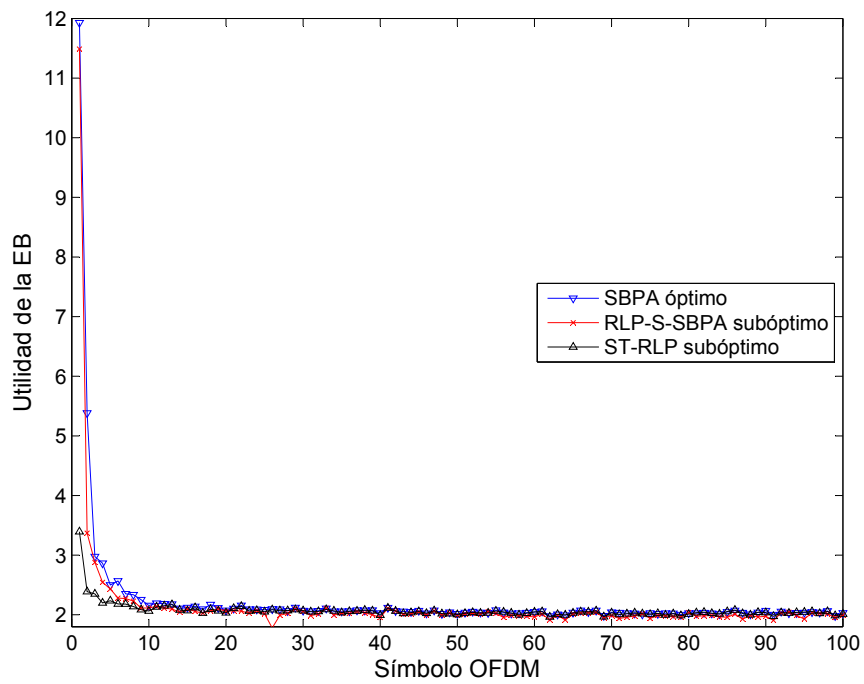


Figura 5.12. Algoritmos óptimo y subóptimos: utilidad obtenida por la estación base con canal de variación rápida, para 3 usuarios con imparcialidad proporcional

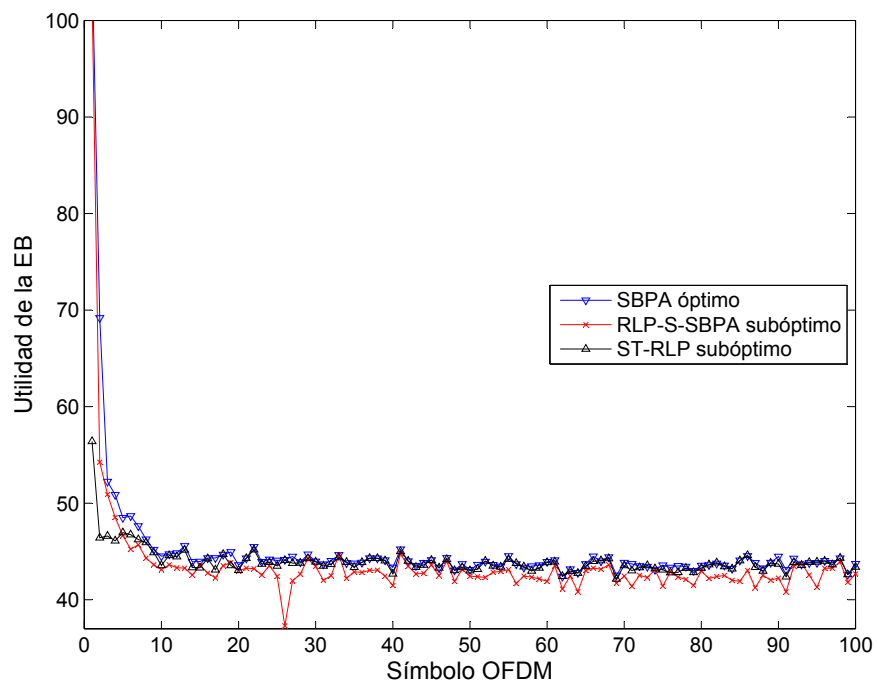


Figura 5.13. Algoritmos óptimo y subóptimos: utilidad obtenida por la estación base con canal de variación rápida, para 3 usuarios con $\alpha = 0,5$

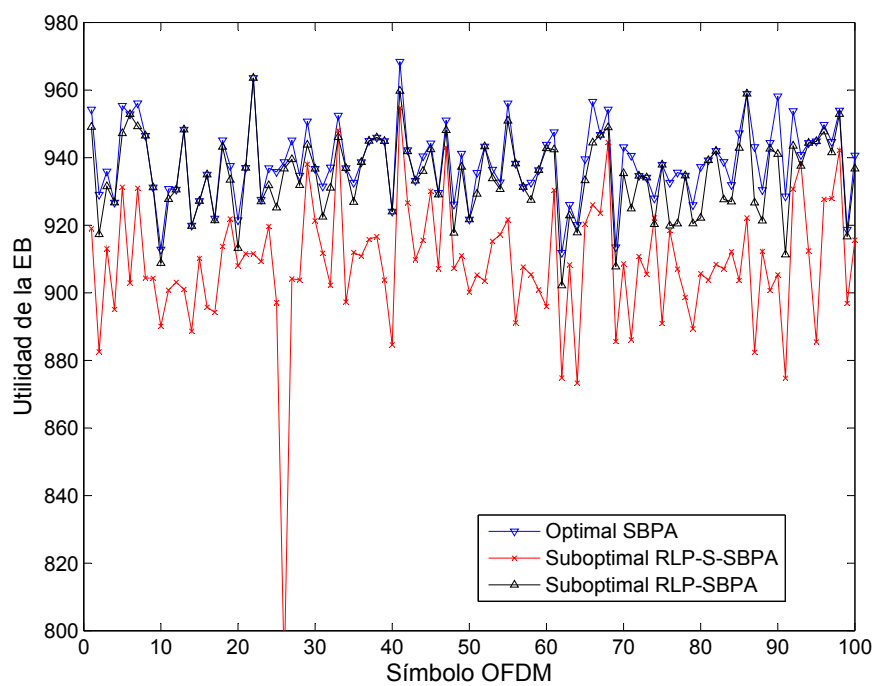


Figura 5.14. Algoritmos óptimo y subóptimos: utilidad obtenida por la estación base con canal de variación rápida, para 3 usuarios con asignación de máximo goodput

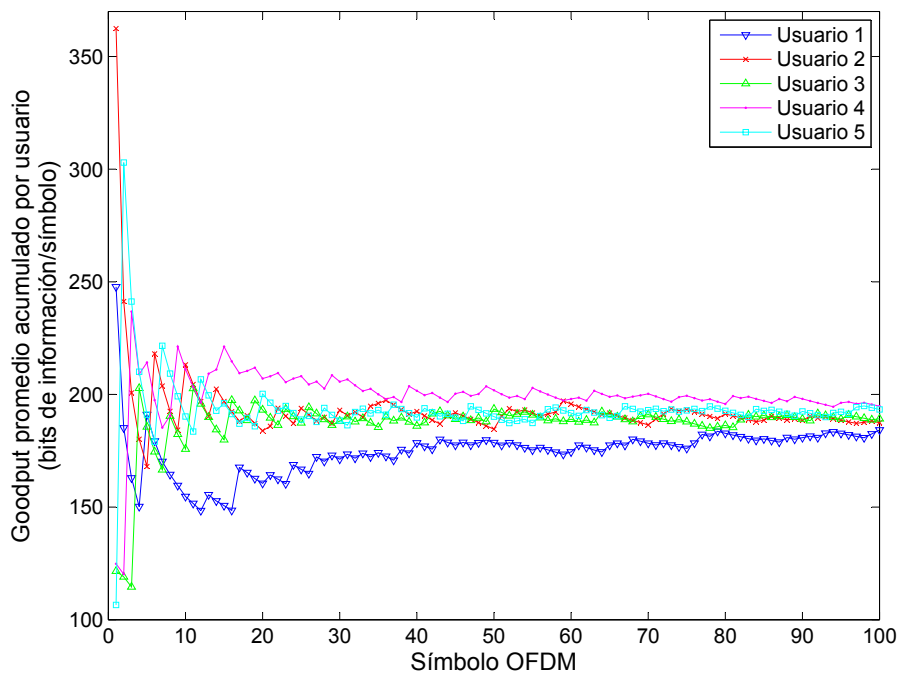


Figura 5.15. Algoritmo ST-RLP, 5 usuarios: goodput acumulado recibido por usuario, con canal de variación rápida, para imparcialidad proporcional

