



UNIVERSIDAD CARLOS III DE MADRID

DEPARTAMENTO DE INGENIERÍA TELEMÁTICA

TESIS DOCTORAL

**ESTRATEGIAS DE CONFIGURACIÓN DE REDES WLAN
IEEE 802.11e EDCA**

Autor: **Pablo Serrano Yáñez-Mingot**
Ingeniero de Telecomunicación

Director: **Albert Banchs Roca**
Doctor Ingeniero de Telecomunicación

Leganés, Septiembre de 2006

ESTRATEGIAS DE CONFIGURACIÓN DE REDES WLAN IEEE 802.11e EDCA

Autor: Pablo Serrano Yáñez-Mingot
Director: Dr. Albert Banchs Roca

Tribunal nombrado por el Mgfco. y Excmo. Sr. Rector de la Universidad Carlos III de Madrid, el día ___ de _____ de _____.

Firma del tribunal calificador:

Firma:

Presidente:

Vocal:

Vocal:

Vocal:

Secretario:

Calificación:

Leganés, ___ de _____ de _____.

“Llamadme Ismael”
– *Moby Dick*, Herman Melville

Agradecimientos

El autor de esta tesis quiere expresar su agradecimiento por el apoyo directo o indirecto para la realización de la misma a: Marina, José Luis, Marcos, y a Rebeca. A Albert, Carlos J., Ignacio, Arturo, María, David, José Ignacio, Manolo, Paco, Marcelo, Alberto, José Félix, Antonio de la O., Richi, Iván, Rubén y/o Ángel, Isaías, Jaime J., Isaac, Carlos G., Antonio C., Carmen, Guillermo, Mónica, y resto del Departamento de Ingeniería Telemática. Y a Pedro, Raúl, Evaristo, Nacho, Antoñete y Esteban, por algún motivo que el autor no es capaz de recordar.

Resumen

Las redes inalámbricas de área local (WLANs) se han convertido en una solución muy común para proporcionar acceso a Internet. El principal mecanismo usado a nivel MAC (Medium Access Control) para las WLANs es el definido por el estándar IEEE 802.11. Recientemente, el grupo de trabajo (Working Group) 802.11 ha aprobado un nuevo estándar, denominado 802.11e, que extiende dicho mecanismo para suministrar calidad de servicio (QoS). Este nuevo estándar se basa en una serie de parámetros a configurar, tema aún pendiente de ser resuelto pues si bien se proporcionan unos valores recomendados para dichos parámetros, estos valores son estáticos por lo que su idoneidad no está garantizada.

La presente tesis aborda el diseño de estrategias de configuración que optimicen el uso del mecanismo EDCA del estándar 802.11e. A tal efecto, se distinguen dos escenarios WLAN bien diferentes, cada uno definiendo una diferente estrategia de configuración para los mencionados parámetros:

- En el primer escenario, denominado *WLAN abierta*, se supone un entorno abierto como, por ejemplo, un *hot-spot*. En dicho escenario no se puede confiar en la caracterización del tráfico y los requisitos de servicio que pudiesen proporcionar los usuarios, dado que estos bien pueden declarar unos valores diferentes a los reales para obtener más recursos de la red. Dichos usuarios pueden incluso saturar su acceso inalámbrico (esto es, tener siempre un paquete en proceso de ser transmitido) para obtener recursos adicionales. Para este escenario se propone una configuración que garantice un rendimiento óptimo incluso en presencia de posibles usuarios saturando sus interfaces. Dada la naturaleza abierta de este escenario, adicionalmente se desarrolla un mecanismo que detecta comportamientos no ajustados a una determinada configuración.
- En el segundo escenario, denominado *WLAN cerrada*, los usuarios son dignos de confianza. Un ejemplo de dicho escenario puede ser un entorno de oficinas, donde sólo los empleados pueden conectarse a la red. En este caso, tanto la caracterización declarada del tráfico, así como los requisitos de servicio, pueden ser utilizados con objeto de calcular la configuración óptima que proporciona el mejor rendimiento. Nótese que, en este escenario, dado que se dispone de más información, sí que es posible escoger la configuración que mejor atiende los requisitos y características de las fuentes de tráfico.

Las principales contribuciones de la presente tesis se resumen a continuación: para el escenario de la WLAN abierta, en primer lugar se analiza el rendimiento de una red saturada,

a partir del cual se propone un algoritmo de configuración que optimiza el rendimiento global de la red, tanto en el modo básico como en el modo *multirate*. También se propone un algoritmo para detectar usuarios maliciosos en la red. Para el caso de la WLAN cerrada, se presentan algoritmos para caracterizar el comportamiento del retardo (valor medio, desviación típica y función de distribución) de la red en condiciones de no saturación. Con base en estos resultados, se presenta un algoritmo que calcula la configuración óptima para un caso de uso de Voz sobre IP. Por último, se describen resultados experimentales centrados también en este caso de uso.

Abstract

Nowadays Wireless LANs (WLANs) have become a very popular technology for Internet access. The Medium Access Control algorithm used by today's WLANs is the one defined by the IEEE 802.11 standard. Recently, the IEEE 802 Working Group has approved a new standard called 802.11e that extends the basic 802.11 algorithm with Quality of Service capabilities. This new standard is based on a number of open parameters the configuration of which is yet an unsolved research issue. Although the standard includes some recommendations for the parameters configuration, these recommendations are statically set and do not guarantee optimized performance.

This thesis takes up the issue of designing optimal configuration guidelines for 802.11e EDCA. To this aim, in this thesis two different WLAN scenarios are considered, each scenario driving to a different configuration strategy for the setting of the parameters:

- In the first scenario, which we call *open WLAN*, an open scenario like e.g. a *hot-spot* is assumed. In such scenario, the network cannot rely on the traffic characterization and requirements declared by the users, as users could potentially declare different values than the real ones thereby gaining additional resources from the network. Such users could even saturate their wireless access (i.e. by having always packets ready for transmission) in order to obtain extra resources. For this scenario, we propose an optimal configuration that guarantees optimal operation even in the presence of users saturating their channel. Given the untrusted nature of this scenario, mechanisms that detect malicious behaviours are devised.
- In the second scenario, the *closed WLAN*, users' behaviour can be trusted. An example of such a scenario could be an office setting in which only trusted employees are allowed to connect to the network. In this case, the source behaviour declared by the users, as well as their Quality of Service requirements, can be used in order to compute an optimal configuration that optimizes the behaviour given this information. Note that with this scenario, given that more information is known, it is possible to choose a parameter configuration that better meets the requirements and characteristics of the sources.

The main contributions of this thesis are summarized as follows. For the open WLAN scenario, we first analyze the performance of a saturated WLAN and then, based on this analysis, we propose configuration algorithms that optimize the overall network performance both in the basic and multirate modes. We also propose an algorithm to detect malicious users in this scenario. For the closed WLAN scenario, we present novel algorithms to compute

the delay performance (average delay, typical deviation and delay distribution) of a non-saturated WLAN. Next, based on these analysis, we present an algorithm for computing the optimal configuration for the use case of VoIP traffic. Finally, we present some experimental results for this use case.

Contenidos

1. Introducción	1
2. WLANs 802.11: Antecedentes y provisión de QoS	3
2.1. Evolución de las redes IEEE 802.11	3
2.1.1. La provisión de QoS y la capa MAC	5
2.1.2. El 802.11e Task Group	7
2.2. El acceso distribuido en 802.11 y 802.11e	8
2.2.1. El modo DCF de 802.11	8
2.2.1.1. Operación RTS/CTS	10
2.2.1.2. Carencias de 802.11	11
2.2.2. El modo EDCA	11
2.3. Estudio de prestaciones mediante simulación	14
2.3.1. 802.11 DCF	15
2.3.2. 802.11e EDCA	15
2.4. Estudio de prestaciones mediante modelo analítico	16
2.4.1. 802.11 DCF	16
2.4.1.1. Modelado en saturación	16
2.4.1.2. Modelado en no saturación	18
2.4.2. 802.11e EDCA	19
2.4.2.1. Modelado en saturación	19
2.4.2.2. Modelado en no saturación	21
2.5. Control de admisión	22
2.5.1. Basado en medición	22
2.5.2. Basado en modelo	23
2.6. Supervisión de usuarios	23
2.7. Resumen	24
3. Configuración de WLAN abiertas	27
3.1. Análisis del modo EDCA en condiciones de saturación	27
3.1.1. Problema de los modelos derivados del análisis de Bianchi	28
3.1.2. Nuevo modelo para el análisis de EDCA	28
3.1.2.1. Demostración de la unicidad de la solución	31
3.1.3. Validación del nuevo modelo	31
3.2. Configuración para modo básico	33
3.2.1. Análisis de las anomalías	33

3.2.1.1.	Tiempo de servicio	33
3.2.1.2.	Estabilidad	35
3.2.1.3.	Ancho de banda	39
3.2.2.	Configuración propuesta	41
3.2.2.1.	Validación de la configuración	42
3.3.	Configuración para modo <i>multirate</i>	44
3.3.1.	El criterio de equidad proporcional	46
3.3.2.	Análisis de los esquemas de asignación	46
3.3.3.	Configuración para el esquema CW	49
3.3.3.1.	Mecanismo centralizado	53
3.3.3.2.	Mecanismo distribuido	54
3.3.4.	Configuración para el esquema TL	55
3.3.5.	Evaluación del rendimiento de los esquemas	57
3.3.5.1.	Asignación de anchos de banda	57
3.3.5.2.	Rendimiento de la red	57
3.3.5.3.	Tiempo de servicio	59
3.3.5.4.	Idoneidad de los esquemas centralizados	60
3.4.	Detección de configuraciones maliciosas	63
3.4.1.	Detección de una configuración maliciosa de CW_{min}	64
3.4.2.	Uso de una configuración óptima	67
3.4.3.	Resultados de simulación	68
3.4.3.1.	Parámetros del algoritmo	69
3.4.3.2.	Configuración 802.11 DCF y efecto captura	70
3.4.3.3.	Configuración 802.11 DCF sin efecto captura	71
3.4.3.4.	Configuración óptima sin efecto captura	71
3.5.	Resumen	72
4.	Configuración de WLAN cerradas	73
4.1.	Análisis del valor medio y desviación típica del tiempo de servicio	74
4.1.1.	Análisis de las probabilidades de transmisión	74
4.1.1.1.	Estudio del ancho de banda	75
4.1.1.2.	Cálculo de la probabilidad de transmisión	78
4.1.2.	Análisis del tiempo medio de servicio	79
4.1.3.	Análisis de la desviación típica del tiempo medio de servicio	85
4.1.4.	Validación del modelo	87
4.1.4.1.	Tiempo medio de servicio	88
4.1.4.2.	Desviación típica	90
4.2.	Análisis de la función de distribución del tiempo de servicio	90
4.2.1.	Modelo	92
4.2.1.1.	Caso de RTS/CTS	94
4.2.1.2.	Caso de longitudes variables de paquetes	94
4.2.2.	Validación	95
4.3.	Caso de estudio: entorno VoIP	95
4.3.1.	Evaluación de la configuración propuesta	97
4.4.	Metodología experimental en un escenario real	98

4.4.1. Escenario	99
4.4.2. Configuración del tráfico <i>real-time</i>	100
4.4.3. Configuración del tráfico no <i>real-time</i>	103
4.4.4. Comparación con la configuración estándar	106
4.5. Resumen	106
5. Conclusiones	109
Referencias	111

Índice de figuras

2.1. Certificación de la WiFi Alliance	4
2.2. Proceso de <i>backoff</i> en el modo DCF	9
2.3. Funcionamiento RTS/CTS	10
2.4. Parámetros del modo EDCA	12
2.5. Diferenciación proporcionada por EDCA	14
3.1. Cadena de Markov para el proceso $\{j, k\}$	29
3.2. Validación del nuevo modelo para saturación	32
3.3. Tiempo medio de servicio	34
3.4. Modelo para el tiempo medio de servicio	36
3.5. Inestabilidad del modo básico	36
3.6. Estado de las colas en el modo básico	37
3.7. Modelo para la inestabilidad de DCF	38
3.8. Anomalía en la tasa recibida	40
3.9. Configuración propuesta para WLAN abiertas	41
3.10. Retardo para la configuración óptima	43
3.11. Estabilidad de la configuración óptima	43
3.12. Ancho de banda con la configuración óptima	44
3.13. Configuración uniforme con estaciones a distinta tasa	45
3.14. Esquema CW, 2 estaciones por grupo, $m_i = 5$	49
3.15. Esquema CW, 10 estaciones por grupo, $m_i = 5$	50
3.16. Esquema CW, 10 estaciones por grupo, $m_i = 0$	50
3.17. Esquema TL, 10 estaciones por grupo	51
3.18. Asignación de anchos de banda para los diferentes esquemas	58
3.19. Valoración de asignaciones, dos tasa de transmisión	58
3.20. Retardo de acceso para los esquemas centralizados	61
3.21. Retardos de acceso para los esquemas distribuidos	61
3.22. Comparación con búsqueda exhaustiva para los esquemas propuestos	62
3.23. Comparación con búsqueda exhaustiva en un esquema combinado	63
3.24. Ancho de banda obtenido en presencia de efecto captura	65
3.25. Ancho de banda obtenido en ausencia de efecto captura	65
3.26. Relación entre P_{FA} , K y T_{obs}	68
3.27. Probabilidad de detección en función de K , $T_{obs} = 5s$	70
3.28. Probabilidad de detección en función de T_{obs} , $K = 2$	71

4.1. k -ranuras y probabilidad de transmisión (ejemplo para $k=2$)	75
4.2. Probabilidad de vacío en una k -ranura	76
4.3. Retardo medio para j retransmisiones	80
4.4. Componentes del retardo (ejemplo para $k = 2$)	80
4.5. Componentes de $T_{inter}^{-i,j}$ (para $A_i = 3$ y $j = 2$)	85
4.6. Tiempo medio de servicio, diferenciación por AIFS	88
4.7. Tiempo medio de servicio, diferenciación por CW	89
4.8. Tiempo medio de servicio, diferenciación combinada	89
4.9. Desviación típica, diferenciación por AIFS	90
4.10. Desviación típica, diferenciación por CW	91
4.11. Desviación típica, diferenciación combinada	91
4.12. Validación del análisis de la función de distribución del tiempo de servicio .	96
4.13. Distribución del retardo para 20 flujos VoIP	98
4.14. Arquitectura empleada en la generación del tráfico VoIP	99
4.15. Retardo de voz para el modo DCF	100
4.16. Retardo de VoIP con el uso de $TXOP_{max}$	101
4.17. Retardo para diferentes valores de AIFS y CW_{min}	104
4.18. Tasa de transmisión de los datos vs. número de flujos VoIP	105
4.19. Configuración para varias estaciones de datos	106
4.20. Comparación con la configuración propuesta por el estándar	107

Índice de tablas

2.1. Task Groups de 802.11	7
2.2. Parámetros de 802.11b	9
2.3. Configuración estándar para EDCA con 802.11b	13
3.1. Parámetros de 802.11b empleados en el simulador	48
3.2. Evaluación del rendimiento de los cuatro esquemas para 4 grupos y 5 estaciones	59
3.3. Probabilidad de falsa alarma	69
3.4. Configuración por defecto 802.11 DCF y efecto captura	71
3.5. Configuración por defecto 802.11 DCF sin efecto captura	72
3.6. Configuración óptima, sin efecto captura	72
4.1. Evaluación del algoritmo de configuración	97
4.2. Obtención de la configuración de CW para VoIP	103

Capítulo 1

Introducción

Las redes inalámbricas son una realidad bien asentada. Sin embargo, dicha realidad proporciona un servicio de cierta calidad únicamente en condiciones de baja carga de tráfico y/o ausencia de requisitos en la entrega del tráfico. El hecho de disponer de un acceso al medio compartido sin capacidad de diferenciación puede ocasionar que una aplicación de voz, por ejemplo, sufra de pérdidas o retardos muy elevado debido a la existencia de una o varias descargas de ficheros en la misma WLAN. Que las redes WLAN puedan ser empleadas con garantías en telefonía IP inalámbrica o cualquier otro tipo de servicio que incluya requisitos de entrega dependerá, en gran medida, de su capacidad para proporcionar calidad de servicio (*Quality of Service*).

Tal y como se explicará en el Capítulo 2, varios son los motivos que han propiciado la gran aceptación de la tecnología inalámbrica, en concreto de los productos derivados del estándar 802.11. Dicho estándar ha sido recientemente ampliado (802.11e) para incluir mecanismos que permitan el suministro de garantías de servicio a las aplicaciones; en dicho capítulo se presentan tales mecanismos, el trabajo existente hasta la fecha en el modelado analítico para predecir el comportamiento de una WLAN, y las herramientas desarrolladas para suministrar QoS a las aplicaciones. Sin embargo, tal y como se analiza en el capítulo, este trabajo existente no presenta unas guías claras para la configuración y supervisión en escenarios reales.

La presente tesis tiene por objetivo presentar los esquemas de configuración que, definido un escenario, permiten maximizar el rendimiento obtenido. Se realiza una clasificación de dichos escenarios en dos bloques: por una parte, lo que se ha denominado *WLAN cerradas*, en las que se supone que los usuarios son dignos de confianza, por lo que puede emplearse tanto la caracterización del tráfico que generan como sus requisitos de entrega de cara a proporcionar la mejor configuración de la WLAN. Dichas redes se corresponden, por ejemplo, con el caso de un entorno de oficinas, donde sólo los empleados harán uso de los recursos inalámbricos. En contraposición, se define el escenario de *WLAN abiertas* como aquél en el que las anteriores premisas no son válidas: puede que existan usuarios que pretendan obtener los máximos recursos de la WLAN, para lo cual pueden proporcionar una caracterización incorrecta de su tráfico, requisitos de entrega no reales o, incluso, modificar los parámetros de su interfaz inalámbrica. En este escenario, que podría corresponderse con un punto de acceso abierto al público en una cafetería, la única posibilidad consiste en suponer condiciones de saturación para todos los usuarios y actuar

en consecuencia.

Dado que los modelos son diametralmente opuestos, resulta claro que las estrategias a seguir para obtener el mayor rendimiento en cada esquema serán diferentes. El Capítulo 3 trata sobre el estudio y optimización de las *WLAN abiertas*, sobre las que se suele suponer condiciones de *saturación*¹. Si bien dichas redes han sido objeto de varios análisis [1–4], estos plantean sistemas de ecuaciones sin unicidad de la solución garantizada. Al inicio del Capítulo 3 se presenta un nuevo análisis con solución única [5]. Por otra parte, es conocido [6] que las redes WLAN pueden presentar en ciertas circunstancias inestabilidades en su comportamiento, lo que debe ser evitado. En dicho capítulo se profundiza en el análisis y modelado de dichas inestabilidades, proponiéndose una configuración que garantiza la estabilidad del sistema (trabajo presentado en [7]). A continuación, se extiende la configuración propuesta para entornos *multirate*, empleándose a tal efecto el criterio de *equidad proporcional* propuesto por Kelly [8]; este trabajo fue presentado en [9]. Por último, dado que los usuarios en redes WLAN abiertas no serán necesariamente bienintencionados, resulta preciso desarrollar un mecanismo de supervisión de usuarios que vigile que se respetan las reglas del protocolo MAC; dicho mecanismo, presentado en [10], permite establecer un compromiso entre las probabilidades de detección y falsa alarma, lo que supone una mejora substancial sobre mecanismos existentes, basados en heurísticos, como DOMINO [11].

El Capítulo 4 se dedica al estudio y configuración de redes *WLAN cerradas*. En dichas redes se supone que se conoce tanto el número de usuarios que van a hacer uso de sus recursos, como la naturaleza del tráfico que va a ser cursado y sus requisitos de QoS. Es por ello que se no se realiza la suposición de saturación, pues al recibir QoS las colas de dichas transmisiones no estarán constantemente con tramas pendientes de ser transmitidas. El capítulo presenta un modelo para la no saturación, con el que se realiza la caracterización del valor medio y desviación típica del tiempo de servicio [12]; también se presenta una caracterización de la función de distribución de dicho tiempo [13]. Con ambas herramientas, es posible realizar el camino inverso y obtener qué configuración permite dar soporte a unos determinados requisitos (bien en términos de retardo medio, de percentil de retardo, etc) dado un escenario para la WLAN; dicho esquema se presenta para un *caso de uso* que considera un escenario de voz sobre IP. Para completar el capítulo se presenta un estudio experimental que permite obtener la configuración más adecuada para proporcionar QoS en un entorno real. En este estudio se sigue una metodología pragmática para encontrar la configuración óptima en un escenario concreto, configuración que se compara con la obtenida a través del modelo analítico. Dicho estudio amplía el trabajo presentado en [14] para un escenario experimental, con una primera versión de dichos resultados disponible en [15].

Las conclusiones, los retos y trabajos futuros que se presentan tras los esquemas desarrollados en la presente tesis son objeto del Capítulo 5.

Otras publicaciones del autor relacionadas con el entorno en que se ha desarrollado la presente tesis doctoral son: [16–20].

¹Las condiciones de saturación consisten en suponer que todas las estaciones de la WLAN siempre tienen un paquete pendiente de ser transmitido. Se trata de una de las condiciones más habituales para analizar el rendimiento de una WLAN.

Capítulo 2

WLANs 802.11: Antecedentes y provisión de QoS

Resulta innegable la gran aceptación de dispositivos WLAN basados en el estándar 802.11. Los interfaces inalámbricos que emplean dicha tecnología no sólo se encuentran en cualquier ordenador portátil reciente, sino que también aparecen en dispositivos de telefonía móvil e, incluso, de entretenimiento (consolas, sistemas de audio, ...). Si bien el estándar planteaba un mecanismo de acceso centralizado para proporcionar ciertas garantías de servicio, éste no ha tenido gran difusión. Sin embargo, debido a la gran penetración de 802.11 y al uso de aplicaciones *real-time* sobre dichas redes, se hace necesario alguna herramienta que permita proporcionar garantías de servicio a las aplicaciones.

El nuevo estándar 802.11e ofrece mecanismos en el nivel MAC que permiten proporcionar diferentes tipos de servicio a aplicaciones que comparten el mismo medio inalámbrico. En este capítulo se presentan dichos mecanismos, para lo cual resultará imprescindible partir del modo de acceso de 802.11. También se presentan aquellos artículos que han abordado el análisis de las prestaciones de 802.11e, bien por simulación bien por modelado analítico, así como aquellas líneas de investigación que han tratado el tema del control de admisión en una red WLAN o la supervisión de usuarios que no se adhieren a la configuración establecida.

2.1. Evolución de las redes IEEE 802.11

En un principio, las redes de área local (Local Area Networks, LAN) fueron creadas para la comunicación de datos a distancias cortas, si bien esta limitación ha ido paliándose con el desarrollo de la tecnología. Una red de área local inalámbrica (Wireless LAN, WLAN) define el mismo concepto, salvo que en la transmisión se emplea una comunicación no cableada (radio o luz), lo que determina una mayor restricción en cuanto a zona de disponibilidad. Es por ello que, hasta hace relativamente poco tiempo, las WLANs se consideraban soluciones muy costosas para entornos muy limitados.

Sin embargo, hoy en día las WLANs son una solución tecnológica adoptada para cada vez más y más situaciones. Para explicar este cambio, resulta adecuado acudir a



Figura 2.1: Certificación de la WiFi Alliance

la disquisición de [21]: de acuerdo con Joseph Pelton¹, el hecho de que una tecnología esté preparada para ser una solución de comunicaciones viene dado por una mejora en cinco áreas clave, que son: el estándar, la regulación, la tecnología, los servicios y las posibilidades económicas. Para el caso de 802.11, la evolución y el estado actual de estos aspectos es:

El estándar. El estándar 802.11 aparece en 1997. Por aquel entonces, las velocidades disponibles eran 1 y 2 Mbps, la tecnología de semiconductores en la banda de 2.4 GHz aún resultaba muy poco asequible y existían problemas de interoperabilidad (por ejemplo, en cifrado y gestión de potencia). Las dos primeras limitaciones (velocidad y precio) fueron atenuadas con la aparición de nuevos estándares (ver Tabla 2.1) y el creciente interés de los fabricantes en dicha tecnología, pero la interoperabilidad seguía siendo un obstáculo. Para subsanar esto, los principales fabricantes constituyeron en 1999 la Wireless Ethernet Compatibility Alliance (WECA, ahora conocida como WiFi Alliance), que crea la marca Wi-Fi² para garantizar que la comunicación entre equipos con dicha marca (Figura 2.1) no presenta problemas de interoperabilidad. En el año 2001 se anunció la creación de la marca Wi-Fi5 para los productos que superen las pruebas de interoperabilidad en la banda de 5 GHz.

La regulación. El espectro disponible para las comunicaciones inalámbricas se divide en dos clases: bandas con licencia, donde es precisa autorización explícita del Regulador para poder transmitir, y bandas no reguladas, en las que basta cumplir una serie de requisitos mínimos —en comparación con los anteriores— para poder usar dichas frecuencias. El sistema de licencias (telefonía móvil, emisoras de radio, televisión) implica la ventaja de un mayor control por parte del Gobierno, pero tiene una serie de contrapartidas: mayores retardos en la implantación de la tecnología (y de mejoras en la misma), causas legales entre operadores, . . . Sólo la cuestión del mecanismo a seguir para obtener una licencia merecería una disquisición aparte (recuérdense las subastas de UMTS). Para el caso de 802.11, las ventajas de la no regulación han superado -de momento- los inconvenientes, si bien es cierto que ahora, con la amplia difusión de dicha tecnología, es cuando comienzan a surgir las preocupaciones por la disminución de prestaciones en condiciones de alta carga, así como por las interferencias (tanto entre redes WLAN como con Bluetooth).

La tecnología. Los avances en el manejo de la tecnología de semiconductores han supuesto

¹Investigador del Instituto para la Investigación Espacial Aplicada de la Universidad George Washington.

²Wi-Fi no significa Wireless Fidelity: véase el artículo a tal respecto en BoingBoing.net <http://www.boingboing.net/2005/11/08/wifi.isnt.short.for..html>.

un enorme impulso para WLAN, debido a la disminución en el precio de los componentes y, por ende, en el producto final. De forma paralela, la escasez de la banda disponible ha fomentado el uso más eficiente de la misma, mejorándose las técnicas de espectro ensanchado para atenuar las interferencias a la par que se aumentaba la velocidad y el alcance. Estas mejoras, como se ha comentado en el punto anterior, se pueden introducir sin excesiva demora dada la regulación de la banda.

Los servicios. El estándar 802.11 se diseñó para poder sustituir a las capas MAC y PHY de 802.3. Las implicaciones de este criterio de diseño son evidentes: todos los servicios disponibles para éste (por citar un ejemplo, IP) están disponibles en la nueva tecnología. La demanda de acceso a Internet en cualquier lugar ha fomentado la aparición de puntos de acceso WLAN públicos (*hotspots*) por doquier: aeropuertos, cafeterías, hoteles... incluso motiva la aparición de nuevos modelos de negocio (por ejemplo, el caso de FON³). No hay que olvidar, asimismo, el prometedor matrimonio de la tecnología WLAN con las PDAs (dispositivos pequeños a la par que muy potentes), que permite el desarrollo de aplicaciones en control de procesos, educación, medicina, etc.

Las posibilidades económicas. Por todo lo anterior, resulta claro que las redes WLAN permiten el desarrollo de modelos de negocio con perspectivas de beneficio económico. Dada la existencia de un mercado de servicios previo, y una tecnología eficaz a un coste razonable, las oportunidades de negocio son varias: por ello, no sólo cualquier ordenador portátil incluye un interfaz WLAN, también las consolas de última generación, los dispositivos PDA y de telefonía VoIP, ... También es de destacar la aparición de operadores de 802.11 para aeropuertos, hoteles, ciudades (aunque en este último caso se encuentran con la competencia por parte de algunas administraciones locales), etc.

2.1.1. La provisión de QoS y la capa MAC

Si bien las redes WLAN son un éxito en cuanto a su implantación, el servicio que ofrecen en la actualidad es únicamente del tipo *best effort*. Dado que la velocidad real ofrecida, a compartir entre los usuarios en la red, sigue siendo muy poco elevada, la alternativa del sobredimensionado para proporcionar QoS no es una opción pues, en un caso típico, una WLAN únicamente proporciona una navegación web a una velocidad aceptable; sin embargo, para otros servicios esto no es así: baste con imaginar, en un streaming de vídeo a través de un canal inalámbrico, los efectos devastadores que tendría la irrupción de una aplicación *peer to peer* en dicho medio.

Las dificultades para proporcionar QoS en una WLAN, por lo tanto, nacen del hecho de disponer de un único canal a compartir con unas tasas de transmisión efectivas no muy elevadas (lo que deriva, en parte, de lo anterior), sin poder distinguir entre aplicaciones que requieren QoS (tiempo real) frente a aplicaciones menos prioritarias (descarga de ficheros). Incluso si se encontrase el mejor mecanismo de compartición del medio, existirá el riesgo

³FON es una iniciativa promovida por Martín Varsavsky que pretende crear una red de usuarios que comparten su cobertura WLAN, bien de forma solidaria, bien mediante transferencia económica. Más información en <http://www.fon.es/>

de que un usuario decidiese no respetar las reglas de la capa MAC en beneficio propio (por ejemplo, para descargar con mayor velocidad un determinado archivo).

Partiendo de lo anterior, las herramientas necesarias para proporcionar QoS se pueden clasificar en tres bloques: las que permiten la diferenciación de recursos, las que se encargan de configurar esta diferenciación, y las que regulan tanto el acceso a la WLAN como el comportamiento de los participantes en la comunicación:

Diferenciación. Diferentes aplicaciones, con diferente importancia y/o diferentes requisitos de calidad de servicio, precisan ser tratadas de forma desigual. Es por ello que se precisan mecanismos que permitan otorgar un trato distinto a aplicaciones con requisitos diferentes: una aplicación de transferencia de ficheros puede exigir un mayor ancho de banda que una aplicación de login remoto, pero ésta última precisa de una entrega más veloz que aquélla. La capa MAC debe proporcionar esta diferenciación, bien sea por mecanismos de prioridad, bien por diferentes parámetros del acceso al medio para diferentes tipos de tráfico.

Configuración. No basta con un mecanismo que permita servir de forma diferente a los distintos tipos de aplicaciones. Un sistema de prioridad estricta, por ejemplo, no sería de utilidad en la ausencia de una tabla de asignación de flujos a prioridades. Para cada mecanismo de diferenciación que exista es preciso proporcionar la configuración de dicho mecanismo en función de las condiciones de operación del sistema. Una configuración fija, independiente del entorno, siempre será subóptima al no explotar al máximo la información disponible sobre la red.

Supervisión. Aún una red adecuadamente configurada para provisionar QoS precisa de la supervisión para su correcto funcionamiento. Dicha supervisión se debe ejercer en dos dimensiones:

- Por una parte, es imprescindible el uso de un **control de admisión**, que determine si una nueva aplicación puede ser atendida (en cuyo caso puede ser preciso iniciar de nuevo el proceso de configuración) o no. Esto es así dado que, una vez que la red opera al límite de sus capacidades (con todos los usuarios recibiendo la QoS que demandan), la entrada de un nuevo flujo no sólo provocaría que dicho flujo no obtuviese las garantías que precisa, sino que además el resto de aplicaciones se viesen seriamente afectadas por esta entrada. En el control de admisión, adicionalmente, se pueden implementar diferentes políticas de comportamiento ante situaciones cercanas a la congestión: se renegocian las garantías ofrecidas o el tráfico cursado, se expulsa algún flujo de menor prioridad que el entrante, etc.
- Por otra parte, el hecho de proporcionar una determinada configuración a los usuarios no implica que estos se vayan a comportar conforme a la misma: suponer eso significaría dejar el rendimiento de la red a la buena fe, lo que es una decisión arriesgada. Si los usuarios tienen la capacidad de alterar el comportamiento de su capa MAC (y existen productos comerciales que así lo permiten), la provisión de QoS en el nivel de enlace debe venir acompañada de un mecanismo de **supervisión** del comportamiento de cada usuario, con

Tabla 2.1: Task Groups de 802.11

Grupo	Nombre	Estado/Publicación
802.11	Estándar original (1 y 2 Mbps)	Julio 1997
802.11a	Tasas de hasta 54 Mbps en la banda de 5 GHz	Octubre 1999
802.11b	Tasas de 5.5 y 11 Mbps en la banda de 2.4 GHz	Octubre 1999
802.11c	Operación de <i>bridge</i>	En 802.1D
802.11d	Soporte a la itinerancia	Julio 2001
802.11e	Soporte de QoS	Noviembre 2005
802.11f	Protocolo de comunicación entre puntos de acceso	Julio 2003
802.11g	Tasas de hasta 54 Mbps en la banda de 2.4 GHz	Junio 2003
802.11h	Mejoras de gestión y transmisión en la banda de 5 GHz	Octubre 2003
802.11i	Mejoras de la seguridad en la capa MAC	Julio 2004
802.11j	Operación en la banda de 5 GHz en Japón	Octubre 2004
802.11k	Medida de recursos radio	Activo
802.11n	Altas tasas de transmisión	Activo
802.11p	Acceso inalámbrico para vehículos	Activo
802.11r	Mecanismos de itinerancia rápida	Activo
802.11s	Extensión para <i>mesh networking</i>	Activo

objeto de garantizar (con cierta probabilidad, dado que el acceso al medio es de naturaleza aleatoria) que dicho usuario se comporta según la configuración esperada, y no intenta obtener un mejor servicio del que debe.

2.1.2. El 802.11e Task Group

Dentro del IEEE, es la *Computer Society* la responsable de la serie de estándares IEEE 802, a través de LAN/MAN Standards Committee (que ahora también incluye las redes de área personal, PAN). Éste se divide en Working Groups (WG) y Technical Advisory Groups (TAG), encargados estos últimos de asesorar a los primeros -que son los que desarrollan los estándares-. El grupo número 11 es el encargado de WLAN. Los WG y TAG puede incluir Task Groups (TG) para el desarrollo de trabajo más especializado, y se denominan a través una letra otorgada de forma secuencial: por ejemplo, la letra *b* en 802.11b implica que dicho grupo se definió después del 802.11a, y no que sea un sucesor de éste (como se puede llegar a creer).

En la Tabla 2.1 se citan los TG que han surgido dentro de 802.11⁴. El primer estándar se centró en la transmisión de datos en un medio inalámbrico, sin abarcar temas de movilidad, seguridad o la calidad de servicio: la ausencia de fiabilidad o de garantías en la entrega se sacrificaban en pos de la eficiencia y facilidad de implementación. Con la toma de conciencia de la necesidad de la provisión de mecanismos de QoS a las aplicaciones, en Marzo del año

⁴La situación actualizada de cada grupo, así como los estándares publicados, se puede obtener en <http://grouper.ieee.org/groups/802/11/Reports/802.11-Timelines.htm>. Resulta preciso matizar que el estándar 802.11b fue corregido en Noviembre de 2001, ampliado con información sobre normativa de frecuencias.

2000 nace el grupo de trabajo 802.11e (inicialmente también con el objetivo de la seguridad, aunque ésta propició el nacimiento de un grupo específico, el 802.11i, al año siguiente).

El estándar 802.11e [22], como se verá a continuación, proporciona los mecanismos necesarios para diferenciar el servicio ofrecido a cada aplicación, pudiéndose priorizar un tipo de tráfico frente a otro a través de los parámetros de la capa MAC. Existe además la capacidad de modificar dichos parámetros de acuerdo con decisiones de gestión de recursos. En el estándar se recomienda una determinada configuración para las diferentes clases de servicio, si bien la idoneidad de esta configuración, estática, no puede ser garantizada.

2.2. El acceso distribuido en 802.11 y 802.11e

En el caso más típico, una red 802.11 está constituida por una serie de estaciones y un punto de acceso (Access Point, AP), que suele ser además pasarela con una red cableada⁵. Para regular el acceso al medio, el primer estándar propone dos mecanismos:

- Un mecanismo centralizado, PCF (Point Coordination Function), que se basa en un mecanismo de *polling* iniciado por el AP. Este mecanismo ha tenido escasa difusión.
- Un mecanismo distribuido, DCF (Distributed Coordination Function), en el que las estaciones emplean un mecanismo del tipo CSMA/CA para compartir el acceso al medio. A diferencia del anterior, éste sí ha conseguido una gran aceptación.

El estándar 802.11e propone, también, dos modos de regular el acceso al medio:

- Un mecanismo centralizado, el HCCA (Hybrid coordination function Controlled Channel Access), sucesor del PCF.
- Un mecanismo distribuido, el EDCA⁶ (Enhanced Distributed Channel Access), sucesor del DCF.

Debido en parte a la escasa difusión del modo PCF, el futuro del HCCA resulta incierto. No existe consenso en la comunidad científica ni entre los fabricantes sobre la bondad de una opción centralizada frente a una distribuida; en la presente tesis se abordará únicamente el estudio del mecanismo de acceso distribuido EDCA.

2.2.1. El modo DCF de 802.11

La operación básica del modo DCF se ilustra en la Figura 2.2. Una estación A, cuando tiene que transmitir una trama del nivel superior, inicia la escucha del medio inalámbrico. Si éste es clasificado como no activo durante un tiempo DIFS (Distributed Interframe Space), la estación considera que el medio se encuentra disponible y transmite la trama. Hecho esto, pasado un tiempo SIFS (Short Interframe Space), la entidad receptora emite un asentimiento.

⁵Una descripción bastante amena del estándar 802.11 [23] se puede encontrar en [24]. Cisco [25] presenta tanto el estándar original como las diferentes extensiones, con un enfoque –como suele ser habitual– muy práctico. Para conocer fabricantes y distribuidores, es muy recomendable [21].

⁶Originariamente conocido como EDCF

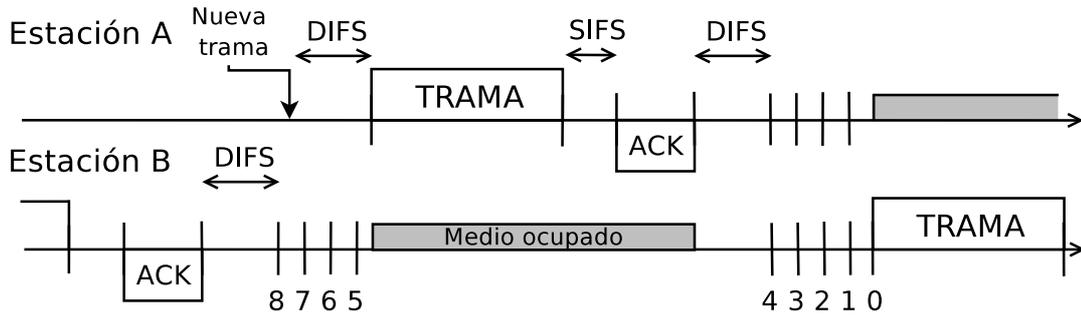


Figura 2.2: Proceso de *backoff* en el modo DCF

Tabla 2.2: Parámetros de 802.11b

Parámetro	Valor
SIFS	10 μ s
T_e	20 μ s
DIFS	50 μ s
CW_{min}	32
CW_{max}	1024

El hecho de que el tiempo SIFS sea menor que el DIFS permite dar prioridad al envío de asentimientos frente al envío de tramas.

Si la estación sondea el medio durante un tiempo DIFS y detecta que éste está ocupado, inicia el **proceso de backoff**. Éste es como sigue: la estación selecciona un número al azar, distribuido entre 0 y un valor dado ($CW_{min} - 1$). Una vez que vuelva a detectar el medio libre durante un tiempo igual a DIFS, decrementará dicho valor en una unidad por cada tiempo T_e que el medio permanezca inactivo (en la figura, estos instantes son cuando la estación B realiza el decremento 8, 7, 6, ...). Si el medio se detecta ocupado (en el ejemplo, durante la transmisión de A), el contador se paraliza, hasta que no se vuelve a detectar libre durante un tiempo DIFS. Una vez que el contador llega a 0, la estación transmite su trama.

Si no se recibe un asentimiento tras la transmisión, se considera que la trama no se entregó correctamente por lo que se vuelve a iniciar el proceso de backoff, pero en este caso duplicando el valor de CW_{min} ⁷. A cada fallo en la transmisión de una misma trama se duplicará dicho valor hasta un máximo dado (CW_{max}). Una vez que la trama es entregada correctamente, o se llega al máximo número de retransmisiones R , CW_{min} recupera su valor original. Para el estándar 802.11b, los valores de los parámetros mencionados son los de la Tabla 2.2.

A la vista de lo anterior, resulta claro que el modo DCF emplea un acceso al medio *ranurado*: una *ranura de tiempo* es el intervalo que pasa entre dos decrementos consecutivos del contador de backoff de una estación. Unas ranuras estarán ocupadas (por transmisiones

⁷Se supone que la transmisión no tuvo éxito al coincidir en el tiempo con, al menos, otra transmisión de otra estación (esto es, se produjo una *colisión*). Para intentar evitar una nueva colisión se aumenta el rango de la elección del número aleatorio: este retroceso exponencial binario sirve para que las estaciones se adapten a situaciones temporales de alta carga.

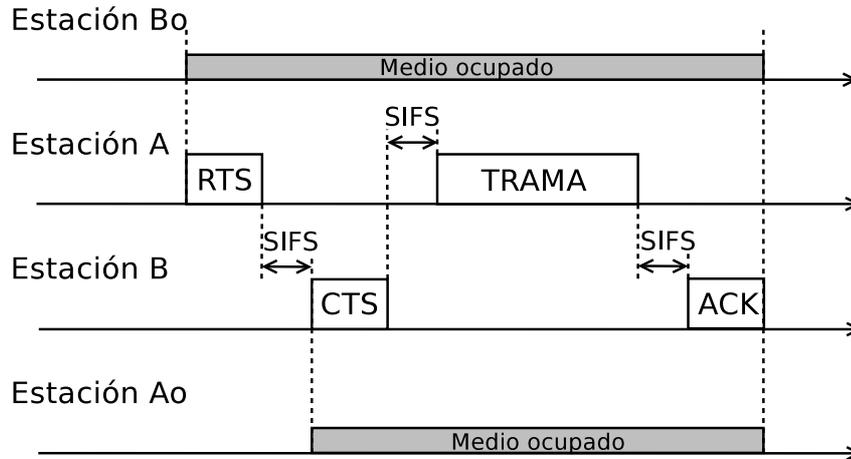


Figura 2.3: Funcionamiento RTS/CTS

y colisiones de otros nodos) y otras estarán vacías (con duración T_e ⁸).

La duración de una ranura que contiene una transmisión exitosa de una trama de tamaño l viene dada por la siguiente expresión:

$$T_s(l) = T_{PLCP}(C) + \frac{H+l}{C} + SIFS + \frac{ACK}{C} + DIFS \quad (2.1)$$

Donde H es la cabecera del nivel de enlace, C es la velocidad de transmisión, ACK es la duración del asentimiento y $T_{PLCP}(C)$ es la cabecera del nivel físico (cuya duración depende de la velocidad a la que se esté transmitiendo⁹). La duración de una ranura que contiene una colisión, donde la mayor longitud que se transmite es l , viene dada por:

$$T_c(l) = T_{PLCP} + \frac{H+l}{C} + EIFS \quad (2.2)$$

Siendo EIFS la suma del tiempo necesario para transmitir una trama de asentimiento y el tiempo de un SIFS.

2.2.1.1. Operación RTS/CTS

El estándar define el uso de un mecanismo, opcional, para usar antes de la transmisión de una trama entre dos estaciones. Dicho mecanismo, ilustrado en la Figura 2.3, consiste en que el transmisor primero envía una trama denominada RTS (Request To Send), a lo que el receptor reacciona, pasado un tiempo SIFS, con el envío de una trama CTS (Clear To Send). De esta forma, se consiguen tres objetivos:

⁸La duración de una ranura vacía viene determinada por el tiempo necesario para detectar la existencia o no de transmisiones en la WLAN, y debe tener en cuenta el retardo de propagación, el tiempo para pasar del estado de recepción al de transmisión, y el tiempo para indicar desde el nivel físico la existencia de transmisiones.

⁹Para 802.11b, todas las estaciones deben soportar un preámbulo de 192 μs , si bien existe la opción de emplear uno más corto (96 μs) para las velocidades superiores a la mínima.

- Por una parte, que un nodo fuera del alcance –esto es, *oculto*– del transmisor (Estación Ao en la figura) o fuera del alcance del receptor (Estación Bo) puedan ser conscientes de la existencia de la comunicación, absteniéndose de transmitir. Para ello, tanto RTS como CTS contienen un campo que define la duración de la comunicación (NAV, Network Allocation Vector).
- Por otra parte, caso de existir colisiones, éstas serían únicamente en las tramas RTS, de longitud relativamente corta. A partir de una cierta carga en la red, la ineficiencia introducida por este diálogo previo se compensa por la menor duración de las colisiones (a este respecto véase, por ejemplo, [26, 27]).
- Por último, el RTS/CTS permite la coexistencia de diferentes niveles físicos (802.11b y 802.11g) en la misma WLAN, debido a que las tramas de control son transmitidas con una modulación entendida por todas las estaciones, que de esta forma son capaces de obtener el NAV y marcar el medio como ocupado (aunque luego no pudiesen detectar que se está llevando a cabo una transmisión).

2.2.1.2. Carencias de 802.11

El mecanismo de acceso DCF proporciona justicia a largo plazo, entendiendo ésta como que todos los nodos tienen la misma probabilidad de ocupar una ranura de tiempo con una transmisión propia. Si bien esa justicia puede parecer deseable en un principio, ello tiene una sutil consecuencia (analizada por Heusse *et al.* [28]): en una red con varios esquemas de modulación y, por tanto, diferentes velocidades de transmisión, basta con que una estación transmita a la velocidad mínima para que todas las estaciones reciban prácticamente el mismo ancho de banda que aquélla¹⁰. Esto se ha denominado *performance anomaly* (anomalía de rendimiento), y supone una degradación de prestaciones muy severa: en una WLAN que dé servicio a una amplia zona, basta con que una estación experimente un canal radio débil para que todas las estaciones disminuyan muy apreciablemente su tasa de transferencia.