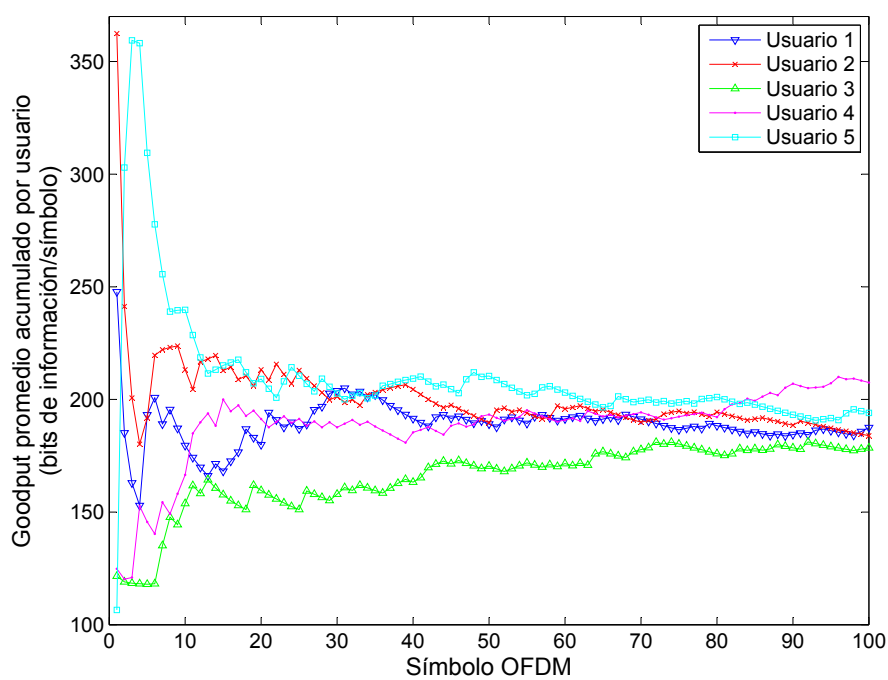


Figura 5.16. Algoritmo ST-RLP, 5 usuarios: goodput acumulado recibido por usuario, con canal de variación rápida, para asignación de máximo goodput

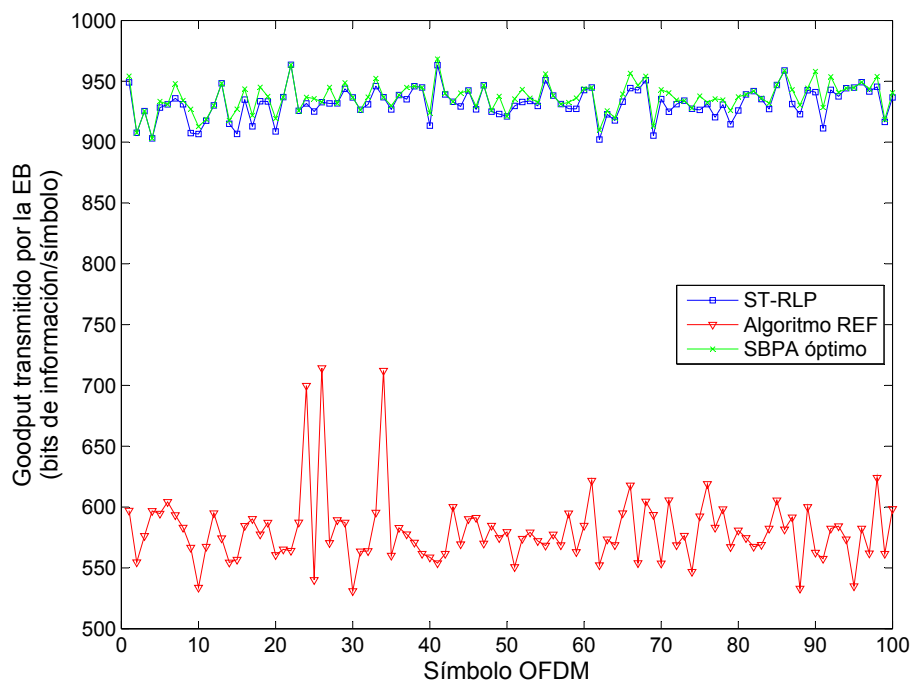


Figura 5.17. Comparación del goodput transmitido por la estación base con canal de variación rápida con ST-RLP, SBPA y el algoritmo REF de [Mohanram y Bhashyam, 2005]

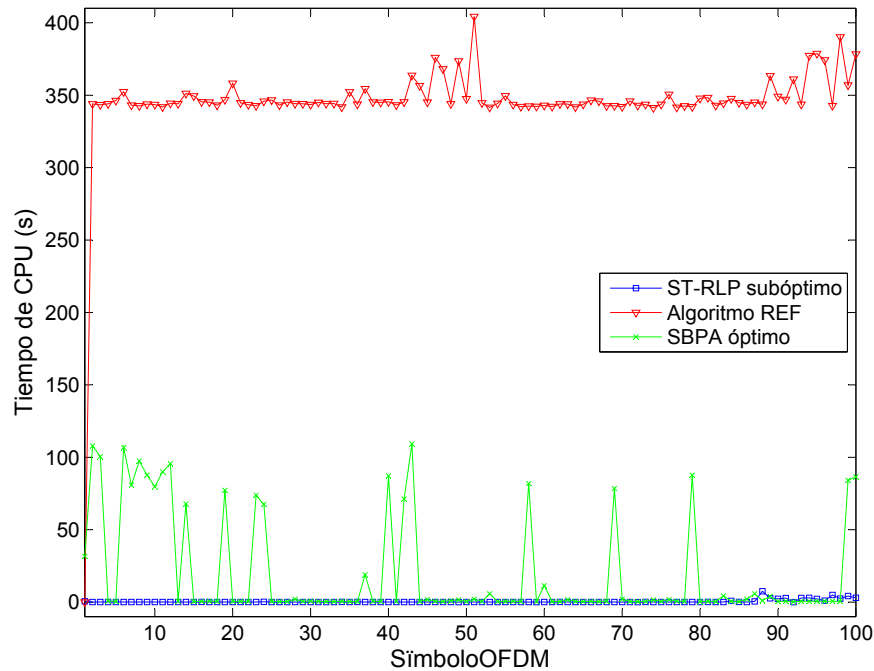


Figura 5.18. Comparación del tiempo de CPU con canal de variación rápida con ST-RLP, SBPA y el algoritmo REF de [Mohanram y Bhashyam, 2005]

de [Mohanram y Bhashyam, 2005]

5.7. Conclusiones

En este capítulo el objetivo es optimizar el *goodput* transmitido por la estación base para comunicaciones de un único salto en sistemas OFDM multiusuario con control de la imparcialidad. Agrupando la subportadoras en subcanales, el problema de asignación de subportadoras y *bit-loading* considerado usualmente es transformado en un problema de asignación de subcanal y *bit-loading*, de manera que simplificando la asignación de ancho de banda se consigue reducir considerablemente la complejidad del problema a resolver.

Para resolver el problema, proponemos el esquema de asignación de recursos propuesto en dos pasos, SBPA, que reduce la complejidad del problema de asignación original y se puede resolver mediante algoritmos de optimización entera estándar, aunque no en tiempo polinomial. Para reducir la complejidad, se propone un planteamiento diferente, considerando que es posible codificar cada uno

de los subcanales independientemente (S-SBPA). Utilizando el S-SBPA, aparte de una considerable disminución de la complejidad, el impacto en el rendimiento del sistema es poco apreciable respecto al óptimo. El problema resultante del S-SBPA es resoluble mediante algoritmos óptimos estándar, como *branch-and-bound*, o subóptimamente mediante relajación lineal; no obstante, esta segunda opción, si bien mucho menos costosa en términos de complejidad que la óptima, tiene el inconveniente de que proporciona soluciones no enteras, que no son aplicables en la práctica. Para evitar la complejidad del *branch-and-bound* y obtener una solución entera, se ha propuesto el algoritmo RLP, que toma como punto de partida la solución proporcionada por la relajación lineal.

Desde el punto de vista de sistemas reales, la codificación independiente plantea problemas de índole práctica, como la implementación de tantas parejas de codificadores/decodificadores como subcanales haya en el sistema. Por ello, hemos propuesto la combinación de las soluciones óptima y subóptima a través del esquema de transmisión ST-RLP, consistente en el envío secuencial en una única trama de los bits del usuario, pero con las asignaciones de subcanal, bit y potencia resultantes de aplicar el RLP. De los resultados obtenidos, observamos lo siguiente:

1. Para variación lenta del canal, el comportamiento del ST-RLP es casi óptimo (figuras 5.3–5.5) y permite el control de la imparcialidad a través del valor de α .
2. Para canal de variación rápida, el comportamiento del ST-RLP es asimismo casi óptimo (figuras 5.12 y 5.14). Esta variación rápida de los coeficientes de canal introduce imparcialidad, pero el algoritmo ST-RLP consigue un cierto control de la imparcialidad del sistema, aunque menos acusado que en el caso de variación lenta del canal.
3. Comparado con un algoritmo de *bit-loading* de altas prestaciones, el algoritmo ST-RLP resulta ser mucho menos exigente desde el punto de vista computacional, a la vez que consigue mucho mejor resultado en términos de *goodput*.

CAPÍTULO 6

CONCLUSIONES Y LINEAS FUTURAS DE INVESTIGACIÓN

Este capítulo presenta las conclusiones generales extraídas de los trabajos contenidos en esta tesis y posibles futuras líneas de investigación a que dan lugar los resultados alcanzados. Alguna de estas líneas ya está siendo explorada como continuación de los estudios aquí expuestos.

6.1. Conclusiones generales

En esta tesis hemos mostrado la importancia de las técnicas de transmisión adaptativa, concretada en la resolución de dos problemas: la minimización de energía en redes inalámbricas y la optimización del *goodput* o eficiencia de transmisión en redes OFDMA. Estos problemas vienen motivados por la existencia de un medio radioeléctrico por un lado cada vez más saturado e interferido, y por otro por la demanda, por parte de los usuarios, de servicios de telecomunicación muy fiables y con alta tasa de transmisión. Estos factores traen consigo la necesidad de diseñar técnicas o esquemas para redes de comunicaciones con asignación dinámica de sus parámetros de transmisión. Estas técnicas se concretan mediante la asignación óptima de recursos, que nos proporciona una herramienta valiosa a la hora de optimizar redes inalámbricas.