Por otra parte, los principales motivos por los que 802.11 no puede proporcionar calidad de servicio (ni en el modo DCF, ni en el PCF) son los siguientes [29]:

- El modo DCF no proporciona mecanismo de diferenciación de tráfico, pues todas las estaciones recibe el mismo tipo de servicio *best-effort*.
- El modo PCF, de escasa difusión (lo que ya constituiría una limitación drástica), tampoco proporciona diferenciación de flujos. El inicio del periodo de polling por parte del AP no es inmediato, siendo difícil limitar el tiempo que dedica una estación a sus transmisiones (si bien esto se ha corregido con HCCA).

2.2.2. El modo EDCA

El acceso distribuido o basado en contienda de 802.11e se denomina EDCA (Enhanced Distributed Channel Access). Parte del modo DCF, pero extendiéndolo en varios aspectos:

¹⁰El principal motivo es el aumento de la duración de las ranuras de tiempo debido a la existencia de estaciones *lentas*. Si bien el contador de backoff sigue siendo el mismo, el canal aparece ocupado durante mucho más tiempo.

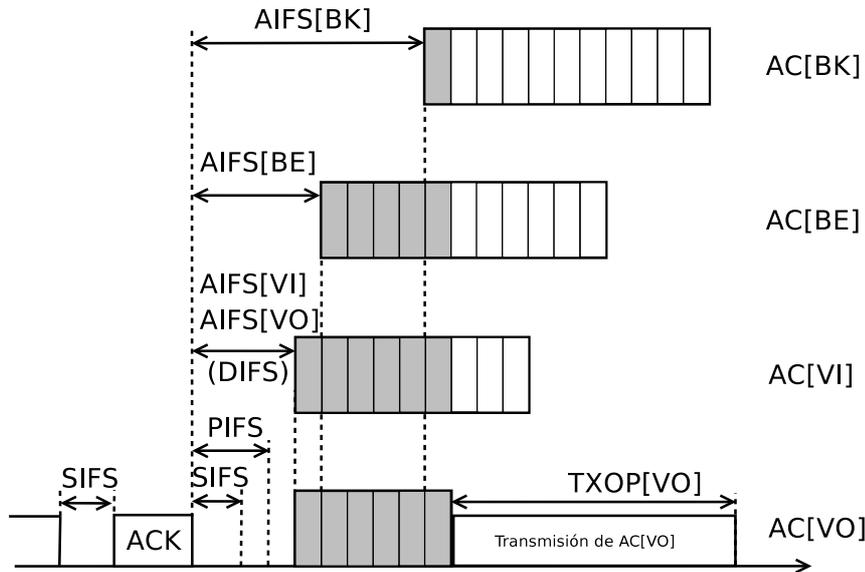


Figura 2.4: Parámetros del modo EDCA

por una parte, cada estación implementa hasta cuatro categorías de acceso (AC)¹¹ independientes, cada una de las cuales asociada a un determinado tipo de tráfico. El nombre que se da a cada AC indica el tipo de aplicación que se supone que ocupará la misma: voz (VO), vídeo (VI), *best effort* (BE) y *background* (BK). Si durante la ejecución del backoff se produjese una colisión interna entre dos AC's de la misma estación (llamada *colisión virtual*), la cola con más prioridad transmitiría la trama al medio físico, mientras que la otra cola se comportaría como si hubiese sufrido una colisión.

Por otra parte, el proceso de backoff también se ve modificado, por la introducción de unos parámetros (diferentes para cada AC) que afectan al comportamiento del mismo. Estos parámetros se ilustran en la Figura 2.4:

TXOP (Transmission Opportunity): Una vez que una estación ha conseguido acceder al medio, tiene derecho a utilizar el canal durante un tiempo menor o igual al TXOP (teniendo en cuenta los asentimientos). Esto permite que se puedan transmitir varias tramas durante la misma ranura de tiempo, existiendo la posibilidad de realizar un asentimiento por cada trama, o un único asentimiento final (opción mucho más eficiente en ausencia de errores, como se muestra en [30]). Si el valor de este parámetro es cero supone que la estación utilizará el canal el tiempo necesario para transmitir una única trama (como en DCF). Si el tiempo necesario para la transmisión de una trama superase el valor indicado por este parámetro, la trama debe ser fragmentada.

CW (Contention Window): El funcionamiento del retroceso exponencial binario es muy

¹¹Se empleará indistintamente los siguientes nombres: cola, clase, categoría de acceso (AC) o función de acceso al canal (CAF).

Tabla 2.3: Configuración estándar para EDCA con 802.11b

Parámetro	AC[VO]	AC[VI]	AC[BA]	AC[BE]
A_i	0	0	1	5
CW_{min}	8	16	32	32
CW_{max}	16	32	1024	1024
TXOP	3 ms	6 ms	3 ms	0

similar al mecanismo DCF, salvo por una diferencia: tanto el valor de CW_{min} como el de CW_{max} puede ser diferente para cada clase de tráfico¹².

AIFS (Arbitration InterFrame Space): En el modo DCF, tras una ranura de tiempo ocupada las estaciones deben esperar DIFS hasta el siguiente decremento de su contador de backoff. En el modo EDCA, el tiempo que debe esperar una estación hasta decrementar dicho contador es diferente para cada cola de tráfico, y viene determinado por este parámetro. Este parámetro toma la forma $AIFS = SIFS + nT_e$, con $n \geq 2$, o bien $AIFS = DIFS + A_iT_e$ con $A_i \geq 0$ (para DCF, $A_i = 0$).

Los valores recomendados para los anteriores parámetros dependen del nivel físico que se esté empleando. Para el caso de 802.11b, se recomiendan los indicados en la Tabla 2.3. El efecto del uso de dichos parámetros es, a grandes rasgos, el siguiente:

TXOP: El términos absolutos, valores elevados del TXOP aumentarán la eficiencia de la red, pues el tiempo dedicado a ocupar una ranura será cada vez menor en comparación con la duración de dicha ranura. Sin embargo, un exceso de dicho valor puede ocasionar un incremento del retardo para el resto de ACs. En términos de diferenciación entre flujos, para dos ACs con idéntica configuración y un nivel de aplicación voraz, la relación entre anchos de banda conseguidos será proporcional al cociente entre valores de TXOPs.

CW: El comportamiento de la CW se puede analizar casi exclusivamente a través del valor de CW_{min} , dado que el efecto del parámetro CW_{max} sólo se aprecia en situaciones de alta congestión (con un número elevado de colisiones). Valores muy bajos de CW_{min} , por una parte, propiciarán altas probabilidades de colisión, por lo que el rendimiento será escaso. Por otra parte, valores elevados de CW_{min} evitarán las colisiones a costa de esperar un número excesivo de ranuras de tiempo: es por ello que existirá, en cada situación, un valor *óptimo* de ventana de contienda.

En cuanto a los efectos relativos en la configuración de CW_{min} , resulta claro que el tiempo que debe esperar una estación antes de transmitir es proporcional al número medio de ranuras de tiempo que debe contar; por ello, la relación entre anchos de

¹²Algunas líneas de investigación propusieron modificar el comportamiento de la ventana de contienda: en primeras versiones del estándar se proponía poder alterar el multiplicador empleado tras una colisión, denominado PF (para DCF, $PF = 2$), si bien su uso fue rápidamente desaconsejado (véase, por ejemplo, [31]). Otros trabajos [32] proponían ajustar tanto el valor máximo de la ventana de contienda utilizado como el valor mínimo –hacerlo mayor que 0–, con objeto de evitar la posibilidad de que estaciones poco prioritarias pudiesen obtener valores bajos de backoff. Ninguna de estas alternativas fue finalmente adoptada en el estándar.

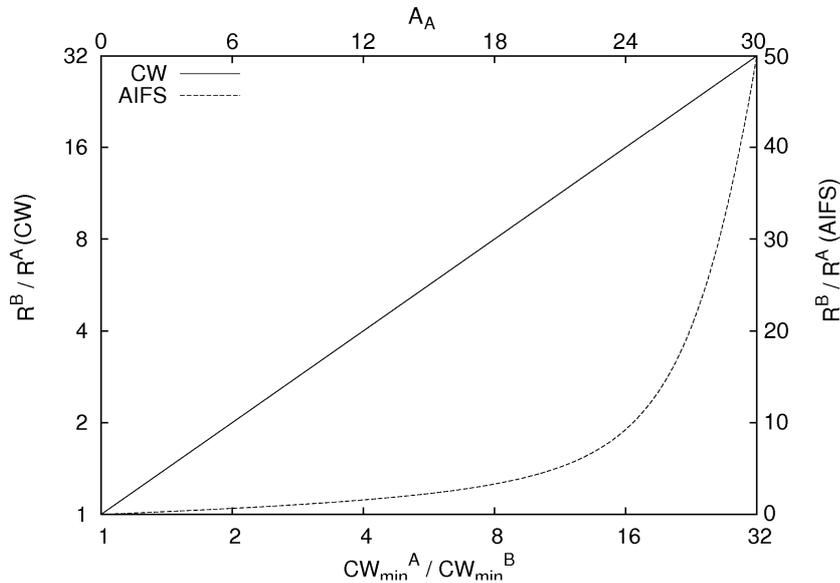


Figura 2.5: Diferenciación proporcionada por EDCA

banda de dos ACs será inversamente proporcional al cociente del valor de CW_{min} de dichas ACs –véase Figura 2.5– (para una demostración analítica puede acudir, por ejemplo, a [33]).

AIFS: El uso de AIFS no proporciona en términos absolutos ninguna ventaja, dado que un valor común y elevado para todas las ACs únicamente aumentaría el tiempo que la WLAN permanece ociosa. En cambio, en términos relativos (para la diferenciación de flujos) su efecto es bien diferente: dado que cada vez que se ocupa el medio una estación debe esperar A_i ranuras vacías antes de volver a decrementar el contador de backoff, en situaciones de baja ocupación su comportamiento será poco apreciable. En cambio, en situaciones de alta ocupación el número de ranuras llenas será elevado, hecho que postergará las transmisiones de los ACs de mayor A_i , lo que a su vez motivará que el resto de ACs vuelva a transmitir con mayor facilidad (produciéndose una realimentación). Es por ello que esta diferenciación, en vez del carácter *lineal* de las anteriores, presenta un carácter *exponencial*. Una ilustración del comportamiento relativo en ancho de banda (esto es, del cociente entre tasas de transmisión obtenidas R^B/R^A) para sólo dos estaciones A y B (con $A_B = 0$) se ilustra también en la Figura 2.5 (tomada de [34]).

2.3. Estudio de prestaciones mediante simulación

En esta sección se presentan los trabajos más relevantes que han abordado la caracterización del DCF y EDCA mediante simulación, paso que permite una primera valoración cualitativa y cuantitativa de los mecanismos de acceso al medio.

2.3.1. 802.11 DCF

Anastasi y Lenzini [26] proporcionan un detallado análisis mediante simulación de las prestaciones del modo DCF de 802.11, proporcionando para diferentes patrones de llegada (modelados según estudios experimentales) los tiempos de servicio y espera en cola, así como el ancho de banda cursado por la red. Abogan por el uso del mecanismo de RTS/CTS (si bien es en el modelado analítico de [27] cuando se establece de forma justificada en qué condiciones aplicarlo), así como por mantener en 15 el número máximo de estaciones en la WLAN.

2.3.2. 802.11e EDCA

Grilo y Nunes [35] realizan una serie de simulaciones, para fuentes de datos según modelos de tráfico, que muestran la adecuación del nuevo estándar para la provisión de servicios diferenciados a las aplicaciones. Su trabajo se centra en realizar una serie de comparaciones entre el funcionamiento del estándar anterior y las mejoras que introduce el 802.11e, para lo que compara tanto PCF+DCF con HCCA+EDCA como DCF con EDCA en la provisión de QoS para sesiones de tiempo real.

Lindgren *et al.* [36] realizan un estudio por simulación del modo PCF y EDCA (para una configuración determinada de ambos), así como de DFS [37]¹³ y de *Blackburst* [38,39]¹⁴. En dicho trabajo, se muestra que si bien es *Blackburst* el mecanismo que proporciona el mejor servicio al tráfico de tiempo real para una carga notable de éste, la alternativa propuesta en el estándar es una opción razonable –si bien no existe un claro vencedor para todas las circunstancias–.

He y Shen [40] realizan un estudio de acuerdo con las primeras recomendaciones del borrador del estándar para la configuración de las diferentes colas. Se muestra que, para esos valores y en situaciones de baja carga de tráfico, el EDCA proporciona un retardo y ancho de banda en media muy similares a aquellos proporcionados por DCF. Sólo a partir de un cierto nivel de ocupación el mecanismo diferenciador de EDCA se hace notar, llegando al punto de producir inanición de los flujos con baja prioridad de acceso.

La contribución de Mangold *et al.* [41] es doblemente notable: por un lado, proporciona una descripción del nuevo estándar amena y completa; por otro, realiza una evaluación de los parámetros de configuración propuestos en el estándar para las diferentes colas, que muestra de forma cuantitativa el grado de diferenciación que se ofrece a los diferentes tipos de servicio (ilustrándose de nuevo la inanición que puede ocasionar el servicio de voz al resto de flujos).

En la misma línea que el anterior se sitúa el trabajo de Xiao [42], aunque bien es cierto que la descripción del estándar resulta algo menos completa, y sólo muestra unos resultados preliminares que ilustran la diferenciación que proporciona EDCA. Sí resulta bien interesante el trabajo de Choi *et al.* [43], en el que se compara el rendimiento de DCF y EDCA para el mismo escenario, compuesto por cuatro estaciones de datos, dos estaciones de vídeo y cuatro estaciones de voz. En estas condiciones (con un tráfico total ofrecido de 7

¹³Propuesta de implementación de *fair queuing* en entornos distribuidos, basado en la modificación del contador de backoff según pesos y tamaños de paquete.

¹⁴Propuesta que intenta que las aplicaciones real-time utilicen el medio de forma TDM, incluyendo mecanismos de transmisión de *pulsos de energía* (a nivel físico) de duración proporcional a la prioridad.

Mbps), se aprecia que, si bien la voz recibe un servicio razonable en ambas configuraciones (mucho más preferente con EDCA), es el vídeo el que se ve altamente beneficiando por la diferenciación ofrecida por 802.11e. También se ilustra, en el mismo trabajo, las ventajas de transmitir varias tramas de forma consecutiva a través del parámetro TXOP.

Una simulación de prestaciones bien interesante, por usar los valores recomendados en el estándar (si bien sólo para tres clases) la realiza Ni [29]. Se muestra la abrupta diferenciación lograda por los mecanismos de AIFS y diferenciación en la CW, así como la eficacia de un algoritmo (Adaptive EDCA [44]) que estima con una media exponencial la probabilidad de colisión de cada estación, y con ella calcula la configuración de la CW más adecuada (para lo cual precisa de un modelo analítico, tema de la siguiente sección).

2.4. Estudio de prestaciones mediante modelo analítico

Ya desde los años setenta el estudio del ancho de banda y retardo de los mecanismos de acceso al medio basados en contienda ha sido objeto de gran interés por la comunidad investigadora. Con la estandarización y rápido despliegue de las redes 802.11 en los años 90, el modelado teórico de su capa MAC de WLAN ha sido un fuerte centro de interés que llega hasta nuestros días.

Con objeto de presentar adecuadamente el modelado analítico de los mecanismos de acceso distribuidos en redes 802.11, se realizarán dos distinciones:

- Por una parte, se distinguirá entre los análisis del modo DCF y los análisis del modo EDCA. Si bien la tesis propuesta no aspira a emplear el modo DCF, la gran mayoría de los análisis del modo EDCA parten de los resultados obtenidos en aquél, por lo que resulta ineludible su presentación.
- Por otra parte, y por motivos similares, se distinguirá entre análisis basados en saturación, y aquellos que no suponen condiciones de saturación. Suponer condiciones de saturación implica, como se ha explicado anteriormente, suponer que todos los nodos siempre tienen un paquete en cola esperando a ser transmitido. Muchos análisis en condiciones de no saturación bien extienden el modelo de saturación, bien emplean resultados obtenidos para la saturación.

2.4.1. 802.11 DCF

2.4.1.1. Modelado en saturación

El trabajo de Chhaya y Gupta [45] analiza y simula el mecanismo DCF. Suponen que el tiempo entre un DIFS y la generación de un paquete por una estación se distribuye de forma exponencial, obteniendo la expresión del ancho de banda total cursado por la red inalámbrica para la transmisión de paquetes de tamaño fijo, con el uso de RTS/CTS y para paquetes de tamaño variable (en tres modelos sustancialmente diferentes). Los resultados analíticos se aproximan razonablemente a los valores de simulación, resultando especialmente destacable la aparición de inestabilidades en el ancho de banda cursado en función del tráfico inyectado a la red. También es destacable el hecho de considerar la existencia de terminales ocultos o el efecto captura.

Bianchi [27, 46] supone uno de los principales trabajos en el modelado del mecanismo de acceso al medio DCF. Su contribución se basa en dos suposiciones: primero, que la probabilidad de transmisión en una determinada ranura de tiempo es constante e idéntica para todas las estaciones, independientemente del estado de su proceso de backoff; segundo, que la probabilidad de colisión de una determinada transmisión también es constante e independiente de cualquier otra variable temporal del modelo. Dadas estas dos suposiciones sobre estas dos variables (la probabilidad de transmisión, τ , y la probabilidad de colisión, p_c), construye una cadena de Markov que modela el comportamiento de una estación en saturación. De esta forma es capaz de expresar τ como función de p_c , y p_c como función de τ , en una ecuación no lineal que se resuelve numéricamente. Una vez obtenidas las probabilidades de transmisión (y colisión), resulta inmediato derivar expresiones para la tasa obtenida por cada estación. Adicionalmente, deriva la configuración de CW que maximiza dicho ancho de banda. El análisis de Bianchi permite obtener el tiempo medio de servicio, calculando el número medio de ranuras de tiempo que debe contar una estación y multiplicándolo por la duración media de éstas (que es lo que hace Chatzimisios *et al.* en [47]).

El trabajo de Calì *et al.* [48, 49] modela el funcionamiento del mecanismo DCF mediante un protocolo de acceso al medio p-persistente, de modo tal que el número de ranuras de backoff se puede aproximar mediante una distribución genérica con parámetro $p = 1/(E[B] + 1)$, siendo $E[B]$ la esperanza del contador de backoff. Su análisis se basa en el *tiempo de transmisión virtual*, siendo éste el tiempo necesario para que se produzca el envío con éxito de una trama en el canal físico. Además, propone una modificación de la capa MAC del modo DCF para aproximarse al máximo teórico de eficiencia en el uso del canal.

Tay y Chua [50] abordan el problema del modelado analítico en saturación de una forma más sencilla que los anteriores, bajo la justificación de que un modelado analítico debería proporcionar expresiones de poca complejidad que permitan ser analizadas fácilmente (si bien es discutible que logren dicho objetivo). Su modelo supone una probabilidad de colisión constante e independiente al igual que [27], pero para el cálculo de la probabilidad de transmisión emplean el valor medio de la ventana de contienda (en vez del modelado mediante de la cadena de Markov). De esta forma se llega a resultados próximos a los obtenidos mediante simulación, así como a valores de configuración que optimizan la tasa de transmisión de la red, resultados más exactos cuanto menor es el número de estaciones en la red.

La aportación de Wu *et al.* [51] es doble: por un lado, extiende el modelo de [27] para tener en cuenta la existencia del número máximo de retransmisiones R , lo que mejora los resultados para un número elevado de nodos en la red (más de 20). Por otra parte, propone un mecanismo (denominado $DCF+$) para mejorar las prestaciones de TCP sobre una red 802.11, siendo éste un mecanismo basado en el método RTS/CTS que pretende dar prioridad a la transmisión del asentimiento del nivel de transporte (dicho mecanismo no ha sido adoptado).

De los anteriores trabajos se puede obtener el ancho de banda disponible para una estación en una red WLAN con todos los nodos saturados. Dado este resultado, obtener el tiempo que pasa en media desde que un paquete inicia el proceso de backoff hasta que resulta transmitido con éxito es trivial (como se realizó en [47]). La aportación de Banchs [52]

consiste en modelar analíticamente la función de distribución de este tiempo de servicio, y no sólo su valor medio, de modo tal que se puede obtener la probabilidad de que dicho tiempo de servicio esté por debajo de un determinado valor, siendo éste un paso importante en la obtención del retardo total de transmisión.

Posteriormente, Bianchi y Tinnirello [53] derivan una nueva forma de analizar el modo DCF sin necesidad de emplear una cadena de Markov, obteniendo que la probabilidad τ sólo depende del valor medio del contador de backoff en cada intento de transmisión. También señalan un error en el modelado del decremento del contador de backoff, que aparece tanto en uno de los primeros en extender el trabajo de Bianchi [54] como en [55]. No obstante, dicho fallo (indicado también por Foh y Tantra [56]) no supone errores significativos en la obtención de las prestaciones del sistema; tanto es así, que los autores lo mantienen, de forma explícita, en [57]. Por último, Kumar *et al.* [58] realizan un nuevo estudio del punto de operación de un protocolo con mecanismo genérico de retroceso, estableciendo las condiciones para que dicho punto sea único.

2.4.1.2. Modelado en no saturación

Para modelar estaciones que no siempre tienen una trama pendiente de ser transmitida, Ahn *et al.* [59] introducen una variable p_{on} que modela la probabilidad (fija y constante) de que una estación esté ejecutando el proceso de backoff. De esta forma, multiplican la τ del modelo de Bianchi por dicha probabilidad, que debe ser estimada por simulación —dado que depende de la carga en el sistema—, lo que reduce enormemente su utilidad.

Con la suposición de llegada de tramas según un proceso de Poisson, Zaki y El-Hadidi [60] proponen un modelo para la no saturación que parte del trabajo de Bianchi, introduciendo un estado de no actividad (*idle*) al que se llega tras la transmisión de una trama. Los autores defienden que dicho modelo sería fácilmente aplicable para considerar otro patrón de llegadas —aunque dicha afirmación sea discutible, dada la simplicidad que supone la falta de memoria de los procesos de Poisson—. Adicionalmente, su modelado se basa en una única longitud de transmisión, así como en el uso del mecanismo de RTS/CTS.

Otra variación del modelo de Bianchi, en la misma línea, consiste en añadir varios estados ociosos a la cadena de Markov, con objeto de representar la espera hasta la llegada de una nueva trama que transmitir. Es el caso de Ergen y Varaiya [61] que, adicionalmente, abogan por la necesidad de un control de admisión que evite la disminución de prestaciones en la red debido a la congestión.

Malone *et al.* [62] realizan un modelado que también parte de la cadena de Bianchi, extendida con el periodo de *post backoff*¹⁵. Se introduce adicionalmente la probabilidad $1 - q$ de que el buffer de transmisión esté vacío al inicio del decremento del contador de backoff (que viene definida para cada estación según su patrón de generación de tráfico). Los resultados obtenidos, para ancho de banda y probabilidad de colisión, se aproximan razonablemente a los valores de la simulación.

El trabajo de Foh y Zukerman [63] emplea una cadena de Markov, pero de forma diferente a los anteriores. La idea básica de su modelo es la siguiente: cada estado de la

¹⁵Según el estándar, una estación que acaba de dar servicio a una trama, debe ejecutar una nueva instancia del proceso de backoff aunque no tenga paquetes en cola, no pudiendo transmitir trama alguna —aunque llegase— hasta que dicho proceso haya finalizado. Esto evita que estaciones con valores elevados de CW compitan en desigualdad de condiciones con estaciones que han logrado una transmisión exitosa.

cadena indica el número de flujos activos en el sistema. El tiempo de servicio se obtiene de la tasa en saturación (para ese número de fuentes) de acuerdo con [27], mientras que la tasa de llegadas dependerá del número de usuarios que estén siendo atendidos (modelo de población finita). Para los valores de la capa física considerada (1 Mbps), muestran que la distribución del tiempo de servicio no se puede aproximar mediante una exponencial (para el caso considerado, la mayor aportación al tiempo de servicio viene dada por el tiempo que se emplea en la transmisión, siendo casi despreciables los efectos del proceso de backoff). Una aplicación inmediata de este trabajo la realizan Patil y Apte [64], calculando para un tráfico de entrada y unas condiciones de red a partir de qué número de usuarios se produciría saturación y, por lo tanto, la red no podría cursar todo el tráfico ofrecido. Los resultados resultan razonablemente ajustados (errores del 13 % en el número máximo de estaciones) para el un tráfico HTTP 1.0.

Los mismos autores presentan en [65] un criterio alternativo para medir las prestaciones de una red 802.11. En vez de analizar la tasa ofrecida en el estado permanente, miden el *tiempo de recuperación*, siendo éste el tiempo medio que se tarda desde una situación en la que todos los nodos comienzan con un único paquete para transmitir, hasta que todo el tráfico ha sido cursado. Para ello, modelan el escenario con una cadena bidimensional, donde cada estado se representa por el número de estaciones que aún mantienen el paquete pendiente de ser transmitido, y por el número de ranuras de tiempo pasadas desde el inicio de la situación.

2.4.2. 802.11e EDCA

2.4.2.1. Modelado en saturación

Los análisis del modo EDCA se puede agrupar según el mecanismo de diferenciación que hayan considerado: bien sólo la CW, bien sólo el AIFS, o bien ambos mecanismos.

Análisis de la CW: El trabajo de Ge y Hou [66] consiste en una extensión del modelo de Cali *et al.* [49] que considera distintas probabilidades de transmisión p_i en función de la prioridad a la que pertenezca el flujo. Este análisis es validado mediante la simulación de una capa MAC que implementa el modelo de p-persistencia.

Qiao y Shin [67] extienden el modelo de Bianchi para considerar diferentes valores de CW_{min} con objeto de proporcionar equidad a través de dicho mecanismo. De esta forma, se llega a que las ventanas deben ser proporcionales a los pesos de cada flujo. Adicionalmente, con objeto de establecer valores concretos para cada ventana, se realiza un análisis que motiva la propuesta de un protocolo de ajuste dinámico de la ventana de contienda en función de el estado de la WLAN.

El análisis de Xiao [68] realiza el modelado de cualquier tipo de comportamiento de la ventana de contienda: el tamaño mínimo, el factor de incremento y el máximo número de retransmisiones. Parte del modelo de Ziouva y Antonakopoulos [54], a los que le añade un número finito de reintentos. Los resultados los presenta para dos clases de tráfico (si bien permite varios niveles de prioridades), con un tamaño fijo de los paquetes y para un número igual de estaciones en cada clase (comenzando por cinco estaciones).

Banchs y Vollero [69] emplean un modelo analítico con diferentes valores de τ para proporcionar la mejor configuración de acuerdo al criterio de *weighted max-min*

*fairness*¹⁶. Demuestran que la diferenciación mediante el parámetro AIFS proporciona unas prestaciones -de acuerdo al criterio establecido- iguales o inferiores a las obtenidas sin necesidad de dicho parámetro, mientras que el valor del TXOP se considera impuesto según algún criterio (por ejemplo, el retardo); de esta forma, tras la validación del modelo, se obtiene la configuración de las ventanas de contienda que optimiza el funcionamiento de la WLAN.

Análisis del AIFS: Bianchi y Tinnirello [71] no emplean cadenas de Markov en el modelado analítico de la diferenciación del parámetro AIFS (lo cual resulta bastante original, como se verá a continuación). Se basan en la definición de la probabilidad acumulada del contador de backoff de cada estación (teniendo en cuenta las diferencias entre cada clase), para lo cual es preciso clasificar el número de ranuras en función de la diferencia Δ entre el AIFS de cada tipo de estación. Se trata de un análisis que proporciona resultados muy ajustados al caso de la simulación, pero que adolece de una cierta complejidad incluso para el caso de solo dos ACs.

Análisis de CW y AIFS: El trabajo de Zhao *et al.* [72] extiende el *clásico* modelado mediante la cadena de Markov para tener en cuenta la diferenciación mediante AIFS. De esta forma, se consideran tres dimensiones: el valor actual de la ventana de contienda, el número de retransmisión para la actual trama considerada, y el número de ranuras de tiempo desde la última transmisión. Con este modelado se obtiene una formulación compleja que proporciona resultados bastante ajustados a la simulación.

Battiti y Li [73] también plantean la extensión de la cadena bidimensional de Markov propuesta por Bianchi [27]. En dicho artículo, se modela la clase con menor AIFS mediante el modelo [27], mientras que para las estaciones con un AIFS mayor se añade una nueva dimensión a la cadena (como en el trabajo anterior), para tener en cuenta las ranuras de tiempo que se introducen en dicha diferenciación. Los resultados obtenidos (para unas condiciones de trabajo con un número de estaciones elevado) se aproximan bastante a los resultantes de la simulación.

Otra aproximación al problema también basada en el modelado mediante cadena de Markov es la realizada por Zhu y Chlamtac [31], que presenta la particularidad de modelar el efecto diferenciador del AIFS mediante un parámetro, λ_n , que define la tasa a la que una determinada estación decremента su contador de backoff (siendo función de la diferencia con el AIFS de la clase de máxima prioridad y de la probabilidad de que ésta ocupe el medio). También analizan el efecto de diferentes políticas en el crecimiento de la ventana de contienda, comprobándose que la diferenciación es similar a la proporcionada por el parámetro CW_{min} .

Hui y Devetsikiotis [4] presentan, en primer lugar, una destacable taxonomía de las principales formas de plantear el modelo de EDCA, distinguiendo tres paradigmas: la cadena de Markov [27], el modelo de p-persistencia [49] y el estudio del valor medio del proceso de renovación [50]. Luego proponen un nuevo modelo analítico para calcular la tasa de

¹⁶La aplicación de este criterio [70] resulta en lo siguiente: dada una situación con un determinado número de estaciones asignadas a diferentes categorías de acceso, y unos pesos según los cuales distribuir el ancho de banda de forma proporcional, se encuentra la configuración de la CW que maximiza el mínimo del cociente r_i/w_i , siendo r_i la tasa de transmisión que recibe una estación de la clase i y w_i el peso de dicha estación.

transmisión y el retardo del modo EDCA. Dicho modelo se basa en dividir la complejidad del modelado por estación según la actuación del parámetro AIFS, para luego integrar las expresiones mediante una matriz de probabilidad de transmisiones y otra de probabilidad de transiciones. De esta forma, extienden la cadena de Markov para tener en cuenta los efectos de un valor diferente de AIFS, si bien en vez de aumentar la dimensionalidad de la cadena emplean diferentes versiones de la misma.

Partiendo del modelo de Bianchi, Robinson y Randhawa [1] construyen un modelo para el modo EDCA en condiciones de saturación. Para ello, suponen la constancia e independencia de la probabilidad de colisión en cada *zona de contienda*, definidas ésta como aquel estado en que sólo participan las estaciones con un determinado AIFS. De esta forma definen otra cadena de Markov para la probabilidad de estar en una determinada zona de contienda, que depende de las probabilidades de transmisión. Sin embargo, en su evaluación de resultados emplean AIFS que no son los sugeridos por el estándar, comparándose únicamente las prestaciones por pares de ACs.

El análisis de Banchs y Vollero [2], al igual que [71], evita la extensión de la cadena de Markov de [27], abordando el problema del modelado del EDCA (parámetros CW_{min} , CW_{max} y AIFS) mediante la definición de k -ranuras (k -slots). Dichas ranuras de tiempo son las separaciones entre decrementos del contador de backoff de las estaciones con un $AIFS = SIFS + kT_e$ (lo que resulta similar a las zonas de contienda de [1]).

2.4.2.2. Modelado en no saturación

Tantra *et al.* [57] parten del resultado de Foh y Zukerman [63] para emplearlo en el modelado de la diferenciación mediante CW (para lo que utilizan las probabilidades de transmisión de Xiao [55]). Empleando la fórmula de Little (ya usada por Bianchi [53]) son capaces de obtener el retardo y ancho de banda para dos prioridades (alta y baja, sin relación con los valores propuestos en el estándar), si bien los resultados lo obtienen para tamaños fijos de tramas por cada clase y con uso de RTS/CTS (lo que mantiene el tamaño fijo de las colisiones).

El análisis para no saturación de Engelstad y Østerbø [74] parte del modelado de Xiao [68], extendiéndolo para cubrir el rango entre no saturación y saturación. Tiene en cuenta la etapa de *post backoff* existente tras la transmisión de un paquete, y considera la diferenciación tanto por CW como por AIFS (para un tamaño fijo de los paquetes). Adicionalmente, proponen un mecanismo para que los APs puedan detectar en qué momento se va a producir inanición (*starvation*) en la red, de modo tal que se puedan tomar las acciones oportunas (reconfiguración de parámetros, control de admisión, etc). Su modelo parte de una cadena de Markov, similar a la de Bianchi pero extendida mediante un factor de utilización ρ , que considera la no saturación de las estaciones. Para el modelado del AIFS, en cambio, introducen la variable A_i en la probabilidad de cambiar de estado en la cadena de Markov. A pesar de esto último el modelado sigue siendo de cierta complejidad, obteniéndose resultados analíticos que siguen el comportamiento mostrado mediante simulación (en la que las cuatro clases inyectan la misma tasa a la red).

2.5. Control de admisión

Con la llegada del nuevo estándar 802.11e y la posibilidad de suministrar QoS a las aplicaciones, surge un creciente interés por el desarrollo de mecanismos de control de acceso al medio. Siguiendo la clasificación de Gao *et al.* [75], se puede distinguir entre esquemas basados en medición y esquemas basados en modelo: en los primeros, las decisiones sobre aceptar o no un nuevo flujo se basan en estimaciones de ancho de banda y retardo, mientras que los segundos construyen ciertas métricas partiendo de modelos analíticos para decidir la admisión de un nuevo flujo.

2.5.1. Basado en medición

Un primer esquema es el denominado DAC (Distributed Admission Control) [76], consistente en anuncios por parte del AP del tiempo de transmisión disponible para cada AC. De esta forma, las estaciones pueden ajustar el tiempo que deben transmitir durante el siguiente intervalo, administrando así el ancho de banda de sus conexiones; las conexiones existentes sólo pueden aumentar su reserva si existe tiempo disponible, precisándose de factores de amortiguamiento y crecimiento. Las principales complicaciones de este esquema son dos: la inherente inestabilidad por el ajuste dinámico del tiempo a transmitir, y la falta de relación directa entre el ajuste del tiempo y los parámetros de QoS a provisionar.

Continuando con el mecanismo DAC, Xiao *et al.* proponen una extensión [77] basada en el establecimiento de dos niveles de protección: uno para mantener la calidad de las conexiones existentes frente a otros flujos que también requieren QoS, y otro para proteger los flujos QoS de los flujos de tipo *best effort* —dado que la presencia de un elevado número de flujos de este último tipo aumentan las colisiones en la red, degradando el funcionamiento general—. El primer mecanismo se basa en las mismas ideas que DAC, mientras que el segundo ajusta el valor de la CW_{min} y $AIFS$ de las estaciones de baja prioridad. Este esquema mantiene las mismas debilidades que el anterior: oscilación, y falta de relación entre parámetros de QoS y mecanismos empleados; adicionalmente, dado que requiere un número mayor de variables para operar, resulta complicado encontrar los parámetros óptimos de funcionamiento.

Otra alternativa es el uso de un nivel MAC *virtual* (VMAC) [78, 79], consistente en el uso de una instancia no real tanto del nivel de aplicación como del nivel de acceso al medio. De esta forma, se estiman las condiciones del servicio que se obtendría, pudiendo realizarse el control de acceso (esto es: se trata de un mecanismo de *probing* sin consumo de recursos en la red). Esto requiere estimar la probabilidad de colisión, caso de que el paquete fuese realmente enviado, para poder ejecutar el mecanismo de retroceso exponencial binario. Si bien este esquema resulta atractivo en términos del consumo de recursos en la red, supone introducir cambios en el nivel MAC, resulta costoso computacionalmente y la validez de los resultados es aproximada.

Iera *et al.* [80] refinan la idea del VMAC, utilizando una media exponencial para estimar el retardo medio que obtendría una determinada AC. Con esta información, proponen un mecanismo de asignación dinámica de flujos a AC, comparándolo con la configuración estática¹⁷ propuesta por el estándar. En un trabajo posterior [82] prueban el uso de dicho

¹⁷La naturaleza estática de la configuración propuesta, así como el efecto de los cambios en el canal radio, es

mecanismo en redes WLAN de varios saltos (*multihop*).

Un ejemplo de esquema basado en umbrales lo proporcionan Gu y Zhang [83], planteando el uso de dos diferentes alternativas: bien usando la ocupación en ancho de banda (con zonas de decisión: admitir una nueva conexión, no hacer nada o posponer la transmisión de un flujo de mínima prioridad), bien usando el valor medio de la tasa de colisión (con el mismo esquema de decisión). En este caso, claramente, las ventajas caen del lado de la simplicidad del esquema, mientras que los inconvenientes son tanto por el criterio para fijar los umbrales, como por la pérdida de prestaciones en escalas cortas de tiempo.

Li *et al.* proponen [84] un esquema heurístico de reasignación de prioridades, basado en la reserva de ancho de banda para las aplicaciones de tiempo real, y con una estimación por simulación del ancho de banda disponible en cada circunstancia. El esquema propuesto se adapta a diferentes algoritmos de planificación, si bien no se presenta una solución cerrada, sino un modelo de arquitectura.

Un último esquema, similar al de Xiao [77], es el propuesto por Zhang y Zeadally, HARMONICA [85]. En él se plantea medir la tasa de pérdidas, retardo y ancho de banda, para las diferentes clases, y con ello ejecutar dos algoritmos en diferentes escalas de tiempo. En el primero se modifican los parámetros de cada AC, con objeto de alcanzar la diferenciación más adecuada (vía CW_{min} , CW_{max} y *AIFS*), mientras que en el segundo, con base en la medida del ancho de banda del sistema, se busca la optimización del funcionamiento global de la WLAN (ajustando la ventana de contienda). El problema estriba en la necesidad de definir cuidadosamente los umbrales de funcionamiento, para evitar tanto las oscilaciones (alta sensibilidad) como los tiempos de respuesta demasiado elevados (baja sensibilidad).

2.5.2. Basado en modelo

El trabajo de Pong y Moors [86] estima las probabilidades de colisión de cada estación mediante una media exponencial. Con éstas, y empleando el análisis de [27], obtienen el ancho de banda que podría recibir cada estación. El algoritmo se ejecuta en el AP, que admite o no un nuevo flujo en función del ancho de banda que resultaría para el resto de estaciones. La principal debilidad de este modelo es asumir condiciones de saturación, lo que sólo resulta válido en muy determinadas circunstancias.

La propuesta de Banchs *et al.* [52] también nace del trabajo de [27], proponiendo un esquema de ajuste del valor de las CW en función de los requisitos de cada tipo de tráfico, optimizándose el uso de la WLAN. Un nuevo flujo, con unos requisitos de ancho de banda dados, sólo es admitido si existe una nueva configuración de CW que mantenga las prestaciones del resto de estaciones. Al igual que el anterior, este trabajo no emplea el mecanismo de diferenciación de *AIFS*, y sólo considera estaciones en saturación.

2.6. Supervisión de usuarios

Con la aparición del modo EDCA, resulta mucho más sencillo para un usuario malintencionado modificar el comportamiento de su interfaz WLAN para obtener un mejor servicio de la red, dado que el estándar requiere que las estaciones puedan modificar los

motivo de análisis por parte de Ramos *et al.* [81].

parámetros de sus ACs. Es por ello que la capacidad para supervisar el comportamiento de los nodos se convierte en una necesidad, cada vez más imperante a medida que los productos 802.11e se van introduciendo en el mercado. Incluso si los usuarios son bienintencionados, es posible que las tarjetas de red no se comporten de acuerdo al estándar, bien por no respetar los tiempos de guarda introducidos, o por no generar según una distribución uniforme el número de ranuras de tiempo a decrementar (véase el trabajo de Di Stefano *et al.* [87]).

Una primera aproximación al problema es la realizada por Kyasanur y Vaidya [88], en el que se propone un mecanismo de detección y castigo del uso fraudulento del MAC: la idea es permitir al receptor asignar valores de backoff al emisor tanto en tramas CTS como en tramas de asentimiento, para posteriormente comprobar que se ajusta al comportamiento; de no ser así, se penaliza a la estación en el siguiente backoff. Los problemas de dicho sistema son obvios: por una parte, exige una modificación del protocolo MAC para el desarrollo del mecanismo; por otra parte, el emisor queda controlado por el receptor, lo que permite la aparición de un nuevo abanico de comportamientos inadecuados.

El trabajo más relevante en la detección de comportamientos fraudulentos es el de Raya *et al.* [11], que presentan *DOMINO*: una batería de pruebas basada en heurísticos diseñados para detectar si una estación está siguiendo las reglas del proceso de backoff. Su trabajo distingue entre dos grupos de técnicas para obtener un mayor ancho de banda: la alteración de tramas de control del protocolo (CTS o ACK) para simular una colisión y forzar un nuevo backoff al resto de estaciones; y la manipulación de los parámetros de funcionamiento del protocolo (SIFS, NAV, CW). Hecha esta clasificación, se propone realizar una serie de pruebas para clasificar una estación como maliciosa o no. El principal problema de *DOMINO* es la ausencia de un mecanismo para fijar criterios en la detección, estando basado en parámetros de tolerancia que se deben fijar a buen criterio del diseñador.

2.7. Resumen

A la vista de lo anteriormente expuesto, el estado de la investigación relacionada con la configuración y provisión de QoS en escenarios EDCA se puede resumir en los siguientes apartados:

Análisis del modo EDCA: El modelado en saturación ha concentrado la atención de la gran mayoría de los análisis. Los modelos de no saturación, por lo general, han supuesto un tiempo entre llegadas exponencial. En muy pocas circunstancias se han considerado distribuciones del tamaño de los paquetes, y pocos artículos han siquiera afrontado la validez de su modelo para los valores de configuración recomendados por el estándar. La mayoría de los análisis, además, se ha centrado en calcular el ancho de banda que recibiría una determinada AC, abandonando cuestiones tan importantes para la provisión del QoS como son el retardo o, incluso, el *jitter*. No se ha afrontado el análisis de un modelo con fuentes no saturadas, en el que se emplee los mecanismos del nuevo estándar para proporcionar una diferenciación que pueda ser caracterizada de forma cuantitativa, mediante un modelado adecuado. Falta, por tanto, un análisis completo de las prestaciones y el funcionamiento de una red WLAN 802.11e en modo EDCA que proporcione una caracterización del tipo de servicio (ancho de banda, retardo, *jitter*) proporcionado a las diferentes categorías de acceso.

Configuración óptima de la WLAN: Debido a las ausencias del anterior apartado, no existe ninguna solución que proporcione la mejor configuración posible del modo EDCA, en función de un tráfico a servir y unos requisitos de QoS a cumplir. La casi totalidad de las soluciones en este ámbito pasan por recurrir a heurísticos basados en estimaciones de la ocupación de la red, lo que, entre otras cosas, provoca oscilaciones en el comportamiento de la WLAN.

Control de admisión: Al igual que los mecanismos de configuración propuestos hasta la fecha, las soluciones para el control de admisión se suelen basar en estimaciones sobre el estado actual de la WLAN. Esto da lugar a que las medidas del rendimiento dependan de las ventanas de tiempo empleadas y del tipo de medición realizada, cuestiones que afectan a los resultados del control ejercido. Así, un esquema basado en medición puede que acepte un elevado número de usuarios en un principio, si estos no inyectan tráfico en exceso (por ejemplo, con bajo índice de actividad de la voz), para luego no proporcionar las garantías de calidad pedidas por las aplicaciones si alguna aumenta la tasa de generación de datos.

Supervisión de usuarios: La posibilidad de modificar los parámetros del nivel MAC del modo EDCA conlleva la posible aparición de usuarios malintencionados que, sin ajustarse a la configuración suministrada por el gestor de la WLAN, actúen en búsqueda de un mejor servicio de forma ilícita (disminuyendo el valor de su CW o su AIFS). Los algoritmos de control existentes hasta la fecha, basados en heurísticos, son claramente insatisfactorios pues, de nuevo, las prestaciones de dichos algoritmos dependen de variables arbitrariamente fijadas.

En la presente tesis se abordan las carencias de los anteriores apartados, realizándose un nuevo análisis en saturación que presenta solución única [5], un análisis del modo EDCA para no saturación [12], un esquema de configuración para redes WLAN abiertas en modo básico [7] y multimodo [9], la caracterización del retardo [13] que permite el desarrollo de un control de admisión y un esquema de supervisión para detectar configuraciones diferentes a las suministradas [10].

Capítulo 3

Configuración de WLAN abiertas

Como se ha explicado en el Capítulo 1, se considera una *WLAN abierta* aquella en la que la configuración no puede basarse en la información que proporcionan los usuarios, debido a la ausencia de una relación de confianza con los mismos. En caso contrario, podría darse la circunstancia de que un usuario solicite una configuración para transmitir VoIP (tráfico muy prioritario pero de baja tasa de transmisión) y que, sin embargo, decida emplear dicha cola para descargar un fichero, provocando la inanición del resto de usuarios (a pesar de cumplir con la configuración de la capa MAC). Debido a esto, la configuración de dichas redes obedecerá exclusivamente a criterios de rendimiento.