La calidad de servicio es un requisito cada día más demandado por los usuarios de servicios de telecomunicación. Esta calidad de servicio se puede formular a través de diferentes parámetros, como la tasa de error de bit y la pérdida de paquetes, o las empleadas en esta tesis: la SER, la tasa de transmisión y el *good-put*. Es importante por tanto considerar estas restricciones en el diseño de sistemas prácticos de comunicaciones, ya que proporcionarán un rendimiento más adecuado a lo esperado por los usuarios. No obstante, como hemos visto, la introducción de estas restricciones en las especificaciones del problema añade una dificultad adicional al diseño del sistema, que ha de ser contemplada para propuestas aplicables a redes reales. Entre los factores que afectan a la solución óptima, destacamos dos. Por un lado, en sistemas prácticos el número de bits por símbolo ha de ser un número entero, e incluso a menudo un valor de un conjunto discreto. Por otro lado, la carga computacional asociada a la resolución de los problemas hace que, en caso de ser alta, los problemas no sean resolubles en tiempo real incluso si hay un nodo dedicado a ello como puede ser la estación base, por lo que la transmisión adaptativa no será efectiva.

El esquema basado en modulación adaptativa eficiente en energía presentado en la tesis, formulado como un problema de minimización de la energía total de transmisión de los nodos y que emplea parámetros de capa física y de capa de enlace, supone un ahorro de energía significativo respecto a TDMA, tanto si el tamaño de los intervalos de transmisión es fijo como si es variable. El problema es abordado garantizando la tasa binaria y la probabilidad de error (SER) requerida por cada sensor de la red, lo que se consigue adaptando el nivel de la modulación. Este esquema resulta particularmente interesante para WSN, dada su sencillez conceptual: la adaptación consiste, esencialmente, en variar el número de símbolos que transmite cada usuario, manteniendo constante el período de símbolo, y adaptar el número de bits por símbolo.

Aunque hemos desarrollado nuestra propuesta para WSN usando MQAM, la utilización de MPSK sería asimismo factible, estando motivada la elección de MQAM por su mayor eficiencia en energía en el entorno estudiado. Hemos analizado la influencia del modelo de pérdidas en el rendimiento del sistema, viendo qué efecto produce la correlación entre los canales, por lo que se refiere al desvanecimiento de pequeña escala; los resultados empíricos muestran que el esquema es robusto frente a esta correlación entre los desvanecimientos Rayleigh, ya que afecta poco al rendimiento del esquema, y estos resultados son refrendados ana-

líticamente. Este esquema además permite utilizar potencias de transmisión del orden de los valores usuales para aplicaciones en redes de sensores, y muy por debajo del límite especificado por el estándar IEEE 802.15.4.

El método propuesto en la tesis para optimizar el *goodput* transmitido en sistemas OFDMA con control de imparcialidad, convierte el problema convencional de asignación de subportadoras y *bit-loading* en un problema de asignación de subcanal y *bit-loading*, de manera que se consigue reducir considerablemente la complejidad del problema a resolver. El esquema de asignación de recursos propuesto, SBPA, consta de dos pasos y se puede resolver mediante algoritmos de optimización entera estándar, aunque no en tiempo polinomial. Para reducir la complejidad, se propone una perspectiva diferente, considerando que es posible codificar cada uno de los subcanales independientemente (S-SBPA). El S-SBPA supone una considerable disminución de la complejidad, a la par que la influencia en el rendimiento del sistema es poco apreciable respecto al óptimo. El problema resultante del S-SBPA es resoluble mediante algoritmos óptimos estándar, como *branch-and-bound*, o subóptimamente mediante relajación lineal; como este último proporciona soluciones no enteras, no aplicables en la práctica, se ha propuesto el algoritmo RLP, que toma como punto de partida la solución de la relajación lineal. Desde el punto de vista de sistemas reales, la codificación independiente plantea problemas de índole práctica. Por ello, hemos propuesto la combinación de las soluciones óptima y subóptima a través del esquema ST-RLP, consistente en el envío secuencial en una única trama de los bits del usuario, con las asignaciones de subcanal, bit y potencia resultantes de aplicar el RLP. De los resultados obtenidos, observamos que:

1. Para variación lenta del canal, el comportamiento del ST-RLP es casi óptimo y permite el control de la imparcialidad a través del valor de α .
2. Para canal de variación rápida, el comportamiento del ST-RLP es asimismo casi óptimo, considerando la propia imparcialidad introducida por la variación rápida de los coeficientes de canal.
3. Comparado con un algoritmo de *bit-loading* de altas prestaciones, el algoritmo ST-RLP resulta ser mucho más eficiente computacionalmente y consigue mucho mejor resultado en términos de *goodput*.

6.2. Líneas futuras de investigación

A lo largo de esta tesis se han ido abriendo distintas líneas de investigación que, tomando como punto de partida la transmisión adaptativa, pueden dar lugar a estudios futuros de interés.

Para el canal ascendente, en sistemas OFDMA los usuarios siempre buscan obtener la máxima tasa de transmisión posible. Una línea que consideramos de interés, en la cual se está trabajando actualmente, es tomar como base el método de optimización del *goodput* y analizar el acceso múltiple en el canal ascendente desde la perspectiva de los mecanismos de subastas, ya que el comportamiento de los usuarios, en términos de imparcialidad, es eminentemente egoísta.

El esquema propuesto para minimización de la energía podría extenderse a redes de salto múltiple, donde los nodos de la red habrán de actuar como repetidores para aquellas comunicaciones que no sean viables con un único enlace. El modelo de energía en este caso cambia sustancialmente, ya que ha de tenerse en cuenta que los nodos consumen energía no sólo al retransmitir la información de otros nodos, sino al recibir tal información. Un tratamiento desde la perspectiva de teoría de juegos puede resultar interesante en esta dirección.

El modelo de canal considerado para el esquema de optimización de energía no contempla la movilidad, asumiendo que los sensores son estáticos o restringidos a un movimiento muy lento. La extensión del esquema eficiente en energía a entornos con movilidad puede resultar adecuado con la inclusión del modelo de canal apropiado, para aplicaciones muy interesantes como pueden ser los sistemas de vigilancia, siempre con velocidades moderadas.

ANEXOS

ANEXO A

EXPRESIONES MATRICIALES DEL CÁLCULO DIFERENCIAL

Dependiendo del contenido del apéndice será conveniente o no incluir un párrafo introductorio como en el caso de los capítulos.

Definición A.1 (Matriz hessiana). *Dada la función de n variables reales $f(\mathbf{x}) = f(x_1, \dots, x_n)$, si existen todas las derivadas parciales segundas de f , se define la matriz hessiana de f respecto a \mathbf{x} como $H_{\mathbf{x}}f$. El elemento (i, j) de la matriz, $H_{\mathbf{x}}f_{i,j}$, viene dado por $\partial^2 f(\mathbf{x}) / \partial x_i \partial x_j$, de manera que la expresión de la matriz hessiana tiene la expresión*

$$H_{\mathbf{x}}f = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

Definición A.2 (Matriz jacobiana). *Dada la función de n variables reales $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$, si existen todas las derivadas parciales de primer orden de g , la matriz*

jacobiana de g $Jg(\mathbf{x})$ se define como

$$Jg(\mathbf{x}) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_n}{\partial x_1} & \cdots & \frac{\partial g_n}{\partial x_m} \end{pmatrix}$$

Definición A.3 (Forma cuadrática). Sea la matriz cuadrada \mathbf{A} de dimensión n y las variables x_1, x_2, \dots, x_n , tal que $\mathbf{x} = (x_1, x_2, \dots, x_n)$. $Q(\mathbf{x})$ es la forma cuadrática asociada a la matriz \mathbf{A} y se determina como $Q(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$.

Definición A.4 (Matriz definida positiva y negativa). Se dice que la matriz cuadrada \mathbf{A} de dimensión n es definida positiva si $\mathbf{x}^T \mathbf{A} \mathbf{x} > 0$, $\forall \mathbf{z} \in \mathbb{R}^n, \mathbf{z} \neq \mathbf{0}$. De manera análoga, \mathbf{A} es definida negativa si $\mathbf{x}^T \mathbf{A} \mathbf{x} < 0$, $\forall \mathbf{z} \in \mathbb{R}^n, \mathbf{z} \neq \mathbf{0}$.

ANEXO B

LISTA DE PUBLICACIONES

En este último anexo se enumeran, desglosadas por capítulos, las publicaciones derivadas de las diferentes contribuciones que constituyen esta tesis.

Capítulo 2:

- J. J. Escudero-Garzás, A. García-Armada, *Emerging telecommunications technologies: cognitive radio*, Handbook of Research on Telecommunications Planning and Management for Business, Chapter L, pp.788–803, Lee, I., 2009, Information Science Publishing.