En primer lugar, se presenta un nuevo estudio de una WLAN en condiciones de saturación mediante un análisis que, a diferencia del trabajo previo, plantea una única solución. A continuación, y teniendo en consideración el caso de la no saturación, se presenta otro modelo que explica el comportamiento inestable que en ocasiones puede mostrar una WLAN, operando ésta en el modo básico. Dicho comportamiento no aparece cuando la WLAN está configurada de forma que se maximice el tráfico cursado en saturación, lo que es una razón poderosa para abogar por dicha configuración. Posteriormente se realiza una extensión de la configuración propuesta para entornos *multirate* en la que, además, se aplicará un criterio de equidad que evite la inanición de las estaciones *lentas* al maximizar el rendimiento de la red. Por último, y dado que en este tipo de entornos no hay que descartar que los usuarios alteren el comportamiento de su nivel MAC, se presentará un algoritmo que permite determinar si un comportamiento se ajusta a las reglas del protocolo MAC o no, con objeto de evitar que determinados usuarios reciban un servicio mejor a costa del resto de usuarios.

3.1. Análisis del modo EDCA en condiciones de saturación

Ya ha sido objeto de mención la considerable cantidad de trabajo (por ejemplo, [1–4]) que ha abordado el análisis del modo EDCA en condiciones de saturación. Realizar dicho análisis es un punto ineludible si se pretende obtener las condiciones de funcionamiento óptimas para una WLAN en dichas circunstancias. Sin embargo, a pesar del esfuerzo existente en este sentido, los trabajos anteriormente citados no garantizan la unicidad de la solución obtenida para el punto de operación de la red, con los problemas que de ello se derivan.

En esta sección se diagnostica dicho problema y se propone un modelo diferente que, a diferencia de los anteriores, sí que presenta una única solución.

3.1.1. Problema de los modelos derivados del análisis de Bianchi

La mayoría de los análisis del modo EDCA en saturación parten del modelo de Bianchi [27] de la siguiente forma: la probabilidad de que una estación perteneciente a una AC i transmita en una ranura de tiempo, dada una probabilidad de colisión c_i , su ventana mínima de contienda W_i y su máximo crecimiento de dicha ventana m_i , es la siguiente:

$$\tau_i = \frac{2}{1 + W_i + c_i W_i \sum_{j=0}^{m_i-1} (2c_i)^j} \quad (3.1)$$

Para el cálculo de c_i , los modelos parten de la suposición de [27] de que cada estación transmite con una probabilidad constante e independiente. Esto lleva a la siguiente expresión para la probabilidad de colisión:

$$c_i = 1 - (1 - \tau_i)^{n_i-1} \prod_{j \neq i} (1 - \tau_j)^{n_j} \quad (3.2)$$

en la que n_i es el número de estaciones pertenecientes a la AC i .

De (3.1) y (3.2) se obtiene un sistema de ecuaciones en las τ_i 's: de (3.2) se puede expresar cada c_i como función de todas las τ_i 's y de (3.1) se puede obtener cada τ_i dadas las c_i 's. Ello lleva a un sistema de N ecuaciones no lineales (con N el número de AC's presentes en la WLAN) que se resuelve, por lo general, mediante técnicas numéricas.

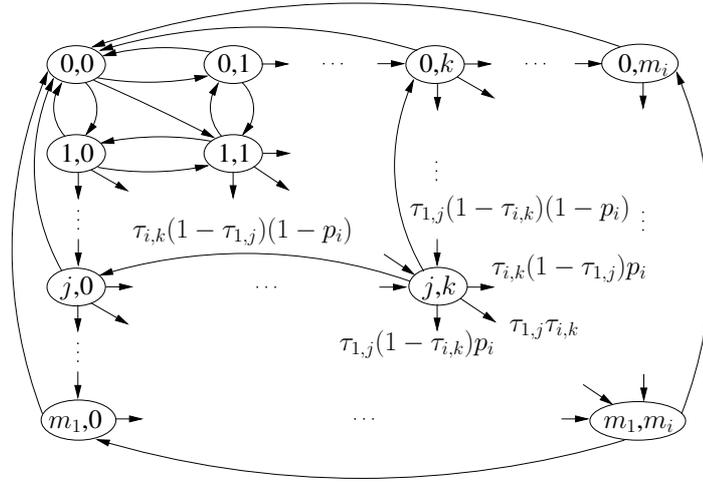
Sin embargo, el anterior sistema no presenta necesariamente una única solución. Para demostrar esto, baste con un contraejemplo: considérese una WLAN con dos AC's diferentes. La AC 1 presenta una configuración $\{W_1 = 2, m_1 = 5\}$ y la AC 2 presenta la configuración $\{W_2 = 2, m_2 = 6\}$. Las parejas de valores de $\{\tau_1, \tau_2\}$ que satisfacen el sistema definido anteriormente son las siguientes: $\{0.237, 0.514\}$, $\{0.318, 0.413\}$ y $\{0.589, 0.142\}$. Si se realiza una simulación para obtener dichos valores, resulta $\{0.411, 0.318\}$, que no se corresponde con ninguno de los anteriores.

Por lo tanto, los análisis del modo EDCA que se basan en un sistema como el formado por (3.1)–(3.2) no tienen solución única garantizada. Esto supone un severo problema, dado que las técnicas numéricas pueden conllevar problemas de convergencia en sistemas con múltiples soluciones; incluso, caso de converger, el resultado obtenido puede que no sea el más cercano al punto de operación del sistema. Lo que es más: en el ejemplo considerado, ninguna de las soluciones que satisfacen el sistema de ecuaciones es próxima al valor obtenido mediante simulación.

3.1.2. Nuevo modelo para el análisis de EDCA

El modelo propuesto parte de la misma suposición que [49]: los contadores de backoff siguen una distribución geométrica. Bajo esta suposición, una estación de la AC i transmite de forma independiente en una ranura de tiempo con la siguiente probabilidad

$$\tau_{i,j} = \frac{2}{2^j W_i + 1} \quad (3.3)$$

Figura 3.1: Cadena de Markov para el proceso $\{j, k\}$

donde j es el número de retransmisiones que lleva para la trama que está intentando transmitir (esto es: el estadio de su contador de backoff, que indica el número de veces que ha duplicado su ventana de contienda para dicha trama).

Considérese ahora una estación de la AC 1 y otra estación de una AC i , y denomínese mediante p_i la probabilidad de que cualquier otra estación diferente de las consideradas transmita en una ranura de tiempo. Se supondrá que dicha probabilidad es constante e independiente. De esta forma, el sistema se puede modelar como un proceso estocástico bidimensional $\{j, k\}$ donde j es el estadio de backoff de la estación de la AC 1 y k es dicho estadio para la AC i . Este proceso se representa en la cadena de Markov de la Figura 3.1.

A continuación se analizan las probabilidades de transición entre estados. Una transmisión con éxito tiene lugar cuando una de las estaciones consideradas transmite, y ni la otra estación ni el resto de estaciones lo hace. Esto lleva a las siguientes expresiones:

$$p_{\{j,k\} \rightarrow \{0,k\}} = \tau_{1,j}(1 - \tau_{i,k})(1 - p_i) \quad (3.4)$$

$$p_{\{j,k\} \rightarrow \{j,0\}} = \tau_{i,k}(1 - \tau_{1,j})(1 - p_i) \quad (3.5)$$

Por otra parte, cuando una de las dos estaciones en cuestión transmite, la otra estación no pero sí que existe otra transmisión en la misma ranura, se ocasiona el incremento del estadio de backoff de la estación que ha transmitido

$$p_{\{j,k\} \rightarrow \{j+1,k\}} = \tau_{1,j}(1 - \tau_{i,k})p_i \quad (3.6)$$

$$p_{\{j,k\} \rightarrow \{j,k+1\}} = \tau_{i,k}(1 - \tau_{1,j})p_i \quad (3.7)$$

Por último, independientemente de lo que haga el resto de las estaciones, cuando las dos estaciones consideradas transmiten a la vez ambas aumentan el valor de su estadio de backoff

$$p_{\{j,k\} \rightarrow \{j+1,k+1\}} = \tau_{1,j}\tau_{i,k} \quad (3.8)$$

Definidas las anteriores probabilidades se puede resolver la cadena de Markov planteada, obteniéndose las probabilidades $P_{j,k}$ de cada estado. Con éstas, las probabilidades de transmisión de las estaciones consideradas se obtienen de forma inmediata

$$\tau_1 = \sum_{j=0}^{m_1} \sum_{k=0}^{m_i} P_{j,k} \tau_{1,j} \quad (3.9)$$

$$\tau_i = \sum_{j=0}^{m_1} \sum_{k=0}^{m_i} P_{j,k} \tau_{i,k} \quad (3.10)$$

Con todo lo anterior, se obtiene que dada una probabilidad p_i de que alguna estación de las no consideradas transmita, se puede resolver la cadena de Markov planteada y, de ésta, calcular las probabilidades de transmisión de las estaciones de la AC 1 y la AC i . Se denominará mediante $\tau_1(p_i)$ y $\tau_i(p_i)$ a estas dos funciones de la probabilidad p_i .

El anterior modelo se puede aplicar para diferentes valores de i , por lo que se puede obtener τ_1 y τ_i como una función de p_i para cualquier $i \in \{2, \dots, N\}$. De esta forma, se plantea un sistema con $N - 1$ incógnitas, que son las diferentes p_i 's, que precisa de $N - 1$ ecuaciones independientes.

Las primeras $N - 2$ ecuaciones se obtienen de forzar a que los valores de τ_1 sean iguales, independientemente de la AC i que se considere para la otra estación del análisis. Estas $N - 2$ ecuaciones se definen, pues, de la siguiente forma:

$$\tau_1(p_i) = \tau_1(p_{i+1}), \quad i \in \{2, \dots, N - 1\} \quad (3.11)$$

Por otra parte, bajo la suposición de que todas las estaciones salvo las dos consideradas transmiten con probabilidad constante e independiente, p_i se puede expresar de la siguiente forma:

$$p_i = 1 - (1 - \tau_1)^{n_1 - 1} (1 - \tau_i)^{n_i - 1} \prod_{k \neq 1, i} (1 - \tau_k)^{n_k} \quad (3.12)$$

Dado que las τ_i 's se pueden expresar como una función de las p_i 's, la anterior expresión es una ecuación en las p_i 's. La $(N - 1)$ -ésima ecuación que falta se corresponde con la expresión resultante de multiplicar todas las partes izquierda y todas las partes derecha de la anterior expresión:

$$\prod_{i=2}^N p_i = \prod_{i=2}^N \left(1 - (1 - \tau_1(p_i))^{n_1 - 1} (1 - \tau_i(p_i))^{n_i - 1} \prod_{k \neq 1, i} (1 - \tau_k(p_k))^{n_k} \right) \quad (3.13)$$

De esta forma se define un sistema de $N - 1$ ecuaciones no lineales, siendo las incógnitas las p_i 's, lo que termina el modelo. Este modelo presenta una complejidad computacional similar los ya existentes, que requieren la resolución de un sistema de N ecuaciones.

La principal ventaja del modelo propuesto, en contraposición con los anteriores, es que el sistema de ecuaciones planteado presenta una única solución –lo que se demostrará a

continuación—. De hecho, con el mismo ejemplo presentado en la sección anterior, la solución para $\{\tau_1, \tau_2\}$ que se obtiene del nuevo modelo es $\{0.416, 0.324\}$, muy similar a la obtenida mediante simulación (que era $\{0.411, 0.318\}$).

3.1.2.1. Demostración de la unicidad de la solución

Teorema 1. *Existe una y sólo una solución para el sistema de $N - 1$ ecuaciones no lineales formado por (3.11)–(3.13).*

Demostración. En primer lugar, se mostrará que, dada una p_i , el resto de p_i 's se puede expresar como una función creciente de dicha p_i que alcanza el valor 1: de la resolución de la cadena de Markov (3.4)–(3.8) se puede apreciar que, para todo i , $\tau_1(p_i)$ y $\tau_i(p_i)$ son funciones decrecientes de p_i ; y que para todo i y j , $\tau_1(p_i)|_{p_i=1} = \tau_1(p_j)|_{p_j=1}$.

Considérese aquella AC k para la que $\tau_1(p_k)|_{p_k=0}$ es mínima. Que todas las p_i 's se pueden derivar de p_k se puede deducir de (3.11) como sigue: dado p_k , $\tau_1(p_k)$ es una constante, mientras que $\tau_1(p_i)$ es una función decreciente de p_i en el intervalo $p_i \in (0, 1)$; dicha función, dada la elección de p_k , empieza en un valor no inferior a $\tau_1(p_k)$ y decrece hasta un valor no superior a $\tau_1(p_k)$. Por lo tanto, existe un único valor que satisface la condición $\tau_1(p_i) = \tau_1(p_k)$. Además, del hecho de que tanto $\tau_1(p_i)$ como $\tau_1(p_k)$ son ambas funciones decrecientes, se deduce que dicho valor de p_i aumenta con el valor de p_k . Dado que $\tau_1(p_i)|_{p_i=1} = \tau_1(p_k)|_{p_k=1}$ se tiene, además, que alcanza el valor 1 cuando $p_k = 1$.

Considérese ahora la expresión (3.13). La parte izquierda de dicha expresión, por lo indicado anteriormente, se puede expresar como una función de p_k que arranca en 0 y llega hasta 1 en el rango $p_k \in (0, 1)$. Dado que τ_i es una función decreciente de p_i , y p_i es creciente con p_k , la parte derecha de (3.13) es una función decreciente de p_k . Se deduce, además, que dicha función arranca en un valor no nulo y decrece hasta un valor inferior a la unidad. Por lo tanto, sólo existe un valor de p_k que satisface la condición de dicha expresión. Con dicho valor, y deshaciéndose los pasos indicados, se obtiene una solución para el sistema. La unicidad de la misma se deriva del hecho de que todas las relaciones son biyectivas y cualquier solución debe cumplir la condición (3.13) que, como se acaba de ver, sólo tiene una solución. \square

3.1.3. Validación del nuevo modelo

Con objeto de validar el nuevo modelo, se calcula el ancho de banda que reciben las estaciones para una WLAN en condiciones de saturación, y se compara con los resultados obtenidos mediante simulación. Para obtener la tasa recibida por las estaciones se puede acudir a [27], donde se presenta la siguiente expresión:

$$r_i = \frac{p(s_i)l}{p(s)T_s + p(c)T_c + p(e)T_e} \quad (3.14)$$

donde $p(s_i)$ es la probabilidad de que una ranura de tiempo contenga un éxito de una estación dada de una AC i , l es la longitud de la trama transmitida, $p(s)$, $p(c)$ y $p(e)$ son las probabilidades de que una ranura contenga un éxito, una colisión o esté vacía y T_s , T_c y T_e son las duraciones en cada caso –véase (2.1) y (2.2)–. Las expresiones de dichas probabilidades son las siguientes:

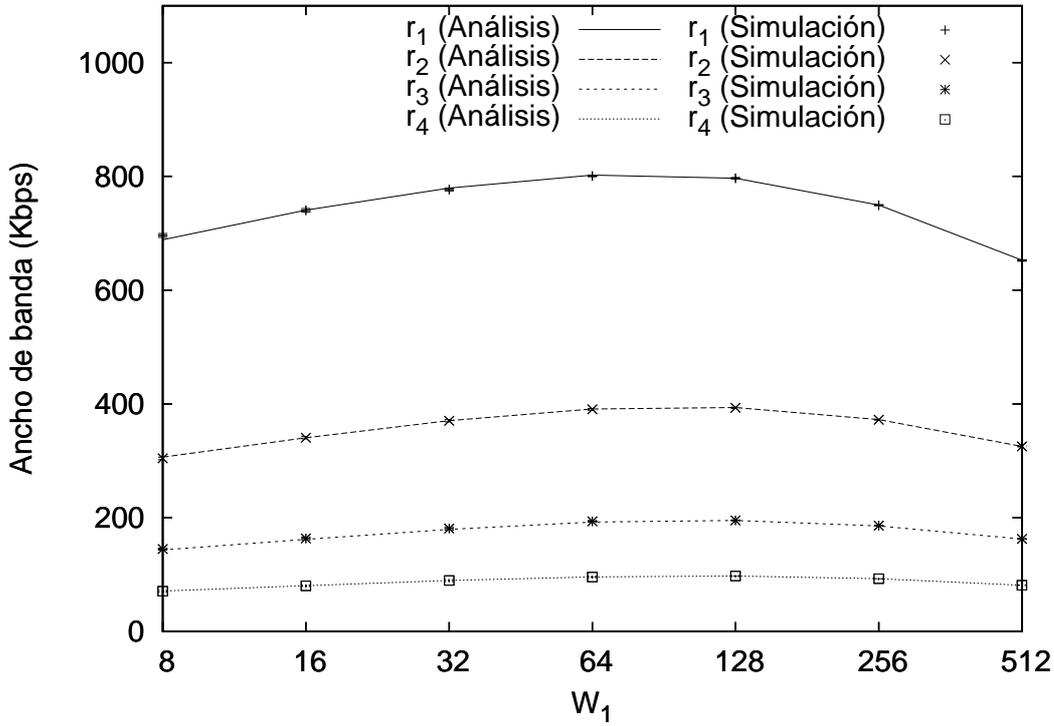


Figura 3.2: Validación del nuevo modelo para saturación

$$p(e) = \prod_i (1 - \tau_i)^{n_i} \quad (3.15)$$

$$p(s_i) = \tau_i (1 - \tau_i)^{n_i - 1} \prod_{j \neq i} (1 - \tau_j)^{n_j} \quad (3.16)$$

$$p(s) = \sum_i n_i p(s_i) \quad (3.17)$$

$$p(c) = 1 - p(e) - p(s) \quad (3.18)$$

Se considera un escenario donde una WLAN en el modo 802.11b opera en condiciones de saturación. Existen cuatro AC's diferentes ($AC\ i \in \{1, 2, 3, 4\}$) configuradas de la siguiente manera: dada una ventana de contienda W_1 para la primera AC, el resto de AC's utilizan el valor de $W_i = W_1 2^{i-1}$. Se fija $m = 5$ para todas las AC's, que cuentan con $n_i = 5$ estaciones cada una. Todas las transmisiones tienen una longitud constante de 1500 octetos. Los resultados se obtienen a través de un simulador orientado a eventos desarrollado en C++, que refleja el comportamiento de cada estación en el mecanismo EDCA.

En la Figura 3.2 se presentan los valores obtenidos mediante el nuevo análisis (líneas) y mediante simulación (puntos). Se realiza un intervalo de confianza al 95 % para la media, si bien su escasa longitud dificulta su apreciación. Los resultados analíticos coinciden con los valores que arroja la simulación, de lo que se deriva la validez del modelo propuesto.

3.2. Configuración para modo básico

Se ha dedicado bastante esfuerzo a encontrar la configuración que optimiza el rendimiento de una WLAN operando en condiciones de saturación (por ejemplo, [27, 67])¹. Sin embargo, el uso de una configuración óptima en situaciones reales, donde es habitual que las estaciones no tengan siempre tráfico pendiente de ser transmitido, resulta una estrategia cuestionable [75, 89]. En esta sección se justifica el uso de dicha configuración óptima para todos los casos, saturados o no, en los que no se tenga más información sobre las condiciones y el servicio que se pretende proporcionar en la WLAN, y ésta opere a una única velocidad de transmisión.

La justificación de la necesidad del uso de la configuración óptima nace del hecho de que una WLAN puede sufrir una serie de anomalías e inestabilidades que es preciso evitar. Si bien algunas de estas anomalías han sido objeto de trabajos anteriores [6, 90], en esta sección se desarrolla y valida un modelo que da explicación a estos comportamientos.

3.2.1. Análisis de las anomalías

3.2.1.1. Tiempo de servicio

Para ilustrar el comportamiento en tiempo de servicio de una WLAN se considera el siguiente escenario: sea una red con 40 estaciones, todas ellas configuradas según el estándar 802.11b. Supónganse infinitas las colas de transmisión, y que todas las tramas son de 1500 octetos siendo generadas a una tasa constante.

Se define *tasa de saturación* (r_{sat}) como el ancho de banda que cada estación obtiene en condiciones de saturación. Se realizan dos pruebas:

- En la primera, todas las estaciones generan tráfico a una tasa (r_{send}) del 101 % la tasa de saturación ($r_{send} = 1,01 r_{sat}$).
- En la segunda, la tasa de generación es el 99 % de la tasa de saturación ($r_{send} = 0,99 r_{sat}$).

En ambas pruebas se mide el tiempo medio de servicio, definido éste como el tiempo desde que una trama alcanza la primera posición de la cola de salida hasta que es transmitida con éxito. La media se realiza en intervalos de un segundo, mostrándose en la Figura 3.3 una gráfica de la evolución de dichas medias a lo largo de cien segundos, para las dos pruebas anteriores.

Los resultados de dicha figura resultan sorprendentes, en tanto que muestran que una pequeña variación en la tasa de generación de tráfico puede provocar un sustancial aumento del tiempo medio de servicio. Si bien este comportamiento puede ser intuitivo para caso del *tiempo total* (incluyendo el tiempo de espera en cola), el hecho de que aparezca esta variación para el caso del tiempo medio de servicio puede resultar sorprendente.

El modelo que explica dicho comportamiento se desarrolla a continuación: sea una WLAN con N estaciones. Una *ranura* es el intervalo de tiempo entre dos decrementos consecutivos del contador de backoff, y la probabilidad de que una estación transmita al

¹De ahora en adelante se emplearán indistintamente los términos *configuración óptima* y *configuración óptima en condiciones de saturación*.

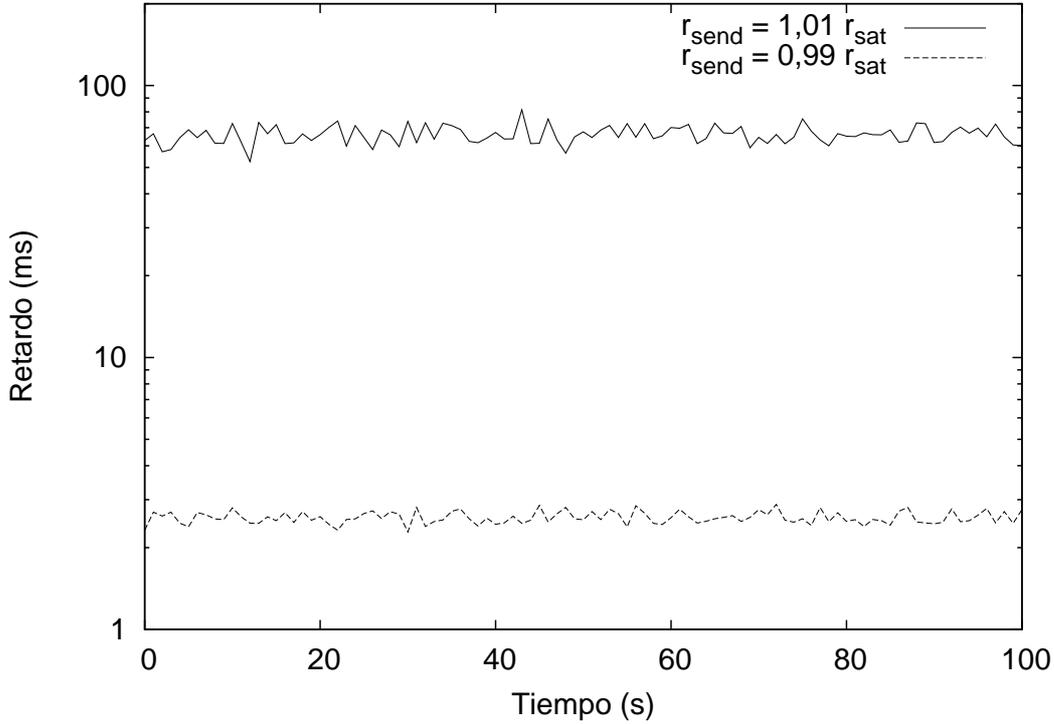


Figura 3.3: Tiempo medio de servicio

inicio de una ranura es τ –se toma misma suposición de constancia e independencia que Bianchi [27]–. Bajo estas suposiciones, se puede calcular la tasa de transmisión como la media de la información transmitida con éxito en una ranura dividida por la longitud media de ésta:

$$r = \frac{p_s l}{T_{slot}} \quad (3.19)$$

Donde p_s es la probabilidad de que una ranura contenga un éxito de la estación considerada, l la longitud de dicho éxito y T_{slot} la duración media de una ranura. La probabilidad de que una ranura contenga un éxito viene dada por el que la estación considerada transmita (lo que sucede, según el modelo, con probabilidad τ) y que el resto de $N - 1$ estaciones no lo haga (cada una con probabilidad $1 - \tau$):

$$p_s = \tau(1 - \tau)^{N-1} \quad (3.20)$$

La duración media de una ranura se calcula con las probabilidades de que dicha ranura contenga un éxito, una colisión o esté vacía:

$$T_{slot} = (1 - \tau)^N T_e + N\tau(1 - \tau)^{N-1} T_s + (1 - (1 - \tau)^N - N\tau(1 - \tau)^{N-1}) T_c \quad (3.21)$$

donde T_s , T_c y T_e son las longitudes de una ranura cuando ésta contiene un éxito, una colisión o permanece sin actividad.

En la Figura 3.4 se representa r en función de τ según la expresión indicada en la Ecuación 3.19, en el intervalo $[0, \tau_{sat}]$. τ_{sat} es la probabilidad de transmisión bajo la condición de que todas las estaciones siempre tienen una trama que transmitir, por lo que representa el máximo valor que puede alcanzar τ . En el punto de operación, bien la tasa de transmisión se corresponde con la tasa de generación de paquetes, o bien ésta supera a aquélla (por lo que la cola de transmisión crece indefinidamente y el sistema se satura). De la figura se deduce que, para los dos anteriores casos, los valores que puede tener τ son los siguientes:

- Para $r_{send} = 0,99 r_{sat}$, el valor que se corresponde con dicha tasa de transmisión es τ_a .
- Para $r_{send} = 1,01 r_{sat}$, la tasa de servicio es menor que la de generación, por lo que las colas nunca se vacían y la red opera en τ_{sat} .

La notable diferencia en los tiempos medios de servicio, por tanto, se explica por el hecho de que, a pesar de que las tasas de generación de tráfico son muy similares en ambos casos, las τ 's de operación son bien diferentes. El hecho de que τ_{sat} sea mucho mayor que τ_a implica un número mayor de colisiones en la red, lo que provoca el abrupto aumento en el tiempo de servicio.

Para comprobar la validez del modelado, se calcula el valor medio del tiempo de servicio para ambos escenarios. Para el caso de no saturación, se resuelve la Ecuación 3.19 para obtener τ_a , mientras que para el caso de saturación se puede acudir a Bianchi [27]. Para cada valor de τ se utiliza el cálculo del tiempo medio de servicio que se detalla en el Capítulo 4 (Sección 4.1.2), obteniéndose los valores de 2.30 y 71.35 ms, respectivamente, que se corresponden con los mostrados en la Figura 3.3.

3.2.1.2. Estabilidad

En el anterior apartado se ha mostrado que, bajo la suposición de colas de transmisión ilimitadas, pequeñas variaciones en la tasa de generación de tráfico pueden causar diferencias sustanciales en el tiempo de servicio. A continuación se examina el caso de colas de tamaño limitado.

Sea, de nuevo, una WLAN con $N = 40$ estaciones, cada una de ellas con una cola de $Q = 50$ paquetes en una primera prueba y de $Q = 10$ paquetes en una segunda prueba. Se genera tráfico al 110 % de la tasa que se obtendría en saturación. El tiempo medio de servicio para cada caso, promediado a intervalos regulares de diez segundos, se representa en la Figura 3.5.

La inestabilidad es manifiesta: para el caso $Q = 50$, el tiempo medio de servicio crece súbitamente hasta un punto en el que permanece estable a lo largo de la simulación. Para $Q = 10$, en cambio, el tiempo de servicio oscila entre dos valores bien definidos, si bien permanece más tiempo en el valor superior. Resulta evidente que este tipo de comportamiento, dada su naturaleza aleatoria, no es el más deseable a la hora de proporcionar un servicio al tráfico que se cursa en la WLAN.

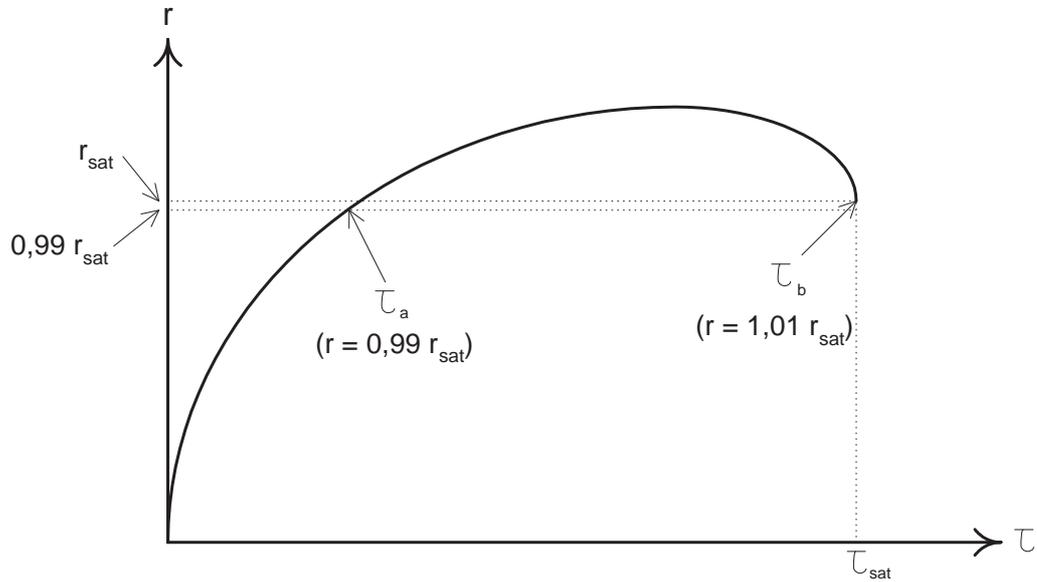


Figura 3.4: Modelo para el tiempo medio de servicio

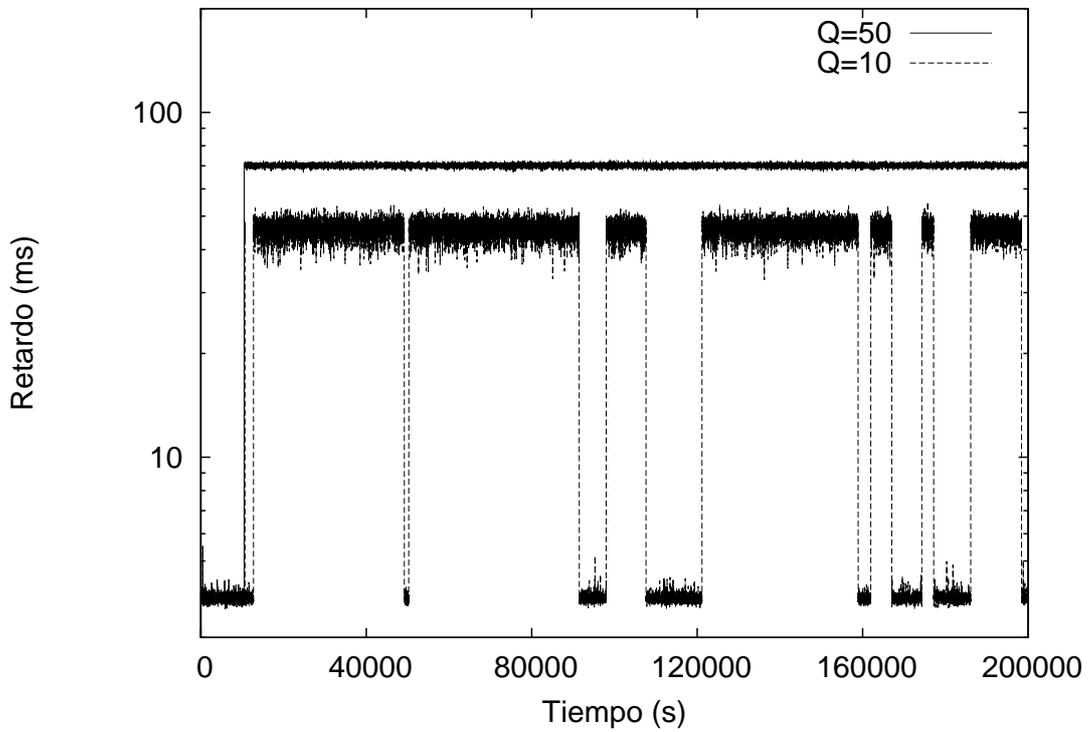


Figura 3.5: Inestabilidad del modo básico

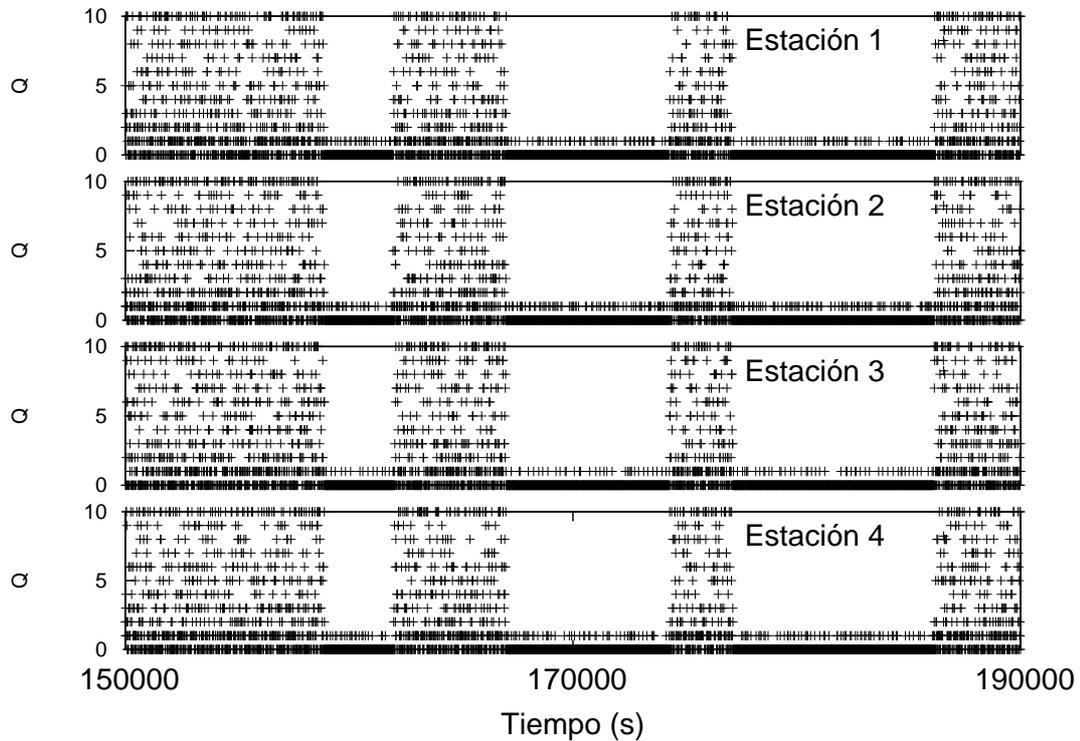


Figura 3.6: Estado de las colas en el modo básico

Con objeto de investigar las causas de dicho comportamiento, se realiza un gráfico para el caso $Q = 10$ de la ocupación de la cola para cuatro estaciones al azar a lo largo del tiempo. Dicho gráfico se muestra en la Figura 3.6. Se aprecia una clara sincronización en la ocupación de dichas colas: cuando los tiempos de servicio son elevados, las colas se encuentran completamente ocupadas, mientras que en los periodos en los que el tiempo de servicio es menor, las colas se encuentran prácticamente vacías.

Para explicar dicho comportamiento se parte de la Figura 3.7. Para una tasa superior a la de saturación, los puntos en los que puede operar la WLAN son los siguientes:

- $\tau = \tau_a$. Para este caso la tasa de generación se corresponde con la tasa de envío en la WLAN.
- $\tau = \tau_b$. Este caso también cumple la condición anterior.
- $\tau = \tau_{sat}$. Para esta situación la tasa de generación de tráfico resulta superior a la tasa que obtiene la estación, por lo que las colas se llenan y la WLAN se mantiene en dicho punto.

Los anteriores puntos, no obstante, no son puntos estáticos de operación, dado que cada uno sufre ligeras perturbaciones debido a la naturaleza dinámica y aleatoria del mecanismo MAC: una serie de éxitos provoca una disminución temporal en el número de colisiones y, por lo tanto, una disminución de τ , mientras que una serie de colisiones provoca un aumento

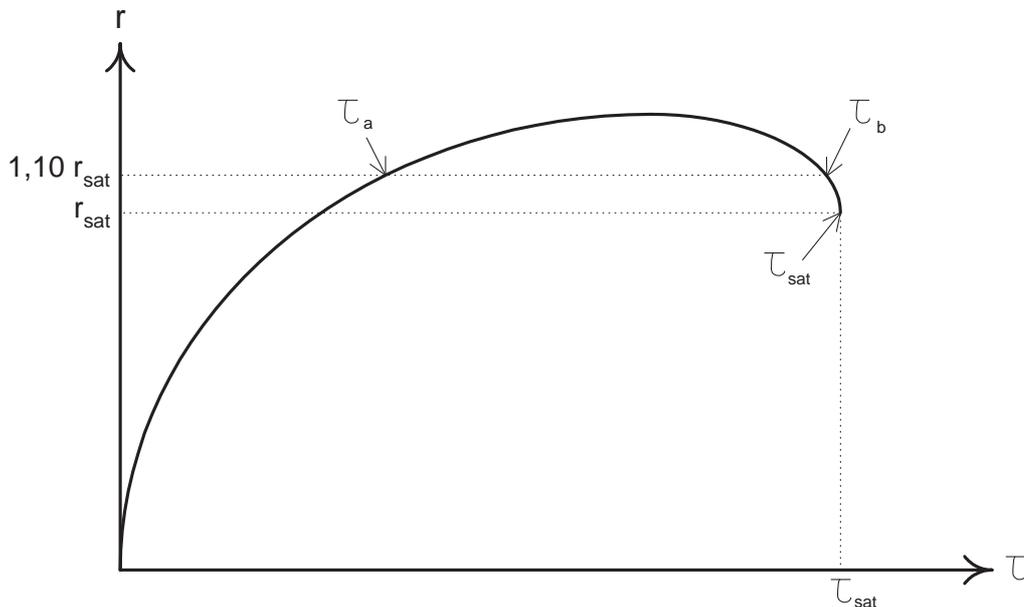


Figura 3.7: Modelo para la inestabilidad de DCF

de las transmisiones y de τ . Esto conlleva un impacto en la estabilidad de los anteriores puntos de operación de la WLAN:

- $\tau = \tau_a$. En este valor de τ , ligeros incrementos en la probabilidad suponen un incremento en la tasa de transmisión. Esto conlleva que se puedan servir un mayor número de tramas, por lo que ocasiona una disminución en τ (el comportamiento es recíproco para disminuciones de τ). Debido a esta realimentación negativa, el punto de operación es *estable*, mientras la magnitud de la perturbación en sentido ascendente no sea demasiado elevada.
- $\tau = \tau_b$. En este punto, un incremento temporal de τ provoca una tasa de servicio menor que la tasa de generación. De esta forma, la ocupación de la cola aumenta al no dar servicio a todos los paquetes generados, lo que realimenta de forma positiva el incremento de τ (un comportamiento similar se obtiene para perturbaciones a la baja de la probabilidad τ). Se concluye, por tanto, que el punto de operación τ_b es *inestable*.
- $\tau = \tau_{sat}$. Una vez operando en este punto, las colas de transmisión se encuentran con una ocupación muy elevada. Dado que la tasa de envío de tramas es menor que la de generación, sólo para perturbaciones muy fuertes en la τ se abandona dicho punto, por lo que se califica de *estable*.

De lo anterior se deduce la siguiente explicación para el comportamiento de la Figura 3.5: al principio de la simulación, el valor de τ se incrementa hasta el punto τ_a , donde opera durante un tiempo, mientras se cumple la condición de que todas las tramas generadas son

servidas. Sin embargo, una serie de colisiones ocasionan un aumento substancia en el valor de τ lo que ocasiona que se opere a un valor mayor que τ_b , provocando un crecimiento en las colas de transmisión. Para la simulación realizada, el valor de $Q = 50$ provoca que en ningún caso se abandone este punto de operación, por lo que la WLAN permanece en saturación. Para el valor de $Q = 10$, en cambio, en ocasiones las colas de transmisión se llegan a vaciar (por una serie de éxitos consecutivos), por lo que se vuelve al punto de operación τ_a ².

Con objeto de validar el modelo, se calculó el retardo medio del tiempo de servicio y se comparó con los valores de la simulación. Se obtuvo un valor de 2.86 ms para la situación de τ_a , y el mismo valor que en la anterior sección (71.35 ms) para τ_{sat} .

3.2.1.3. Ancho de banda

En este caso, considérese de nuevo una WLAN con $N = 40$ estaciones, todas ellas con colas de transmisión de tamaño infinito. En un primer experimento, una de dichas estaciones se encuentra en saturación, mientras que las 39 restantes generan tráfico a un 99 % de la tasa r_{sat} . En la Figura 3.8 se muestra el ancho de banda que obtiene la primera estación (99 %, no saturada) y una estación de las restantes (99 %, saturada), promediado en ventanas de 1 segundo. El segundo experimento, también ilustrado en la misma figura, muestra el ancho de banda que obtiene una de las estaciones cuando todas ellas generan tráfico al 101 % de r_{sat} (101 %).

A la vista de los resultados, se comprueba que la tasa de servicio que recibe una estación puede sufrir un severo deterioro cuando el resto de estaciones aumenta mínimamente su tasa de generación de tráfico: el ancho de banda de las 39 estaciones se mantiene al mismo nivel, mientras que el ancho de banda de la estación en saturación disminuye de forma drástica.

A continuación se presenta un modelo para explicar este cambio en la tasa de servicio. Sea τ_{sat} la probabilidad de transmisión para cada una de las 39 estaciones, y τ_{nonsat} dicha probabilidad para la estación restante. De acuerdo con [27], la probabilidad de que la estación saturada transmita viene fijada por la siguiente expresión:

$$\tau_{sat} = \frac{2}{1 + CW_{min} + p(c)_{sat} CW_{min} \sum_{j=0}^{m-1} (2p(c)_{sat})^j} \quad (3.22)$$

Donde $CW_{max} = 2^m CW_{min}$ y $p(c)_{sat}$ es la probabilidad de que una transmisión de la estación saturada colisione. Ello ocurre si una o más del resto de estaciones no se queda sin transmitir, por lo que

$$p(c)_{sat} = 1 - (1 - \tau_{nonsat})^{N-1} \quad (3.23)$$

El ancho de banda que reciben las estaciones no saturadas es igual a su tasa de generación de tráfico (que es el 99 % de la tasa en el caso de que las N estaciones estuviesen saturadas), lo que implica la siguiente ecuación

$$0,99 r_{sat,N} = \frac{\tau_{nonsat} (1 - \tau_{nonsat})^{N-2} (1 - \tau_{sat}) l}{T_{slot}} \quad (3.24)$$

²Se realizaron otros experimentos para $Q = 50$ y $N = 10$ estaciones, obteniéndose un comportamiento similar al anterior. Esto se debe a que un menor número de estaciones facilita la posibilidad de que todas las colas se vacíen, aunque éstas sean de un tamaño elevado.

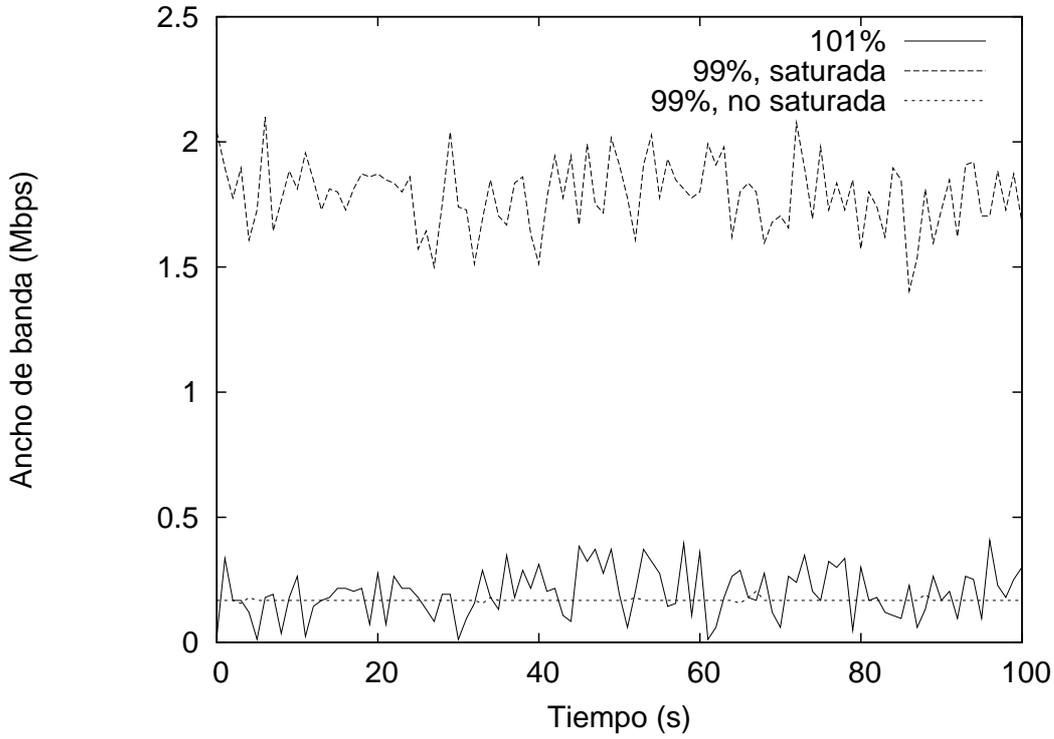


Figura 3.8: Anomalía en la tasa recibida

Donde $r_{sat,N}$ es el ancho de banda recibido por cada estación en el caso de que las N estaciones estuviesen saturadas, mientras que T_{slot} es la duración media de una ranura de tiempo, que viene dada por la siguientes expresiones:

$$T_{slot} = p_e T_e + p_s T_s + (1 - p_e - p_s) T_c \quad (3.25)$$

$$p_e = (1 - \tau_{sat})(1 - \tau_{nonsat})^{N-1} \quad (3.26)$$

$$p_s = (N - 1)\tau_{nonsat}(1 - \tau_{sat})(1 - \tau_{nonsat})^{N-2} + \tau_{sat}(1 - \tau_{nonsat})^{N-1} \quad (3.27)$$

Con las anteriores expresiones se forma un sistema de ecuaciones para τ_{sat} y τ_{nonsat} que presenta una solución en los rangos $\tau_{sat} \in (\tau_{sat,N}, 1)$ y $\tau_{nonsat} \in (0, \tau_{sat,N})$. Resolviendo dicho sistema se puede calcular el ancho de banda recibido por cada tipo de estación, dado por las siguientes expresiones:

$$r_{sat} = \frac{\tau_{sat}(1 - \tau_{nonsat})^{N-1}l}{T_{slot}} \quad (3.28)$$

$$r_{nonsat} = \frac{\tau_{nonsat}(1 - \tau_{sat})(1 - \tau_{nonsat})^{N-2}l}{T_{slot}} \quad (3.29)$$

De la resolución del anterior sistema y el cálculo de dichas expresiones, se obtiene un ancho de banda de 1.79 Mbps para la estación saturada y de 0.16 Mbps para el resto de

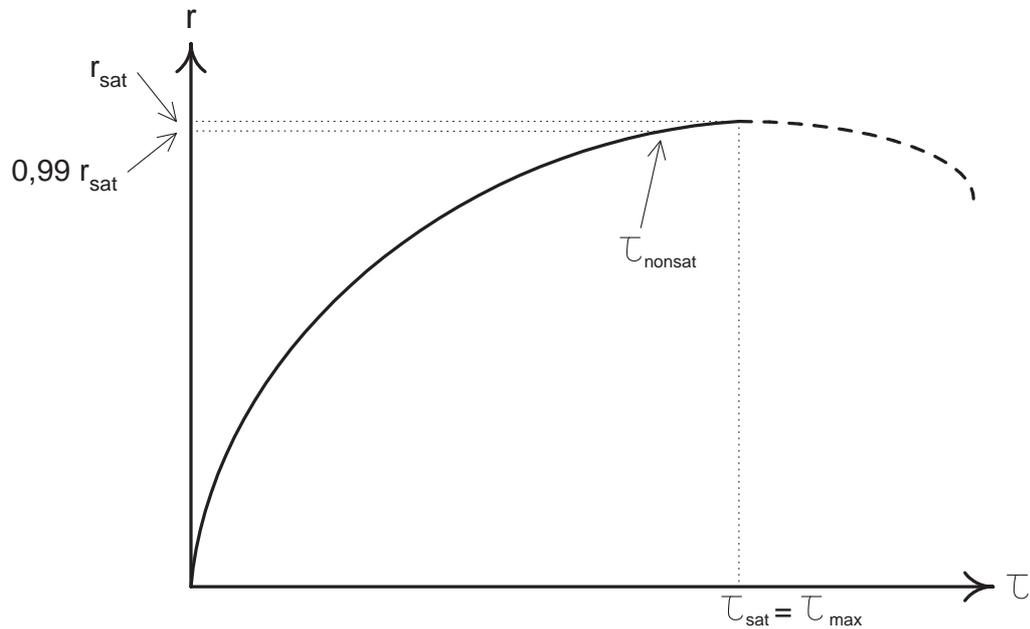


Figura 3.9: Configuración propuesta para WLAN abiertas

estaciones (para el primer caso). Para el segundo caso, se obtiene (siguiendo el modelo de [27]) un ancho de banda de 0.16 Mbps para cada estación. Estos valores se corresponden con los representados en la Figura 3.8, lo que valida del modelo desarrollado.

3.2.2. Configuración propuesta

Dada la existencia de las anomalías anteriormente descritas y diagnosticadas, se aborda el problema de plantear una configuración que evite la aparición de las mismas. La causa última de dichas anomalías se debe a la forma de campana de la curva que representa la ecuación 3.19 (Figuras 3.4 y 3.7). Esta forma permite que dada una tasa de transmisión (o bien dos tasas muy próximas), existan varios posibles puntos de operación de las τ 's notablemente espaciados. De hecho, todas las anomalías observadas derivan de la posibilidad de la WLAN de operar en cualquiera de los dos puntos que cumplen la condición sobre el ancho de banda.

Con objeto de evitar dichas anomalías, se propone una configuración que evite la forma de campana de dicha figura, mediante el establecimiento de una CW_{min} y CW_{max} que fuerce a que el punto τ_{sat} se corresponda con el máximo de la función $r(\tau)$. Este criterio se ilustra en la Figura 3.9, y al valor de τ_{sat} con la configuración óptima, τ_{max} .

Con la configuración propuesta se limita la función $r(\tau)$ a ser estrictamente creciente, existiendo un único punto de operación de τ , estable, para una determinada tasa de generación de tráfico –lo que elimina las causas de inestabilidades y anomalías–. De esta forma, la configuración propuesta consigue que la transición desde la no saturación a la saturación sea suave, de modo que las τ 's correspondientes a las tasas de generación $0,99r_{sat}$

y r_{sat} se encuentren muy próximas (compárese la Figuras 3.4 con la Figura 3.9).

El valor de τ_{max} es el valor que maximiza la tasa de la WLAN en condiciones de saturación. La configuración que se propone, por lo tanto, coincide con la de trabajos ya existentes; sin embargo, la diferencia fundamental es la motivación que subyace a la propuesta de dicha configuración: la motivación de los trabajos previos es maximizar el rendimiento de la WLAN cuando opera en condiciones de saturación –lo que no es una suposición realista–, mientras que en el caso considerado lo que se persigue es evitar la aparición de anomalías e inestabilidades al operar en condiciones de no saturación. En este caso, por lo tanto, se suponen tasas finitas de transmisión, lo que se corresponde con un modelo más realista que los anteriores.

Para calcular el punto de configuración óptima, se anula la derivada de la expresión (3.19) con respecto a la probabilidad τ , lo que lleva al siguiente punto de operación [27]:

$$\tau_{opt} \approx \frac{1}{N} \sqrt{\frac{2T_e}{T_c}} \quad (3.30)$$

Obtenido el valor de τ_{opt} , se puede emplear el modelo de la Sección 3.1.2 para obtener la configuración óptima de CW_{min} y CW_{max} : con el valor de la probabilidad de transmisión se obtiene la p_i mediante (3.12), y definido un valor de m se tiene una ecuación para W_i que se resuelve de forma numérica.

3.2.2.1. Validación de la configuración

Para $N = 40$ estaciones, los valores resultantes para la configuración propuesta son $CW_{min} = 386$ y $CW_{max} = 12352^3$. Con esta configuración se repiten los experimentos de los anteriores apartados, con objeto de mostrar la ausencia de las anomalías.

Tiempo de servicio. Se repite el experimento de la Sección 3.2.1.1, obteniéndose los resultados de la Figura 3.10. Los resultados analíticos arrojan valores de 58.62 y 58.32 ms, que se corresponden con los mostrados en la figura. Se comprueba que el tiempo medio de servicio es muy similar para los dos casos ($1,01r_{sat}$ y $0,99r_{sat}$), por lo que no aparece el abrupto incremento de dicho tiempo ante un mínimo aumento de la tasa de generación de tráfico. Se observa que los retardos son mayores que los de la Sección 3.2.1.1; esto se debe a que, con la configuración óptima, la tasa obtenida en saturación (y por lo tanto las tasas de envío) es notablemente mayor que la del caso anterior, por lo que la WLAN opera en una situación de mayor carga que implica un aumento de dichos tiempos.

Estabilidad. Repitiéndose la simulación de la Sección 3.2.1.2, se comprueba (Figura 3.11) que la WLAN opera de forma estable en un único punto durante todo el tiempo, con independencia del tamaño de la cola de transmisión. Este punto de operación se corresponde al de saturación y proporciona un tiempo medio de servicio de 58.62 ms, que se corresponde con el obtenido por simulación.

³Con ánimo de mantener la naturaleza adaptativa del mecanismo de retroceso exponencial binario, se mantiene el valor $m = 5$ del modo DCF.

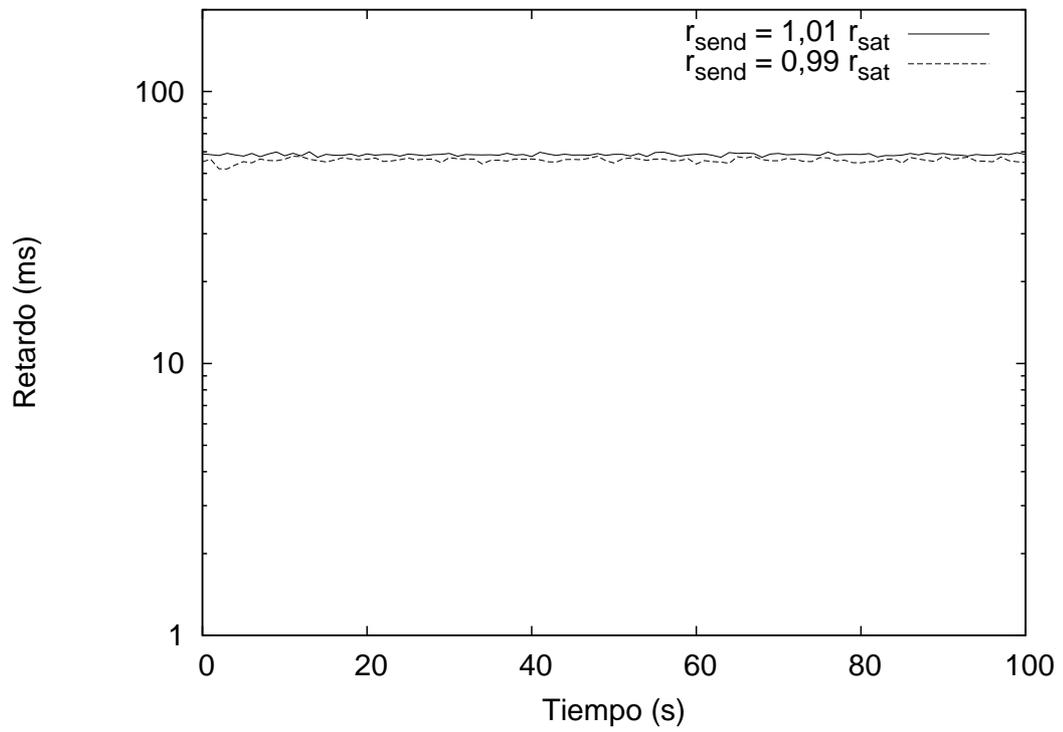


Figura 3.10: Retardo para la configuración óptima

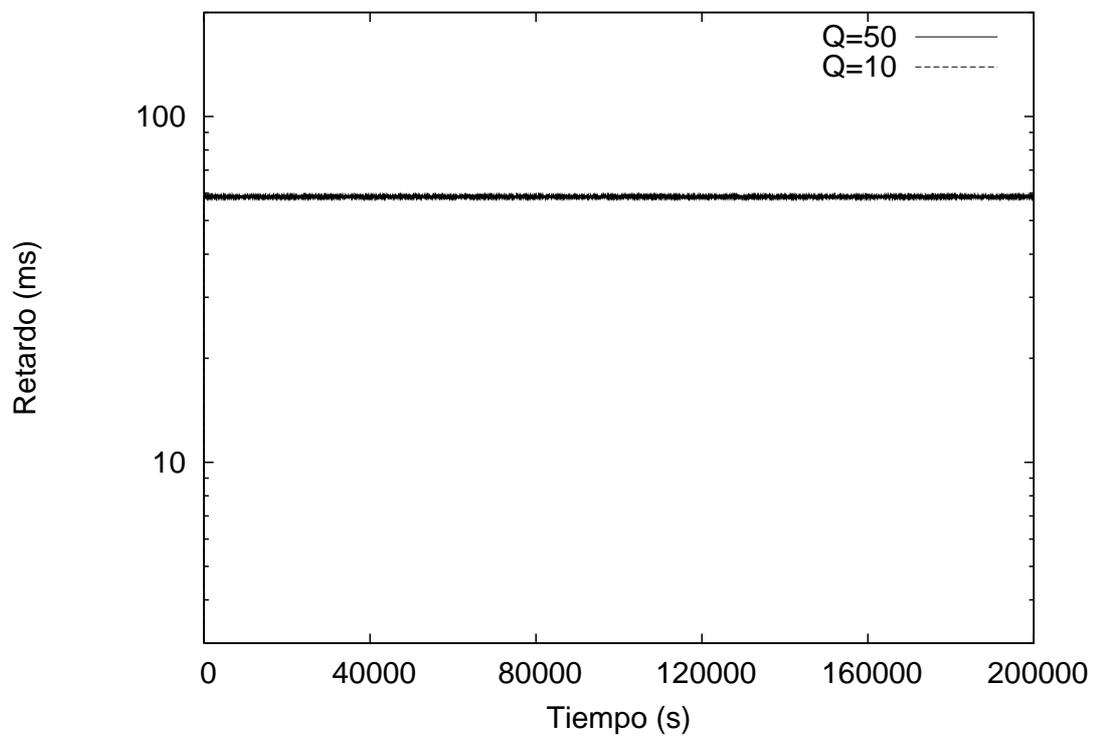


Figura 3.11: Estabilidad de la configuración óptima

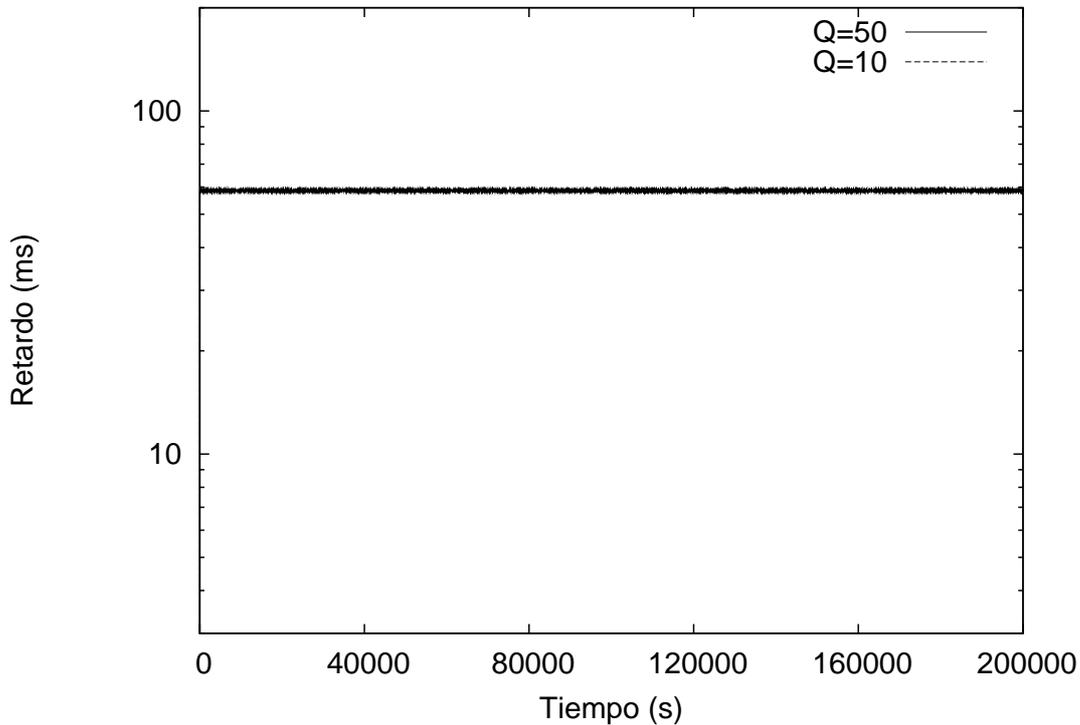


Figura 3.12: Ancho de banda con la configuración óptima

Tasa de transmisión. Por último, se repite la simulación de la Sección 3.2.1.3 para obtener la Figura 3.12. No se aprecia anomalía alguna, ajustándose los resultados al modelo analítico (que proporciona unas tasa de 0.21 y 0.20 Mbps⁴). La tasa de servicio recibida, por lo tanto, no se ve modificada de forma drástica por un cambio mínimo en las tasas de generación de tráfico.

Se confirma, a la luz de los resultados obtenidos, lo predicho por el modelado analítico: la configuración óptima de la WLAN impide que aparezcan inestabilidades o anomalías en el comportamiento de la red.

3.3. Configuración para modo *multirate*

En el anterior apartado se ha obtenido que, con objeto de evitar un comportamiento indeseable en una WLAN en condiciones de no saturación, dicha red debe estar configurada de modo tal que se consiga el mayor rendimiento de la misma en condiciones de saturación. Dicho resultado ha sido obtenido para una WLAN operando en el modo básico, en el que todas las estaciones transmiten a la misma velocidad.

Sin embargo, con la existencia de diferentes velocidades, las estaciones pueden modificar la tasa a la que inyectan datos al medio con objeto de fijar un compromiso entre

⁴Resulta adecuado indicar que, con la configuración óptima, el ancho de banda total obtenido (8.1 Mbps) es similar al de la Sección 3.8 para el caso de 99 % r_{sat} (8.03 Mbps) y muy superior al del caso de 101 % r_{sat} (6.4 Mbps).

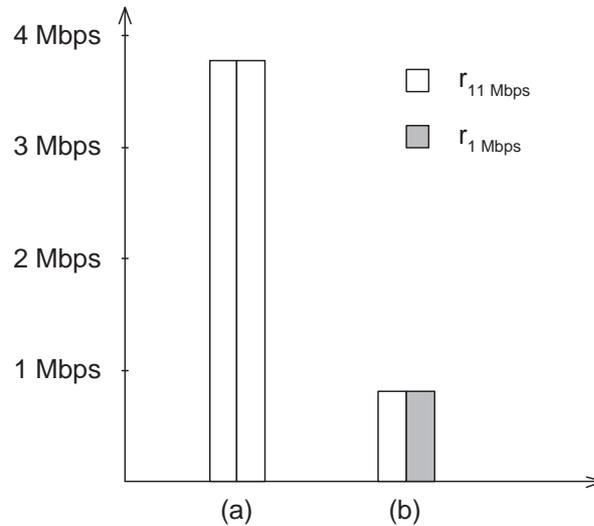


Figura 3.13: Configuración uniforme con estaciones a distinta tasa

velocidad y fiabilidad. Esto puede llevar a lo que Heusse *et al.* denominaron *anomalía en el rendimiento* [28], pues si todas las estaciones acceden al canal con idéntica probabilidad, las estaciones más lentas *arrastrarán* consigo a las estaciones más rápidas⁵.

La causa de dicha anomalía, como se explicó en el Capítulo 2, se encuentra en el mecanismo CSMA/CD, que garantiza –a largo plazo– la misma probabilidad de acceso a una ranura para todas las estaciones en la WLAN, independientemente de su tasa de transmisión. De esta forma, las estaciones con menor tasa de transmisión ocupan más tiempo el canal que aquellas con mejores condiciones de radio. En la Figura 3.13 se muestra el ancho de banda recibido por dos estaciones en una simulación que considera dos escenarios: en el escenario *a*), las dos estaciones en saturación transmiten a una tasa de pico de 11 Mbps; en el escenario *b*), una de las estaciones disminuye su tasa a 1 Mbps.

Por otro lado, la mera aplicación del principio “maximizar la eficiencia en condiciones de saturación” provocaría la inanición de las estaciones lentas, pues son las que disminuyen el rendimiento de la WLAN. Si se desea aplicar el mismo principio de optimización que en la sección anterior, resulta imprescindible definir previamente un criterio de *equidad* entre estaciones que garantice un servicio mínimo a las que empleen los recursos de forma menos eficiente que el resto.

Para hacer frente a esto se emplean los mecanismos disponibles en el modo EDCA. Para ello, en primer lugar se planteará el criterio de equidad a emplear en la WLAN, para luego realizar un modelado analítico del mecanismo MAC de acceso que permita obtener las expresiones para la configuración necesaria a establecer en la WLAN. El hecho de que la configuración para el modo *multirate* propuesta mantenga como objetivo el maximizar el ancho de banda de la WLAN en condiciones de saturación (si bien las estaciones serán ponderadas por diferentes pesos), hace que se eviten las anomalías identificadas en la sección anterior.

⁵Según describen Pan *et al.* [91], en presencia de errores en la transmisión, las estaciones que transmiten a la tasa menor pueden obtener incluso anchos de banda mayores que las de tasa más elevada.

3.3.1. El criterio de equidad proporcional

Por lo anterior, resulta claro que el modo DCF no alcanza un uso óptimo de los recursos de la WLAN, pues supone una pérdida de prestaciones para aquellas estaciones que mantengan un canal radio de calidad. Ello no debe implicar el otro extremo, dado que la maximización del caudal cursado por la red sin atender a otro criterio provocaría que las estaciones a velocidades bajas no tuviesen acceso al recurso inalámbrico.

Para establecer un compromiso entre ambos extremos se acude al criterio de *equidad proporcional* (propuesto por Kelly [92]): dada un vector de asignación de anchos de banda $\{r_1, r_2, \dots, r_n\}$ (donde r_i es la tasa recibida por la estación i), dicho vector es proporcionalmente equitativo si es factible y, para cualquier otra posible asignación $\{r_1^*, \dots, r_n^*\}$, la suma agregada de diferencias no es positiva:

$$\sum_i \frac{r_i^* - r_i}{r_i} \leq 0 \quad (3.31)$$

Para ilustrar las implicaciones de esta definición, considérese el siguiente ejemplo: sea un escenario con dos estaciones. Dado un vector de asignación de anchos de banda, otro vector con un ancho de banda un 10 % menor para la primera estación debería otorgar a la segunda una mejora de más del 10 % para ser una alternativa *mejor* que el primero. Así se alcanza el compromiso entre la ineficiencia de un caso en el que todos reciben el mismo servicio con la injusticia del caso en que sólo se sirve a las estaciones con el mejor canal. Este criterio de equidad proporcional, que será el que se use a lo largo de esta sección, se propuso originariamente en el entorno de las redes cableadas, habiendo sido utilizado para afrontar varios problemas relativos a la asignación de recursos [8, 93, 94].

Para profundizar en el el criterio establecido, considérese el vector de equidad proporcional y una pequeña perturbación en su entorno, $r_i \rightarrow r_i + dr_i$. La Ecuación 3.31 queda como sigue

$$\sum_i \frac{dr_i}{r_i} \leq 0 \quad (3.32)$$

que puede ser expresada como

$$\sum_i (\log(r_i))' dr_i \leq 0 \quad (3.33)$$

Por lo tanto, el punto de asignación equitativa supone un máximo de la función $\sum_i \log(r_i)$. Por ser una función cóncava se deduce que sólo tiene un máximo, por lo que un máximo local también es el máximo absoluto⁶. Siguiendo este resultado, de ahora en adelante se perseguirá aquella asignación que maximice la suma $\sum_i \log(r_i)$.

3.3.2. Análisis de los esquemas de asignación

Para cumplir con el criterio de equidad propuesto, es preciso hacer uso de los mecanismos de diferenciación de la capa MAC. De los tres mecanismos del modo EDCA (el uso del parámetro AIFS, la configuración de la ventana de contienda o el TXOP), el primero

⁶Véase el Teorema 1.2.3 de [95] para una demostración formal.

no resulta especialmente adecuado (véase [52]), por lo que no será objeto de consideración. Se analizarán, por tanto, los siguientes mecanismos:

Esquema CW: Basado en la configuración de los parámetros CW_{min} y CW_{max} de cada estación. De esta forma se puede regular el número de ocasiones en que se accede al canal.

Esquema TL: En este caso, dada una velocidad de transmisión, se fija el tiempo que la estación puede dedicar al envío de sus tramas, por lo que se limita el tamaño máximo de la trama que se puede enviar en un acceso al canal.

Dada la estación i , sean CW_{min}^i y CW_{max}^i los valores de configuración de su ventana de contienda, y sea m_i tal que $CW_{max}^i = 2^{m_i} CW_{min}^i$. Sea asimismo l_i el tamaño de la trama que se intenta transmitir al acceder al canal, R_i la tasa de pico a la que se envía y n el número total de estaciones en la WLAN.

Assumiendo condiciones de saturación, se puede calcular (véase la anterior sección) la probabilidad τ de que una determinada estación transmita tras un decremento de su contador de backoff. Una vez se han obtenido las τ_i 's de todas las estaciones, se procede a calcular el ancho de banda que reciben de la siguiente forma (a continuación se extiende el análisis de la Sección 3.1.3 para considerar el caso *multirate*):

$$r_i = \frac{p(s_i) l_i}{p(s)T_s + p(c)T_c + p(e)T_e} \quad (3.34)$$

En este caso, es preciso realizar el cálculo de la duración media de un éxito (T_s) y de una colisión (T_c), tal y como sigue. La duración media de un éxito se calcula por ponderación de las longitudes de cada estación a su probabilidad de éxito en una determinada ranura de tiempo:

$$T_s = \sum_{i=1}^n \frac{p_{s,i}}{p_s} T_s^i \quad (3.35)$$

Para calcular la duración media de una colisión, es preciso tener en cuenta las longitudes de las transmisiones que intervienen en la misma. A tal efecto, se numeran las estaciones en la WLAN por orden creciente de la duración de su transmisión, y se calcula

$$T_c = \sum_{i=2}^n \frac{p_{c,i}}{p_c} T_c^i \quad (3.36)$$

donde $p_{c,i}$ es la probabilidad de que una ranura de tiempo contenga una colisión en la que intervenga la estación i y ninguna con longitud mayor, y T_c^i es la duración de la colisión en dicho caso.

La probabilidad $p_{c,i}$ se calcula por el producto de la probabilidad de que i transmita, la probabilidad de que ninguna estación de índice superior lo haga y la probabilidad de que al menos una de índice inferior no permanezca sin transmitir:

$$p_{c,i} = \tau_i \prod_{j=i+1}^n (1 - \tau_j) \left(1 - \prod_{j=1}^{i-1} (1 - \tau_j) \right) \quad (3.37)$$

Tabla 3.1: Parámetros de 802.11b empleados en el simulador

Parámetro	1 Mbps	2 Mbps	5.5 Mbps	11 Mbps
R	1 Mbps	2 Mbps	5.5 Mbps	11 Mbps
T_{PCLP}	192 μs	96 μs	96 μs	96 μs
H	34 bytes	34 bytes	34 bytes	34 bytes
ACK	14 bytes	14 bytes	14 bytes	14 bytes
T_e	20 μs	20 μs	20 μs	20 μs
SIFS	10 μs	10 μs	10 μs	10 μs
DIFS	50 μs	50 μs	50 μs	50 μs

Por otro lado, la duración media de una transmisión con éxito de la estación i se calcula mediante las expresiones (2.1) y (2.2) (de nuevo se repiten algunas expresiones con objeto de facilitar la lectura):

$$T_s^i = T_{PCLP}^i + \frac{H + l_i}{R_i} + SIFS + T_{PCLP}^i + \frac{ACK}{R_i} + DIFS \quad (3.38)$$

en la que T_{PCLP}^i es la duración del preámbulo cuando lo transmite la estación i , R_i su tasa de transmisión, l_i el tamaño de las tramas que inyecta a la red y H y ACK el tamaño de la cabecera y del asentimiento, respectivamente.

De forma similar, para el caso de una colisión donde interviene una transmisión de la estación i y ésta resulta ser la de mayor longitud, dicha colisión tiene por tamaño

$$T_c^i = T_{PCLP}^i + \frac{H + l_i}{R_i} + EIFS \quad (3.39)$$

El modelo presentado se valida empleándose los parámetros mostrados en la Tabla 3.1. En primer lugar se valida el modelo para el uso del esquema CW. Sean cuatro grupos de estaciones, con tasas de 11 Mbps, 5.5 Mbps, 2 Mbps y 1 Mbps, con la siguiente configuración: el valor de CW_{min}^{11Mbps} se mantiene fijo a 32, mientras que para un valor de CW_{min}^{1Mbps} (que variará en los experimentos), las otras dos se fijan a los siguientes valores⁷:

$$CW_{min}^{5.5Mbps} = \frac{1}{\frac{1}{CW_{min}^{11Mbps}} + \frac{1}{3} \left(\frac{1}{CW_{min}^{1Mbps}} - \frac{1}{CW_{min}^{11Mbps}} \right)} \quad (3.40)$$

y

$$CW_{min}^{2Mbps} = \frac{1}{\frac{1}{CW_{min}^{11Mbps}} + \frac{2}{3} \left(\frac{1}{CW_{min}^{1Mbps}} - \frac{1}{CW_{min}^{11Mbps}} \right)} \quad (3.41)$$

En las Figuras 3.14, 3.15 y 3.16 se presentan los resultados obtenidos para una variación de CW_{min}^{1Mbps} entre 32 y 1024, en los siguientes casos: 2 estaciones por grupo y $m_i = 5$, 10 estaciones por grupo y $m_i = 5$, y 10 estaciones por grupo y $m_i = 0$, respectivamente.

⁷La elección de estos valores persigue que los anchos de banda recibidos permitan para mostrar adecuadamente el ajuste del análisis a los resultados.

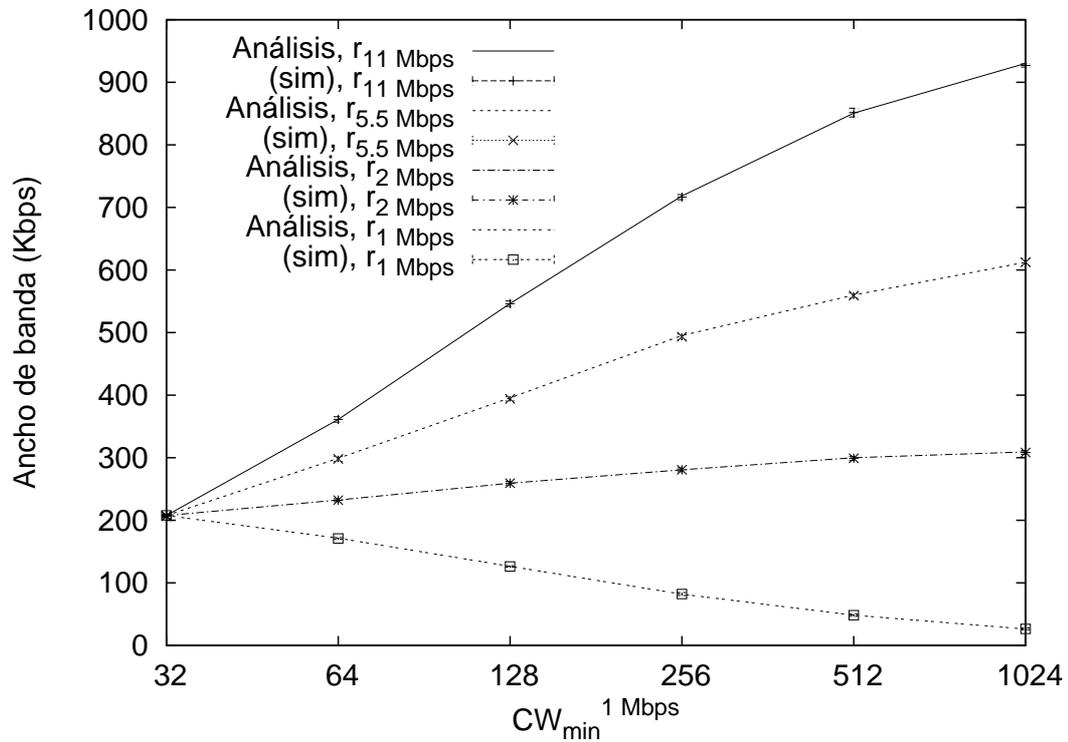


Figura 3.14: Esquema CW, 2 estaciones por grupo, $m_i = 5$

La longitud de las transmisiones se fija a $l_i = 1500$ octetos para todas las estaciones. Los resultados analíticos prácticamente coinciden por los arrojados por la simulación, para los que se han obtenido intervalos de confianza al 95 % para la media.

Para la validación el esquema TL, se emplea el mismo grupo de estaciones. Con 10 estaciones por grupo, se fija que la longitud l_i de las transmisiones de cada grupo sea, respectivamente, de 1500, $1500 - L$, $1500 - 2L$ y $1500 - 3L$ (de menor longitud la de menor velocidad de transmisión), siendo L un valor que se modifica en cada experimento. La configuración del resto de parámetros se corresponde con el estándar DCF. Los resultados de esta validación se muestran en la Figura 3.17.

A la vista de los resultados de simulación, se confirma la validez del modelo propuesto para el análisis de los esquemas CW y TL con los que se perseguirá la consecución de una asignación proporcionalmente equitativa.

3.3.3. Configuración para el esquema CW

Una vez modelado satisfactoriamente el comportamiento de la WLAN, se afronta determinar la configuración de la red que proporciona una asignación de recursos proporcionalmente equitativa. En esta sección, a tal efecto, se hará uso del mecanismo de diferenciación proporcionado por CW. En primer lugar, se mostrará que una condición necesaria es que el tiempo disponible en el canal se comparta de forma equitativa, independientemente de la tasa de cada estación.

De acuerdo con la Sección 3.3.1, el objetivo consiste en encontrar la configuración que

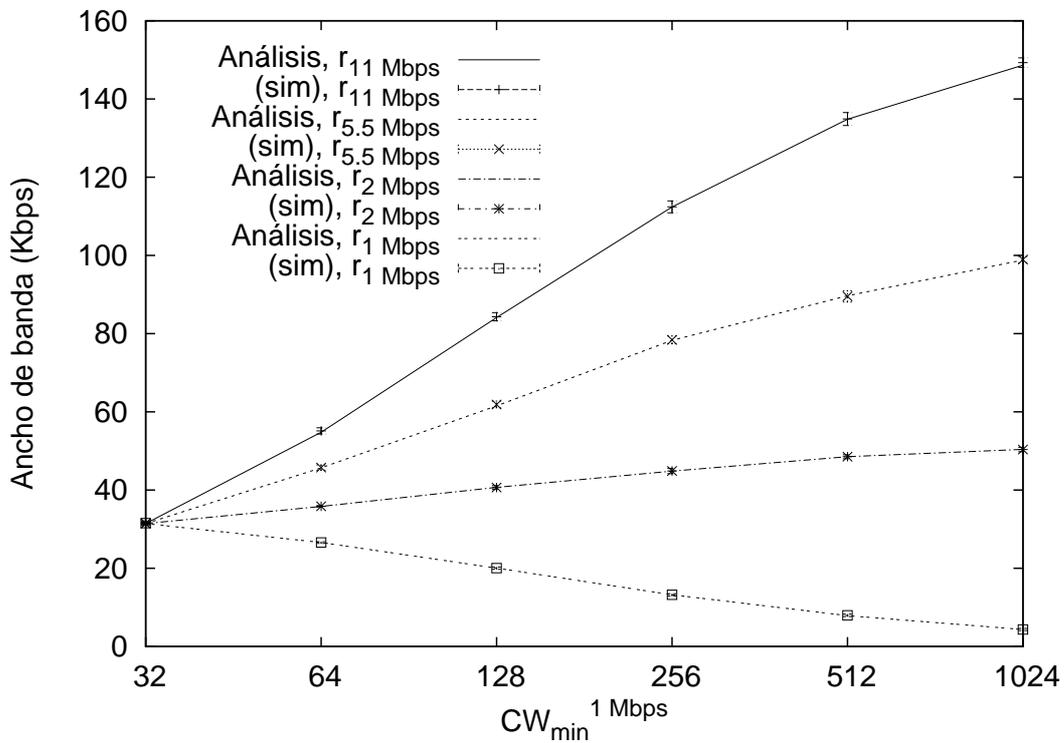


Figura 3.15: Esquema CW, 10 estaciones por grupo, $m_i = 5$

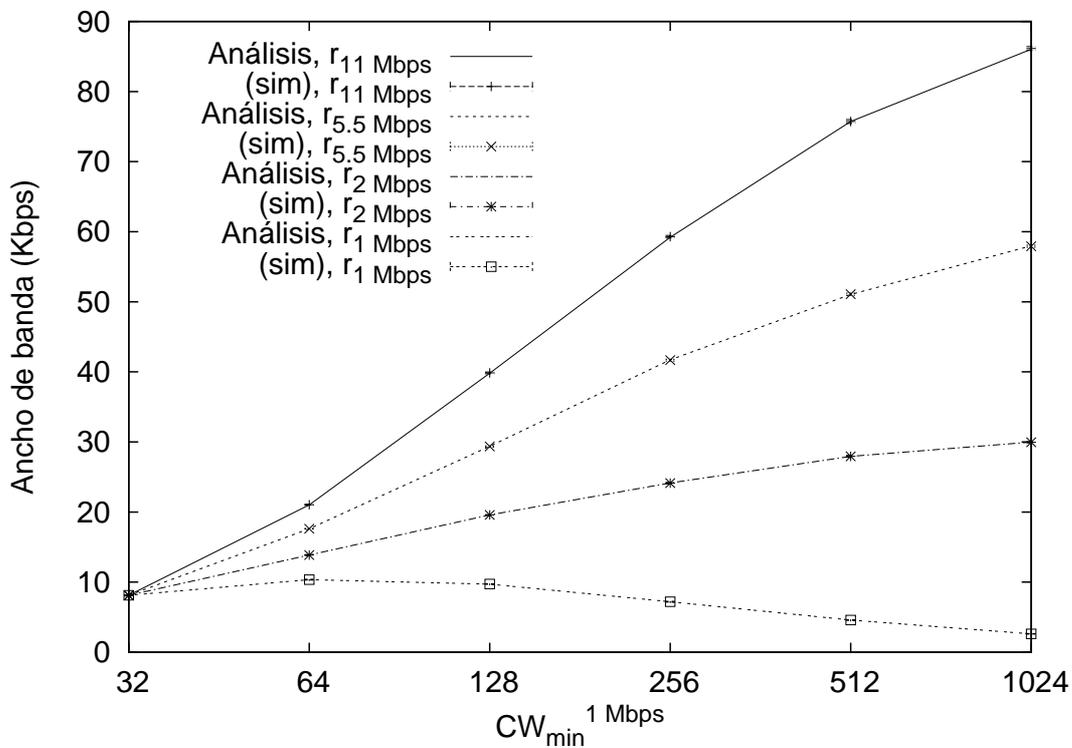


Figura 3.16: Esquema CW, 10 estaciones por grupo, $m_i = 0$

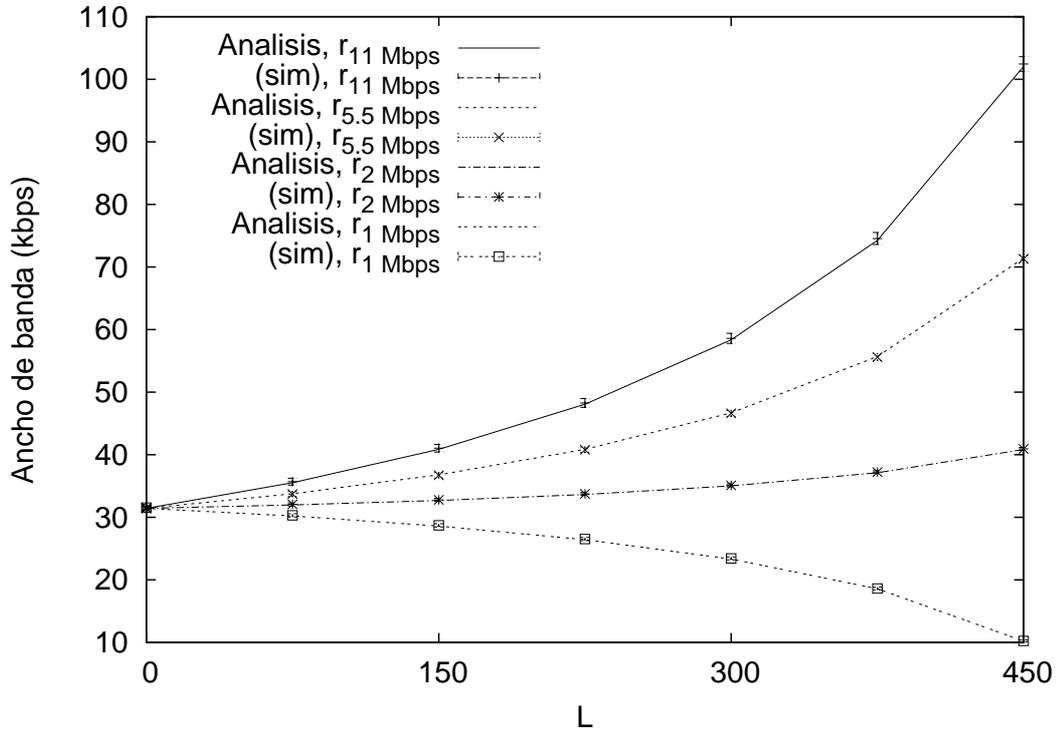


Figura 3.17: Esquema TL, 10 estaciones por grupo

proporciona

$$\max \left(\sum_{i=1}^n \log(r_i) \right) \quad (3.42)$$

siendo r_i el ancho de banda que obtiene la estación i . Ello se traduce en un problema de optimización cuya solución puede obtenerse de

$$\frac{\partial}{\partial \tau_j} \left(\sum_{i=1}^n \log(r_i) \right) = 0 \quad \forall j \quad (3.43)$$

De la Ecuación 3.34 se tiene que r_i se puede expresar de la siguiente manera:

$$r_i = \frac{p_{s,i} l}{E(T_{slot})} \quad (3.44)$$

$E(T_{slot})$ es la duración media de una ranura de tiempo y l es la longitud de una transmisión, que se considera constante para todas las estaciones.

Sustituyendo la anterior en la Ecuación 3.43 y realizando las derivadas se obtiene

$$\left(\frac{1}{\tau_j} - \sum_{i \in \{1, \dots, n\} \setminus j} \frac{1}{1 - \tau_i} \right) E(T_{slot}) - n \frac{\partial E(T_{slot})}{\partial \tau_j} = 0 \quad \forall j \quad (3.45)$$

con n el número de estaciones en la WLAN.

Para $\tau_i \ll 1 \forall i$ se puede aproximar la anterior expresión por⁸

$$E(T_{slot}) - n\tau_j \frac{\partial E(T_{slot})}{\partial \tau_j} = 0 \quad \forall j \quad (3.46)$$

Sumando la k -ésima ecuación con la j -ésima, para todo $j \neq k$, se llega a la siguiente condición necesaria para la equidad:

$$n\tau_j \frac{\partial E(T_{slot})}{\partial \tau_j} - n\tau_k \frac{\partial E(T_{slot})}{\partial \tau_k} = 0 \quad \forall j \neq k \quad (3.47)$$

Despreciándose la probabilidad de colisiones en las que intervengan más de dos estaciones, $E(T_{slot})$ se puede expresar de la siguiente forma:

$$\begin{aligned} E(T_{slot}) = & \sum_{i=1}^n \tau_i \prod_{j \in \{1, \dots, n\} \setminus i} (1 - \tau_j) T_s^i + \\ & + \prod_{i=1}^n (1 - \tau_i) T_e + \sum_{i=1}^n \sum_{j=i+1}^n \tau_i \tau_j \prod_{k \in \{1, \dots, n\} \setminus i, j} (1 - \tau_k) T_c^{i,j} \end{aligned} \quad (3.48)$$

en la que $T_c^{i,j}$ es la longitud de una colisión entre las estaciones i y j . Despreciando los términos de segundo orden y superiores para τ_i , $E(T_{slot})$ se puede expresar como

$$E(T_{slot}) \approx \sum_{i=1}^n \tau_i T_s^i + \left(1 - \sum_{i=1}^n \tau_i\right) T_e \quad (3.49)$$

Lo que lleva a

$$\frac{\partial E(T_{slot})}{\partial \tau_i} = T_s^i - T_e \approx T_s^i \quad (3.50)$$

Por último, se sustituye la anterior expresión en la Ecuación 3.47

$$\frac{\tau_i}{\tau_j} = \frac{T_s^j}{T_s^i} \quad (3.51)$$

Que resulta aproximadamente equivalente a

$$\frac{r_i}{r_j} = \frac{T_s^j}{T_s^i} \quad (3.52)$$

Se concluye, de lo anterior, que el cociente entre anchos de banda expresados en la Ecuación 3.52 es una condición necesaria para proporcionar equidad. Dicha condición consiste en que la tasa a la que una estación transmite con éxito es inversamente proporcional a la duración de sus transmisiones, lo que lleva a que todas las estaciones ocupan el canal durante el mismo tiempo.

A continuación, dependiendo de las herramientas disponibles en la WLAN, se presentan dos alternativas para desarrollar el algoritmo que asigne los recursos de forma equitativa.

⁸Nótese que la suposición de $\tau_i \ll 1$ resulta bastante razonable, dado que valores altos de τ_i provocarían altas probabilidades de colisión y, por lo tanto, a un uso ineficiente de los recursos de la WLAN.

3.3.3.1. Mecanismo centralizado

El mecanismo centralizado precisa de un servidor para la configuración que, mediante algún mecanismo, conoce el número de estaciones en la WLAN y su tasa actual de transmisión. Dicha información puede ser obtenida bien por señalización explícita, bien por monitorización de la actividad en el canal⁹. Con esta información, y con el algoritmo que a continuación se presenta, el servidor suministra a cada estación la configuración $\{CW_{min}, CW_{max}\}$ que debe emplear.