Capítulo 4:

- J. J. Escudero-Garzás, C. Bousoño-Calzón, A. García-Armada, *An energy-efficient adaptive modulation suitable for wireless sensor networks with SER and throughput constraints*, EURASIP Journal on Wireless Communications and Networking, vol.2007, no.1, pp.1–7, 2007.
- J. J. Escudero-Garzás, A. García-Armada, *Transmisión eficiente con QoS para redes de sensores inalámbricas*, XXI Simposium Nacional de la Unión Científica Internacional de Radio URSI 2006. Oviedo. Septiembre 2006.
- J. J. Escudero-Garzás, A. García-Armada, *QoS-Aware Scheduling for Wireless Sensor Networks*, 15th IST Mobile & Wireless Communications Summit

2006, Mykonos. Junio 2006.

- J. J. Escudero-Garzás, A. García-Armada, *Analysis of the performance of energy-efficient scheduling for wireless sensor networks using SNR gap*, International ITG/IEEE Workshop on Smart Antennas, WSA 2006. Ulm. Marzo 2006.

Capítulo 5:

- J. J. Escudero-Garzás, B. Devillers, A. García-Armada, *Fairness-Adaptive Goodput Optimization with Subchannel, Bit and Power Allocation in Multi-user OFDM systems*, IEEE Transactions on Wireless Communications, en proceso de revisión.
- J. J. Escudero-Garzás, B. Devillers, L. Vanderdorpe, Ana García-Armada, *Subchannel, bit and power allocation in multiuser OFDM systems for goodput optimization with fairness*, 12th International Symposium on Wireless Personal Multimedia Communications 2009. Sendai. Septiembre 2009
- J. Joaquin Escudero-Garzás, Bertrand Devillers, Ana García-Armada, *Asignación de recursos radio para optimización del goodput en redes OFDM multiusuario*, XXIV Simposium Nacional de la Unión Científica Internacional de Radio URSI 2009. Santander. Septiembre 2009.

Bibliografía

- [802, 2004] “IEEE Standard for Local and Metropolitan Area Networks Part 16: Air Interface for Fixed Broadband Wireless Access Systems”, *IEEE Std 802.16-2004 (Revision of IEEE Std 802.16-2001)*, págs. 1–857, 2004.
- [802, 2006a] “IEEE Standard for Information technology- Telecommunications and information exchange between systems- Local and metropolitan area networks- Specific requirements Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (WPANs)”, *IEEE Std 802.15.4-2006 (Revision of IEEE Std 802.15.4-2003)*, págs. 1–305, 2006a.
- [802, 2006b] “IEEE Standard for Local and metropolitan area networks Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems Amendment 2: Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands and Corrigendum 1”, *IEEE Std 802.16e-2005 and IEEE Std 802.16-2004/Cor 1-2005 (Amendment and Corrigendum to IEEE Std 802.16-2004)*, págs. 1–822, 2006b.
- [IEE, 2009] “IEEE Draft Standard for Information technology- Telecommunications and information exchange between systems-Local and metropolitan area networks-Specific requirements-Part 22.1: Standard to Enhance Harmful Interference Protection for Low Power Licensed Devices Operating in the TV Broadcast Bands”, *IEEE Unapproved Draft Std P802.22.1/D6, Feb 2009*, 2009.
- [Acharya y Yates, 2007] J. ACHARYA y R. YATES, “A Framework for Dynamic Spectrum Sharing Between Cognitive Radios”, en *IEEE International Conference on Communications, 2007. ICC'07*, págs. 5166–5171, 2007.

- [Akyildiz et al., 2006] I. AKYILDIZ, W. LEE, M. VURAN y S. MOHANTY, “NeXt generation/dynamic spectrum access/cognitive radio wireless networks: a survey”, *Computer Networks*, 50(13):2127–2159, 2006.
- [Akyildiz et al., 2002] I. AKYILDIZ, W. SU, Y. SANKARASUBRAMANIAM y E. CAYIRCI, “Wireless sensor networks: a survey”, *Computer Networks*, 38(4):393–422, 2002.
- [Alouini y Goldsmith, 1999] M.-S. ALOUINI y A. GOLDSMITH, “Capacity of Rayleigh fading channels under different adaptive transmission and diversity-combining techniques”, *IEEE Transactions on Vehicular Technology*, 48(4):1165–1181, Jul 1999.
- [Althaus et al., 2003] E. ALTHAUS, G. CALINESCU, I. MANDOIU, S. PRASAD, N. TCHERVENSKI y A. ZELIKOVSKY, “Power efficient range assignment in ad-hoc wireless networks”, *2003 IEEE Wireless Communications and Networking, 2003. WCNC 2003*, 3, 2003.
- [Atakan y Akan, 2007] B. ATAKAN y O. AKAN, “Biologically-inspired spectrum sharing in cognitive radio networks”, en *IEEE Wireless Communications and Networking Conference, 2007. WCNC 2007*, págs. 43–48, 2007.
- [Baamrani et al., 2006] K. BAAMRANI, A. OUAHMAN, V. JIMÉNEZ, A. ARMADA y S. ALLAKI, “Subcarrier and power allocation for the downlink of multiuser OFDM transmission”, *Wireless Personal Communications*, 39(4):457–465, 2006.
- [Bansal et al., 2007] G. BANSAL, M. HOSSAIN y V. BHARGAVA, “Adaptive Power Loading for OFDM-Based Cognitive Radio Systems”, en *IEEE International Conference on Communications, 2007. ICC’07*, págs. 5137–5142, 2007.
- [Bater et al., 2007] J. BATER, H. TAN, K. BROWN y L. DOYLE, “Modelling Interference Temperature Constraints for Spectrum Access in Cognitive Radio Networks”, en *IEEE International Conference on Communications, 2007. ICC’07*, págs. 6493–6498, 2007.
- [Bazaraa et al., 2006] M. BAZARAA, H. SHERALI y C. SHETTY, *Nonlinear programming: theory and algorithms*, Wiley-Interscience, 2006.

- [Bertsekas, 1999] D. BERTSEKAS, *Nonlinear programming*, Athena Scientific, Belmont, MA, 1999.
- [Bertsimas y Weismantel, 2005] D. BERTSIMAS y R. WEISMANTEL, *Optimization over integers*, Dynamic Ideas Belmont, MA, 2005.
- [Bhushan et al., 2006] N. BHUSHAN, C. LOTT, P. BLACK, R. ATTAR, Y.-C. JOU, M. FAN, D. GHOSH y J. AU, “CDMA2000 1xEV-DO revision a: a physical layer and MAC layer overview”, *IEEE Communications Magazine*, 44(2):37–49, Feb. 2006.
- [Bingham, 1990] J. BINGHAM, “Multicarrier modulation for data transmission: an idea whose time has come”, *IEEE Communications Magazine*, 28(5):5–14, May 1990.
- [Boyd y Vandenberghe, 2004] S. BOYD y L. VANDENBERGHE, *Convex optimization*, Cambridge University Press, 2004.
- [Brah et al., 2009] F. BRAH, J. LOUVEAUX y L. VANDENDORPE, “CDIT-based constrained resource allocation for mobile WiMAX systems”, *EURASIP Journal on Wireless Communications and Networking*, 2009, 2009.
- [Budiarjo et al., 2006] I. BUDIARJO, H. NIKOOKAR y L. LIGTHART, “Combined spectrum pooling and adaptive bit loading for Cognitive Radio OFDM based system”, en *2006 Symposium on Communications and Vehicular Technology*, págs. 73–76, 2006.
- [Caire y Kumar, 2007] G. CAIRE y K. KUMAR, “Information theoretic foundations of adaptive coded modulation”, *Proceedings of the IEEE*, 95(12):2274–2298, 2007.
- [Campello, 1999] J. CAMPELLO, “Practical bit loading for DMT”, en *1999 IEEE International Conference on Communications, 1999. ICC'99*, volumen 2, 1999.
- [Chang, 1966] R. CHANG, “Synthesis of band-limited orthogonal signals for multichannel data transmission”, *Bell Syst. Tech. J.*, 45(10):1775–1796, 1966.
- [Changho y Jeonghoon, 2008] S. CHANGHO y M. JEONGHOON, “Resource Allocation for Multicast Services in Multicarrier Wireless Communications”, *IEEE Transactions on Wireless Communications*, 2008.

- [Chiang y Bell, 2004] M. CHIANG y J. BELL, “Balancing supply and demand of bandwidth in wireless cellular networks: utility maximization over powers and rates”, en *IEEE INFOCOM*, volumen 4, págs. 2800–2811, 2004.
- [Chiang et al., 2007] R. CHIANG, G. ROWE y K. SOWERBY, “A Quantitative Analysis of Spectral Occupancy Measurements for Cognitive Radio”, en *IEEE 65th Vehicular Technology Conference, 2007. VTC2007-Spring*, págs. 3016–3020, 2007.
- [Choi y Bahk, 2008] Y. CHOI y S. BAHK, “Multichannel wireless scheduling under limited terminal capability”, *IEEE Transactions on Wireless Communications*, 7(2):611–617, 2008.
- [Chou et al., 2007] C. CHOU, H. KIM y K. SHIN, “What and how much to gain by spectrum agility”, *IEEE Journal on Selected Areas in Communications*, 25(3):576–588, 2007.
- [Chow, 1993] P. CHOW, *Bandwidth optimize digital transmission techniques for spectrally shaped channels*, Tesis Doctoral, Stanford University, 1993.
- [Chung y Goldsmith, 2001] S. CHUNG y A. GOLDSMITH, “Degrees of freedom in adaptive modulation: a unified view”, *IEEE Transactions on Communications*, 49(9):1561–1571, 2001.
- [Cioffi et al., 1995] J. CIOFFI, G. DUDEVOIR, M. EYUBOGLU y G. FORNEY JR, “MMSE decision-feedback equalizers and coding. II. Coding results”, *IEEE Transactions on Communications*, 43(10):2595–2604, 1995.
- [Clemens y Rose, 2005] N. CLEMENS y C. ROSE, “Intelligent power allocation strategies in an unlicensed spectrum”, en *IEEE First International Symposium on New Frontiers in Dynamic Spectrum Access Networks, 2005. DySPAN 2005*, págs. 37–42, 2005.
- [Conti et al., 2007] A. CONTI, M. Z. WIN y M. CHIANI, “Slow Adaptive M - QAM With Diversity in Fast Fading and Shadowing”, *IEEE Transactions on Communications*, 55(5):895–905, May 2007.
- [Cover y Thomas, 1991] T. COVER y J. THOMAS, *Elements of information theory*, Wiley, 1991.