Con la restricción introducida por la Ecuación 3.52, la expresión (3.42) se traduce en:

$$r_i = \frac{1/T_s^i}{\sum_{j=1}^n 1/T_s^j} r \quad (3.53)$$

donde r es el ancho de banda total de la WLAN. El problema de maximización se transforma, por lo tanto, en obtener el mayor valor de la siguiente expresión

$$\sum_{i=1}^n \log \left(\frac{1/T_s^i}{\sum_{j=1}^n 1/T_s^j} \right) + \sum_{i=1}^n \log(r) \quad (3.54)$$

sujeto a la restricción de (3.52). Sea la siguiente expresión para r

$$r = \frac{p_s l}{p_s T_s + p_c T_c + p_e T_e} \quad (3.55)$$

De (3.51) se puede expresar cualquier τ_i en función de τ_1 (que se toma como referencia), como sigue

$$\tau_i = \frac{T_s^1}{T_s^i} \tau_1 = w_i \tau_1 \quad (3.56)$$

donde se define w_i como T_s^1/T_s^i .

Sustituyendo lo anterior en la Ecuación 3.55 y con la suposición de $\tau_1 \ll 1$, se puede aproximar r por

$$r \approx \frac{a\tau_1 - b\tau_1^2}{c\tau_1 + d} \quad (3.57)$$

donde

$$a = \sum_{i=1}^n w_i \quad (3.58)$$

$$b = \sum_{i=1}^n \sum_{j \in \{i+1, \dots, n\}} w_i w_j \quad (3.59)$$

$$c = \sum_{i=1}^n w_i (T_s^i - T_e) \quad (3.60)$$

⁹Si bien existen propuestas para estimar, con información local, el número de estaciones en la WLAN – véase a tal ejemplo [49, 96]–, dichos mecanismos no son aplicables para el caso en que consideramos, dado que se supone que (1) todas las estaciones están configuradas de idéntica manera y (2) no es necesario conocer la velocidad a la que transmite cada estación.

$$d = T_e \quad (3.61)$$

El valor de τ_1 que maximiza la expresión para r , denominado τ_1^* , se obtiene de la anulación de la derivada para r

$$\begin{aligned} \frac{dr}{d\tau_1} \Big|_{\tau_1=\tau_1^*} &= 0 \\ \implies bc(\tau_1^*)^2 + 2bd\tau_1^* - ad &= 0 \\ \implies \tau_1^* &= \frac{\sqrt{(bd)^2 + abcd} - bd}{bc} \end{aligned} \quad (3.62)$$

Obtenido τ_1^* , el resto de las τ_i 's se obtiene mediante (3.56).

De los posibles valores de CW_{min}^i y CW_{max}^i para obtener las τ_i 's se decide fijar, por sencillez, $m_i = 0$ por lo que la expresión para τ_i queda de la siguiente forma:

$$\tau_i = \frac{2}{1 + CW_{min}^i} \quad (3.63)$$

Por lo tanto, y teniendo en cuenta que los valores de CW deben ser números enteros, la configuración óptima que se suministra a cada estación viene determinada por

$$CW_i^* = \text{round int} \left(\frac{2}{\tau_i^*} - 1 \right) \quad (3.64)$$

Por último, resulta adecuado señalar que el cálculo de la configuración no implica complejidad computacional destacable, al requerir la realización de pocas operaciones básicas.

3.3.3.2. Mecanismo distribuido

Suponer la existencia de un servidor centralizado, con la información necesaria para ejecutar el algoritmo de la sección anterior, implica una complejidad no despreciable para el uso del algoritmo propuesto. Es por ello que se plantea un mecanismo alternativo que, si bien proporcionará un funcionamiento subóptimo, tendrá un coste de desarrollo menor.

Dado que no existe información global disponible, cada estación debe fijar su configuración de CW con base a la información local disponible, que es la tasa de transmisión. Para este caso se mantiene el retroceso exponencial binario con $m_i = 5$, lo que permite un funcionamiento que se adapte a las condiciones del entorno.

Para las estaciones a 11 Mbps, se mantiene la configuración del estándar. De esta forma, y con el uso de 802.11b, $CW_{min}^{11Mbps} = 32$. Con la suposición de que todas las estaciones colisionan con la misma probabilidad, el ancho de banda que recibe una estación resulta ser inversamente proporcional a la CW_{min} que tenga configurada. Por lo tanto, para cumplir con la condición fijada en (3.52), se establece el parámetro de CW_{min}^i para el resto de estaciones como sigue:

$$CW_{min}^i = \frac{T_s^i}{T_s^{11Mbps}} CW_{min}^{11Mbps} \quad (3.65)$$

3.3.4. Configuración para el esquema TL

Mediante el uso del esquema TL, las únicas variables que se pueden modificar son las longitudes l_i que utilizan las estaciones para transmitir sus tramas una vez tienen acceso al canal, que deberán ser configuradas para maximizar la suma $\sum_i \log(r_i)$. De nuevo, se procede como en la sección anterior:

$$\frac{\partial}{\partial l_j} \left(\sum_{i=1}^n \log(r_i) \right) = 0 \quad \forall j \quad (3.66)$$

De lo que se deriva

$$\frac{1}{r_j} \frac{\partial r_j}{\partial l_j} + \sum_{i \in \{1, \dots, n\} \setminus j} \frac{1}{r_i} \frac{\partial r_i}{\partial l_j} = 0 \quad \forall j \quad (3.67)$$

El ancho de banda que recibe una estación i se expresa como

$$r_i = \frac{p_{s,i} l_i}{E(T_{slot})} \quad (3.68)$$

lo que lleva a las siguientes expresiones:

$$\frac{\partial r_j}{\partial l_j} = \frac{p_{s,j} E(T_{slot}) - p_{s,j} l_j \frac{\partial E(T_{slot})}{\partial l_j}}{E(T_{slot})^2} \quad (3.69)$$

y

$$\frac{\partial r_i}{\partial l_j} = - \frac{p_{s,i} l_i}{E(T_{slot})^2} \frac{\partial E(T_{slot})}{\partial l_i} \quad (3.70)$$

Por otra parte, $E(T_{slot})$ se puede expresar como suma de las diferentes longitudes y probabilidades de éxito:

$$E(T_{slot}) = \sum_{i=1}^n p_{s,i} T_s^i + p_c T_c + p_e T_e \quad (3.71)$$

con

$$T_{s,i} = K + \frac{l_i}{R_i} \quad (3.72)$$

siendo K una constante que engloba a los términos no dependientes de l_i (véase la Ecuación 3.38).

Con la anterior relación entre $T_{s,i}$, y despreciando el efecto del término $p_c T_c$ ¹⁰, se tiene que

$$\frac{\partial E(T_{slot})}{\partial l_j} = \frac{p_{s,i}}{R_i} \quad (3.73)$$

¹⁰Esta aproximación es similar a la realizada en la sección anterior, donde se suponía $\tau_i \ll 1$ y, por tanto, se despreciaron los términos de segundo orden en τ_i . Dichos términos son los principales componentes de la probabilidad de colisión.

Sustituyendo las expresiones (3.69) y (3.73) en (3.67) se obtiene

$$\frac{1}{r_j} E(T_{slot}) - \frac{1}{R_j} \sum_{\forall i} \frac{1}{r_i} p_{s,i} l_i = 0 \quad \forall j \quad (3.74)$$

de lo que se deduce

$$\frac{r_j}{R_j} = \frac{E(T_{slot})}{p_s \sum \frac{l_i}{r_i}} \quad \forall j \quad (3.75)$$

Dado que la parte derecha de la anterior ecuación es una constante, se deriva que la condición necesaria para la configuración óptima es la siguiente

$$\frac{r_i}{R_i} = \frac{r_j}{R_j} \quad \forall i, j \quad (3.76)$$

En el esquema TL, la diferenciación se obtiene a través de configurar las longitudes que puede transmitir cada estación. La configuración de la ventana de contienda es la misma para todas las estaciones, por lo que $p_{s,i} = p_{s,j}$ lo que, aplicado en (3.68), lleva a lo siguiente:

$$\frac{l_i}{l_j} = \frac{R_i}{R_j} \quad \forall i, j \quad (3.77)$$

Esta condición necesaria para la equidad implica que las longitudes que se transmitan deben ser inversamente proporcionales a las tasas que se empleen para hacerlo. Esto lleva a que las estaciones ocupen, aproximadamente, el canal durante el mismo tiempo: se llega al resultado de la sección anterior.

Dada una longitud para las transmisiones de una estación de referencia, el resto de longitudes se calculan de la forma indicada. Falta, por tanto, fijar dicha longitud de referencia. Fijar una longitud de trama supone llegar a un compromiso entre retardo y eficiencia: a mayores longitudes de transmisión, menor es el impacto de los datos de control introducidos por lo que mayor es el caudal efectivo que se puede transmitir; sin embargo, el hecho de ocupar el canal durante más tiempo provoca un aumento del tiempo que debe esperar el resto de estaciones antes de poder transmitir su tráfico.

Por lo anterior, se decide que las estaciones que transmitan a 11 Mbps empleen el tamaño propuesto por el estándar, que se supone fija un compromiso aceptable entre los anteriores extremos. De esta forma, cuando una estación disminuya su tasa de transmisión, seguirá ocupando el canal durante el mismo tiempo, por lo que no se influirá en el servicio recibido por el resto de estaciones. Por ello, las estaciones a 11 Mbps transmitirán un tamaño $l_{default}$ mientras que el resto fijarán su tamaño l_i al valor $l_{default} R_i / 11 Mbps$ (según la Ecuación 3.77).

Sólo resta la configuración de los demás parámetros. Con el esquema TL los valores de CW son idénticos para todas las estaciones; para fijar un valor, se puede emplear dos mecanismos diferentes:

- Un mecanismo centralizado, donde se calcula la CW óptima con las expresiones de la anterior sección, suponiéndose que $w_i = 1$ y $T_s^i = T_s \forall i$.
- Un mecanismo distribuido donde, en ausencia de información global, se utilicen los parámetros propuestos por el estándar.

3.3.5. Evaluación del rendimiento de los esquemas

A continuación se analiza el rendimiento de los cuatro esquemas propuestos: el esquema TL y el CW, cada uno para el caso centralizado y el distribuido. En primer lugar, se comparará su rendimiento frente al modo sin diferenciación. A continuación, se evaluará las prestaciones en retardo frente al caso en que todas las estaciones transmiten a la mayor tasa. Por último, se estudiará si existe alguna otra configuración del tipo CW, TL o mixta que mejore los esquemas centralizados propuestos.

La valoración de los esquemas se hará con base en los resultados proporcionados por el modelado analítico de la Sección 3.3.2. Dado que dicho modelo ha sido validado para una variedad de escenarios, proporcionando para todos ellos resultados muy cercanos a los obtenidos mediante simulación, emplear esta estrategia permite realizar los cálculos en tiempos de computación asequibles: un barrido exhaustivo en simulación requeriría de una capacidad de cálculo muy elevada. En todos los casos se ha considerado un tamaño de l de 1500 octetos, y los resultados para el retardo se han obtenido del modelo que se presenta en el Capítulo 4.

3.3.5.1. Asignación de anchos de banda

Se repite el experimento de la Figura 3.13 para el esquema CW centralizado (*c*), el distribuido (*d*), el esquema TL centralizado (*e*) y el distribuido (*f*), con los resultados que aparecen en la Figura 3.18. Se aprecia que los cuatro esquemas propuestos proporcionan, básicamente, la misma asignación. En todos ellos, las estaciones que transmiten a la velocidad elevada no son severamente afectadas en comparación con el escenario en que todas transmiten a la velocidad más alta (*a*), al contrario de lo que sucede con el esquema por defecto (*b*). Se concluye que con una asignación proporcionalmente equitativa, cuando una estación transmite a una velocidad menor sólo el ancho de banda de dicha estación se ve afectado, mientras que el resto de estaciones mantiene aproximadamente la misma tasa de transmisión. Esto sucede con independencia del mecanismo empleado para proporcionar la equidad.

3.3.5.2. Rendimiento de la red

De la definición de proporcionalidad equitativa se sigue que el rendimiento de una determinada asignación es medido en términos de $\sum_i \log(r_i)$: cuanto mayor sea dicha suma, mejor será la asignación. De ahora en adelante se utilizará esta métrica para la valoración del rendimiento de los esquemas.

Sea un escenario WLAN en el que la mitad del número total de estaciones n transmite a 11 Mbps, y la otra mitad a 1 Mbps. En la Figura 3.19 se muestra dicha métrica en función del número de estaciones en la WLAN para los cuatro esquemas propuestos, además de para la configuración uniforme—fijada como referencia—.

Todas las propuestas mejoran sustancialmente el rendimiento original, lo que indica el beneficio de utilizar los esquemas propuestos. Los esquemas centralizados, claramente, proporcionan mejores resultados que los distribuidos, lo que era un resultado a esperar habida cuenta de la mayor cantidad de información que se maneja: existe un compromiso entre el rendimiento y la complejidad de implementación. Otro resultado que se puede sacar

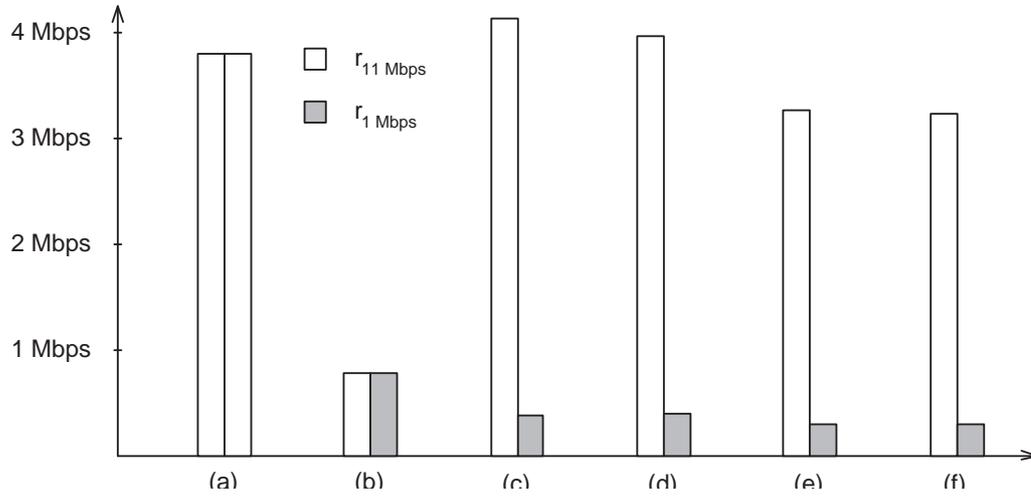


Figura 3.18: Asignación de anchos de banda para los diferentes esquemas

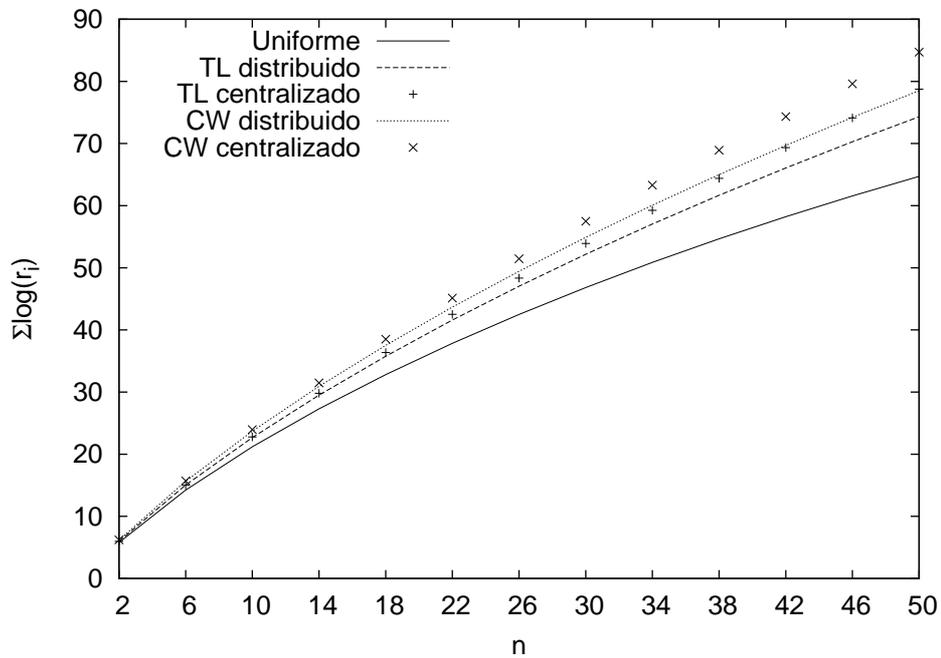


Figura 3.19: Valoración de asignaciones, dos tasa de transmisión

Tabla 3.2: Evaluación del rendimiento de los cuatro esquemas para 4 grupos y 5 estaciones

Esquema	Uniforme	CW centralizado	CW distribuido	TL centralizado	TL distribuido
r_{11Mbps} (Kbps)	71.68	400.65	357.74	328.52	293.61
(CW_{min}, CW_{max})	(32,1024)	(213,213)	(32,1024)	(383,383)	(32,1024)
l (bytes)	1500	1500	1500	1500	1500
$r_{5.5Mbps}$ (Kbps)	71.68	201.27	185.34	164.26	146.81
(CW_{min}, CW_{max})	(32,1024)	(424,424)	(58,1856)	(383,383)	(32,1024)
l (bytes)	1500	1500	1500	750	750
r_{2Mbps} (Kbps)	71.68	78.01	70.17	59.79	53.44
(CW_{min}, CW_{max})	(32,1024)	(1094,1094)	(150,4800)	(383,383)	(32,1024)
l (bytes)	1500	1500	1500	273	273
r_{1Mbps} (Kbps)	71.68	42.90	35.09	29.79	26.62
(CW_{min}, CW_{max})	(32,1024)	(1989,1989)	(298,9536)	(383,383)	(32,1024)
l (bytes)	1500	1500	1500	136	136
$\sum_i \log(r_i)$	37.11	42.16	41.06	39.91	38.94

de la Figura 3.19 es que las configuraciones centralizada y distribuida del esquema CW proporcionan mejores resultados que las del esquema TL. Esta afirmación, sin embargo, no tiene en cuenta las prestaciones en término de retardo, lo que será objeto de análisis en la siguiente sección.

Las alternativas distribuidas proporcionan resultados similares a las centralizadas para valores bajos de n , si bien su rendimiento empeora a medida que n aumenta. El motivo de este comportamiento es evidente: los valores de CW_{min} escogidos para las opciones distribuidas ($CW_{min}^i = 32$ para el esquema TL, y $CW_{min}^i = 32T_s^i/T_s^{11Mbps}$ para el CW) resultan adecuados para valores bajos de n , no así para los altos. Varios algoritmos para adaptar el valor de la CW en función del estado de la red (para determinar n) han sido propuestos para tratar este problema (a tal efecto véase, por ejemplo, [44, 97, 98]); estos algoritmos se podrían utilizar en los esquemas distribuidos propuestos.

En la Tabla 3.2 se muestra el ancho de banda recibido por cada estación, la configuración y $\sum_i \log(r_i)$ para el caso en que el canal se comparte por 4 grupos de 5 estaciones cada uno, transmitiéndose a las cuatro diferentes velocidades (1 Mbps, 2 Mbps, 5.5 Mbps, 11 Mbps) del estándar. Los resultados proporcionan las mismas conclusiones que las anteriormente descritas.

3.3.5.3. Tiempo de servicio

De los resultados de la sección anterior, podría deducirse que los esquemas CW son *preferibles* a los esquemas TL, dado que la métrica considerada ($\sum_i \log(r_i)$) arroja un resultado mayor para aquellos que para estos. Sin embargo, y como se verá en esta sección, las prestaciones que ofrecen en tiempo de servicio son diferentes, por lo que no ha lugar a una comparación en dichos términos entre ambos.

Considérense los esquemas centralizados. En la Figura 3.20 se muestra el tiempo medio de acceso para un escenario con idéntico número de estaciones a 1 Mbps y a 11 Mbps. Se muestra, como referencia, el tiempo medio de servicio para el caso en que todas las

estaciones transmiten a la máxima velocidad. Los resultados indican que la naturaleza del servicio y de las prestaciones de la red son bien diferentes para cada uno de los esquemas:

- Por una parte, el esquema TL proporciona el mismo tiempo medio de acceso al canal para todas las estaciones, independientemente de la velocidad a la que transmitan. Esto se debe, obviamente, a que todas las estaciones son configuradas con el mismo valor de la ventana de contienda, por lo que en media deberán esperar aproximadamente el mismo número de ranuras. Dichas ranuras, por configuración, tienen aproximadamente la misma duración independientemente de la velocidad a la que se transmita.
- Por otra parte, el esquema CW proporciona valores bien diferentes en el tiempo medio de servicio para cada grupo de estaciones. El grupo que transmite a 11 Mbps mantiene un tiempo medio de acceso similar al caso de referencia, mientras que el grupo que transmite a 1 Mbps sufre un considerable aumento de dicho tiempo. El motivo para esto es que, si bien ambos grupos ocupan el canal durante el mismo tiempo, en este esquema esto se consigue con valores elevados de la ventana de contienda para las estaciones que transmiten a 1 Mbps.

Por lo tanto, al comparar los esquemas TL y CW, se obtiene que ninguno es *mejor* que el otro. Las prestaciones en la métrica considerada son mejores para el caso CW, si bien esto se logra mediante una clara diferenciación en los tiempos medio de acceso al canal para cada grupo de estaciones. El esquema TL, en cambio, permite un retardo de acceso al canal similar para ambos grupos de estaciones, si bien sus prestaciones en términos de equidad proporcional resultan ligeramente inferiores.

En la Figura 3.21 se muestra el retardo medio de acceso que suministran los esquemas distribuidos, de nuevo comparados con un caso de referencia en el que todas las estaciones transmiten a la máxima velocidad. Los resultados muestran las mismas conclusiones indicadas anteriormente.

3.3.5.4. Idoneidad de los esquemas centralizados

Los esquemas centralizados persiguen obtener la mejor configuración posible bajo la condición de equidad proporcional. Con objeto de investigar si existen configuraciones mejores que las propuestas, en esta sección se compara el rendimiento de los esquemas descritos frente a una búsqueda exhaustiva sobre todas las configuraciones posibles: se realiza una búsqueda sobre todo el espacio de configuración (para todas las estaciones) y se escoge aquella que proporciona el mejor rendimiento. Nótese que este proceder resultaría poco práctico para una situación real, y que sus fines son únicamente comparativos.

Esquemas independientes. Para el esquema CW se realiza una búsqueda exhaustiva en los parámetros CW_{min}^i y CW_{max}^i de todas las estaciones, escogiéndose aquella configuración que proporciona un mejor resultado (es decir, el mayor valor de $\sum_i \log(r_i)$), sujeto a la condición de que las prestaciones de retardo de las estaciones a 11 Mbps no se reduzcan. De esta forma se mantiene el objetivo de que el esquema CW sigue proporcionando el

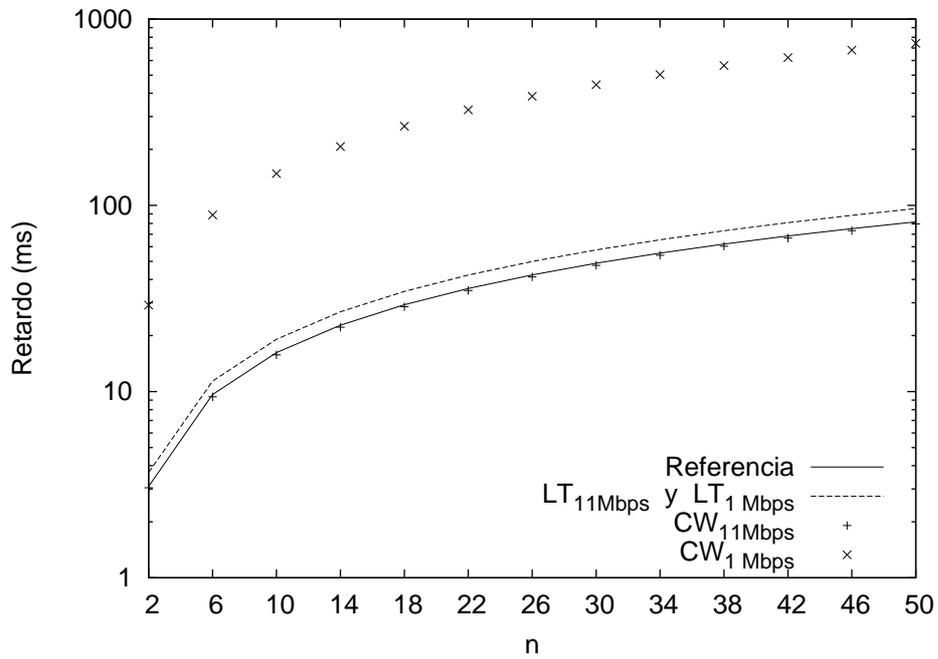


Figura 3.20: Retardo de acceso para los esquemas centralizados

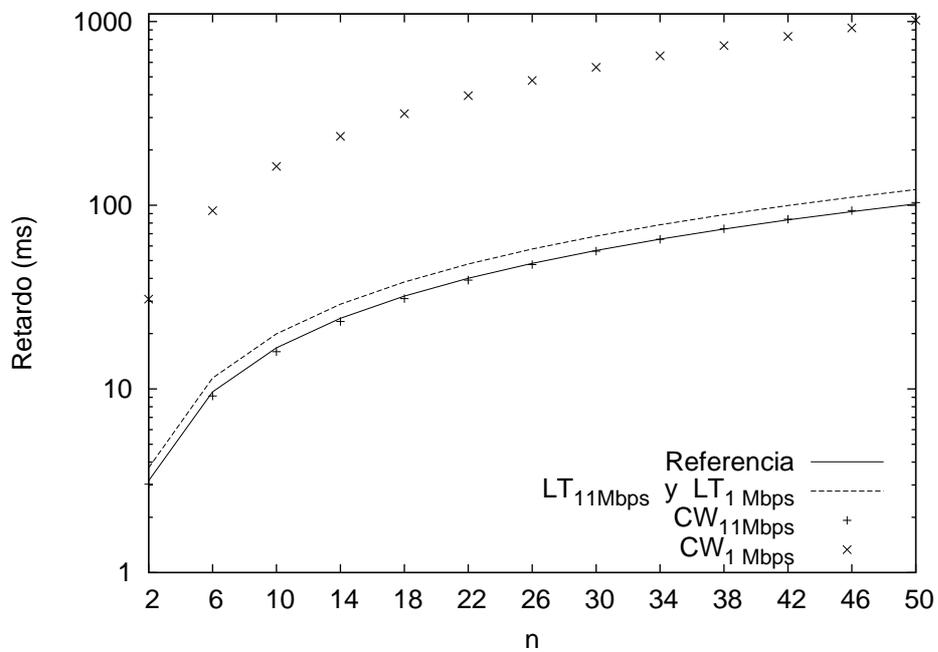


Figura 3.21: Retardos de acceso para los esquemas distribuidos

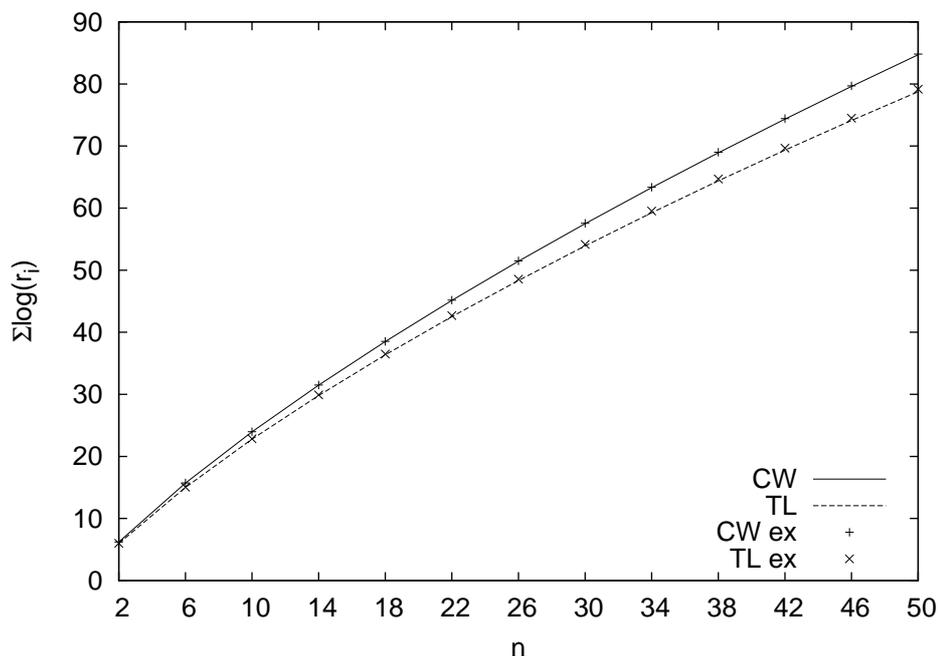


Figura 3.22: Comparación con búsqueda exhaustiva para los esquemas propuestos

mismo tiempo medio de acceso al canal para las estaciones que no disminuyen su tasa de transmisión¹¹.

Para el esquema TL, se realiza un barrido en el espacio de l_i , CW_{min} y CW_{max} (estos últimos idénticos para todas las estaciones), escogiéndose la configuración que proporciona el mejor rendimiento, sujeto a la condición que el retardo medio de acceso de todas las estaciones no se degrada, por motivos similares.

En la Figura 3.22 se muestra el resultado de la comparación de las alternativas centralizadas propuestas ('CW' y 'TL'), frente a los barridos exhaustivos para dichas alternativas ('CW ex' y 'TL ex'). Los resultados validan la configuración propuesta, al no existir configuración alternativa para los esquemas TL y CW que, manteniendo las prestaciones en cuanto a retardo, proporcione una mejora significativa en el rendimiento de la WLAN en comparación con los esquemas propuestos.

Esquema combinado. En la sección anterior se ha restringido la búsqueda bien al modelo del esquema CW, bien al modelo de TL, pero en ningún caso se ha permitido emplear ambos mecanismos a la vez. En este apartado, en cambio, se realiza una búsqueda exhaustiva en todo el posible espacio de configuración; esto es, cada grupo de estaciones puede tener tanto diferente configuración de CW_{min}^i y CW_{max}^i como de l_i .

A la hora de realizar el barrido, se fijan dos criterios diferentes para el retardo medio: tomando como referencia la situación en la que todas las estaciones transmiten a 11 Mbps,

¹¹Concretamente, la condición impuesta en la búsqueda exhaustiva es que el retardo de las estaciones a 11 Mbps no sea mayor que en el esquema CW. Esta condición garantiza que ambos esquemas cumplen con el objetivo de mantener el tiempo medio de acceso de las estaciones a 11 Mbps al mismo nivel.

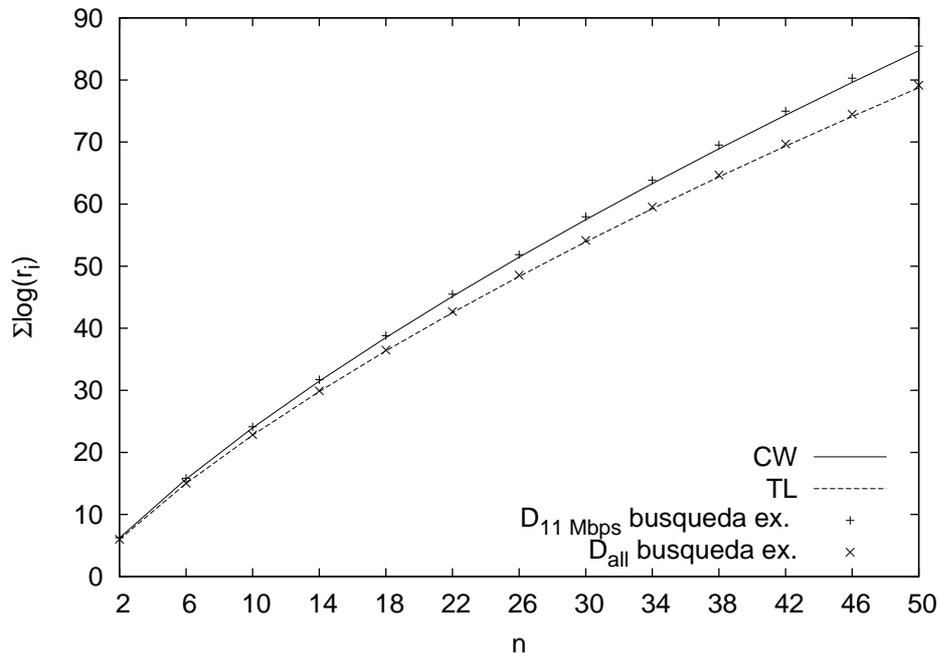


Figura 3.23: Comparación con búsqueda exhaustiva en un esquema combinado

los criterios son 1) que el retardo medio de las estaciones a 11 Mbps se mantenga a pesar del cambio de tasa de las otras estaciones; o 2) que el retardo medio de acceso se mantenga para todas las estaciones (lo que se representa por $D_{11\text{ Mbps}}$ y D_{all} , respectivamente).

En la Figura 3.23 se muestra el resultado de comparar los esquemas centralizados propuestos ('CW' y 'TL') con las dos búsquedas exhaustivas definidas anteriormente. Se comprueba que, con los criterios definidos, no existe una configuración que mejore significativamente los esquemas presentados. Se concluye, por tanto, que dependiendo del comportamiento que se desee en la provisión de tiempos medios de acceso, se debe emplear el esquema CW o el TL, dado que no existe alternativa que mejore los resultados de las anteriores propuestas.

3.4. Detección de configuraciones maliciosas

En las anteriores secciones se ha definido la configuración más adecuada para redes WLAN abiertas en las que, por la propia naturaleza del escenario considerado, se ha decidido maximizar el rendimiento bajo un criterio de equidad entre usuarios. Todos los usuarios son tratados sin privilegios, salvo por cambios en su tasa de transmisión con objeto de mejorar las prestaciones globales de la WLAN. Sin embargo, suponer que en tal escenario todos los usuarios se ajustarán a la configuración que se les suministra, sin intentar obtener un servicio *un poco mejor*, resulta muy arriesgado. De hecho, como se mencionó en el Capítulo 2, existen tarjetas inalámbricas 802.11e que no se ajustan al comportamiento definido por el estándar y/o permiten modificar los parámetros de cada AC a los propios usuarios.

En esta sección se abordará la definición y validación de un algoritmo de detección de

configuraciones maliciosas del modo EDCA, entendiéndose por *configuración maliciosa* toda aquella que persigue obtener un servicio mejor que el que le corresponde a un determinado usuario, por modificación de los parámetros EDCA. El proceso de detección se basará en la observación, a intervalos regulares, del comportamiento de un usuario, con obtención de métricas que permitirán estimar con una determinada verosimilitud si dicho usuario ha configurado su tarjeta de una forma diferente a la proporcionada por la red.

Si bien el modo EDCA define varios parámetros susceptibles de reconfiguración por parte de un usuario, sólo será preciso centrar la atención en lo referente a la ventana de contienda (CW). Esto es así ya que cambios maliciosos (esto es: con objeto de obtener un mejor servicio) tanto del tanto el parámetro $AIFS$ como del parámetro $TXOP_{limit}$ son fácilmente detectables por su naturaleza determinista.

El caso de la ventana de contienda, sin embargo, es radicalmente diferente a los anteriores, debido al comportamiento aleatorio de dicho parámetro: es por ello que la clasificación de una determinada estación como “maliciosa” o “no maliciosa” irá aparejada en todos los casos de una determinada probabilidad de error; esto es así ya que las instancias de *backoff* se ejecutan a partir de una variable aleatoria uniformemente distribuida entre 0 y $CW_{min} - 1$, y siempre existirá la probabilidad de que una estación bien configurada obtenga de forma consecutiva valores muy bajos de dicha variable aleatoria y, por lo tanto, un tiempo de servicio menor que el de otra estación (durante un determinado intervalo de observación).

Por otra parte, un efecto a considerar en este caso es el denominado *efecto captura*. Éste se produce cuando, tras una colisión entre dos (o más) tramas, la que llega al receptor con mayor potencia es bien recibida, a pesar de dicha colisión –por lo que se dice que *captura* el canal–. Es preciso hacer notar que, en este caso, la colisión se traduce en una transmisión con éxito de la estación que captura el canal, por lo que su CW no crece para el siguiente proceso de *backoff*.

Con objeto de valorar cuantitativamente el impacto del efecto captura, se tiene el siguiente resultado [34]: dado un punto de acceso, se sitúa una estación a 2 metros del mismo y otra estación a 4 metros. Ambas estaciones se encuentran en saturación, siempre con un paquete UDP pendiente de ser transmitido. Se mide el ancho de banda obtenido por cada estación, para tres valores de CW_{min} . A continuación, se colocan ambas estaciones a la misma distancia –4 metros–, repitiendo el experimento. Los resultados obtenidos en cada caso se presentan en las Figuras 3.24 y 3.25.

En dichas figuras también se compara el resultado obtenido con los análisis de [99] y [27], comprobándose que los valores experimentales coinciden con los teóricos. Lo que es más importante, se concluye que para distancias no demasiado elevadas el efecto captura puede ocurrir con notable frecuencia en una WLAN y que, por lo tanto, debe ser tenido en cuenta en el algoritmo de detección que se desarrolle: sólo en el caso en que un usuario obtenga más ancho de banda que aquel que le correspondería si siempre se beneficiase del efecto captura podremos afirmar que dicho usuario ha modificado, para obtener un mejor servicio, la configuración de su tarjeta inalámbrica.

3.4.1. Detección de una configuración maliciosa de CW_{min}

Como se ha explicado en la sección anterior, el principal reto en la detección de configuraciones maliciosas de los parámetros del modo EDCA es discernir si la CW de

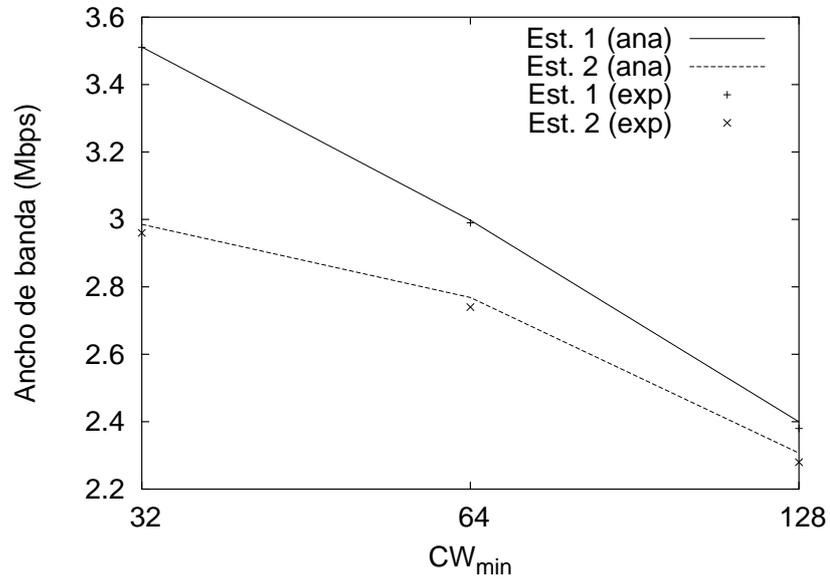


Figura 3.24: Ancho de banda obtenido en presencia de efecto captura

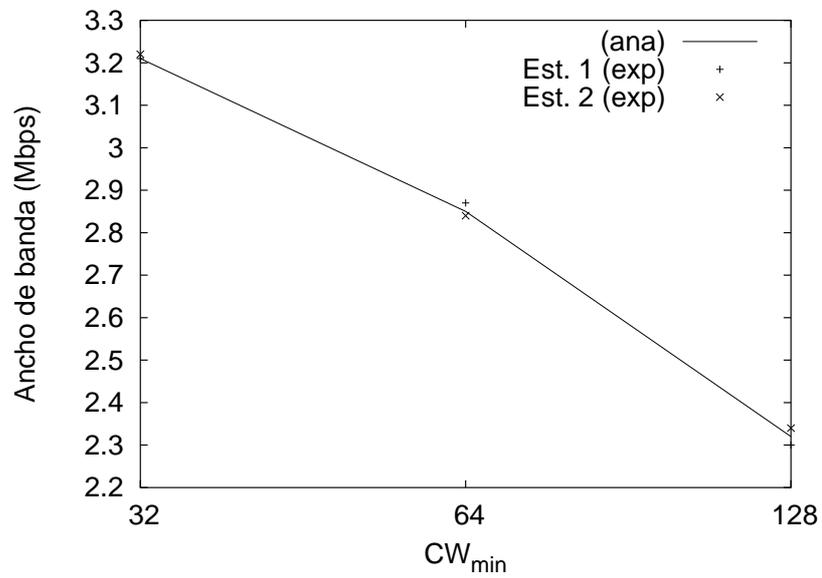


Figura 3.25: Ancho de banda obtenido en ausencia de efecto captura

un determinado usuario, que siempre capture el canal en caso de colisión, está configurada según los parámetros determinados por la red, o bien dicha configuración ha sido modificada por el usuario con objeto de obtener un servicio mejor.

El algoritmo aquí presentado puede ser ejecutado por cualquier estación en el rango de la WLAN en consideración, y se puede controlar cualquier estación perteneciente a la misma. De forma específica, dicho algoritmo puede ser ejecutado en el punto de acceso para controlar a todas las estaciones de la red. De ahora en adelante se asume, sin pérdida de generalidad, que existe una única estación que ejecuta el algoritmo y una única estación supervisada por aquélla.

El algoritmo funciona de la siguiente forma: el canal es monitorizado durante intervalos de observación de una duración configurable T_{obs} , en los cuales se intentará averiguar a partir del comportamiento de la estación controlada si su configuración es inadecuada. De cara a conseguir este objetivo, el algoritmo mide las siguientes variables durante dicho intervalo:

N el número de *ranuras de tiempo* contenidas en el intervalo (esto es, el número de veces que dicha estación debería haber decrementado el contador en dicho intervalo). El seguimiento de dicha variable no supone gran complejidad, ya que cualquier estación de la WLAN tiene por definición la capacidad de ejecutar dicho algoritmo para todas las configuraciones posibles.

S el número de transmisiones con éxito que ha realizado la estación controlada durante T_{obs} .

Entre dos transmisiones exitosas consecutivas –asumiendo condiciones de saturación–, una estación debe esperar como mínimo un número de ranuras de tiempo uniformemente distribuidas en el intervalo $(0, CW_{min} - 1)$. Por ello, en las condiciones más favorables, el número total de ranuras de tiempo contadas cuando la estación controlada ha realizado S transmisiones exitosas debe seguir la siguiente variable aleatoria:

$$N_{total} = \sum_{i=1}^S Unif(1, CW_{min}) \quad (3.78)$$

por lo que el número medio de ranuras de tiempo entre un éxito y el siguiente seguirá la siguiente variable aleatoria:

$$N_{avg} = \frac{1}{S} \sum_{i=1}^S Unif(1, CW_{min}) \quad (3.79)$$

Teniendo en cuenta que N_{avg} es la media de un número *elevado* de variables aleatorias independientes, según el Teorema Central del Límite se puede modelar como una distribución normal, cuya media y varianza se calcula de la siguiente forma:

$$m = \sum_{i=1}^{CW_{min}} \frac{i}{CW_{min}} = \frac{CW_{min} + 1}{2} \quad (3.80)$$

$$\sigma^2 = \frac{1}{S} \left(\sum_{i=1}^{CW_{min}} \frac{i^2}{CW_{min}} - \left(\frac{CW_{min} + 1}{2} \right)^2 \right) \quad (3.81)$$

Si la CW_{min} de una estación está configurada adecuadamente, el número de ranuras contabilizado en el intervalo de tiempo dividido por el número de éxitos de la estación controlada, N/S , será una muestra de la variable aleatoria N_{avg} . La probabilidad de que esta muestra tenga un valor inferior a un determinado umbral $m - K\sigma$ se calcula de forma inmediata:

$$P(N_{avg} < m - K\sigma) = 0,5 \operatorname{erfc} \left(-\frac{K}{\sqrt{2}} \right) \quad (3.82)$$

Un usuario que haya configurado su CW_{min} a un valor menor del que le correspondiese, con objeto de obtener un mayor ancho de banda, mostrará típicamente valores de N/S menores del esperado. De acuerdo con esto, el algoritmo procede de la siguiente forma: dado S , se calcula m y σ de acuerdo a (3.80) y (3.81). A continuación, se compara el cociente N/S con el valor $m - K\sigma$. Si $N/S < m - K\sigma$, el algoritmo clasifica al usuario como malicioso. Si no, se considera que su comportamiento se corresponde con la configuración que debe tener.

El algoritmo descrito se basa en los siguientes parámetros:

P_{FA} Probabilidad de falsa alarma, que es la probabilidad de clasificar como malicioso un usuario con un comportamiento correcto.

T_{obs} Duración del intervalo de confianza, que se corresponde con la *granularidad temporal*. Que este valor sea bajo permite detectar en poco tiempo un comportamiento avaricioso por parte de los usuarios.

K : este parámetro se corresponde con la *probabilidad de detección* del algoritmo; cuando menor sea su valor, menos ancho de banda puede obtener un usuario sin ser detectado.

Los tres parámetros se encuentran relacionados, de modo tal que fijados dos de ellos, el tercero está dado. De esta forma:

- Cuanto menor sea la probabilidad de falsa alarma, menos sensibilidad presenta el algoritmo. Esto es así dado que, para disminuir aquella, K debe tener un valor elevado por lo que el ancho de banda adicional que puede obtener un usuario malicioso es mayor. Este compromiso se ilustra en la Figura 3.26.
- Cuanto mayor sea el intervalo de observación, también es mayor la probabilidad de detección. De hecho, con un incremento de T_{obs} mayor es también S por lo que menor es σ (véase Ecuación 3.81), lo que lleva a una mayor probabilidad de detección (para la misma probabilidad de falsa alarma). Esto también está representado en la Figura 3.26.