- [Cui et al., 2004] S. CUI, A. GOLDSMITH y A. BAHAI, “Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks”, *IEEE Journal on Selected Areas in Communications*, 22(6):1089–1098, 2004.
- [Cui et al., 2005a] S. CUI, A. GOLDSMITH y A. BAHAI, “Energy-constrained modulation optimization”, *IEEE Transactions on Wireless Communications*, 4(5):2349–2360, 2005a.
- [Cui et al., 2005b] S. CUI, R. MADAN, A. GOLDSMITH y S. LALL, “Joint routing, MAC, and link layer optimization in sensor networks with energy constraints”, en *2005 IEEE International Conference on Communications, 2005. ICC’05*, volumen 2, págs. 725–729, May 2005b.
- [Dardari, 2004] D. DARDARI, “Ordered subcarrier selection algorithm for OFDM-based high-speed WLANs”, *IEEE Transactions on Wireless Communications*, 3(5):1452–1458, Sep. 2004.
- [Devillers et al., 2008] B. DEVILLERS, J. LOUVEAUX y L. VANDENDORPE, “Bit and power allocation for goodput optimization in coded parallel subchannels with ARQ”, *IEEE Transactions on Signal Processing*, 56(8 Part 1):3652–3661, 2008.
- [Einhaus et al., 2008] M. EINHAUS, B. WOLZ y B. WALKE, “The Influence of Subchannel Diversity on the Performance of OFDMA Systems Based on IEEE 802.16”, en *IEEE 4th International Conference on Circuits and Systems for Communications, 2008. ICCSC 2008*, págs. 20–24, May 2008.
- [Ergen et al., 2003] M. ERGEN, S. COLERI y P. VARAIYA, “QoS aware adaptive resource allocation techniques for fair scheduling in OFDMA based broadband wireless access systems”, *IEEE Transactions on Broadcasting*, 49(4):362–370, Dec. 2003.
- [Eriksson y Ottosson, 2007] T. ERIKSSON y T. OTTOSSON, “Compression of feedback for adaptive transmission and scheduling”, *Proceedings of the IEEE*, 95(12):2314–2321, 2007.
- [Ermolova y Makarevitch, 2007] N. Y. ERMOLOVA y B. MAKAREVITCH, “Low complexity adaptive power and subcarrier allocation for OFDMA”, *IEEE Transactions on Wireless Communications*, 6(2):433–437, Feb. 2007.

- [Falahati et al., 2004] S. FALAHATI, A. SVENSSON, T. EKMAN, M. STERNAD et al., “Adaptive modulation systems for predicted wireless channels”, *IEEE Transactions on Communications*, 52(2):307–316, 2004.
- [FCC, 2005a] FCC, “Establishment of an Interference Temperature Metric to Quantify and Manage Interference and to Expand Available Unlicensed Operation in Certain Fixed, Mobile and Satellite Frequency Bands”, *ET Docket No. 03-237, Notice of Inquiry and Notice of Proposed Rulemaking, FCC 03-289*, págs. 1–42, March 2005a.
- [FCC, 2005b] FCC, “Facilitating opportunities for flexible, efficient, and reliable spectrum use employing cognitive radio technologies”, *ET Docket No. 03-108, Report and Order, FCC 05-57*, págs. 1–42, March 2005b.
- [Feiten et al., 2008] A. FEITEN, R. MATHAR y M. REYER, “Rate and Power Allocation for Multiuser OFDM: An Effective Heuristic Verified by Branch-and-Bound”, *IEEE Transactions on Wireless Communications*, 7(1):60–64, 2008.
- [Fu et al., 2008] L. FU, S. LIEW y J. HUANG, “Joint Power Control and Link Scheduling in Wireless Networks for Throughput Optimization”, en *IEEE International Conference on Communications, 2008. ICC’08*, págs. 3066–3072, 2008.
- [Furuskar et al., 1999] A. FURUSKAR, S. MAZUR, F. MULLER y H. OLOFSSON, “EDGE: Enhanced data rates for GSM and TDMA/136 evolution”, *IEEE Personal Communications*, 6(3):56–66, 1999.
- [Ganesan y Ye, 2007] G. GANESAN y L. YE, “Cooperative Spectrum Sensing in Cognitive Radio, Part II: Multiuser Networks”, *IEEE Transactions on Wireless Communications*, 6(6):2214–2222, 2007.
- [Gao et al., 2007] S. GAO, L. QIAN, D. VAMAN y Q. QU, “Energy Efficient Adaptive Modulation in Wireless Cognitive Radio Sensor Networks”, en *IEEE International Conference on Communications, 2007. ICC’07*, págs. 3980–3986, 2007.
- [Garcia-Armada, 2006] A. GARCIA-ARMADA, “SNR gap approximation for M-PSK-based bit loading”, *IEEE Transactions on Wireless Communications*, 5(1):57–60, 2006.

- [Gault et al., 2007] S. GAULT, W. HACHEM y P. CIBLAT, “Performance analysis of an OFDMA transmission system in a multicell environment”, *IEEE Transactions on Communications*, 55(4):740–751, 2007.
- [Goldsmith, 2005] A. GOLDSMITH, *Wireless communications*, Cambridge University Press, 2005.
- [Guerrero-Casas, 1994] F. GUERRERO-CASAS, *Curso de optimización*, Ariel Economía, 1994.
- [Haapola et al., 2005] J. HAAPOLA, Z. SHELBY, C. POMALAZA-RAEZ y P. MAHONEN, “Cross-layer energy analysis of multihop wireless sensor networks”, en *Second European Workshop on Wireless Sensor Networks, 2005*, págs. 33–44, 2005.
- [Han et al., 2005] Z. HAN, Z. JI y K. LIU, “Fair Multiuser Channel Allocation for OFDMA Networks Using Nash Bargaining Solutions and Coalitions”, *IEEE Transactions on Communications*, 53(8):1366–1376, Aug. 2005.
- [Hanzo et al., 2000] L. HANZO, W. WEBB y T. KELLER, *Single-and Multi-carrier Quadrature Amplitude Modulation: Principles and Applications for Personal Communications, WLANs and Broadcasting*, John Wiley & Sons Inc, 2000.
- [Hayes, 1968] J. HAYES, “Adaptive Feedback Communications”, *IEEE Transactions on Communication Technology*, 16(1):29–34, Feb. 1968.
- [Haykin, 2005] S. HAYKIN, “Cognitive radio: brain-empowered wireless communications”, *IEEE Journal on Selected Areas in Communications*, 23(2):201–220, 2005.
- [Hoo et al., 1998] L. HOO, J. TELLADO y J. CIOFFI, “Dual QoS loading algorithms for DMT systems offering CBR and VBR services”, en *IEEE Global Telecommunications Conference, 1998. GLOBECOM'98*, volumen 1, págs. 25–30, 1998.
- [Hosein, 2006] P. HOSEIN, “On the optimal allocation of downlink resources in OFDM-based wireless networks”, *Lecture Notes in Computer Science*, 3970:202, 2006.

- [Hosein y Lee, 2007] P. HOSEIN y J. W. LEE, “Quantifying the Diversity of OFDMA Subchannels”, en *IEEE Wireless Communications and Networking Conference, 2007.WCNC 2007*, págs. 1996–2000, Mar. 2007.
- [Hughes-Hartogs, 1989] D. HUGHES-HARTOGS, *Ensemble modem structure for imperfect transmission media*, US Patent 4,833,706, May 1989.
- [Jang y Lee, 2003] J. JANG y K. LEE, “Transmit power adaptation for multi-user OFDM systems”, *IEEE Journal on Selected Areas in Communications*, 21(2):171–178, 2003.
- [Janssen et al., 1996] G. JANSSEN, P. STIGTER y R. PRASAD, “Wideband indoor channel measurements and BER analysis of frequencyselective multipath channels at 2.4, 4.75, and 11.5 GHz”, *IEEE Transactions on Communications*, 44(10):1272–1288, 1996.
- [Ji y Liu, 2007] Z. JI y K. LIU, “Dynamic Spectrum Sharing: A Game Theoretical Overview”, *IEEE Communications Magazine*, págs. 88–94, 2007.
- [Johansson, 2004] M. JOHANSSON, “Diversity-enhanced equal access-considerable throughput gains with 1-bit feedback”, en *2004 IEEE 5th Workshop on Signal Processing Advances in Wireless Communications*, págs. 6–10, 2004.
- [Kamath y Goeckel, 2004] K. KAMATH y D. GOECKEL, “Adaptive-modulation schemes for minimum outage probability in wireless systems”, *IEEE Transactions on Communications*, 52(10):1632–1635, Oct. 2004.
- [Kamio et al., 1995] Y. KAMIO, S. SAMPEI, H. SASAOKA y N. MORINAGA, “Performance of modulation-level-controlled adaptive-modulation under limited transmission delay time for land mobile communications”, en *1995 IEEE 45th Vehicular Technology Conference*, volumen 1, 1995.
- [Kaneko et al., 2008] M. KANEKO, P. POPOVSKI y J. DAHL, “Proportional fairness in multi-carrier system with multi-slot frames: upper bound and user multiplexing algorithms”, *IEEE Transactions on Wireless Communications*, 7(1):22–26, Jan. 2008.
- [Karush, 1939] W. KARUSH, *Minima of functions of several variables with inequalities as side conditions*, Tesis Doctoral, University of Chicago, Department of Mathematics, 1939.

- [Karvonen y Pomalaza-Raez, 2006] H. KARVONEN y C. POMALAZA-RAEZ, “A cross layer design of coding and awake/sleep periods in WSNS”, en *2006 IEEE 17th International Symposium on Personal, Indoor and Mobile Radio Communications*, págs. 1–5, 2006.
- [Kelly et al., 1998] F. KELLY, A. MAULLOO y D. TAN, “Rate control for communication networks: shadow prices, proportional fairness and stability”, *The Journal of the Operational Research Society*, 49(3):237–252, 1998.
- [Kim et al., 2006] I. KIM, I. PARK y Y. LEE, “Use of linear programming for dynamic subcarrier and bit allocation in multiuser OFDM”, *IEEE Transactions on Vehicular Technology*, 55(4):1195–1207, 2006.
- [Kivanc et al., 2003] D. KIVANC, G. LI y H. LIU, “Computationally efficient bandwidth allocation and power control for OFDMA”, *IEEE Transactions on Wireless Communications*, 2(6):1150–1158, 2003.
- [Knopp y Humblet, 1995] R. KNOPP y P. HUMBLET, “Information capacity and power control in single-cell multiuser communications”, en *IEEE International Conference on Communications, 1995. ICC'95*, volumen 1, págs. 331–335, 1995.
- [Ko et al., 2007] Y. KO, H. YANG, S. EOM y M. ALOUINI, “Adaptive modulation with diversity combining based on output-threshold MRC”, *IEEE Transactions on Wireless Communications*, 6(10):3728–3737, 2007.
- [Kuhn, 1955] H. KUHN, “The Hungarian method for the assignment problem”, *Naval Research Logistic Quaterly*, 2:83–97, 1955.
- [Kuhn y Tucker, 1951] H. KUHN y A. TUCKER, “Non-linear Programming”, en *Second Berkeley Symposium on Mathematical Statistics and Probability*, págs. 481–492, 1951.
- [Kumar et al., 2006] S. KUMAR, V. RAGHAVAN y J. DENG, “Medium Access Control protocols for ad hoc wireless networks: A survey”, *Ad Hoc Networks*, 4(3):326–358, 2006.
- [Kwon et al., 2006] H. KWON, T. KIM, S. CHOI y B. LEE, “A cross-layer strategy for energy-efficient reliable delivery in wireless sensor networks”, *IEEE Transactions on Wireless Communications*, 5(12):3689–3699, 2006.

- [Kwon et al., 2005] T. KWON, H. LEE, S. CHOI, J. KIM, D. CHO, S. CHO, S. YUN, W. PARK y K. KIM, “Design and implementation of a simulator based on a cross-layer protocol between MAC and PHY layers in a Wi-Bro Compatible. IEEE 802.16e OFDMA system”, *IEEE Communications Magazine*, 43(12):136–146, 2005.
- [Lau et al., 2008] V. LAU, W. K. NG y D. HUI, “Asymptotic tradeoff between cross-layer goodput gain and outage diversity in OFDMA systems with slow fading and delayed CSIT”, *IEEE Transactions on Wireless Communications*, 7(7):2732–2739, 2008.
- [Lee y Leung, 2006] K. LEE y V. LEUNG, “Fair allocation of subcarrier and power in an OFDMA wireless mesh network”, *IEEE Journal on Selected Areas in Communications*, 24(11):2051–2060, 2006.
- [Lengoumbi et al., 2007] C. LENGOUMBI, P. GODLEWSKI y P. MARTINS, “Comparison of Different Subchannelization Modes for OFDMA”, en *IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications, 2007. PIMRC 2007*, págs. 1–5, Sep. 2007.
- [Li y Hui, 2005] G. LI y L. HUI, *OFDM-based broadband wireless networks: design and optimization*, Wiley-Blackwell, 2005.
- [Li y Liu, 2005] G. LI y H. LIU, “On the optimality of the OFDMA network”, *IEEE Communications Letters*, 9(5):438–440, 2005.
- [Li y Liu, 2006] G. LI y H. LIU, “Downlink radio resource allocation for multi-cell OFDMA system”, *IEEE Transactions on Wireless Communications*, 5(12):3451–3459, 2006.
- [Li y Liu, 2007] H. LI y H. LIU, “An analysis of uplink OFDMA optimality”, *IEEE Transactions on Wireless Communications*, 6(8):2972–2983, 2007.
- [Liu et al., 2008] J. LIU, Y. HOU, Y. SHI y H. SHERALI, “Cross-Layer Optimization for MIMO-Based Wireless Ad Hoc Networks: Routing, Power Allocation, and Bandwidth Allocation”, *IEEE Journal on Selected Areas in Communications*, 26(6):913–926, 2008.

- [Liu et al., 2004] Q. LIU, S. ZHOU y G. GIANNAKIS, “Cross-layer combining of adaptive modulation and coding with truncated ARQ over wireless links”, *IEEE Transactions on Wireless Communications*, 3(5):1746–1755, 2004.
- [Luenberger, 2003] D. LUENBERGER, *Linear and nonlinear programming*, Springer, 2003.
- [Luo y Yu, 2006] Z. LUO y W. YU, “An introduction to convex optimization for communications and signal processing”, *IEEE Journal on Selected Areas in Communications*, 24(8):1426–1438, 2006.
- [Luo y Zhang, 2008] Z. LUO y S. ZHANG, “Dynamic spectrum management: Complexity and duality”, *IEEE Journal on Selected Topics in Signal Processing*, 2(1):57–73, 2008.
- [Madan et al., 2006] R. MADAN, S. CUI, S. LALL y N. GOLDSMITH, “Cross-layer design for lifetime maximization in interference-limited wireless sensor networks”, *IEEE Transactions on Wireless Communications*, 5(11):3142–3152, 2006.
- [Malan et al., 2004] D. MALAN, T. FULFORD-JONES, M. WELSH y S. MOULTON, “Codeblue: An ad hoc sensor network infrastructure for emergency medical care”, en *International Workshop on Wearable and Implantable Body Sensor Networks*, volumen 5, 2004.
- [Manzoli y Merani, 2006] U. MANZOLI y M. MERANI, “Goodput and delay analysis of a radio link control protocol operating over a multicarrier DS-CDMA architecture”, *IEEE Transactions on Wireless Communications*, 5(6):1313–1321, 2006.
- [Mao y Wang, 2008] Z. MAO y X. WANG, “Efficient optimal and suboptimal radio resource allocation in OFDMA system”, *IEEE Transactions on Wireless Communications*, 7(2):440–445, 2008.
- [Mehrotra, 1992] S. MEHROTRA, “On the implementation of a primal-dual interior point method”, *SIAM Journal on Optimization*, 2:575–601, 1992.
- [Mitola III y Maguire Jr, 1999] J. MITOLA III y G. MAGUIRE JR, “Cognitive radio: making software radios more personal”, *IEEE Personal Communications*, 6(4):13–18, 1999.

- [Mitran et al., 2008] P. MITRAN, L. LE, C. ROSENBERG y A. GIRARD, “Resource Allocation for Downlink Spectrum Sharing in Cognitive Radio Networks”, en *IEEE 68th Vehicular Technology Conference, 2008. VTC 2008-Fall*, págs. 1–5, Sep. 2008.
- [Mohanram y Bhashyam, 2005] C. MOHANRAM y S. BHASHYAM, “A sub-optimal joint subcarrier and power allocation algorithm for multiuser OFDM”, *IEEE Communications Letters*, 9(8):685–687, 2005.
- [Nam et al., 2008] H. NAM, Y. KO y M. ALOUINI, “Performance analysis of joint switched diversity and adaptive modulation”, *IEEE Transactions on Wireless Communications*, 7(10):3780–3790, 2008.
- [Nanda et al., 2000] S. NANDA, K. BALACHANDRAN y S. KUMAR, “Adaptation techniques in wireless packet data services”, *IEEE Communications Magazine*, 38(1):54–64, 2000.
- [Park et al., 2004] I.-S. PARK, Y.-H. JUNG, I. MIN y Y. LEE, “Dynamic sub-channel and bit allocation in multiuser MIMO/OFDMA systems”, en *IEEE 59th Vehicular Technology Conference, 2004. VTC 2004-Spring*, volumen 2, págs. 884–888, May 2004.
- [Pfletschinger et al., 2007] S. PFLETSCHINGER, A. PIATYSZEK y S. STIGLMAYR, “Frequency-Selective Link Adaptation using Duo-Binary Turbo Codes in OFDM Systems”, en *IST 16th Mobile and Wireless Communications Summit, 2007*, págs. 1–5, Jul. 2007.
- [Pollin et al., 2008] S. POLLIN, R. MANGHARAM, B. BOUGARD, L. VAN DER PERRE, I. MOERMAN, R. RAJKUMAR, F. CATTHOOR y L. IMEC, “MEE-RA: Cross-Layer Methodology for Energy Efficient Resource Allocation in Wireless Networks”, *IEEE Transactions on Wireless Communications*, 7(1):98–109, 2008.
- [Prabhakar et al., 2001] B. PRABHAKAR, E. UYSAL BIYIKOGLU y A. EL GAMAL, “Energy-efficient transmission over a wireless link via lazy packet scheduling”, en *IEEE Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies. INFOCOM 2001*, volumen 1, págs. 386–394, 2001.