3.4.2. Uso de una configuración óptima

El problema de suponer que una estación, en caso de colisión, siempre captura el canal es el siguiente: si la configuración es tal que las colisiones son frecuentes, un usuario cuya tarjeta está configurada de forma maliciosa pero que no capture el canal puede no

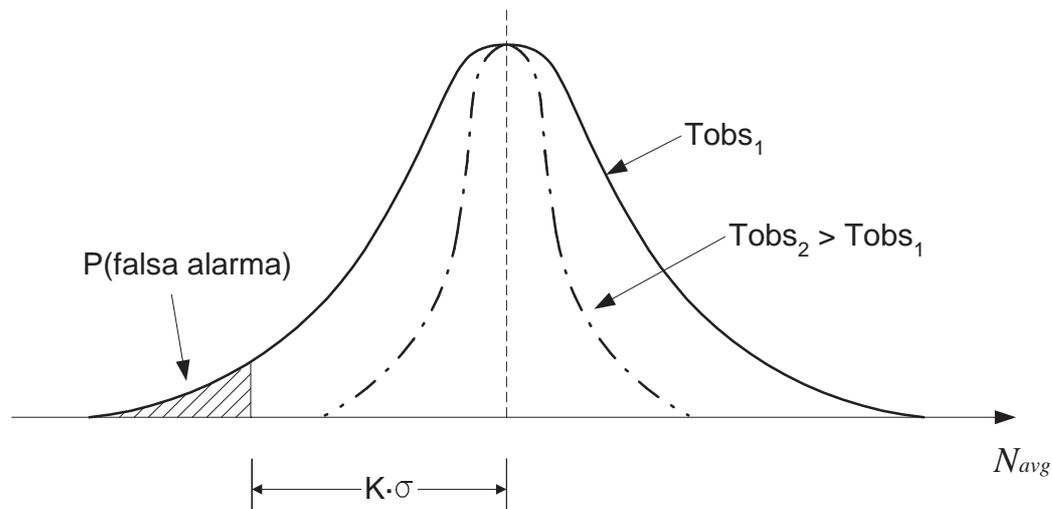


Figura 3.26: Relación entre P_{FA} , K y T_{obs}

ser detectado, debido a que obtiene un ancho de banda similar a la que obtendría con la configuración correcta pero capturando el canal.

Con objeto de evitar este problema, se emplea la configuración de la Sección 3.2.2 para los parámetros CW_{min} y CW_{max} ¹², si bien en este caso se fija $m = 0$ ¹³. Esta configuración se obtiene de la maximización del ancho de banda en la red inalámbrica, y se caracteriza por una probabilidad de colisión muy baja, lo que reduce el impacto del efecto captura. De esta forma, el comportamiento malicioso de un usuario no queda enmascarado por el efecto captura, por lo que no puede obtener sustanciales incrementos en el ancho de banda recibido y pasar desapercibido.

3.4.3. Resultados de simulación

En esta sección se evalúa, mediante simulación, la efectividad del algoritmo en la detección de usuarios maliciosos. Se modela el estándar 802.11b, con 10 estaciones que transmiten paquetes de 1500 octetos.

Existen múltiples maneras mediante las que un usuario malicioso puede proceder para así recibir un mejor servicio del que debiera; la que se ha decidido implementar es la siguiente: el usuario malicioso siempre tiene paquetes para transmitir, lo que proporciona el mayor ancho de banda. De los parámetros de funcionamiento del EDCA, fija CW_{min} y CW_{max} al mismo valor (esto es: la ventana de contienda siempre tiene el mismo valor, a pesar de las colisiones), siendo éste menor que la CW_{min} del resto de estaciones. Dado que el algoritmo aquí expuesto no depende del comportamiento específico de un usuario sino del

¹²La configuración óptima aquí empleada asume, por simplicidad, que todas las estaciones deben recibir el mismo ancho de banda. No obstante, el algoritmo es igualmente aplicable para configuraciones asimétricas de ancho de banda.

¹³El uso de $m = 0$ evita con más eficacia que un comportamiento malicioso quede enmascarado por el efecto captura, si bien redundante en un comportamiento peor para el retardo (Sección 3.2.2.1). Se tiene, por tanto, un compromiso a ser fijado por el administrador de la WLAN.

Tabla 3.3: Probabilidad de falsa alarma

K	T_{obs}	P_{FA} analítica	P_{FA} simulación
1	5	0.1587	0.1567
2	1	0.0228	0.0218
”	5	”	0.0224
”	10	”	0.0227
3	5	0.0013	0.0012

número total de transmisiones con éxito en un intervalo, este comportamiento (simple pero realista) sirve para validar el funcionamiento del algoritmo.

El resto de estaciones configuran de forma correcta los parámetros del modo EDCA. También se supone que todas las estaciones siempre tienen paquetes pendientes de ser transmitidos; sin embargo, el algoritmo es independiente del comportamiento del resto de estaciones, por lo que dicha suposición -esto es, condiciones de saturación- no supone en modo alguno una limitación.

Para cada uno de los escenarios estudiados, se ejecuta el algoritmo durante 100.000 intervalos de observación. La probabilidad de detección se estima midiendo en cuántos de dichos 100.000 intervalos el algoritmo clasifica el comportamiento de la estación observada como malicioso.

3.4.3.1. Parámetros del algoritmo

Con objeto de una mejor comprensión del impacto de los diferentes parámetros del algoritmo, se realizan varios experimentos para diferentes valores de K y de T_{obs} . En estos experimentos, todas las estaciones salvo la controlada emplean la configuración por defecto para la CW de 802.11 DCF ($CW_{min} = 32$ y $CW_{max} = 1024$). La estación bajo control siempre captura el canal, e implementa el comportamiento explicado con anterioridad. Se estudia la probabilidad de detección en función de la CW_{min} de la estación del usuario malicioso. Es preciso hacer notar que, en el caso $CW_{min} = CW_{max} = 32$, el comportamiento del usuario bajo control sí es el adecuado, debido a que una estación que siempre captura el canal en caso de colisión nunca duplica su CW .

En la Tabla 3.3 se presenta la probabilidad de que un usuario bien configurado sea clasificado como malicioso, para distintos valores de K y T , calculado según simulación y análisis (Ecuación 3.82). Los resultados teóricos se aproximan con gran exactitud a los valores obtenidos mediante simulación, lo que confirma que el algoritmo permite ajustar la probabilidad de falsa alarma. Las discrepancias entre dichos valores, además, se reducen conforme aumenta la duración del intervalo de observación. Esto es así ya que, cuanto mayor sea T_{obs} , más sumandos se añaden en la Ecuación 3.79 y, según el Teorema Central del Límite, más ajustada resulta la aproximación mediante una distribución normal.

En la Figura 3.27 se analiza el comportamiento de la probabilidad de detección en función de la CW_{min} de la estación bajo control, para distintos valores de K , con un valor fijo del intervalo de observación $T_{obs} = 5s$. En este caso, se aprecia claramente el compromiso entre la probabilidad de falsa alarma y la probabilidad de detección: para un valor bajo de K , la P_{FA} (caso $CW_{min} = 32$) es elevada pero la de detección también lo es, debido al incremento abrupto de ésta con el crecimiento de CW_{min} . Por el otro lado, un

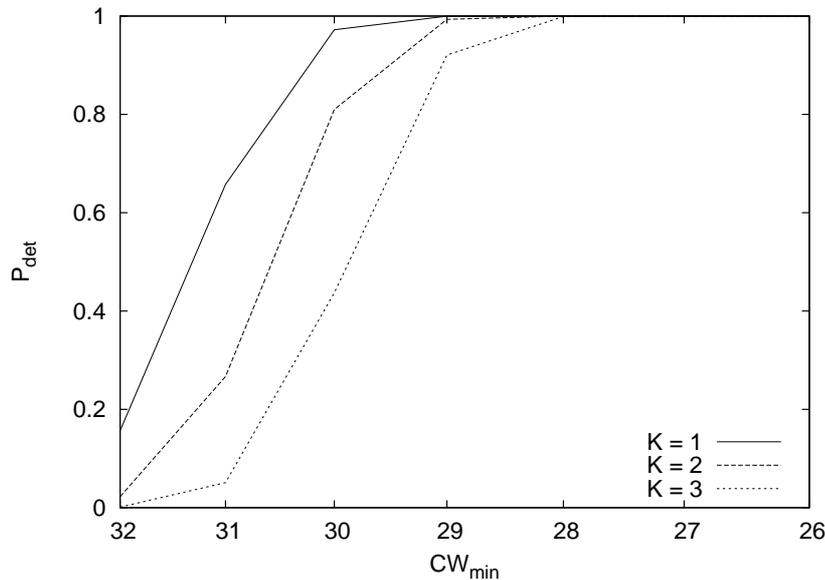


Figura 3.27: Probabilidad de detección en función de K , $T_{obs} = 5s$

valor elevado de K disminuye tanto la probabilidad de falsa alarma como la probabilidad de detección.

En la Figura 3.28 se presenta la probabilidad de detección para diferentes valores del intervalo de observación, con el valor de $K = 2$. En este caso se ilustra el compromiso entre detección y granularidad temporal: cuanto mayores son los intervalos de observación mayor fiabilidad en la detección es obtenida, mientras con intervalos de detección menores (lo que equivaldría a una mayor velocidad de reacción) la detección es menos ajustada.

Durante el resto de este apartado se fijará $K = 2$ y $T = 5s$, realizando una evaluación del algoritmo en tres escenarios diferentes: cuando la estación maliciosa disfruta del efecto captura y con la configuración por defecto de 802.11 DCF, cuando no se beneficia del efecto captura y con la configuración por defecto de 802.11 DCF, y con la configuración óptima sin beneficio por el efecto captura.

3.4.3.2. Configuración 802.11 DCF y efecto captura

A continuación se evalúa la ejecución del algoritmo en el caso en que la estación controlada captura siempre el canal, y todas las demás estaciones emplean la configuración por defecto. Los resultados obtenidos se muestran en la Tabla 3.4. Cuando la estación bajo observación fija su valor de CW_{min} a 32 (y, por lo tanto, no se comporta de forma fraudulenta), el ancho de banda que obtiene ($r_{malicioso}$) es 2,4 veces superior que el del resto de las estaciones (r_{resto}), como consecuencia del efecto captura. Cuando cambia su CW_{min} a un valor de 27 o menor (siendo detectado este comportamiento malicioso en todos los casos), está incrementando su ancho de banda 2,9 veces sobre el resto de los usuarios. Por lo tanto, antes de ser detectado siempre por el algoritmo, un usuario malicioso se puede llevar del orden de un 20 % más de ancho de banda que un usuario con la configuración adecuada pero que siempre capture el canal -aunque, en términos absolutos, se lleva casi el

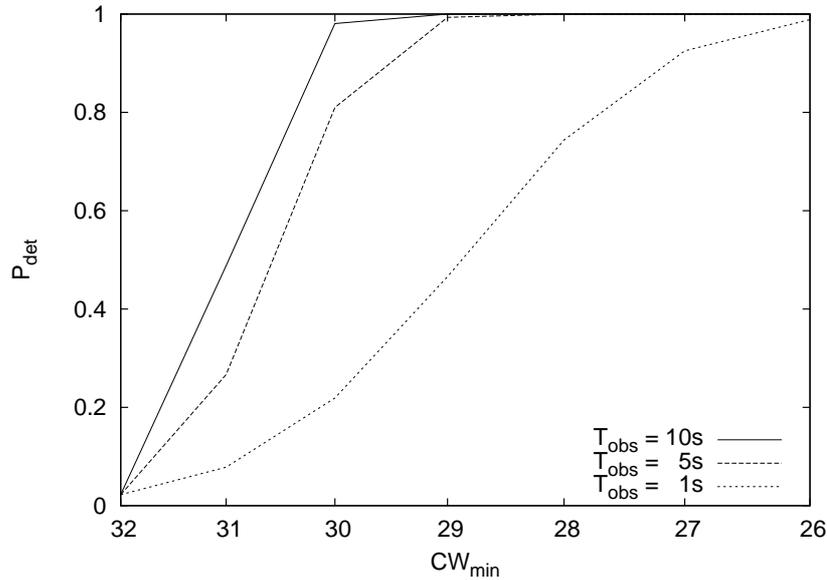
Figura 3.28: Probabilidad de detección en función de T_{obs} , $K = 2$

Tabla 3.4: Configuración por defecto 802.11 DCF y efecto captura

CW_{min}	P_{detect}	$BW_{malicioso}$	BW_{rest}
32	0.0224	1572.67	656.03
31	0.2666	1617.71	652.13
30	0.8101	1665.95	647.88
29	0.9934	1717.12	643.36
28	0.9999	1770.96	638.67
27	1	1828.40	633.62

triple del ancho de banda que el resto de los usuarios-.

3.4.3.3. Configuración 802.11 DCF sin efecto captura

En este caso, el usuario malicioso no captura el canal en caso de colisión, lo que hace más difícil detectar un comportamiento destinado a aprovecharse del resto de los usuarios. Los resultados de la simulación se muestran en la Tabla 3.5; en este caso, es posible decrementar el valor de CW_{min} a un valor menor que el del apartado anterior, debido a que el algoritmo se basa en el número de transmisiones con éxito que ha obtenido una estación (las colisiones ocultan el comportamiento inadecuado). Dado que la tasa $r_{malicioso}$ resulta tres veces mayor que r_{resto} antes de ser detectado en todas las ocasiones, se puede concluir que el algoritmo no resulta demasiado efectivo en este caso.

3.4.3.4. Configuración óptima sin efecto captura

En el caso anterior, la no existencia del efecto captura ocultaba el comportamiento inadecuado del usuario que desea obtener más ancho de banda que el resto de nodos de

Tabla 3.5: Configuración por defecto 802.11 DCF sin efecto captura

CW_{min}	P_{detect}	$BW_{malicious}$	BW_{rest}
32	0	1127.61	656.04
23	0.0396	1522.15	609.49
22	0.4273	1583.87	602.28
21	0.9319	1650.69	594.49
20	0.9994	1722.84	586.14
19	1	1802.29	576.96

Tabla 3.6: Configuración óptima, sin efecto captura

CW_{min}	P_{detect}	$BW_{mal}(kbps)$	$BW_{rest}(kbps)$
174	0	738.01	738.05
153	0.1267	829.32	728.57
147	0.5158	859.51	725.43
141	0.9029	892.37	722.01
130	0.9998	959.27	715.09
129	1	965.79	714.42

la red inalámbrica. En este caso, de nuevo, el usuario malicioso no se beneficiará del efecto captura, pero todas las estaciones estarán configuradas de forma óptima. Los resultados se muestran en la Tabla 3.6, y se comprueba que las posibilidades de obtener un incremento de ancho de banda sin ser detectado son mínimas: siempre se clasifica como malicioso en cuanto el usuario obtiene un incremento del 35 % sobre r_{resto} .

3.5. Resumen

A pesar del trabajo existente en el análisis del modo EDCA en condiciones de saturación, se ha comprobado que dichos análisis no capturaban adecuadamente el comportamiento de la red, lo que se ha intentado subsanar con un nuevo análisis. La configuración que optimiza del rendimiento en condiciones de saturación, por otra parte, resulta ser la configuración más adecuada para el modo básico incluso en situaciones de no saturación, como se ha comprobado mediante el modelado de las anomalías e inestabilidades que pueden darse lugar en una WLAN. La extensión de la configuración propuesta para el modo *multirate*, no obstante, precisa de un criterio de equidad para evitar la inanición de las estaciones que transmiten a una tasa lenta. Dado que en las redes WLAN abiertas pueden transmitir usuarios malintencionados que modifiquen el comportamiento del nivel MAC para obtener un mejor servicio a costa del resto de usuarios, la configuración en estos escenarios se ha complementado con la definición de un algoritmo de detección de comportamientos que no se ajusten a las configuraciones propuestas.

Capítulo 4

Configuración de WLAN cerradas

Una *WLAN cerrada* es una red inalámbrica que da servicio a un conjunto de usuarios con los que se tiene una relación previa de confianza. La implicación de esta relación previa es que, en este caso, y frente al caso de la WLAN abierta, los usuarios seguirán las reglas de configuración de la capa MAC (no será preciso un algoritmo de supervisión del comportamiento). De igual forma, las especificaciones sobre las garantías de servicio que precisen, así como el patrón de generación de tráfico, podrán ser empleadas para optimizar el funcionamiento de la red.

Proporcionar QoS al tráfico cursado en una WLAN requiere, en principio, ser capaz de modelar el comportamiento de dicha WLAN en las condiciones de trabajo (tema de la primera parte de este capítulo). En este capítulo se realiza un estudio del valor medio del tiempo de servicio para el modo EDCA, así como de su varianza. Dichos análisis, que predicen el comportamiento en retardo de cada AC de la WLAN, permiten la construcción de algoritmos que, dado un escenario y unos requisitos para el retardo, proporcionen la configuración que cumple con dichas exigencias temporales. También se modela la función de distribución del tiempo de servicio, lo que posibilita una mejor caracterización de la forma en que se sirve el tráfico. De igual forma, con dicho modelo se permite el desarrollo de herramientas de configuración para requisitos aún más sofisticados.

Posteriormente se ilustra el manejo de las anteriores herramientas mediante la configuración de un escenario VoIP, sobre el que se desarrolla un algoritmo de configuración que proporciona unos requisitos de valor medio y varianza del retardo, proporcionándose la función de distribución para los esquemas obtenidos. El capítulo se completa con el desarrollo de un estudio experimental, que define una metodología para obtener la configuración óptima por vía de la realización de pruebas. La realización de este tipo de estudios experimentales resulta esencial, dado que es la única forma de capturar comportamientos no ideales que no pueden ser reflejados ni por análisis ni mediante herramientas de simulación.

4.1. Análisis del valor medio y desviación típica del tiempo de servicio

En un entorno en el que se suministre QoS a las aplicaciones, las colas de transmisión se encontrarán, con gran frecuencia, desocupadas. Esto se debe a que, si se quiere otorgar ciertas garantías de servicio, una primera condición es que el tiempo medio de servicio sea bastante menor que el tiempo medio entre llegadas. Adicionalmente, conforme la garantía de QoS sea más estricta, se tendrá que la ocupación de las colas resulta incluso menor. Por lo tanto, el mecanismo de TXOP (empleado en el Capítulo 3 para proporcionar diferenciación) no será objeto de estudio a lo largo de esta sección, puesto que con la configuración propuesta las veces en que sea utilizado serán muy infrecuentes¹.

Por el mismo motivo, en una red que proporcione QoS existirán estaciones no saturadas, lo que requiere realizar un análisis de la WLAN en dichas condiciones. Este análisis ampliará el planteamiento de la Sección 3.1.2, extendiéndolo para el caso de no saturación. Una vez analizado el punto de operación de la red, se puede obtener el valor medio del tiempo de servicio (para lo que se extiende el modelo de [2]), así como su varianza. El hecho de disponer de la varianza de una variable aleatoria además de su valor medio proporciona gran información pues permite, incluso, proporcionar unas garantías básicas de servicio².

4.1.1. Análisis de las probabilidades de transmisión

Sea una red WLAN operando en el modo EDCA, con cada estación perteneciendo a una categoría de acceso i ($i \in \{1, \dots, N\}$). Sea la configuración de la AC i la determinada por los parámetros $\{CW_i^{min}, CW_i^{max}, AIFS_i\}$, teniendo dicha categoría n_i estaciones. La tasa de generación de tráfico es de ρ_i , pudiendo cada estación transmitir tramas de diferente longitud l_j con una probabilidad determinada. Se define m_i como $CW_i^{max} = 2^{m_i} CW_i^{min}$ y A_i según $AIFS_i = DIFS + A_i T_e$. Adicionalmente, S_k es el conjunto de estaciones que pertenecen a un AC j con $A_j \leq k$ y Δ es el mayor A_i en la WLAN.

Se define como *ranura de tiempo* (o *ranura*) la duración del intervalo entre dos decrementos consecutivos del contador de backoff de una estación con $AIFS$ mínimo (esto es, $DIFS$). Sea τ_i la probabilidad de que una estación perteneciente a la AC i transmita tras un decremento de su contador, y sea una k -ranura una ranura precedida por k o más ranuras vacías (véase la Figura 4.1)³.

Una estación de la WLAN puede estar *saturada*, si su tasa de generación de tráfico es inferior al ancho de banda r_i que experimenta, o bien *no saturada* si todo su tráfico generado es servido. Se hace por ello necesario abordar el análisis de r_i .

¹Para las estaciones de datos, sin requisitos de entrega, se fijará TXOP de forma que en cada acceso al medio transmitan una única trama, pues de esta forma se minimiza el impacto de dicho tráfico sobre el que sí precisa requisitos de entrega.

²Así, la desigualdad de Chebyshev establece que la probabilidad de que una variable aleatoria con desviación típica σ se aleje de su media un valor $k\sigma$ es menor o igual que $1/k^2$.

³Dado que, una estación con $A_i = k$ reactiva el decremento de su contador sólo tras k ranuras vacías tras una ranura no vacía, los decrementos del contador de backoff de dicha estación coinciden con los intervalos de las k -ranuras. Por lo tanto, una estación de la AC i con $A_i = k$ transmite en una k -ranura con una probabilidad τ_i , y no transmite en ninguna otra ranura.

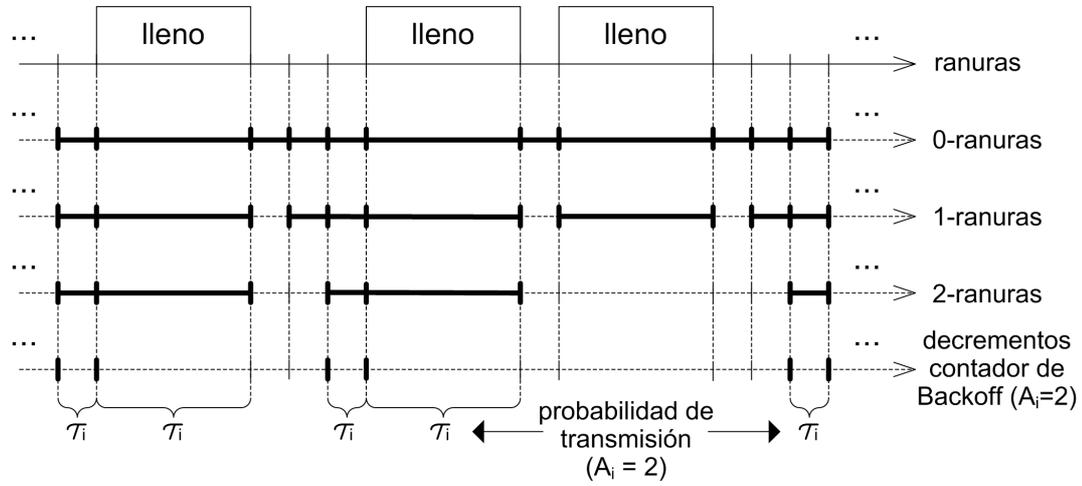


Figura 4.1: *k*-ranuras y probabilidad de transmisión (ejemplo para *k*=2)

4.1.1.1. Estudio del ancho de banda

El ancho de banda que recibe una estación viene dado por la cantidad media de información transmitida en una ranura dividida por la longitud media de ésta,

$$r_i = \frac{p(s_i)l_i}{p(s)T_s + p(c)T_c + p(e)T_e} \quad (4.1)$$

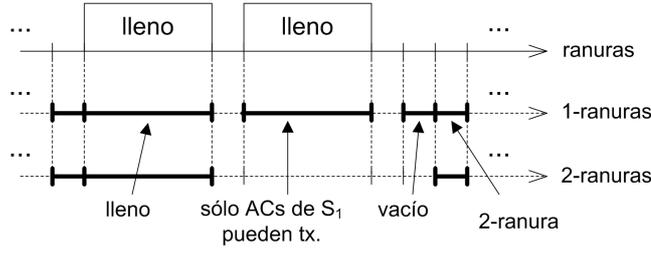
donde l_i es la longitud media de una transmisión de la AC i , $p(s_i)$ es la probabilidad de que una ranura contenga una transmisión con éxito de la AC i , $p(s)$, $p(c)$ y $p(e)$ son las probabilidades de que una ranura contenga un éxito, una colisión o esté vacía, y T_s , T_c y T_e son las duraciones medias de una ranura para los anteriores casos.

Para realizar el cálculo, es preciso en primer lugar obtener la probabilidad de que una determinada *k*-ranura esté vacía ($p(e_k)$). A tal efecto, se procede de forma recursiva: la probabilidad de que un *k*-ranura esté vacía dependerá del estado de la anterior *k*-ranura. Si ésta estuvo llena, sólo aquellas AC's con $A_i \leq k$ podrán transmitir (es decir, las estaciones pertenecientes a S_k), cosa que harán con probabilidad τ_i . Si la anterior *k*-ranura estaba vacía, eso significa que la ranura presente viene precedida por $k + 1$ ranuras vacías, que es precisamente la definición de una $(k+1)$ -ranura por lo que su probabilidad de estar vacía viene dada por $p(e_{k+1})$. Con este razonamiento (ilustrado en la Figura 4.2), se tiene que $p(e_k)$ se calcula de la siguiente forma:

$$p(e_k) = (1 - p(e_k)) \prod_{j \in S_k} (1 - \tau_j) + p(e_k)p(e_{k+1}) \quad (4.2)$$

Dado que en una Δ -ranura, por definición, todas las estaciones pueden transmitir, se tiene que su probabilidad de estar desocupada viene dada por la expresión

$$p(e_\Delta) = \prod_{j \in S_\Delta} (1 - \tau_j) \quad (4.3)$$

Figura 4.2: Probabilidad de vacío en una k -ranura

Dadas unas τ_i 's, mediante la anterior expresión es posible calcular $p(e_\Delta)$. Hecho esto, se puede calcular $p(e_k)$ para el caso $k = \Delta - 1$ mediante (4.2). Siguiendo con este procedimiento, se puede obtener $p(e_k) \forall k$.

La probabilidad $p(e)$ es, por definición, $p(e_0)$, dado que todas las ranuras de tiempo son 0-ranuras. Sea p_k la probabilidad de que una ranura sea una k -ranura. Dicha probabilidad se corresponderá con la probabilidad de que la ranura anterior sea una $(k-1)$ -ranura y esté vacía, por lo que vendrá dada por

$$p_k = p_{k-1}p(e_{k-1}) \quad (4.4)$$

Por definición, $p_0 = 1$, por lo que se puede calcular cualquier probabilidad de la siguiente forma

$$p_k = \prod_{j=0}^{k-1} p(e_j) \quad (4.5)$$

Para el cálculo de la probabilidad de que una ranura contenga un éxito de una estación perteneciente a la AC i se utiliza la siguiente expresión:

$$p(s_i) = \sum_{k=A_i}^{\Delta} p(S_k)p(s_i | S_k) \quad (4.6)$$

siendo $p(S_k)$ la probabilidad de que en una determinada ranura sólo puedan transmitir las estaciones pertenecientes al conjunto S_k , y $p(s_i | S_k)$ es la probabilidad de que, en dichas circunstancias, la ranura contenga un éxito de una determinada estación de la AC i .

Que en una determinada ranura sólo puedan transmitir las estaciones del conjunto S_k vendrá determinada porque la ranura sea una k -ranura pero no una $(k+1)$ -ranura. De esta forma, se deduce que

$$p(S_k) = \begin{cases} p_k - p_{k+1}, & \text{Para } k < \Delta \\ p_\Delta, & \text{Para } k = \Delta \end{cases} \quad (4.7)$$

La probabilidad $p(s_i | S_k)$ viene dada por la probabilidad de que una estación de la AC i transmita y ninguna otra de S_k lo haga

$$p(s_i | S_k) = \tau_i(1 - \tau_i)^{n_i-1} \prod_{j \in S_k \setminus i} (1 - \tau_j)^{n_j} \quad (4.8)$$

La probabilidad de que una ranura contenga un éxito es la suma de la probabilidad de éxito para cada estación:

$$p(s) = \sum_{i \in S_{\Delta}} n_i p(s_i) \quad (4.9)$$

La duración media de un éxito se obtiene al ponderar las diferentes duraciones medias de cada AC i por su probabilidad

$$T_s = \sum_{i \in S_{\Delta}} \frac{n_i p(s_i)}{p(s)} T_s^i \quad (4.10)$$

Donde T_s^i es, de nuevo, la duración de una ranura de tiempo ocupada por la transmisión de una trama de tamaño l_i

$$T_s^i = T_{PCLP} + \frac{H + l_i}{C} + SIFS + T_{PLCP} + \frac{ACK}{C} + DIFS \quad (4.11)$$

La probabilidad de que una ranura contenga una colisión viene dada por la ausencia de éxitos o de ranuras vacías

$$p(c) = 1 - p(e) - p(s) \quad (4.12)$$

Para el cálculo de la duración media de una colisión es preciso tener en cuenta que la trama de mayor longitud determinará la extensión de la misma. De esta forma, se puede ponderar cada duración como sigue:

$$T_c = \sum_{l \in L} \frac{p(c_l)}{p(c)} T_c^l \quad (4.13)$$

En la anterior expresión, L es el conjunto de las posibles longitudes de trama, $p(c_l)$ es la probabilidad de que una ranura contenga una colisión con la trama de mayor tamaño de longitud l , y T_c^l es la duración en dicho caso. Esta última viene dada por

$$T_c^l = T_{PLCP} + \frac{H + l}{C} + EIFS \quad (4.14)$$

$p(c_l)$ se puede calcular de la siguiente forma

$$p(c_l) = \sum_{k=0}^{\Delta} p(S_k) p(c_l | S_k) \quad (4.15)$$

Para calcular $p(c_l | S_k)$, la probabilidad de que en una ranura donde sólo transmiten las estaciones de S_k haya una colisión con la mayor trama involucrada de tamaño l , se procede de la siguiente forma: sea τ_j la probabilidad de que transmita la estación j y $p_j(tx = l)$ la probabilidad de que dicha estación transmita una trama de longitud l . De esta forma

$$p(c_l | S_k) = \sum_{j \in S_k} \tau_j p_j(tx = l) p_{k,-j}(tx \neq l) p_{k,-j}(tx < l) \quad (4.16)$$

Donde $p_{k,\neg j}(tx \not> l)$ es la probabilidad de que ninguna estación de S_k (diferente de j) transmita una trama mayor que l , y $p_{k,\neg j}(tx \mid tx < l)$ es la probabilidad de que, sabiendo que no se transmite un paquete de longitud mayor que l , al menos existe otra transmisión.

Para calcular las anteriores probabilidades se ordenan las estaciones del conjunto S_k , denominándose con $S_{k,j}$ al conjunto de estaciones con índice menor que j^4 . De esta forma, se tiene que

$$p_{k,\neg j}(tx \not> l) = \prod_{m \in S_{k,j}} (1 - \tau_m p_m(tx \geq l)) \prod_{m \in S_k \setminus S_{k,j} \cup j} (1 - \tau_m p_m(tx > l)) \quad (4.17)$$

Las probabilidades $p_m(tx \geq l)$ y $p_m(tx > l)$ indican la probabilidad de que una transmisión de la estación m sea mayor o igual o estrictamente mayor que l , respectivamente.

$p_{k,\neg j}(tx \mid tx < l)$ se calcula como:

$$p_{k,\neg j}(tx \mid tx < l) = 1 - \prod_{m \in S_{k,j}} \frac{1 - \tau_m}{1 - \tau_m p_m(tx \geq l)} \prod_{m \in S_k \setminus S_{k,j} \cup j} \frac{1 - \tau_m}{1 - \tau_m p_m(tx > l)} \quad (4.18)$$

Lo que termina el análisis de r_i .

4.1.1.2. Cálculo de la probabilidad de transmisión

Con objeto de calcular las probabilidades de transmisión τ_i 's de cada AC, será preciso distinguir entre aquellas AC's que se encuentran en saturación y aquellas AC's cuyo tráfico es servido en su totalidad. El discriminante para tal clasificación será la comparación entre la tasa ρ_i generada por cada estación y la tasa r_i que obtienen en la WLAN.

En un principio, se considerará que todas las AC's se encuentran en saturación. En dichas circunstancias, se puede aplicar el análisis de la Sección 3.1.2 para obtener las probabilidades de transmisión de las estaciones.

Una vez calculadas las probabilidades de transmisión, se puede obtener el ancho de banda recibido por cada estación con las expresiones de la Sección 4.1.1.1. De esta forma, una determinada AC puede ser clasificada como *saturada*, si genera más tráfico que dicho ancho de banda, o bien *no saturada* en caso contrario. Si todas las AC's son clasificadas como saturadas, el cálculo ha finalizado.

Si, en cambio, alguna AC pasa a ser clasificada como no saturada, es preciso recalcular las probabilidades de transmisión pues dicha AC no se podrá modelar de la misma forma. Para obtener la probabilidad de transmisión de una AC no saturada se aplica la condición de que todo el tráfico generado es servido:

$$\rho_i(1 - p_i^{R+1}) = r_i \quad (4.19)$$

donde p_i^{R+1} es la probabilidad de que un paquete se descarte por superación del número máximo de retransmisiones R para una trama. Teniendo esto en cuenta, es posible obtener la τ_i de una AC no saturada (τ_{nonsat}) de acuerdo a la siguiente expresión:

⁴Ordenar las estaciones evita contar dos veces un mismo caso para las colisiones.

$$\tau_i^{nonsat} = \frac{\rho_i(p(s)T_s + p(e)T_e + p(c)T_c)}{l_i \sum_{k=A_i}^{\Delta} p(S_k) \prod_{m \in S_k \setminus i} (1 - \tau_m)} \quad (4.20)$$

Para el caso de las AC's que permanezcan saturadas, se aplica un esquema similar al esquema definido anteriormente⁵ para obtener las probabilidades de transición de la cadena de Markov de la Figura 3.1. De esta forma se puede definir un sistema de ecuaciones en las τ_i 's que se resuelve de forma numérica. Una vez obtenidas las τ_i 's de operación, se comprueba si todas las estaciones saturadas reciben menos ancho de banda que el tráfico generado. De no ser así, la AC correspondiente sería clasificada como no saturada por lo que habría que repetir el anterior cálculo, hasta que todas las estaciones que se encuentran clasificadas como saturadas generasen más tráfico que el que pueden transmitir⁶.

El número de iteraciones a ejecutar para calcular las probabilidades de transmisión está acotado por N , siendo éste el número de AC's en la WLAN. Esto es así dado que, a cada iteración, en el peor de los casos sólo una AC pasa a ser clasificada como no saturada.

4.1.2. Análisis del tiempo medio de servicio

Una vez que, mediante el análisis de la sección anterior, se ha obtenido los valores de τ_i , el valor medio del tiempo de servicio que recibe una trama no descartada de la AC i , $E[d_i]$, se puede calcular de la siguiente manera:

$$E[d_i] = \frac{1}{P_{tx,i}} \sum_{j=0}^R E[d_i | j \text{ col}] P_{tx,i}(j \text{ col}) \quad (4.21)$$

donde $P_{tx,i}$ y $P_{tx,i}(j \text{ col})$ son las probabilidades de que una trama perteneciente a una estación de la AC i no sea descartada y de que sea transmitida con éxito después de j intentos, respectivamente. Estas probabilidades vienen dadas por las siguientes expresiones:

$$P_{tx,i} = \sum_{j=0}^R (1 - p_i) p_i^j, \quad P_{tx,i}(j \text{ col}) = (1 - p_i) p_i^j \quad (4.22)$$

La probabilidad p_i de que una estación de la AC i colisione se deriva de la probabilidad que de una A_i -ranura esté vacía como sigue:

$$p(e_{A_i}) = (1 - \tau_i)(1 - p_i) \quad (4.23)$$

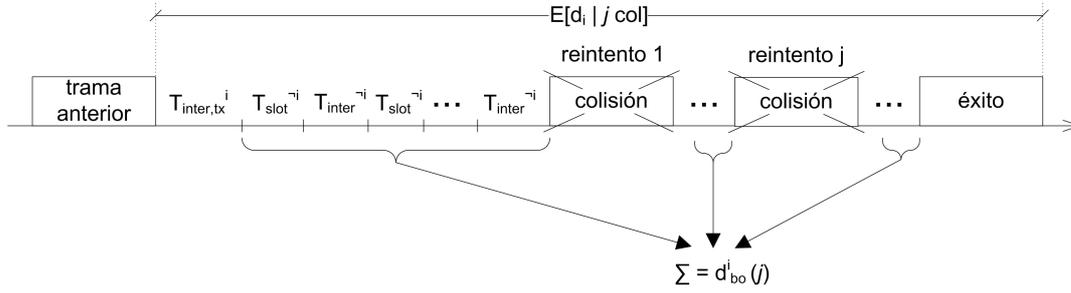
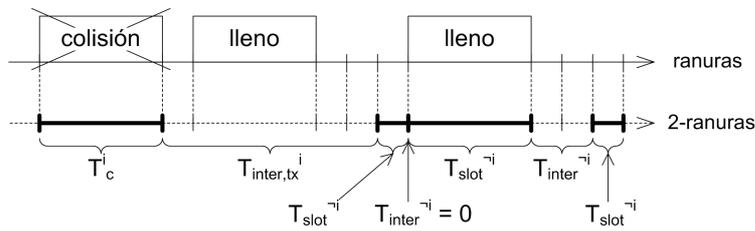
El retardo medio en caso de j retransmisiones viene dado por (véase Figura 4.3)

$$E[d_i | j \text{ col}] = E[T_s^i] + jE[T_c^i] + jE[T_{inter,tx}^i] + E[d_{bo}^i(j)] \quad (4.24)$$

Donde $E[T_s^i]$ es la duración media de un éxito de una estación de la AC i , $E[T_c^i]$ es la duración media de una colisión donde interviene dicha estación, $E[T_{inter,tx}^i]$ es la media del

⁵En este caso, el conjunto de ecuaciones se plantea sobre las τ_i 's y no sobre las p_i 's de la Sección 3.1.2. Para ello, se debe emplear ecuaciones similares a las de la Sección 4.1.1.1

⁶Una vez que una AC deja de estar saturada no puede volver a ser clasificada como saturada, dado que, en una nueva iteración, otras estaciones también habrán dejado de estar saturadas por lo que las condiciones para la transmisión son aún más favorables que en la ocasión en la que dejó de ser clasificada como saturada

Figura 4.3: Retardo medio para j retransmisionesFigura 4.4: Componentes del retardo (ejemplo para $k = 2$)

tiempo que pasa entre una transmisión de dicha estación y la siguiente A_i -ranura, y $E[d_{bo}^i(j)]$ es el tiempo medio que aguarda la estación ejecutando todos los procesos de backoff en la j -ésima colisión de una misma trama. Estos tiempos, tras una colisión, se ilustran en la Figura 4.4⁷.

La duración media de una colisión en la que participa una estación de la AC i viene dada por⁸:

$$T_c^i = \frac{\sum_l p_{col,i}(l) T_c(l)}{\sum_l p_{col,i}(l)} \quad (4.25)$$

en la que $p_{col,i}(l)$ es la probabilidad de que, dado que la estación transmite y hay una colisión, la trama de mayor tamaño que colisiona tiene por longitud l . Esto se puede obtener de la siguiente expresión:

$$p_{col,i}(l) = \sum_{k=A_i}^{\Delta} p_{col,i}(l | S_k) p_{col,i}(S_k) \quad (4.26)$$

La probabilidad $p_{col,i}(S_k)$ es la probabilidad de que, dada una colisión en la que la estación participa, en dicha ranura sólo las estaciones de S_k puedan transmitir; $p_{col,i}(l | S_k)$ es la probabilidad de que, en esas circunstancias, la colisión sea de tamaño l . Para calcular aquélla se aplica el Teorema de Bayes de la siguiente forma

⁷Se supondrá que una estación siempre ejecuta un proceso de backoff en cuanto tiene una nueva trama en su cola de transmisión.

⁸La condición de que la estación de la AC i transmite se mantiene en todas las expresiones desde (4.25) hasta (4.34).

$$p_{col,i}(S_k) = \frac{p(col, i | S_k)p(S_k)}{\sum_{k=A_i}^{\Delta} p(col, i | S_k)p(S_k)} \quad (4.27)$$

donde la probabilidad de que se produzca una colisión donde participe la estación, dado que sólo puede transmitir el conjunto S_k es

$$p(col, i | S_k) = \tau_i \left(1 - \prod_{m \in S_k \setminus i} (1 - \tau_m)\right) \quad (4.28)$$

La probabilidad de que en esas circunstancias se produzca una colisión de tamaño l se obtiene de condicionar la probabilidad de que se produzca una colisión de tamaño l con la estación involucrada a la probabilidad de que se produzca una colisión con dicha estación:

$$p_{col,i}(l | S_k) = \frac{p_i(col\ l | S_k)}{p_i(col | S_k)} \quad (4.29)$$

La probabilidad de colisión de la AC i es la probabilidad complementaria del éxito (esto es, que ninguna estación de S_k transmita):

$$p_i(col | S_k) = 1 - \prod_{m \in S_k \setminus i} (1 - \tau_m) \quad (4.30)$$

Para calcular la probabilidad de que se produzca una colisión de tamaño l sabiendo que i transmite y que sólo las estaciones de S_k pueden hacerlo, se distingue los casos en los que la estación transmite una longitud menor que l del caso en que la estación transmite una trama de dicho tamaño:

$$\begin{aligned} p_i(col\ l | S_k) &= \sum_{l_j < l} p_i(tx = l_j) p_{k,-i}(tx \neq l) p_{k,-i}(tx = l | tx \neq l) + \\ &+ p_i(tx = l) p_{k,-i}(tx \neq l) p_{k,-i}(tx \leq l | tx \neq l) \end{aligned} \quad (4.31)$$

Donde las probabilidades son

$$p_{k,-i}(tx \neq l) = \prod_{m \in S_k \setminus i} (1 - \tau_m t_{m>l}) \quad (4.32)$$

$$p_{k,-i}(tx = l | tx \neq l) = 1 - \prod_{m \in S_k \setminus i} \frac{1 - \tau_m t_{m \geq l}}{1 - \tau_m t_{m>l}} \quad (4.33)$$

$$p_{k,-i}(tx \leq l | tx \neq l) = 1 - \prod_{m \in S_k \setminus i} \frac{1 - \tau_m}{1 - \tau_m t_{m>l}} \quad (4.34)$$

El tiempo medio total que la estación considerada dedica a decrementar su contador de backoff cuando sufre j colisiones vienen dado por la siguiente expresión

$$E[d_{bo}^i(j)] = \sum_n p_{bo}^{i,j}(n) E \left[\underbrace{(T_{slot}^{-i} + T_{inter}^{-i}) + \dots + (T_{slot}^{-i} + T_{inter}^{-i})}_{n\ \text{veces}} \right] \quad (4.35)$$

Donde $p_{bo}^{i,j}$ es la distribución de probabilidad del número total de A_i ranuras que tiene que contar una estación de la AC i para transmitir una trama a lo largo de j colisiones. Dicha distribución se expresa mediante la convolución de las distribuciones uniformes entre 0 y $CW - 1$, siendo CW la ventana de contienda a emplear en cada ejecución del backoff:

$$p_{bo}^{i,j} = U(0, CW_i^{min} - 1) * \dots * U(0, 2^{min(j, m_i)} CW_i^{min} - 1) \quad (4.36)$$

T_{slot}^{-i} es la duración media de una A_i -ranura en la que la estación considerada de la AC i no transmite. Se obtiene por la media de dicha ranura para el caso de un éxito, una colisión y un vacío:

$$T_{slot}^{-i} = p_{-i}(s)T_s^{-i} + p_{-i}(e)T_e + (1 - p_{-i}(s) - p_{-i}(e))T_c^{-i} \quad (4.37)$$

Donde $p_{-i}(e)$ y $p_{-i}(s)$ son las probabilidades de que una A_i -ranura en la que la estación no transmite⁹ esté vacío y contenga un éxito, respectivamente. La probabilidad de una ranura vacía sabiendo que la estación de i no transmite es

$$p_{-i}(e) = \frac{p(e_{A_i})}{1 - \tau_i} \quad (4.38)$$

Para el cálculo de $p_{-i}(s)$ se aplica el teorema de la probabilidad total en el número de A_i -ranuras vacías que preceden a la ranura considerada:

$$p_{-i}(s) = \sum_{j=0}^{\Delta-1-A_i} p_{-i}(s | j_{(e)})p_{-i}(j_{(e)}) + p_{-i}(s | (\Delta - A_i)_{(e)})p_{-i}((\Delta - A_i)_{(e)}) \quad (4.39)$$

Donde $p_{-i}(j_{(e)})$ es la probabilidad de que una A_i -ranura esté precedida exactamente por j A_i -ranuras vacías, $p_{-i}(s | j_{(e)})$ es la probabilidad de que dicha ranura contenga un éxito, $p_{-i}((\Delta - A_i)_{(e)})$ es la probabilidad de que una A_i ranura esté precedida por $\Delta - A_i$ o más A_i ranuras vacías y $p_{-i}(s | (\Delta - A_i)_{(e)})$ es la probabilidad de que, en ese caso, la ranura contenga un éxito.