- [Qian et al., 2007] L. QIAN, X. LI, J. ATTIA y Z. GAJIC, “Power Control for Cognitive Radio Ad Hoc Networks”, en *IEEE 15th Workshop on Local & Metropolitan Area Networks, 2007. LANMAN 2007*, págs. 7–12, 2007.
- [Qin et al., 2007] C. QIN, G. YU, Z. ZHANG, H. JIA y A. HUANG, “Power Reservation-Based Admission Control Scheme for IEEE 802.16e OFDMA Systems”, en *IEEE Wireless Communications and Networking Conference, 2007.WCNC 2007*, págs. 1831–1835, Mar. 2007.
- [Qin y Leung, 2007] T. QIN y C. LEUNG, “A cost minimization algorithm for a multiuser OFDM cognitive radio system”, en *IEEE Pacific Rim Conference on Communications, Computers and Signal Processing, 2007. PacRim 2007*, págs. 518–521, 2007.
- [Qu et al., 2008] Q. QU, L. MILSTEIN y D. VAMAN, “Cognitive Radio Based Multi-User Resource Allocation in Mobile Ad Hoc Networks Using Multi-Carrier CDMA Modulation”, *IEEE Journal on Selected Areas in Communications*, 26(1):70–82, 2008.
- [Rappaport, 2002] T. RAPPAPORT, *Wireless communications*, Prentice Hall PTR New Jersey, 2002.
- [Rhee y Cioffi, 2000] W. RHEE y J. CIOFFI, “Increase in capacity of multiuser OFDM system using dynamic subchannel allocation”, en *IEEE 51st Vehicular Technology Conference Proceedings, 2000. VTC 2000-Spring Tokyo*, volumen 2, págs. 1085–1089, 2000.
- [Sahin et al., 2007] M. SAHIN, S. AHMED y H. ARSLAN, “The Roles of Ultra Wideband in Cognitive Networks”, en *IEEE International Conference on Ultra-Wideband, 2007. ICUWB 2007*, págs. 247–252, 2007.
- [Schurgers et al., 2001] C. SCHURGERS, O. ABERTHORNE y M. SRIVASTAVA, “Modulation scaling for energy aware communication systems”, en *2001 International Symposium on Low Power Electronics and Design*, págs. 96–99, ACM New York, NY, USA, 2001.
- [Seong et al., 2006] K. SEONG, M. MOHSENI y J. CIOFFI, “Optimal Resource Allocation for OFDMA Downlink Systems”, en *2006 IEEE International Symposium on Information Theory*, págs. 1394–1398, Jul. 2006.

- [Setoodeh y Haykin, 2009] P. SETOODEH y S. HAYKIN, “Robust transmit power control for cognitive radio”, *Proceedings of the IEEE*, 97(5):915–939, 2009.
- [Shannon, 1948] C. SHANNON, “A mathematical theory of communication”, *Bell Syst. Tech. J.*, 27:379–423, 1948.
- [Sharp et al., 2005] C. SHARP, S. SCHAFFERT, A. WOO, N. SASTRY, C. KARLOF, S. SASTRY y D. CULLER, “Design and implementation of a sensor network system for vehicle tracking and autonomous interception”, en *Second European Workshop on Wireless Sensor Networks, 2005*, págs. 93–107, 2005.
- [Shelby et al., 2005] Z. SHELBY, C. POMALAZA-RAEZ, H. KARVONEN y J. HAA-POLA, “Energy optimization in multihop wireless embedded and sensor networks”, *International Journal of Wireless Information Networks*, 12(1):11–21, 2005.
- [Shen et al., 2006] J. SHEN, X. YUAN, L. SHEN, D. YAO y H. LIU, “An energy-efficient node for wireless sensor networks using software radio”, en *2006 IEEE International Conference on Electro/information Technology*, págs. 245–249, 2006.
- [Shen et al., 2005] Z. SHEN, J. ANDREWS y B. EVANS, “Adaptive resource allocation in multiuser OFDM systems with proportional rate constraints”, *IEEE Transactions on Wireless Communications*, 4(6):2726–2737, 2005.
- [Shih et al., 2001] E. SHIH, S. CHO, N. ICKES, R. MIN, A. SINHA, A. WANG y A. CHANDRAKASAN, “Physical layer driven protocol and algorithm design for energy-efficient wireless sensor networks”, en *7th Annual International Conference on Mobile computing and Networking*, págs. 272–287, 2001.
- [Sichitiu, 2004] M. SICHITIU, “Cross-layer scheduling for power efficiency in wireless sensor networks”, en *Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies. INFOCOM 2004*, volumen 3, 2004.
- [Simeone et al., 2007] O. SIMEONE, J. GAMBINI, Y. BAR-NESS y U. SPAGNOLINI, “Cooperation and Cognitive Radio”, en *IEEE International Conference on Communications, 2007. ICC’07*, págs. 6511–6515, 2007.

- [Song y Li, 2005a] G. SONG y Y. LI, “Cross-layer optimization for OFDM wireless networks-part I: theoretical framework”, *IEEE Transactions on Wireless Communications*, 4(2):614–624, Mar. 2005a.
- [Song y Li, 2005b] G. SONG y Y. LI, “Cross-layer optimization for OFDM wireless networks-part II: algorithm development”, *IEEE Transactions on Wireless Communications*, 4(2):625–634, Mar. 2005b.
- [Song et al., 2004] K. SONG, A. EKBAL, S. CHUNG y J. CIOFFI, “Adaptive modulation and coding (AMC) for bit-interleaved coded OFDM (BIC-OFDM)”, en *2004 IEEE International Conference on Communications. ICC'04*, volumen 6, 2004.
- [Sternad et al., 2007] M. STERNAD, T. SVENSSON, T. OTTOSSON, A. AHLÉN, A. SVENSSON y A. BRUNSTROM, “Towards systems beyond 3G based on adaptive OFDMA transmission”, *Proceedings of the IEEE*, 95(12):2432–2455, 2007.
- [Svensson, 2007] A. SVENSSON, “An introduction to adaptive QAM modulation schemes for known and predicted channels”, *Proceedings of the IEEE*, 95(12):2322–2336, 2007.
- [Szewczyk et al., 2004] R. SZEWCZYK, E. OSTERWEIL, J. POLASTRE, M. HAMILTON, A. MAINWARING y D. ESTRIN, “Habitat monitoring with sensor networks”, *Communications of the ACM*, 47(6):34–40, 2004.
- [Timmons y Scanlon, 2004] N. TIMMONS y W. SCANLON, “Analysis of the performance of IEEE 802.15. 4 for medical sensor body area networking”, en *2004 First Annual IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks, 2004. IEEE SECON 2004*, págs. 1266–1275, 2004.
- [Verdu, 2002] S. VERDU, “Spectral efficiency in the wideband regime”, *IEEE Transactions on Information Theory*, 48(6):1319–1343, 2002.
- [Vuran y Akyildiz, 2007] M. VURAN y I. AKYILDIZ, “A-MAC: adaptive medium access control for next generation wireless terminals”, *IEEE/ACM Transactions on Networking*, 15(3):574–587, 2007.

- [Wan et al., 2007] L. WAN, W. MA y Z. GUO, “A Cross-layer Packet Scheduling and Subchannel Allocation Scheme in 802.16e OFDMA System”, en *IEEE Wireless Communications and Networking Conference, 2007. WCNC 2007*, págs. 1865–1870, Mar. 2007.
- [Wang et al., 2001] A. WANG, S. CHO, C. SODINI y A. CHANDRAKASAN, “Energy efficient modulation and MAC for asymmetric RF microsensor systems”, en *2001 International Symposium on Low Power Electronics and Design*, págs. 106–111, 2001.
- [Wang et al., 2007a] W. WANG, Y. CUI, T. PENG y W. WANG, “Noncooperative Power Control Game with Exponential Pricing for Cognitive Radio Network”, en *IEEE 65th Vehicular Technology Conference, 2007. VTC2007-Spring*, págs. 3125–3129, 2007a.
- [Wang et al., 2007b] W. WANG, T. PENG y W. WANG, “Optimal power control under interference temperature constraints in cognitive radio network”, en *IEEE Wireless Communications and Networking Conference, 2007. WCNC 2007*, págs. 116–120, 2007b.
- [Webb y Steele, 1995] W. WEBB y R. STEELE, “Variable rate QAM for mobile radio”, *IEEE Transactions on Communications*, 43(7):2223–2230, 1995.
- [Weinstein y Ebert, 1971] S. WEINSTEIN y P. M. EBERT, “Data Transmission by Frequency Division Multiplexing Using the Discrete Fourier Transform”, *IEEE Transactions on Communications Technology*, 19(5):628–634, 1971.
- [Werner-Allen et al., 2005] G. WERNER-ALLEN, J. JOHNSON, M. RUIZ, J. LEES y M. WELSH, “Monitoring volcanic eruptions with a wireless sensor network”, en *Second European Workshop on Wireless Sensor Networks, 2005*, págs. 108–120, 2005.
- [Wieselthier et al., 2000] J. WIESELTHIER, G. NGUYEN y A. EPHREMIDES, “On the construction of energy-efficient broadcast and multicast trees in wireless networks”, en *IEEE Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. INFOCOM 2000*, volumen 2, 2000.
- [Wild y Ramchandran, 2005] B. WILD y K. RAMCHANDRAN, “Detecting primary receivers for cognitive radio applications”, en *IEEE First International*