$p_{-i}(j_{(e)})$ se corresponde con la probabilidad de tener una A_i ranura llena y, a continuación, j ranuras vacías, en las que sólo pueden transmitir las estaciones de los conjuntos $S_{A_i}, S_{A_i+1}, \dots, S_{A_i+j-1}$

$$p_{-i}(j_{(e)}) = (1 - p_{-i}(e)) \prod_{l=A_i}^{A_i+j-1} \prod_{m \in S_l \setminus i} (1 - \tau_m) \quad (4.40)$$

Para calcular $p_{-i}(s | j_{(e)})$ basta con tener en cuenta que sólo las estaciones del conjunto S_{A_i+j} pueden transmitir en dichas ranuras, por lo que resulta

$$p_{-i}(s | j_{(e)}) = \sum_{l \in S_{A_i+j} \setminus i} \tau_l \prod_{m \in S_{A_i+j} \setminus l, i} (1 - \tau_m) \quad (4.41)$$

El cálculo de $p_{-i}((\Delta - A_i)_{(e)})$ y $p_{-i}(s | (\Delta - A_i)_{(e)})$ se realiza con base en las anteriores expresiones

⁹La condición de que la estación considerada no transmita se mantiene en (4.38)–(4.59).

$$p_{\neg i}(\Delta - A_{i(e)}) = 1 - \sum_{j=0}^{\Delta - A_i - 1} p_{\neg i}(j_{(e)}) \quad (4.42)$$

$$p_{\neg i}(s \mid (\Delta - A_i)_{(e)}) = \sum_{l \in S_{\Delta} \setminus i} \tau_l \prod_{m \in S_{\Delta} \setminus l, i} (1 - \tau_m) \quad (4.43)$$

Para calcular la duración media de un éxito contenido en una ranura en la que la estación considerada no transmite, se pondera las longitudes medias de los éxitos de todas las estaciones menos la considerada

$$T_s^{\neg i} = \frac{\sum_{m \setminus i} p_{s, \neg i}(m) T_s(m)}{\sum_l p_{s, \neg i}(m)} \quad (4.44)$$

donde $p_{s, \neg i}(m)$ es la probabilidad de que en una A_i ranura con un éxito sea la estación m la que transmita, y $T_s(m)$ es la duración media de sus transmisiones.

$$p_{s, \neg i}(m) = \sum_{k=A_i}^{\Delta} \frac{p_{s, \neg i}(m \mid S_k)}{\sum_{n \setminus i} p_{s, \neg i}(n \mid S_k)} p(S_k \mid s, \neg i) \quad (4.45)$$

$$p_{s, \neg i}(m \mid S_k) = \tau_m \prod_{n \in S_k \setminus i, m} (1 - \tau_n) \quad (4.46)$$

$p(S_k \mid s, \neg i)$ es la probabilidad de que una ranura sólo permita la transmisión de las estaciones del conjunto S_k , sabiendo que se ha producido un éxito y la estación i no transmite:

$$p(S_k \mid s, \neg i) = \frac{p(s \mid S_k, \neg i) p(S_k)}{\sum_{k=A_i}^{\Delta} p(s \mid S_k, \neg i) p(S_k)} \quad (4.47)$$

donde la probabilidad de que la ranura tenga un éxito se obtiene por suma de la probabilidad de las estaciones (menos la considerada) transmitan con éxito

$$p(s \mid S_k, \neg i) = \sum_{m \in S_k \setminus i} \tau_m \prod_{n \in S_k \setminus i, m} (1 - \tau_n) \quad (4.48)$$

En el caso de que una A_i -ranura contenga una colisión, su duración media vendrá determinada por la siguiente expresión:

$$T_c^{\neg i} = \frac{\sum_l p_{col, \neg i}(l) T_c(l)}{\sum_l p_{col, \neg i}(l)} \quad (4.49)$$

donde la probabilidad de que una colisión tenga tamaño l sabiendo que se ha producido una colisión en la que no ha transmitido la estación de la AC i viene dada por aplicación del teorema de la probabilidad total, condicionando dicho evento a que la ranura sólo permita la transmisión de las estaciones del grupo S_k :

$$p_{col, \neg i}(l) = \sum_{k=A_i}^{\Delta} p_{col, \neg i}(l \mid S_k) p(S_k \mid col, \neg i) \quad (4.50)$$

Para obtener la probabilidad anterior se aplica, de nuevo, el Teorema de Bayes como sigue

$$p(S_k | col, \neg i) = \frac{p(col, \neg i | S_k)p(S_k)}{\sum_{k=A_i}^{\Delta} p(col, \neg i | S_k)p(S_k)} \quad (4.51)$$

Bajo la condición de que en la ranura sólo las estaciones del conjunto S_k puede transmitir, la probabilidad de que se produzca una colisión y la estación de la AC i no intervenga es el caso complementario a que se produzca un éxito o la ranura permanezca vacía:

$$p(col, \neg i | S_k) = (1 - \tau_i) \left(1 - \prod_{m \in S_k \setminus i} (1 - \tau_m) - \sum_{m \in S_k \setminus i} \tau_m \prod_{n \in S_k \setminus i, m} (1 - \tau_n) \right) \quad (4.52)$$

Para el cálculo de la probabilidad de una colisión de tamaño l , sabiendo que la estación del AC i no transmite y que se produce una colisión (esto es, $p_{col, \neg i}(l | S_k)$) se utiliza una expresión similar a la Ecuación 4.29 (modificándose para tener en cuenta el hecho de que la estación considerada no transmite).

Por otra parte, el tiempo $T_{inter}^{\neg i}$ entre que transcurre entre entre dos A_i ranuras bien puede ser nulo (si la ranura está vacía, con probabilidad $p_{\neg i}(e)$), bien puede ser mayor que cero. En este último caso, se procede condicionando dicho tiempo al número de ranuras vacías que preceden la siguiente A_i -ranura:

$$T_{inter}^{\neg i} = (1 - p_{\neg i}(e)) \sum_{j=0}^{A_i} p_{\neg i}(j_{(e)})^* T_{inter}^{\neg i, j} \quad (4.53)$$

Para el cálculo del tiempo hasta la próxima A_i -ranura tras una transmisión ya se conoce que la inmediatamente anterior A_i -ranura estaba ocupada, por lo que

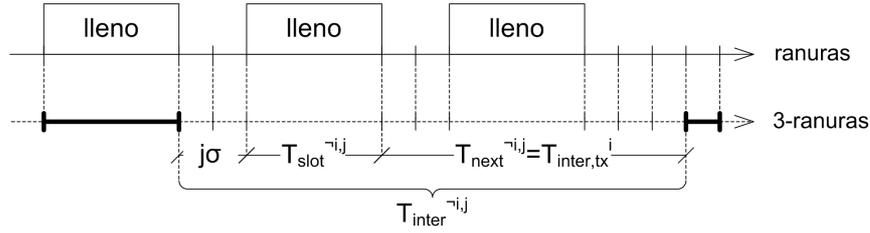
$$T_{inter, tx}^i = \sum_{j=0}^{A_i} p_{\neg i}(j_{(e)})^* T_{inter}^{\neg i, j} \quad (4.54)$$

La probabilidad $p_{\neg i}(j_{(e)})^*$ es la probabilidad de que el tiempo entre una A_i -ranura llena y la siguiente comience con j ranuras vacías, mientras que $T_{inter}^{\neg i, j}$ es la duración media en dicho caso. Para el caso en que haya exactamente A_i -ranuras vacías, el cálculo resulta el siguiente (similar a 4.40):

$$p_{\neg i}(A_i(e))^* = \prod_{l=0}^{A_i-1} \prod_{m \in S_l \setminus i} (1 - \tau_m), \quad T_{inter}^{\neg i, A_i} = A_i T_e \quad (4.55)$$

Mientras que para $j < A_i$ es preciso calcular la duración de las ranuras intermedias (véase la Figura 4.5),

$$p_{\neg i}(j_{(e)})^* = \left(1 - \prod_{m \in S_j \setminus i} (1 - \tau_m) \right) \prod_{l=0}^{j-1} \prod_{m \in S_l \setminus i} (1 - \tau_m) \quad (4.56)$$


 Figura 4.5: Componentes de $T_{inter}^{i,j}$ (para $A_i = 3$ y $j = 2$)

$$T_{inter}^{i,j} = jT_e + T_{slot}^{j,-i} + T_{next}^{i,j} \quad (4.57)$$

donde $T_{slot}^{j,-i}$ es la duración media de una ranura llena –con una colisión o un éxito– precedida por una A_i -ranura llena y j ranuras vacías,

$$T_{slot}^{j,-i} = \left(1 - \frac{\sum_{m \in S_j \setminus i} \tau_m \prod_{n \in S_j \setminus m, i} (1 - \tau_n)}{1 - \prod_{m \in S_j \setminus i} (1 - \tau_m)} \right) \cdot T_c^{j,-i} + \left(\frac{\sum_{m \in S_j \setminus i} \tau_m \prod_{n \in S_j \setminus i, m} (1 - \tau_n)}{1 - \prod_{m \in S_j \setminus i} (1 - \tau_m)} \right) \cdot T_s^{j,-i} \quad (4.58)$$

El cálculo de $T_c^{j,-i}$ y $T_s^{j,-i}$ se realiza con expresiones similares a (4.73) y (4.49), pero teniendo en cuenta que sólo el conjunto S_j puede transmitir. Por último, $T_{next}^{i,j}$ es la duración media entre el fin de la ranura llena y la siguiente A_i -ranura lo que se corresponde con la definición de $T_{inter,tx}^i$ (pues una ranura llena da inicio a una 0-ranura):

$$T_{next}^{i,j} = T_{inter,tx}^i \quad (4.59)$$

Con (4.54)–(4.59) se puede formar un sistema de primer grado del que obtener la variable $T_{inter,tx}^i$, lo que termina el análisis del valor medio del tiempo de servicio.

4.1.3. Análisis de la desviación típica del tiempo medio de servicio

Para calcular la desviación típica del tiempo de servicio se emplea el momento de primer orden (obtenido anteriormente) y el momento de segundo orden:

$$\sigma_{d_i}^2 = E[(d_i)^2] - E^2[d_i] \quad (4.60)$$

Para el cálculo del momento de segundo orden se parte de (4.21)

$$E[(d_i)^2] = \frac{1}{P_{tx,i}} \sum_{j=0}^R P_{tx,i}(j) E[(d_{i,j})^2] \quad (4.61)$$

Mientras que para obtener $E[(d_{i,j})^2]$ se procede a calcular su varianza

$$E[(d_{i,j})^2] = E^2[d_{i,j}] + \sigma_{d_{i,j}}^2 \quad (4.62)$$

La varianza de la anterior distribución, por independencia, se calcula mediante la suma de las varianzas de las variables aleatorias que intervienen

$$\sigma_{d_{i,j}}^2 = j\sigma_{T_c^i}^2 + \sigma_{T_s^i}^2 + (j+1)\sigma_{T_{inter,tx}^i}^2 + \sigma_{d_{bo}(j)}^2 \quad (4.63)$$

Para las varianzas de los éxitos y colisiones se procede como sigue:

$$\sigma_{T_c^i}^2 = E[(T_c^i)^2] - E^2[T_c^i] \quad , \quad \sigma_{T_s^i}^2 = E[(T_s^i)^2] - E^2[T_s^i] \quad (4.64)$$

Para lo que se precisa calcular

$$E[(T_c^i)^2] = \frac{\sum_l p_{col,i}(l)T_c(l)^2}{\sum_l p_{col,i}(l)} \quad , \quad E[(T_s^i)^2] = \sum_l p_i(tx=l)T_s(l)^2 \quad (4.65)$$

De igual forma se procede para el tiempo entre una transmisión y la siguiente A_i -ranura

$$\sigma_{T_{inter,tx}^i}^2 = E[(T_{inter,tx}^i)^2] - E^2[T_{inter,tx}^i] \quad (4.66)$$

Para lo que es preciso resolver la ecuación formada por la siguiente expresión en $E[(T_{inter}^{-i,j})^2]$.

$$E[(T_{inter,tx}^i)^2] = p_{-i}(A_{i(e)})(A_i T_e)^2 + \sum_{j=0}^{A_i-1} p_{-i}(j_{(e)})E[(T_{inter}^{-i,j})^2] \quad (4.67)$$

$$\sigma_{T_{inter}^{-i,j}}^2 + E^2[T_{inter}^{-i,j}] = p_{-i}(A_{i(e)})(A_i T_e)^2 + \sum_{j=0}^{k-1} p_{-i}(j_{(e)}) \left(\sigma_{T_{slot}^{-i}}^2 + \sigma_{T_{inter,tx}^i}^2 + E^2[T_{inter}^{-i,j}] \right) \quad (4.68)$$

El cálculo de $\sigma_{d_{bo}(j)}^2$ se realiza de forma similar

$$\sigma_{d_{bo}(j)}^2 = E[(d_{bo}^i(j))^2] - E^2[d_{bo}^i(j)] \quad (4.69)$$

con

$$E[(d_{bo}^i(j))^2] = \sum_n p_{bo}^{i,j}(n) E \left[\left(\underbrace{(T_{slot}^{-i} + T_{inter}^{-i}) + \dots + (T_{slot}^{-i} + T_{inter}^{-i})}_{n \text{ veces}} \right)^2 \right] \quad (4.70)$$

Para calcular el momento de segundo orden de la suma de n términos del tipo $(T_{slot}^{-i} + T_{inter}^{-i})$ se realiza la suma de sus momentos de primer orden al cuadrado y su varianza. Para completar el cálculo se precisa de las siguientes expresiones:

$$\sigma_{T_{slot}^{-i}}^2 = E[(T_{slot}^{-i})^2] - E^2[T_{slot}^{-i}] \quad (4.71)$$

siendo

$$E[(T_{slot}^{-i})^2] = p_{-i}(s)E[(T_s^{-i})^2] + p_{-i}(e)(T_e)^2 + (1 - p_{-i}(s) - p_{-i}(e))E[(T_c^{-i})^2] \quad (4.72)$$

con

$$E[(T_s^{-i})^2] = \frac{\sum p_{s,-i}(m)E[(T_s^i(m))^2]}{\sum p_{s,-i}(m)} \quad (4.73)$$

$$E[(T_s^i(m))^2] = \sum_l p_i(tx = l)T_s(l)^2 \quad (4.74)$$

Para las colisiones

$$E[(T_c^{-i})^2] = \frac{\sum p_{col,-i}(l)T_c(l)^2}{\sum p_{col,-i}(l)} \quad (4.75)$$

Por último, queda

$$\sigma_{T_{inter}^{-i}}^2 = E[(T_{inter}^{-i})^2] - E^2[T_{inter}^{-i}] \quad (4.76)$$

$$E[(T_{inter}^{-i})^2] = (1 - p_{-i}(e))E[(T_{inter,tx}^i)^2] \quad (4.77)$$

Lo que termina el análisis de la desviación típica del tiempo de servicio.

4.1.4. Validación del modelo

Se valida el modelo comparando los resultados arrojados por éste con los obtenidos mediante simulación. La validación se realiza para dos grupos de tráfico, ambos con el mismo número de estaciones:

- Un grupo transmite paquetes de 80 octetos cada 10 milisegundos, modelando una fuente de tráfico VoIP a 64 Kbps.
- El otro grupo transmite un tráfico según un proceso de Poisson de media 500 Kbps, con la longitud de los paquetes derivada de las medidas de [100].

Para validar el modelo, se realizan los siguientes experimentos:

- En primer lugar, un experimento que valide el análisis del efecto diferenciador del parámetro AIFS. Para ello, ambos grupos de estaciones tendrán la misma configuración para la ventana de contienda $\{CW_{min} = 32, m = 5\}$ diferenciándose únicamente en el parámetro A_i . Para el caso de voz dicho parámetro será siempre nulo, mientras que para los datos en un caso será $A_i = 1$ (poca diferenciación) y en otro $A_i = 5$ (mayor diferenciación).
- De similar forma, se validará que el modelo captura el efecto diferenciador de la ventana de contienda de la siguiente forma: con una misma configuración $A_i = 0$ para los dos grupos de estaciones, y la configuración $\{CW_{min} = 16, m = 1\}$ para la voz, los datos tendrán $\{CW_{min} = 32, m = 4\}$ en el primer caso y $\{CW_{min} = 64, m = 4\}$ en el segundo.

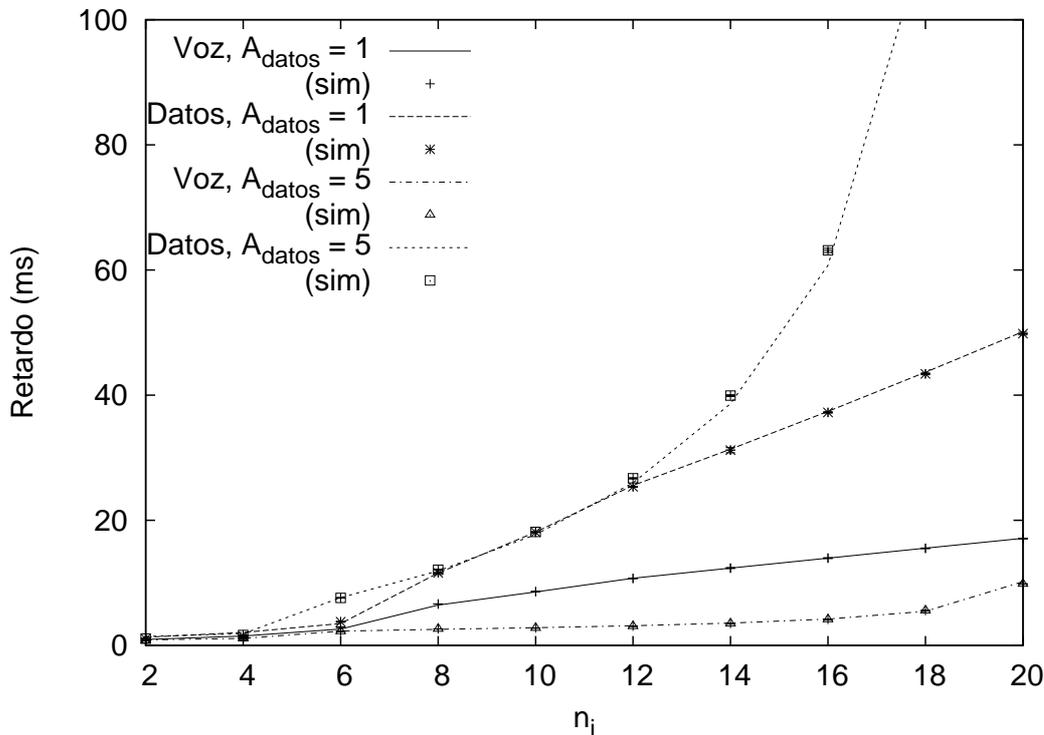


Figura 4.6: Tiempo medio de servicio, diferenciación por AIFS

- Por último, se comprobará la validez del modelo para una diferenciación combinada tanto en AIFS como en CW. Las estaciones de voz tendrán la configuración $\{CW_{\min} = 16, m = 1, A_i = 0\}$ en todos los casos, mientras que las de datos tendrán la configuración $\{CW_{\min} = 32, m = 5, A_i = 1\}$ en un primer caso y la configuración $\{CW_{\min} = 64, m = 4, A_i = 5\}$ en un segundo caso.

4.1.4.1. Tiempo medio de servicio

Los resultados para la validación del cálculo del tiempo medio de servicio se muestran en las Figuras 4.6, 4.7 y 4.8. Los resultados de simulación son presentados con intervalos de confianza para la media al 95%. Dado el ajuste entre el cálculo analítico y el valor de simulación, se concluye la validez del modelo propuesto.

Resulta interesante comprobar la naturaleza de la diferenciación de cada mecanismo: la diferenciación por AIFS (Figura 4.6) se muestra más efectiva conforme aumenta el número n_i de estaciones en cada grupo: si bien para valores bajos de n_i el tipo de servicio es más o menos parecido, a partir de un punto las prestaciones en retardo divergen de forma notable (dicho punto dependiendo del valor de A_i de los datos).

El caso de la diferenciación por CW (Figura 4.7) resulta ligeramente más abrupto, siendo mayor la distinción que se realiza entre estaciones de uno y otro grupo incluso para un número bajo de estaciones en la WLAN. El caso de la diferenciación combinada (Figura 4.8), como cabría esperar, es el que proporciona la mayor prioridad al flujo de voz.

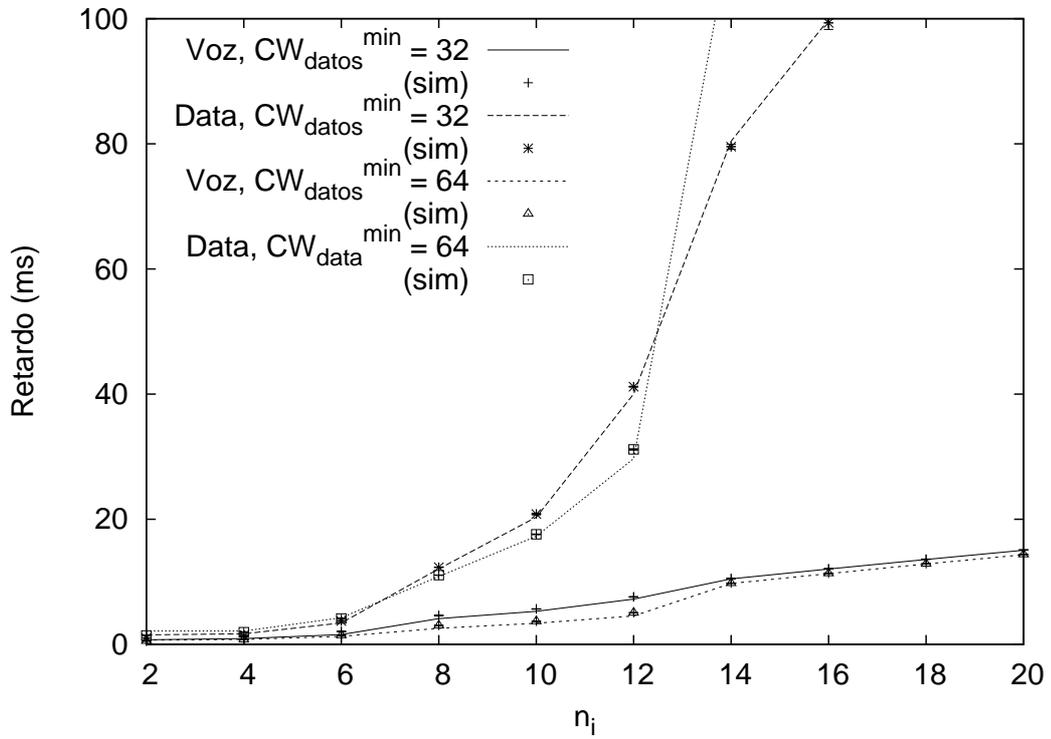


Figura 4.7: Tiempo medio de servicio, diferenciación por CW

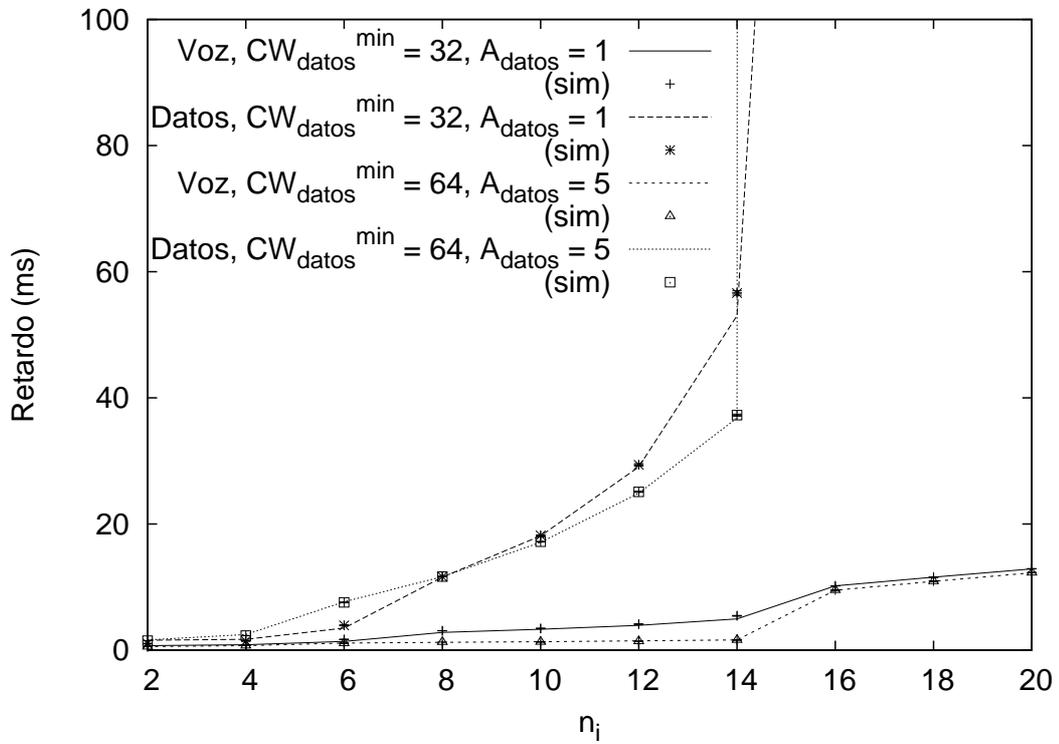


Figura 4.8: Tiempo medio de servicio, diferenciación combinada

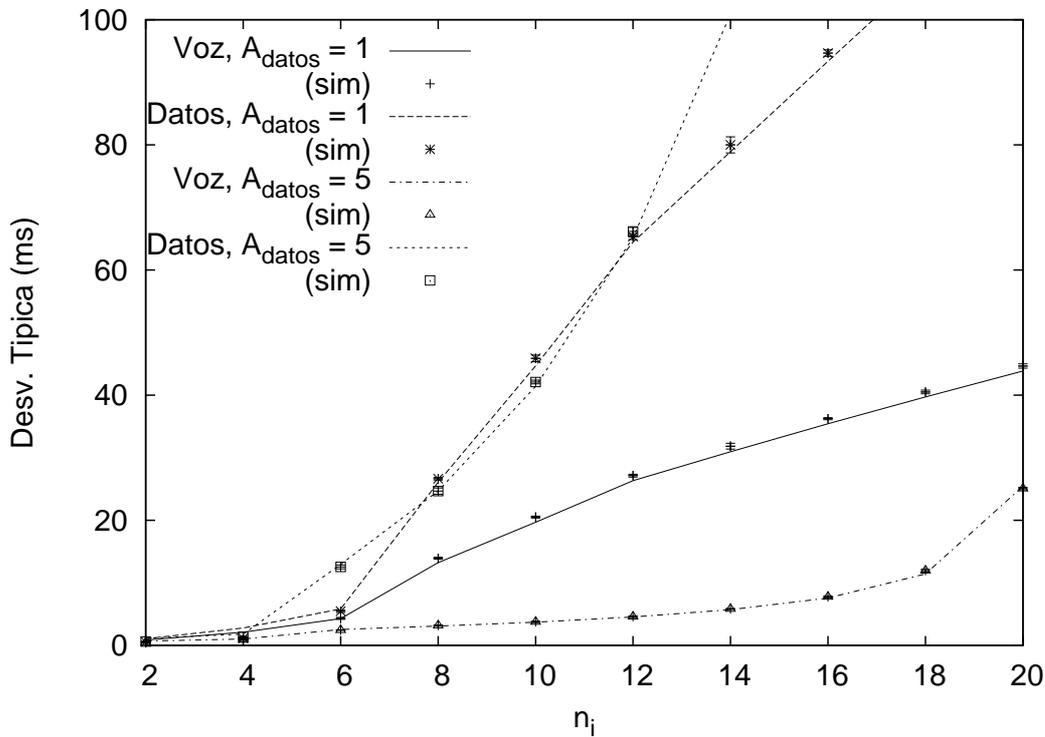


Figura 4.9: Desviación típica, diferenciación por AIFS

4.1.4.2. Desviación típica

De la validación del análisis de la desviación típica del retardo (Figuras 4.9, 4.10 y 4.11) se concluyen los mismos resultados que para el caso anterior, lo que completa la caracterización del momento de primer y segundo orden del tiempo de servicio.

4.2. Análisis de la función de distribución del tiempo de servicio

A continuación, se analiza la función de distribución del tiempo de acceso al medio, tiempo de servicio o tiempo de *backoff*, entendido éste como el tiempo que pasa desde que un paquete alcanza la primera posición de la cola de transmisión hasta que es finalmente transmitido¹⁰. Dicho análisis se realiza para condiciones de no saturación. Se supondrá que la estación cuyo tiempo de servicio va a ser analizado es la más exigente en términos de tiempo de servicio, por lo que estará configurada con AIFS mínimo. También se supondrá, por sencillez, que el resto de estaciones no se diferencian en esta configuración de AIFS, y que todas las estaciones de la WLAN transmiten en cada acceso al medio una única trama, siempre del mismo tamaño¹¹.

¹⁰En el caso en que el paquete es descartado, se considera que su retardo de backoff es ∞ .

¹¹Si bien se podría analizar un caso más general, ello implicaría utilizar expresiones similares a las de la Sección 4.1 lo que aumentaría en gran medida la complejidad del desarrollo. De hecho, en el caso que se estudiará en la Sección 4.3, se cumplen las suposiciones realizadas.

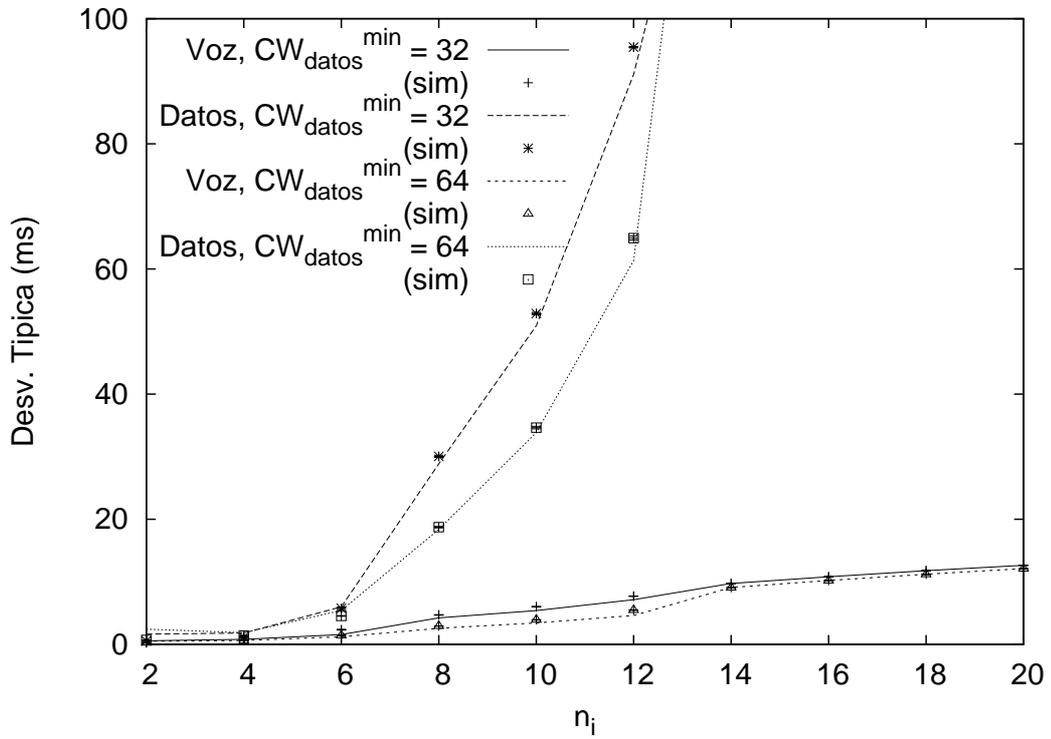


Figura 4.10: Desviación típica, diferenciación por CW

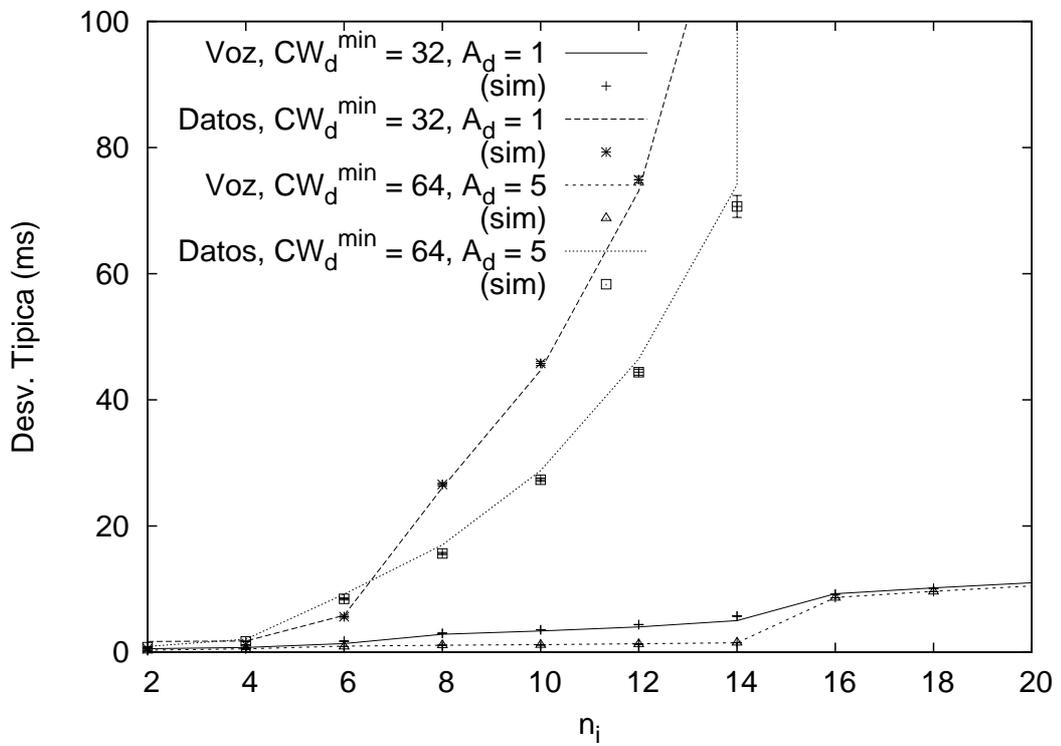


Figura 4.11: Desviación típica, diferenciación combinada

4.2.1. Modelo

Supóngase que una estación i transmite tras un decremento de su contador de backoff con probabilidad constante e independiente τ_i . El cálculo de dicha probabilidad ya ha sido objeto de análisis en la Sección 4.1.1. El objetivo que se persigue es, dado un tiempo D , calcular la probabilidad de que el tiempo de servicio sea menor que D (esto es, siendo d la variable aleatoria que modela el tiempo dedicado al proceso de backoff, obtener $p(d < D)$).

Dicha probabilidad se condiciona al número de colisiones que ha sufrido el paquete que ha iniciado el proceso de backoff

$$P(d < D) = \sum_{i=0}^R P(d < D \mid i \text{ col})P(i \text{ col}) \quad (4.78)$$

$P(i \text{ col})$ es la probabilidad de que un paquete sufra i colisiones antes de ser transmitido con éxito, y R es el máximo número de retransmisiones posibles para una determinada trama. Se puede aplicar el teorema de la probabilidad total al número de ranuras de tiempo que debe decrementar la estación en consideración

$$P(d < D) = \sum_{i=0}^R \sum_{j=0}^{W_i} P(d < D \mid i \text{ col}, j \text{ slots})P(j \text{ slots} \mid i \text{ col})P(i \text{ col}) \quad (4.79)$$

donde $W_i = \sum_{k=0}^i CW_k - 1$, siendo $CW_k = \min(2^k CW_{min}, CW_{max})$, y $P(j \text{ slots} \mid i \text{ col})$ es la probabilidad de que la suma de los $i + 1$ contadores de backoff sea igual a j ,

$$P(j \text{ slots} \mid i \text{ col}) = P\left(\sum_{k=0}^i \text{unif}(0, CW_k - 1) = j\right) \quad (4.80)$$

donde la expresión $\text{unif}(0, C)$ representa una variable aleatoria discreta uniformemente distribuida en el rango $\{0, 1, \dots, C\}$.

La probabilidad de que la suma de una serie de variables aleatorias dé un valor viene dada por la convolución de las mismas

$$P(j \text{ slots} \mid i \text{ col}) = (f_0 * f_1 * \dots * f_i)_j \quad (4.81)$$

siendo f_k la función de probabilidad de una distribución uniforme $\text{unif}(0, CW_{k-1})$. Para el cálculo de dicha convolución se puede hacer uso de transformadas rápidas de Fourier (FFT's), lo que disminuye el tiempo de cómputo.

La probabilidad de que la estación considerada sufra i colisiones antes de transmitir una trama con éxito viene dada por

$$P(i \text{ col}) = p^i(1 - p) \quad (4.82)$$

Siendo p la probabilidad de que una transmisión de una estación colisione, lo que viene dado por la expresión¹²

$$p = 1 - (1 - \tau)^{N-1} \quad (4.83)$$

¹²De nuevo, por sencillez en la exposición, se supondrá que todas las estaciones están configuradas de idéntica manera, por lo que tienen el mismo valor de τ .

La segunda suposición que se realiza es la siguiente: el retardo de backoff, dadas i colisiones y j ranuras de tiempo, se distribuye según una variable aleatoria gaussiana¹³ que se designará por d_{ij} . Con la suposición de independencia entre las diferentes ranuras de tiempo, un número *suficiente* de ranuras de tiempo garantizará la aplicabilidad del Teorema Central del Límite.

Con lo anterior, basta con conocer la media y desviación típica de d_{ij} (esto es, m_{ij} y σ_{ij} , respectivamente) para calcular $P(d < D \mid i \text{ col, } j \text{ slots})$,

$$P(d < D \mid i \text{ col, } j \text{ slots}) = \begin{cases} 0,5 + 0,5 \operatorname{erf} \left(\frac{D - m_{ij}}{\sqrt{2}\sigma_{ij}} \right), & \frac{D - m_{ij}}{\sigma_{ij}} \geq 0 \\ 0,5 \operatorname{erfc} \left(-\frac{D - m_{ij}}{\sqrt{2}\sigma_{ij}} \right), & \frac{D - m_{ij}}{\sigma_{ij}} < 0 \end{cases} \quad (4.84)$$

Con la suposición de independencia entre ranuras, m_{ij} se puede calcular como la suma de la duración de las ranuras de tiempo que intervienen en d_{ij}

$$m_{ij} = j m_n + i T_c + T_s \quad (4.85)$$

donde m_n es la duración media de una ranura en la que la estación considerada no transmite, T_c es la duración de una colisión y T_s es la duración de un éxito.

La duración media de una ranura en que la estación no transmite viene dada por

$$m_n = P_{s,n} T_s + P_{c,n} T_c + P_{e,n} T_e \quad (4.86)$$

donde $P_{s,n}$, $P_{c,n}$ y $P_{e,n}$ son las probabilidades de que la ranura contenga un éxito, una colisión o esté vacía, que se obtienen de forma inmediata de los valores de τ y N :

$$P_{s,n} = (N - 1)\tau(1 - \tau)^{N-2} \quad (4.87)$$

$$P_{e,n} = (1 - \tau)^{N-1} \quad (4.88)$$

$$P_{c,n} = 1 - P_{s,n} - P_{e,n} \quad (4.89)$$

Bajo la suposición de independencia, la desviación típica se puede calcular como

$$\sigma_{ij}^2 = j \sigma_n^2 \quad (4.90)$$

donde σ_n es la desviación típica de la longitud de una ranura de tiempo en la que la estación considerada no transmite,

$$\sigma_n^2 = P_{s,n} T_s^2 + P_{c,n} T_c^2 + P_{e,n} T_e^2 - m_n^2 \quad (4.91)$$

¹³Esta suposición es la principal diferencia entre el modelo presentado y los análisis de [101, 102]. Con dicha suposición, sólo es preciso calcular la media y varianza de la variable aleatoria gaussiana, lo que no implica una complejidad computacional elevada. En los análisis referenciados, en cambio, se calculan todos los posibles valores de d_{ij} y su probabilidad, lo que resulta muy costoso en términos computacionales dado el elevado número de casos a considerar.

4.2.1.1. Caso de RTS/CTS

Si se emplea el mecanismo de RTS/CTS, las transmisiones exitosas de tramas vienen precedidas por el intercambio RTS/CTS. Las colisiones, por otra parte, sólo ocurren en la transmisión de tramas RTS. De esta forma, basta con modificar las longitudes de T_s y T_c del modelo para tener en cuenta el uso de dicho mecanismo:

$$T_s = T_{PLCP} + \frac{RTS}{C} + SIFS + T_{PLCP} + \frac{CTS}{C} + SIFS + T_{PLCP} + \frac{H+l}{C} + SIFS + T_{PLCP} + \frac{ACK}{C} + DIFS \quad (4.92)$$

$$T_c = T_{PLCP} + \frac{RTS}{C} + EIFS \quad (4.93)$$

4.2.1.2. Caso de longitudes variables de paquetes

A continuación, se relaja la condición de que las longitudes de transmisión sean todas iguales. Supóngase que una transmisión es de un tamaño l (considerando un conjunto de longitudes L) con probabilidad P_l . Se supondrá, de nuevo por sencillez, que todas las estaciones transmiten con la misma función de probabilidad.

Para considerar las diferentes longitudes de transmisión, es preciso modificar las expresiones para obtener tanto m_{ij} como σ_{ij} . m_{ij} se calcula como

$$m_{ij} = j m_n + i m_c + m_s \quad (4.94)$$

donde m_c es la longitud media de una colisión en la que participa la estación considerada y m_s es la longitud media de un éxito de dicha estación.

La longitud media de una ranura en la que no transmite la estación se calcula, esta vez, como

$$m_n = \sum_{l \in L} P_{s,l,n} T_{s,l} + \sum_{l \in L} P_{c,l,n} T_{c,l} + P_{e,n} T_e \quad (4.95)$$

donde $P_{s,l,n}$ es la probabilidad de que una ranura en la que la estación no transmite contenga una transmisión de tamaño l , $P_{c,l,n}$ es la probabilidad de que contenga una colisión con la trama de mayor longitud de tamaño l , y $T_{s,l}$ y $T_{c,l}$ son las longitudes de las ranuras para cada uno de esos casos

Las probabilidades $P_{s,l,n}$ y $P_{c,l,n}$ se obtienen de:

$$P_{s,l,n} = (N-1)\tau(1-\tau)^{N-2}P_l \quad (4.96)$$

$$P_{c,l,n} = (1 - P_{s,l,n} - P_{e,n}) P_{c,l} \quad (4.97)$$

donde $P_{c,l}$ es la probabilidad de que la mayor longitud involucrada en una colisión es l . Despreciando las colisiones de más de dos estaciones, se puede obtener como

$$P_{c,l} = 2 P_l \sum_{k \in L_l} P_k - P_l^2 \quad (4.98)$$

siendo L_l el conjunto de longitudes menores o iguales que l .

El cálculo de $T_{s,l}$ y $T_{c,l}$ resulta inmediato de las expresiones anteriormente descritas. El cálculo de m_s y m_c se realiza como sigue

$$m_s = \sum_{l \in L} P_l T_{s,l} \quad (4.99)$$

$$m_c = \sum_{l \in L} P_{c,l} T_{c,l} \quad (4.100)$$

Por último, la desviación típica σ_{ij} para el caso de longitud variable de paquetes se realiza de la siguiente forma:

$$\sigma_{ij}^2 = j \sigma_n^2 + i \sigma_c^2 + \sigma_s^2 \quad (4.101)$$

con

$$\sigma_n^2 = \sum_{l \in L} P_{s,l,n} T_{s,l}^2 + \sum_{l \in L} P_{c,l,n} T_{c,l}^2 + P_{e,n} T_e^2 - m_n^2 \quad (4.102)$$

En este caso, se ha considerado que el mecanismo RTS/CTS se encuentra desactivado. Sin embargo, sería trivial realizar la extensión para tener en cuenta dicho mecanismo, que añadiría una longitud constante a las longitudes de los éxitos, y fijaría un tamaño para la duración de las colisiones.

4.2.2. Validación

Se compara los resultados del modelo con los obtenidos mediante simulación. Sea una WLAN sirviendo a N flujos de VoIP, generando cada flujo paquetes de 80 octetos cada 10 ms. En la Figura 4.12 se muestra el valor de la función de distribución para 10 y 30 estaciones, con tres valores diferentes de la CW_{min} , a saber: 32, 128 y 512.

Se comprueba la validez del modelo, tanto para el caso de mayor carga en la red $N = 30$ (con las estaciones en saturación) como para el caso en el que el número de flujos no es demasiado elevado (no saturación).

4.3. Caso de estudio: entorno VoIP

A continuación, se muestra la aplicación de los anteriores análisis, empleándolos como herramientas para la configuración de una WLAN cerrada. Se supone que únicamente se transporta tráfico de VoIP. Se utilizará el análisis del valor medio y varianza del tiempo de servicio para desarrollar un mecanismo de control de admisión y configuración de la WLAN, y se validará que en dichas circunstancias también se puede caracterizar el comportamiento aleatorio del tiempo de servicio.

Sea un número de estaciones N que transmite, cada una, un tráfico VoIP de ρ Kbps. Dado un máximo retardo medio D y una máxima desviación estándar σ_D admisibles, partiendo del análisis de la Sección 4.1 se desarrolla el siguiente algoritmo de configuración y control de admisión:

- Calcular el punto de operación óptima de la red τ_{opt} (Sección 3.2.2) y obtener la tasa máxima de transmisión que puede proporcionarse a cada flujo. Si dicha tasa es menor que ρ , no existe configuración alguna que permita proporcionar el servicio pedido.