- Symposium on New Frontiers in Dynamic Spectrum Access Networks, 2005. DySPAN 2005.*, págs. 124–130, 2005.
- [Wolsey, 1998] L. WOLSEY, *Integer programming*, Wiley-Interscience Series In Discrete Mathematics and Optimization, John Wiley & Sons, NY, 1998.
- [Wong et al., 1999] C. WONG, R. CHENG, K. LATAIEF y R. MURCH, “Multiuser OFDM with adaptive subcarrier, bit, and power allocation”, *IEEE Journal on Selected Areas in Communications*, 17(10):1747–1758, 1999.
- [Yang et al., 2007] H. YANG, N. BELHAJ y M. ALOUINI, “Performance analysis of joint adaptive modulation and diversity combining over fading channels”, *IEEE Transactions on Communications*, 55(3):520–528, 2007.
- [Yang y Wang, 2008] K. YANG y X. WANG, “Battery-Aware Adaptive Modulation Based on Large-Scale MDP”, *IEEE Transactions on Wireless Communications*, 7(1):72–77, 2008.
- [Yao y Giannakis, 2005] Y. YAO y G. GIANNAKIS, “Energy-efficient scheduling for wireless sensor networks”, *IEEE Transactions on Communications*, 53(8):1333–1342, 2005.
- [Yin y Liu, 2000] H. YIN y H. LIU, “An efficient multiuser loading algorithm for OFDM-based broadband wireless systems”, en *IEEE Global Telecommunications Conference, 2000. GLOBECOM '00*, volumen 1, págs. 103–107, 2000.
- [Yu et al., 2006] G. YU, Z. ZHANG, Y. CHEN, P. CHENG y P. QIU, “Subcarrier and bit allocation for OFDMA systems with proportional fairness”, en *IEEE Wireless Communications and Networking Conference, 2006. WCNC 2006*, volumen 3, págs. 1717–1722, Apr. 2006.
- [Yu y Cioffi, 2006] W. YU y J. CIOFFI, “Constant-power waterfilling: performance bound and low-complexity implementation”, *IEEE Transactions on Communications*, 54(1):23–28, 2006.
- [Yu y Lui, 2006] W. YU y R. LUI, “Dual methods for nonconvex spectrum optimization of multicarrier systems”, *IEEE Transactions on Communications*, 54(7):1310–1322, 2006.

- [Yu et al., 2004] Y. YU, B. KRISHNAMACHARI y V. PRASANNA, “Energy-latency tradeoffs for data gathering in wireless sensor networks”, en *Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies. INFOCOM 2004*, volumen 1, 2004.
- [YuMing et al., 2006] Z. YUMING, G. JINGBO, W. ZANJI y X. YI, “A Novel Sub-carrier and Power Allocation Algorithm for Multi-user OFDM Systems”, en *IEEE International Conference on Communications, 2006. ICC'06*, volumen 12, 2006.
- [Zhang et al., 2007] L. ZHANG, Y. LIANG y Y. XIN, “Joint Admission Control and Power Allocation for Cognitive Radio Networks”, en *IEEE International Conference on Acoustics, Speech and Signal Processing, 2007. ICASSP 2007*, volumen 3, 2007.
- [Zhang y Letaief, 2004] Y. ZHANG y K. LETAIEF, “Multiuser adaptive subcarrier-and-bit allocation with adaptive cell selection for OFDM systems”, *IEEE Transactions on Wireless Communications*, 3(5):1566–1575, 2004.
- [Zhang y Letaief, 2006] Y. ZHANG y K. LETAIEF, “Cross-layer adaptive resource management for wireless packet networks with OFDM signaling”, *IEEE Transactions on Wireless Communications*, 5(11):3244–3254, 2006.

Índice de Figuras

1.1. Relación entre los capítulos	5
2.1. Diagrama de bloques del sistema de transmisión adaptativa	13
2.2. Modulador OFDM para tiempo continuo	20
2.3. Modulador/demodulador OFDM para tiempo discreto	20
2.4. Símbolo OFDM de N muestras con prefijo cíclico de N_g muestras	21
2.5. Rejilla tiempo-frecuencia de OFDM	22
2.6. Modelo de canales paralelos	23
2.7. Asignación de potencia <i>waterfilling</i>	24
2.8. Canalización para: (a) TDMA; (b) FDMA; (c) OFDMA	28
2.9. Asignación de subportadoras para sistemas multiusuario OFDMA con diversidad multiusuario	30
2.10. Estructura de la rejilla tiempo-frecuencia utilizando agrupaciones de subportadoras e intervalos de tiempo	42
2.11. Tareas correspondientes a la capacidad cognitiva	46
2.12. Huecos espectrales	48
2.13. Acceso al espectro en redes cognitivas	49
4.1. Modelo del sistema	78
4.2. Tramas TDMA con longitud de los intervalos de transmisión fija y variable	79
4.3. Tramas TDMA con longitud de los intervalos de transmisión variable en número de símbolos	83
4.4. Variación de energía (dB) del esquema adaptativo con N_n redondeado respecto al valor real, para escenario IEEE.802.15.4	91
4.5. Ganancia de energía (dB) del esquema adaptativo respecto a TDMA-F y TDMA-V con modulación 16QAM, para escenario IEEE.802.15.4	92

4.6. Ganancia de energía (dB) del esquema adaptativo respecto a TDMA-V con modulación 64QAM, para escenario IEEE.802.15.4	93
4.7. Ganancia de energía (dB) del esquema adaptativo respecto a TDMA-F y TDMA-V con modulación 16QAM, para escenario Bluetooth	94
4.8. Ganancia de energía (dB) del esquema adaptativo respecto a TDMA-V con modulación 64QAM, para escenario Bluetooth . . .	95
5.1. Modelo del sistema	101
5.2. Modelo del sistema	109
5.3. Algoritmos óptimo y subóptimos: utilidad obtenida por la estación base con canal de variación lenta, para 3 usuarios con imparcialidad proporcional	122
5.4. Algoritmos óptimo y subóptimos: utilidad obtenida por la estación base con canal de variación lenta, para 3 usuarios con $\alpha = 0,5$	123
5.5. Algoritmos óptimo y subóptimos: utilidad obtenida por la estación base con canal de variación lenta, para 3 usuarios con asignación de máximo <i>goodput</i>	124
5.6. Algoritmo ST-RLP, 3 usuarios: <i>goodput</i> acumulado recibido por usuario, con canal de variación lenta, con imparcialidad proporcional	125
5.7. Algoritmo ST-RLP, 3 usuarios: <i>goodput</i> acumulado recibido por usuario, con canal de variación lenta, con $\alpha = 0,5$	126
5.8. Algoritmo ST-RLP, 3 usuarios: <i>goodput</i> acumulado recibido por usuario, con canal de variación lenta, con máximo <i>goodput</i>	127
5.9. Algoritmo ST-RLP, 7 usuarios: <i>goodput</i> acumulado recibido por usuario, con canal de variación lenta, con imparcialidad proporcional	128
5.10. Algoritmo ST-RLP, 7 usuarios: <i>goodput</i> acumulado recibido por usuario, con canal de variación lenta, con $\alpha = 0,5$	129
5.11. Algoritmo ST-RLP, 7 usuarios: <i>goodput</i> acumulado recibido por usuario, con canal de variación lenta, con máximo <i>goodput</i>	130
5.12. Algoritmos óptimo y subóptimos: utilidad obtenida por la estación base con canal de variación rápida, para 3 usuarios con imparcialidad proporcional	131
5.13. Algoritmos óptimo y subóptimos: utilidad obtenida por la estación base con canal de variación rápida, para 3 usuarios con $\alpha = 0,5$	132

5.14. Algoritmos óptimo y subóptimos: utilidad obtenida por la estación base con canal de variación rápida, para 3 usuarios con asignación de máximo goodput	133
5.15. Algoritmo ST-RLP, 5 usuarios: goodput acumulado recibido por usuario, con canal de variación rápida, para imparcialidad proporcional	134
5.16. Algoritmo ST-RLP, 5 usuarios: goodput acumulado recibido por usuario, con canal de variación rápida, para asignación de máximo goodput	135
5.17. Comparación del goodput transmitido por la estación base con canal de variación rápida con ST-RLP, SBPA y el algoritmo REF de [Mohanram y Bhashyam, 2005]	136
5.18. Comparación del tiempo de CPU con canal de variación rápida con ST-RLP, SBPA y el algoritmo REF de [Mohanram y Bhashyam, 2005]	137

Índice de Tablas

4.1. Ganancia de energía con respecto a TDMA-F y TDMA-V, con correlación entre canales σ^2 , para escenario IEEE.802.16.4	96
--	----