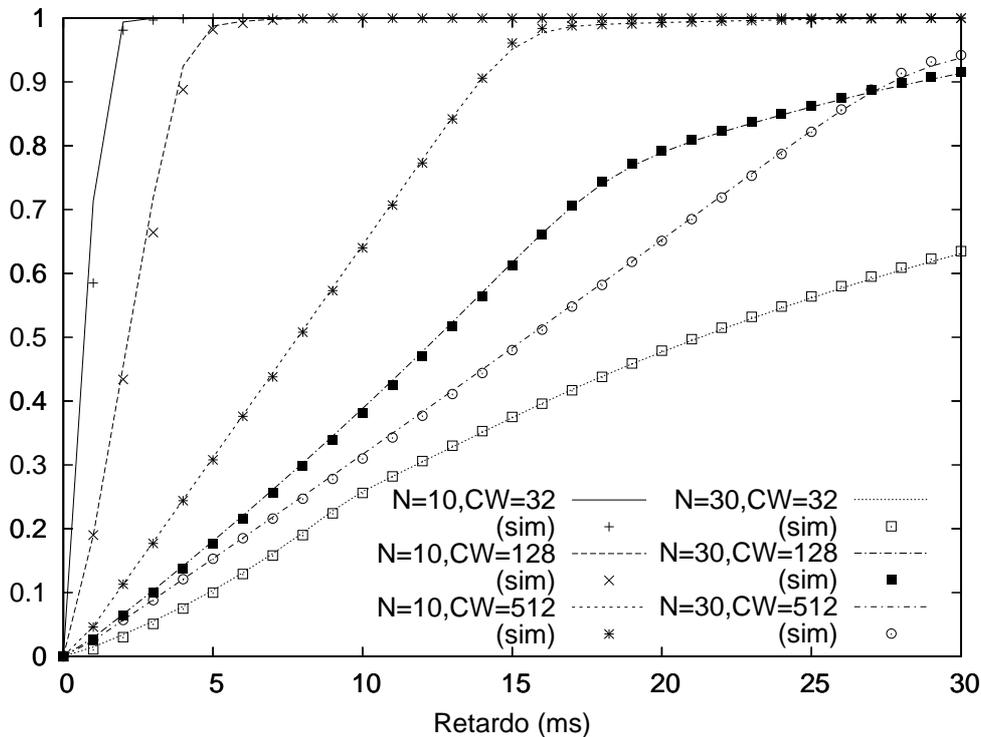


Figura 4.12: Validación del análisis de la función de distribución del tiempo de servicio

- En caso de que la tasa en esas circunstancias sea mayor que ρ , se obtiene la τ de operación de la WLAN, igualando el tráfico generado y el tráfico servido (Sección 4.1.1).
- Calcular, mediante la ecuación que se puede definir a través de (4.21)–(4.59) (esto es, la Sección 4.1.2), la CW^{min} mayor que cumple con el criterio de retardo medio menor que D .
- Calcular mediante las ecuaciones (4.60)–(4.77) (Sección 4.1.3) la mayor CW^{min} que cumple con el criterio de desviación típica σ_D .
- La configuración será es la CW menor de los anteriores valores, si se comprueba que, en condiciones de saturación, con dicha configuración las N estaciones cursan un tráfico mayor que ρ .

El escoger la mayor CW posible, así como realizar la última comprobación sobre el ancho de banda de la red en saturación, resulta preciso para evitar que la WLAN opere en un punto cercano a la τ de saturación y que, caso de llegar a éste, permanezca en él (lo que implicaría la aparición de los problemas descritos en la Sección 3.2).

El algoritmo propuesto de configuración óptima puede emplearse, además, como mecanismo de control de admisión: si para una WLAN operando con N flujos un nuevo flujo solicita admisión, y no existe configuración posible que, con los $N + 1$ flujos, proporcione las mismas garantías de QoS, dicho flujo no podrá ser admitido en la WLAN (si bien otras

Tabla 4.1: Evaluación del algoritmo de configuración

D	σ	n	CW_{ana}	D_{ana}	σ_{ana}	CW_{srch}	D_{srch}	σ_{srch}
5 ms	5 ms	10	314	4.95	2.78	317	4.99	2.82
		15	225	4.91	2.87	229	4.99	2.92
		20	118	4.72	3.02	125	4.99	3.25
		21	-	-	-	-	-	-
6 ms	4 ms	10	381	5.95	3.38	383	5.98	3.39
		15	273	5.91	3.48	277	5.99	3.53
		20	144	5.73	3.71	151	5.98	3.87
		21	-	-	-	-	-	-
7 ms	3 ms	10	330	5.18	2.92	338	5.30	2.99
		15	226	4.92	2.88	236	5.12	2.99
		20	109	4.34	2.81	115	4.59	2.98
		21	-	-	-	-	-	-

opciones, como expulsar algún flujo existente o reducir las garantías ofrecidas, también podrían ser tomadas: ello dependerá de las políticas definidas en la WLAN).

4.3.1. Evaluación de la configuración propuesta

Se valora la efectividad de la configuración propuesta para un total de nueve escenarios, resultado de establecer, para 10, 15 y 20 estaciones, los tres requisitos siguientes sobre el retardo¹⁴:

- $D = 5 \text{ ms}$ y $\sigma_D = 5 \text{ ms}$
- $D = 6 \text{ ms}$ y $\sigma_D = 4 \text{ ms}$
- $D = 7 \text{ ms}$ y $\sigma_D = 3 \text{ ms}$

Los resultados se muestran en la Tabla 4.1. En la columna CW_{ana} se muestra la configuración que arroja el algoritmo propuesto en la anterior sección. D_{ana} y σ_{ana} son los valores de retardo y desviación típica obtenidos mediante simulación con la configuración propuesta CW_{ana} . Además, se realiza un barrido exhaustivo en CW en búsqueda del mayor valor de dicho parámetro que cumple con los criterios establecidos. Dicho valor se muestra en la columna CW_{srch} , y el retardo medio y desviación típica que se obtiene mediante simulación se indica por D_{srch} y σ_{srch} , respectivamente.

De los resultados de la tabla se aprecia que, para los dos primeros criterios de QoS, es el retardo medio el factor que más limita a la hora de fijar la configuración óptima. En el último criterio, el bajo valor de σ_D fijado resulta ser el determinante. Se comprueba, además, que los valores obtenidos mediante búsqueda exhaustiva no difieren significativamente de los propuestos por el algoritmo.

¹⁴Nótese que, si se suministran dichos requisitos de calidad, la cola de transmisión de las tramas -que se generan cada 10 ms- se encontrará en la mayoría de los casos vacía o con un único paquete pendiente de ser enviado.

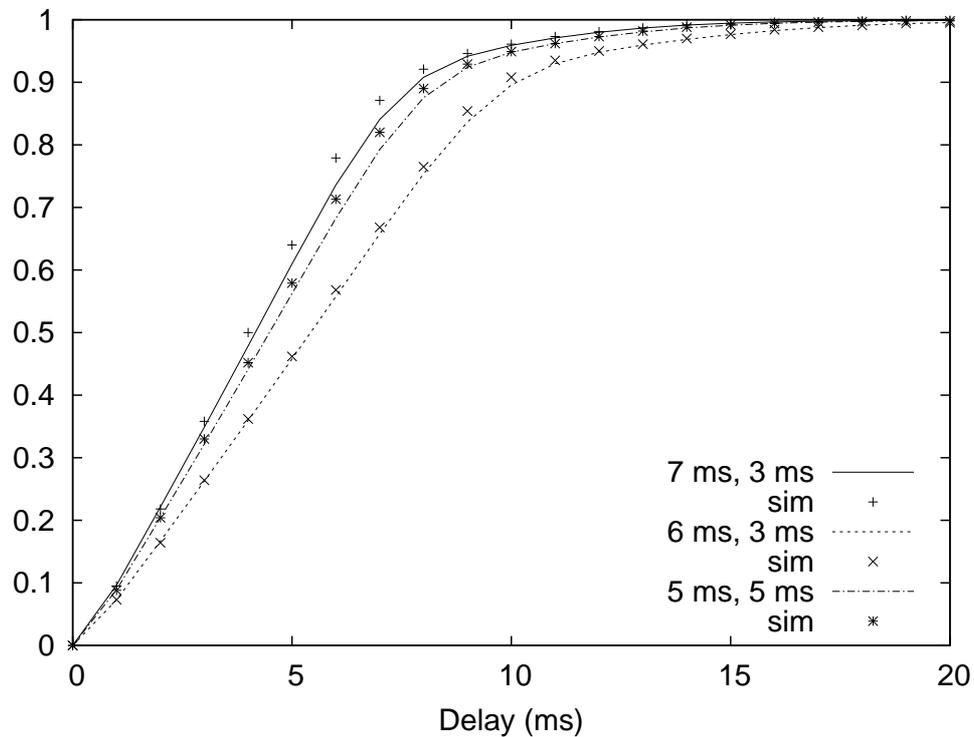


Figura 4.13: Distribución del retardo para 20 flujos VoIP

Con objeto de comprobar la eficiencia proporcionada por el algoritmo, se obtuvo el máximo número de flujos que pueden transmitirse con el criterio de calidad fijado. Para hacer esto se incrementó el valor de N hasta que según el algoritmo propuesto no se encontró ninguna configuración que cumpliera con los requisitos. Dicho N coincide con la búsqueda exhaustiva realizada mediante simulación, por lo que el algoritmo no desaprovecha los recursos disponibles en la WLAN.

En la Figura 4.13 se muestra, para el número máximo de flujos que pueden ser cursados por la WLAN, la función de distribución del tiempo de servicio. Se comprueba el ajuste entre los valores analíticos y los obtenidos a través de la simulación. La existencia de un modelo para la función de distribución permite, por una parte, obtener más información acerca de las prestaciones que se obtienen con la configuración que se ha seleccionado con base en el retardo medio y el *jitter*; por otra parte, permite también desarrollar un control de admisión con garantías de servicio más específicas (como se hizo en [13]), así como el desarrollo de mecanismos de configuración que, no obstante, conllevan mayor complejidad computacional que el realizado para el caso de estudio.

4.4. Metodología experimental en un escenario real

Se desarrolla en esta sección una metodología experimental para obtener la configuración óptima de una WLAN operando en modo EDCA. Se considerará un caso con dos tipos de tráfico: tráfico de voz, con requisitos de entrega; y tráfico de datos, que

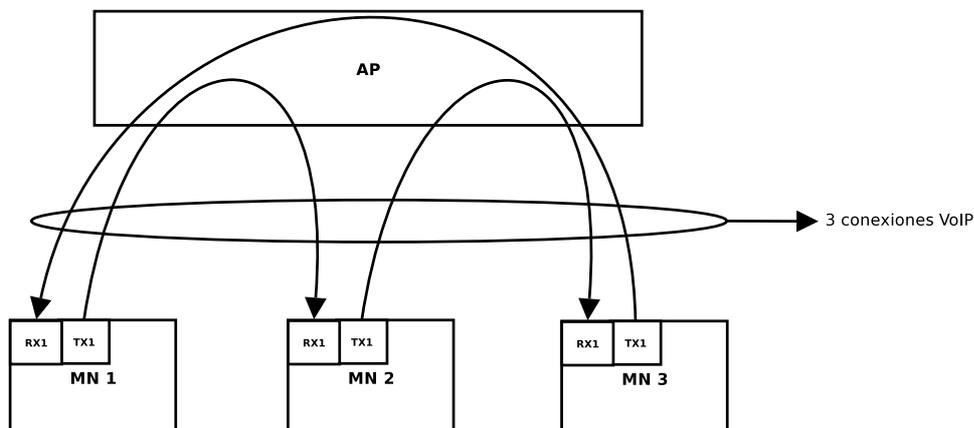


Figura 4.14: Arquitectura empleada en la generación del tráfico VoIP

no precisa de provisión de QoS. La prioridad de la red será admitir el mayor número de conversaciones de VoIP, pero proporcionando el mejor servicio a los datos en la medida de lo posible.

Con objeto de que la configuración obtenida pueda ser aplicada a otros escenarios, se establece el siguiente modelo para cada tipo de flujo:

- El tráfico de voz se modelará como un proceso generador de tramas de 80 octetos a intervalos regulares de 10 ms; de esta forma, con la configuración propuesta otras codificaciones más eficientes (mayor tiempo entre llegadas, menor índice de actividad) no obtendrán un servicio peor.
- Los flujos de datos siempre tendrán una trama de 1500 octetos pendiente de ser transmitida; al igual que en el caso de la voz, cualquier aplicación menos agresiva recibirá un servicio no inferior al recibido para el caso de este modelo.

4.4.1. Escenario

Se instaló un escenario con 15 máquinas ejecutando el sistema operativo Linux (distribución Mandrake 10.0 con un kernel 2.6.11), empleándose tarjetas inalámbricas con un chipset de Atheros que se controlaban a través del *driver* Madwifi (Multiband Atheros Driver for WiFi)¹⁵. La generación de tráfico se realizó con la herramienta MGEN¹⁶, siendo preciso utilizar clientes y un servidor de NTP (Network Time Protocol) para poder realizar medidas en cada sentido de la transmisión. Dadas las limitaciones del planificador de eventos de Linux, y a pesar de las alternativas estudiadas por Gurtov [103], la generación de paquetes más fidedigna al modelo se consiguió empleando una arquitectura similar a la representada por la Figura 4.14, en la que cada máquina sólo genera y recibe un flujo de VoIP y el punto de acceso actúa únicamente para reenviar las tramas.

Una descripción completa y bien detallada del escenario de pruebas desplegado se puede encontrar en [15].

¹⁵<http://madwifi.org/>

¹⁶<http://pf.itd.nrl.navy.mil/mgen>

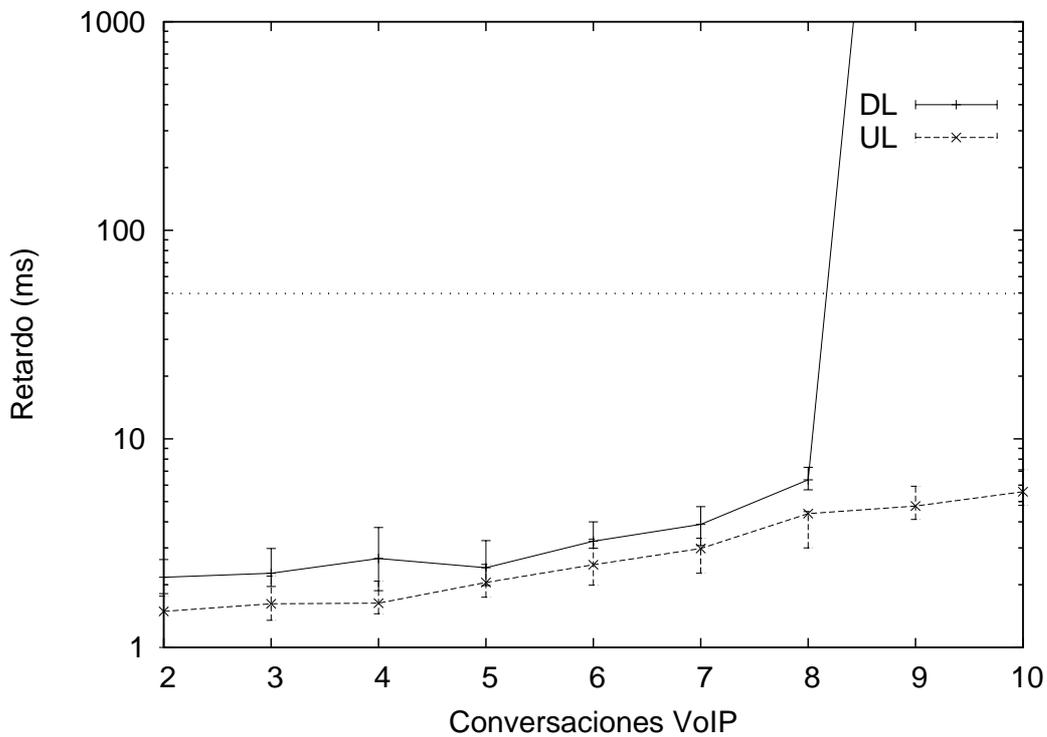


Figura 4.15: Retardo de voz para el modo DCF

4.4.2. Configuración del tráfico *real-time*

Acudiéndose a la recomendación ITU-T G.114 [104], se tiene que si se quiere proporcionar un buen servicio de voz es preciso garantizar un retardo menor de 150 ms por sentido con unas pérdidas no superiores al 5%. Suponiéndose que un tráfico VoIP puede atravesar hasta dos redes WLAN además de una red cableada, se divide como sigue los anteriores requisitos para las tecnologías involucradas: la red cableada debe proporcionar un retardo menor de 50 ms para el 99% de los paquetes, mientras que la tecnología inalámbrica debe garantizar una cota de 50 ms para el 98% de los paquetes que transmite. De esta forma, se tiene que el retardo será de $50 + 50 + 50 = 150 \text{ ms}$ con una probabilidad de $98\% \times 99\% \times 98\% = 95\%$ ¹⁷. Los 50 ms de cota para la tecnología inalámbrica, a diferencia de los retardos de secciones anteriores, incluyen tanto el tiempo de servicio como el tiempo de espera en cola.

En primer lugar se mide, para el modo DCF (con el estándar 802.11b), el percentil .98 para el retardo en cada sentido. Dicha medida se realiza en cinco ocasiones, con objeto de proporcionar cierta información sobre la variabilidad de los resultados obtenidos: se presenta tanto la media como el máximo y mínimo de los 5 experimentos. En la Figura 4.15 se muestra la variación de dicha cota con el número de conversaciones en la WLAN.

Se aprecia en la figura que es el punto de acceso el cuello de botella. Dicho resultado era de esperar: dado que el AP se comporta, en términos de acceso al medio, como cualquier

¹⁷Se supone que las cotas de retardo son independientes entre sí.

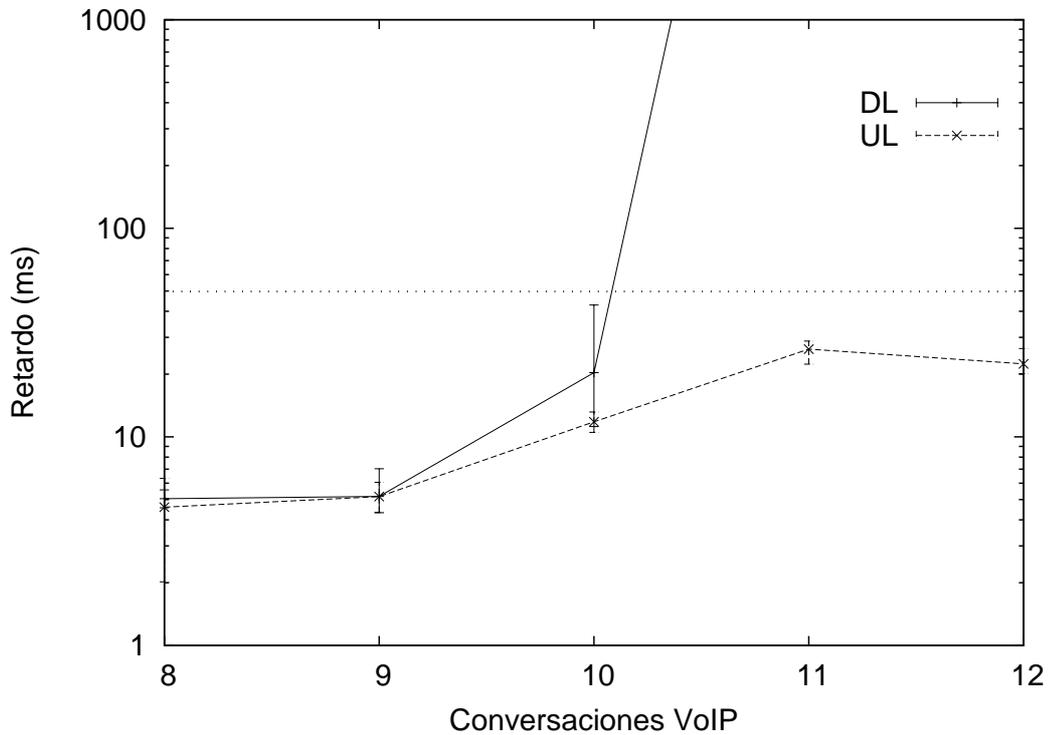


Figura 4.16: Retardo de VoIP con el uso de $TXOP_{max}$

otra estación de la WLAN (no ha sido configurado de forma diferenciada), según aumenta el número de conversaciones se encuentra con mayor tráfico que atender, hasta que no es capaz de proporcionar las garantías requeridas. El número máximo de conversaciones que puede cursar en estas circunstancias y para los requisitos establecidos una WLAN operando en el modo DCF es, por lo tanto, ocho.

Considérese a continuación el uso del modo EDCA. El parámetro TXOP de dicho modo permite la transmisión de varias tramas en un mismo acceso al medio; el uso de este mecanismo resulta del todo aconsejable, dado que, si tras la transmisión de un paquete queda otro en cola pendiente de ser transmitido, este paquete no tendrá que esperar un nuevo proceso de backoff para ser enviado. En caso contrario, el paquete que aguarda en cola (que lleva más de 10 ms esperando a ser transmitido) provocará un aumento considerable en el percentil del retardo para este flujo, frente a la alternativa de provocar un ligero aumento en el tiempo de transmisión del resto de paquetes de la red.

Si se realiza el mismo experimento que el de la Figura 4.15 pero fijando el valor de $TXOP_{voice}$ a su valor máximo, se obtiene como resultado la Figura 4.16: permitir que más de un paquete pueda ser transmitido en el mismo acceso al canal consigue, con el criterio de QoS establecido, que la WLAN pueda dar servicio a dos conversaciones más de VoIP.

Por lo anterior, se propone la siguiente regla de configuración para la WLAN:

Regla 1: $TXOP_{voice} = TXOP_{max}$

Por otra parte, para el tráfico de VoIP la diferenciación mediante AIFS no resulta de utilidad, dado que ésta realiza una distinción bastante grosera entre los flujos en que es

aplicada y sólo para intensidades de tráfico elevadas (véase [34] y [105]). Por ello se propone, para la transmisión de tráfico VoIP, fijar dicho parámetro al mínimo valor:

Regla 2: $AIFS_{voice} = AIFS_{min} = 2$

Resta encontrar la configuración más adecuada para los parámetros que regulan la ventana de contienda. El parámetro de CW_{max} se fijará al mismo valor que CW_{min} , dado que para la configuración óptima no será necesario un mecanismo que se adapte a la tasa de colisiones en la red, y el incremento del contador de backoff supondría aumentar el retardo de un paquete que ya ha sufrido una colisión, lo que llevaría a unas prestaciones aún peores en tiempos de acceso al canal. Ello lleva a la siguiente regla

Regla 3: $CW_{min}^{voice} = CW_{max}^{voice}$

A continuación, se realiza un barrido en el espacio de configuraciones posibles de la ventana de contienda. Dado un valor n del número de flujos de voz, se realiza una búsqueda en el espacio (CW_{MN}, CW_{AP}) para los posibles valores de CW_{min} que cumplan el criterio de retardo impuesto. Si existe alguna configuración que proporcione dicho criterio, se aumenta n en una unidad y se vuelve a realizar la búsqueda, hasta que ninguna configuración proporcione la QoS requerida. De esta forma, se tiene que el máximo número de conversaciones VoIP que se pueden soportar es de 11.

Regla 4: No aceptar más de 11 conversaciones en un mismo AP

En la Tabla 4.2 se muestran los resultados del barrido realizado para $n = 11$ flujos de voz¹⁸. Por cada posible configuración de (CW_{AP}, CW_{MN}) se obtiene la cota al 98 % del retardo, indicándose en la tabla la mayor de la media dichas cotas para cada sentido, así como el valor mínimo y máximo de la serie de experimentos que se realizaron para obtener dicho valor. No se midieron cotas de retardo mayores de un segundo.

Se comprueba que sólo para un rango de configuración de la CW todos los experimentos proporcionan un valor menor de la cota de retardo requerida. Dicho rango se sitúa para valores de $CW_{AP} = \{4, 8, 16, 32\}$ y $CW_{MN} = \{32, 64, 128, 256\}$. La explicación para este comportamiento es bien sencilla: valores bajos de CW_{MN} provocan un número de colisiones elevado, lo que provoca que ni las estaciones ni el punto de acceso puedan transmitir el tráfico con las garantías requeridas; por otra parte, valores demasiado elevados alejan el comportamiento de la red de su punto óptimo, al aumentar en exceso el número de ranuras que debe decrementar una estación antes de transmitir (el comportamiento es similar para el caso de CW_{AP}). Por lo tanto, se tiene que la configuración óptima para el caso considerado (resaltada en la Tabla 4.2) es la siguiente:

Regla 5: Para voz, $CW_{MN} = 64$ y $CW_{AP} = 32$

¹⁸Valores de CW_{AP} mayores a 256 proporcionaban también un retardo de más de un segundo.

Tabla 4.2: Obtención de la configuración de CW para VoIP

CW_{MN}	CW_{AP}						
	4	8	16	32	64	128	256
4	>1 seg (163.54,>1 seg)	>1 seg (>1 seg,>1 seg)	>1 seg (>1 seg,>1 seg)	>1 seg (>1 seg,>1 seg)	>1 seg (>1 seg,>1 seg)	>1 seg (>1 seg,>1 seg)	>1 seg (>1 seg,>1 seg)
8	278.55 (248.16,325.19)	>1 seg (462.34,>1 seg)	>1 seg (>1 seg,>1 seg)	>1 seg (>1 seg,>1 seg)	>1 seg (>1 seg,>1 seg)	>1 seg (>1 seg,>1 seg)	>1 seg (>1 seg,>1 seg)
16	105.79 (76.38,126.24)	102.96 (89.7,128.88)	>1 seg (76.81,>1 seg)	>1 seg (>1 seg,>1 seg)	>1 seg (>1 seg,>1 seg)	>1 seg (>1 seg,>1 seg)	>1 seg (>1 seg,>1 seg)
32	33.61 (27.29,44.74)	45.87 (32.42,67.02)	36.3 (27.7,42.44)	69.63 (14.76,189.57)	>1 seg (97.36,>1 seg)	>1 seg (>1 seg,>1 seg)	>1 seg (>1 seg,>1 seg)
64	23.4 (21.67,25.4)	29 (24.75,32.56)	25.08 (19.77,24.89)	22.56 (20.01,26.59)	>1 seg (20.82,>1 seg)	>1 seg (507.47)	>1 seg (>1 seg,>1 seg)
128	28.08 (23.52,31.76)	29.53 (27.34,32.11)	29.61 (28.38,30.88)	26.94 (23.93,29.98)	>1 seg (20.05,>1 seg)	>1 seg (>1 seg,>1 seg)	>1 seg (>1 seg,>1 seg)
256	39.76 (38,40.88)	44.97 (41.29,48.08)	47.77 (41.65,53.94)	43.48 (39.7,45.75)	42.67 (65.73,>1 seg)	>1 seg (>1 seg,>1 seg)	>1 seg (>1 seg,>1 seg)
512	69.18 (61.44,73.91)	68.54 (61.26,75.42)	76.43 (71.59,93.54)	73.45 (69.51,76.74)	>1 seg (24.47,>1 seg)	>1 seg (184.83,>1 seg)	>1 seg (>1 seg,>1 seg)

4.4.3. Configuración del tráfico no *real-time*

A continuación se aborda la obtención de la configuración de las estaciones de datos que proporcionan un mejor rendimiento de la WLAN.

En primer lugar, por la naturaleza del parámetro TXOP, resulta claro que un uso de dicho parámetro por parte de las transmisiones de datos provocaría mayores retardos en el tráfico de voz, dado que las estaciones de datos podrían causar esperas aún mayores e impredecibles en la entrega de tramas de VoIP (lo que no sería muy recomendable, siendo el servicio para los datos de tipo *best effort*). Es por ello que sólo se permite a las estaciones de datos transmitir una única trama por cada acceso al medio:

$$\text{Regla 6: } TXOP_{data} = 0$$

Por otra parte, dado que el servicio que se proporcionará a las estaciones de datos no define garantías de retardo en la entrega, el uso del mecanismo de retroceso exponencial resulta aconsejable. En efecto, no sólo dicho mecanismo permite a la red adaptarse en condiciones de alta carga de tráfico, sino que además, caso de una colisión con una estación de voz, se facilita que ésta pueda volver a transmitir sin la participación de la estación con la que acaba de colisionar. Por esto, se decide emplear el mismo esquema de crecimiento de la ventana que DCF, que permite que se duplique hasta 5 veces para la transmisión de una misma trama.

$$\text{Regla 7: } CW_{max}^{data} = 2^5 CW_{min}^{data}$$

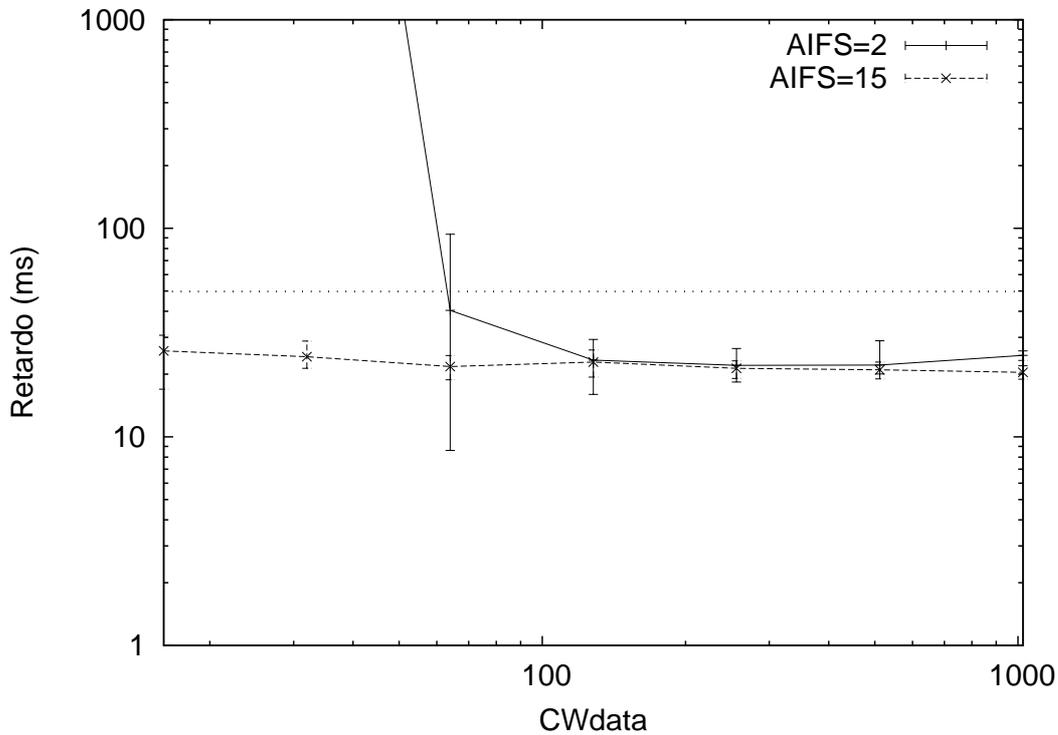


Figura 4.17: Retardo para diferentes valores de AIFS y CW_{min}

Queda por especificar la configuración de los parámetros AIFS y CW_{min} para los flujos de datos. A tal efecto, se realiza el siguiente experimento: sea la WLAN operando con el máximo número de flujos de VoIP que puede soportar; para el valor de $AIFS^{data}$ mínimo, se realiza un barrido en la CW_{min} de los datos (incrementando su valor) midiéndose el máximo de la cota al 98% para el retardo de VoIP en cada sentido. Se repite el mismo experimento, pero con el valor de $AIFS^{data}$ máximo. El resultado de dicho experimento se muestra en la Figura 4.17.

El comportamiento mostrado en la figura resulta fácilmente explicable: la configuración con $AIFS_{min}$ sólo permite mantener la cota en el retardo para valores elevados de CW_{min} , que reduce la frecuencia de las transmisiones de las estaciones saturadas en el medio. En cambio, con la configuración de $AIFS_{max}$, producir un decremento en el contador de backoff en situaciones de cierta carga resulta bastante costoso (pues hay un número relativamente elevado de ranuras no vacías), por lo que en todo el rango de CW_{min} considerado ($\{16,1024\}$) se mantienen las prestaciones de QoS requeridas.

Para distinguir entre ambas posibles configuraciones, se realiza un nuevo experimento: para ambas configuraciones, se realiza un barrido en el número de conversaciones de voz que tienen lugar en la WLAN, midiéndose el ancho de banda que recibe un único flujo de datos (a la par que se comprueba que las garantías de QoS se mantienen). Los resultados se muestran en la Figura 4.18.

El comportamiento que se observa era de esperar: en situaciones de baja carga, la estación con mínimo AIFS no obtiene ventaja del bajo número de ranuras que tiene que

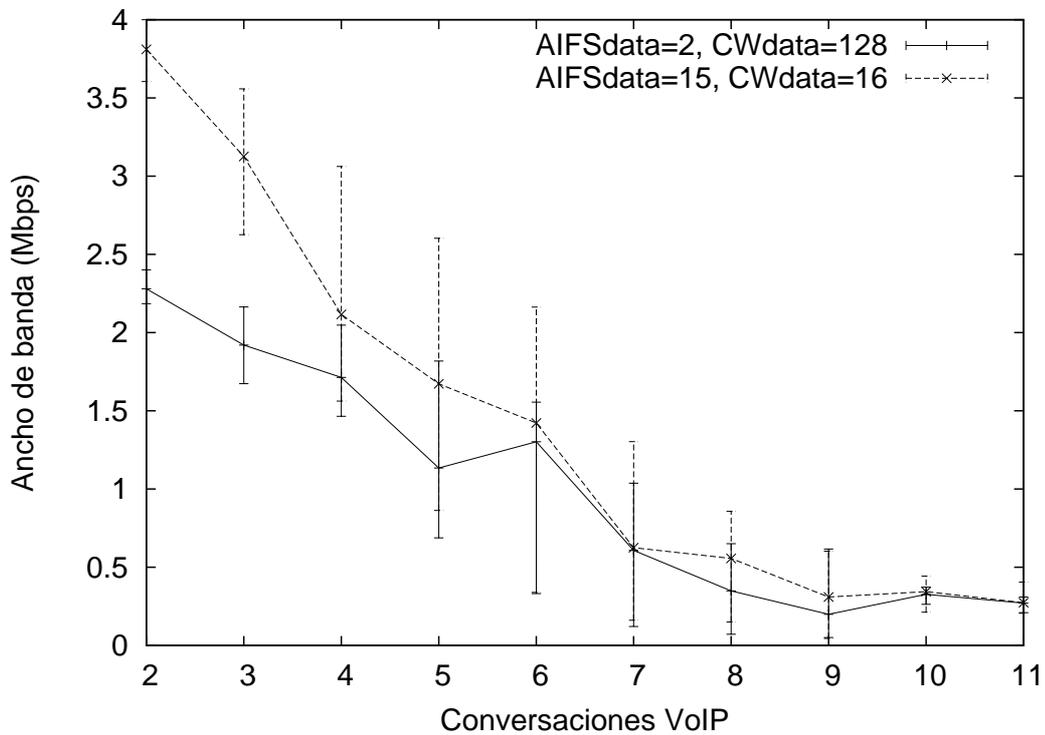


Figura 4.18: Tasa de transmisión de los datos vs. número de flujos VoIP

esperar antes de decrementar su contador de backoff, pues el valor medio de éste es mucho mayor que el de la estación con AIFS elevado. En este caso, en cambio, si bien tras una transmisión tiene que esperar un cierto número de ranuras, luego compensa sobradamente dicho tiempo al arrojar valores medios de backoff menores, por lo que recibe una mayor tasa que en el caso anterior. A medida que aumenta la carga de tráfico prioritario, el comportamiento de ambas alternativas se asemeja (si bien la que emplea $AIFS_{max}$ siempre proporciona mayor ancho de banda), hasta que en la situación límite los recursos destinados a la transmisión de datos resulta prácticamente el mismo, independientemente del mecanismo que empleen para acceder.

Por lo anterior, se recomienda la siguiente configuración para una estación de datos:

Regla 8: Para datos, $AIFS = 15$ y $CW^{min} = 16$

Por último, en el caso de tener más estaciones de datos, será preciso suministrar una configuración tal que su comportamiento sea similar al proporcionado mediante la regla anterior. A tal efecto, se plantea que si n es el número de estaciones de datos, cada estación deberá configurar su AC para los datos con un valor n veces superior al de la regla para una única estación de datos. Para confirmar la validez de esta última regla, se realiza la medición de la cota del retardo para un escenario con diferente número de estaciones de datos conformes a dicha configuración. Los resultados, que validan dicha regla, se muestran en la Figura 4.19.

Regla 9: Con n estaciones de datos, $CW^{min} = n \times 16$

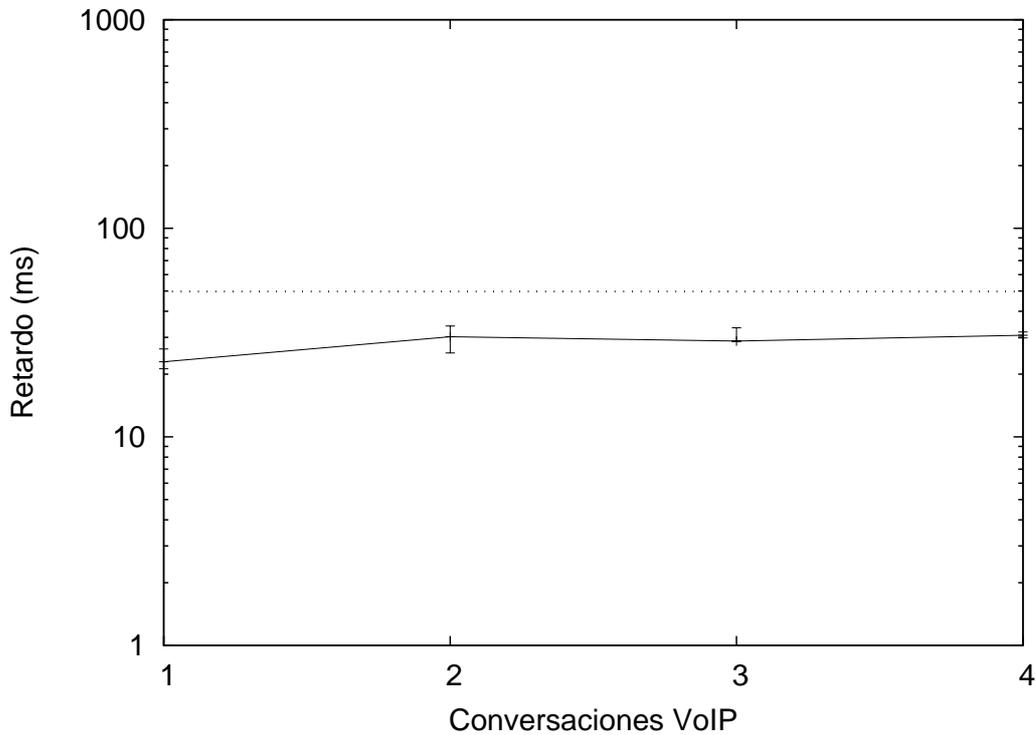


Figura 4.19: Configuración para varias estaciones de datos

4.4.4. Comparación con la configuración estándar

La configuración propuesta se compara, en un último experimento, con los valores de configuración propuestos por el estándar. Para las estaciones y el punto de acceso se escoge los valores de las reglas anteriores en la configuración propuesta y se mide el máximo del retardo frente al número de estaciones en la WLAN (sólo tráfico VoIP). Posteriormente se repite el experimento, pero con el valor de los parámetros propuesto por el estándar para la AC[VO]. Los resultados son los mostrados en la Figura 4.20.

Resulta interesante comprobar que, para un número de conversaciones menor de 9, la configuración del estándar proporciona mejores prestaciones en retardo que la configuración propuesta. Sin embargo, con 10 conversaciones de voz la WLAN se encuentra al límite de su capacidad (no cumpliéndose los requisitos fijados de QoS), disparándose el retardo para 11 conversaciones. La configuración propuesta, en cambio, si bien no presenta unos valores de cota de retardo tan reducidos para n bajos como la configuración estándar de 802.11e, sí que permite dar servicio a un mayor número de conversaciones en la WLAN.

4.5. Resumen

La gran mayoría del trabajo existente hasta la fecha ha consistido en poner de manifiesto el carácter diferenciador de los parámetros del modo EDCA en condiciones de saturación. En este capítulo se ha abordado, para la no saturación, la caracterización del tiempo de servicio en una red EDCA, lo que sirve de base para el desarrollo de mecanismos de

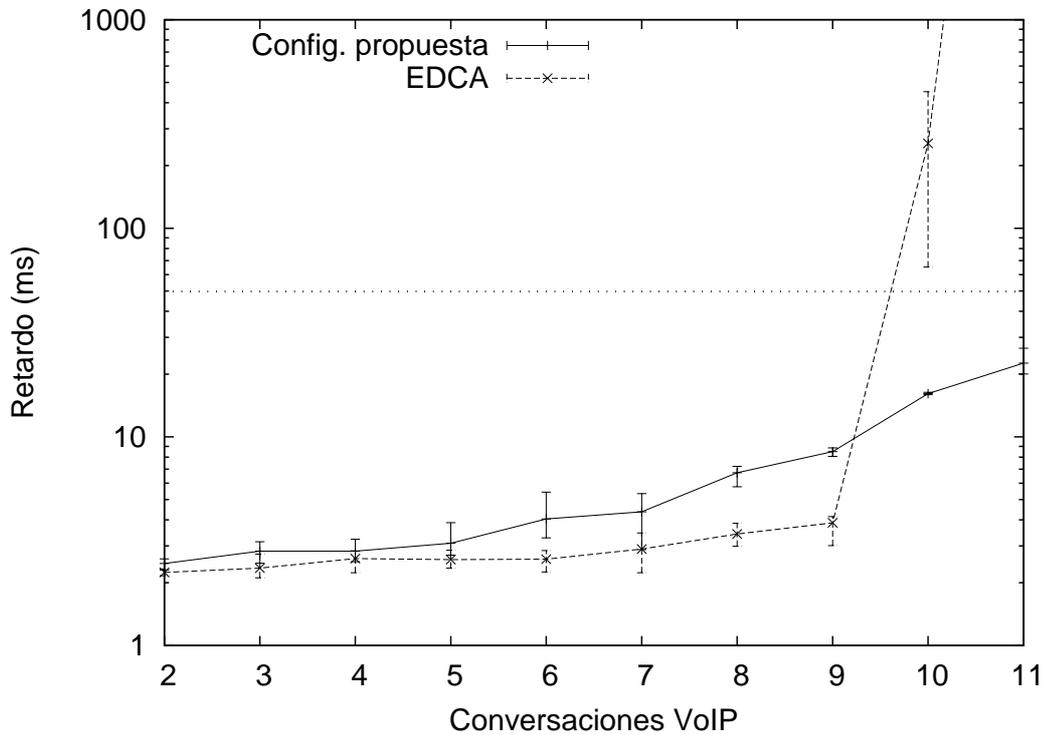


Figura 4.20: Comparación con la configuración propuesta por el estándar

configuración. Dicho análisis ha resuelto la caracterización del tiempo medio de servicio, así como su varianza, e incluso se ha desarrollado un mecanismo para el cálculo de la función de distribución de dicho tiempo. Se ha ilustrado, para un caso de estudio, cómo dicho análisis posibilita el desarrollo de herramientas de configuración y de control de admisión, fijándose un tiempo medio de servicio y *jitter* y comprobándose que la función de distribución de dicho tiempo se ajusta al modelo descrito.

De forma paralela, se ha desarrollado una metodología sobre un entorno real de trabajo para derivar reglas empíricas de configuración de una WLAN, dados unos requisitos y un tráfico de entrada. Si bien en un entorno real las diferencias con los modelos analíticos pueden ser importantes (entre otras causas, por errores en la transmisión, efecto captura, interferencias), y los requisitos que se han establecido para cada caso también son diferentes (en un caso, valores medios del tiempo de servicio de 5 ms; en el otro, percentiles del 98 % para el retardo total de 50 ms), resulta interesante comprobar que los resultados son razonablemente ajustados: para el caso analítico, la CW para 20 flujos de voz se sitúa en un valor próximo a 100; para las medidas experimentales, para 11 conversaciones de VoIP (esto es, 22 flujos) se obtienen valores de $CW = 64$. Realizar el contraste entre los resultados experimentales y el modelado analítico para un mismo escenario constituye un inmediato trabajo futuro.

Capítulo 5

Conclusiones

Tras el innegable éxito de las redes 802.11, y una vez que éstas soportan tasas de transmisión comparables a las tecnologías cableadas, uno de los retos pendientes es el manejo de los mecanismos de diferenciación introducidos por el estándar 802.11e para obtener la mejor configuración de la red. Dicha necesidad nace del hecho de que, a diferencia de lo que sucede en las tecnologías cableadas, la naturaleza compartida del medio obliga a prestar especial atención a la forma en que se reparten los recursos disponibles.

En la presente tesis se ha abordado el uso de las herramientas que ofrece el modo EDCA para proporcionar la *mejor* configuración posible de una WLAN. La valoración de la bondad de una determinada configuración dependerá de los requisitos que se establezcan en el escenario considerado; es por ello que se ha dividido el conjunto de posibles escenarios en dos paradigmas:

WLAN abiertas: Tras la revisión de los análisis existentes, la configuración propuesta coincide, en un primer caso, con la que maximiza el rendimiento de la red, si bien la motivación es radicalmente diferente a la de anteriores trabajos; en este caso, dicha configuración garantiza la ausencia de inestabilidades temporales. También se ha propuesto el uso de los mecanismos EDCA para entornos *multirate*, añadiendo el requisito de suministrar equidad proporcional. Trabajos previos que defienden configuraciones similares no se basan en un desarrollo bien fundamentado como el que se ha presentado. Por lo tanto, partiendo de un nuevo análisis en condiciones de saturación que captura mejor el comportamiento de la WLAN, en la presente tesis se ha profundizado en el funcionamiento y el criterio a manejar en la configuración de una WLAN abierta.

La existencia del modo EDCA, sobre el que se desarrollan los anteriores mecanismos, exige un control sobre el comportamiento de los usuarios; a tal efecto se ha definido un esquema de detección de comportamientos maliciosos. En una época donde muchos usuarios pretenden acceder a los recursos WLAN sin importar su naturaleza (con, incluso, variedad de programas para burlar la protección WEP o el uso de “portales prisioneros” en los aeropuertos), este tipo de mecanismos se convierten en una necesidad.

WLAN cerradas: En el otro extremo se encuentran las redes cerradas, donde el objetivo prioritario es cumplir con los requisitos del servicio que se establecen. En este caso, se

ha avanzado significativamente en la caracterización de la forma en que una red EDCA proporciona el servicio; los valores medios suponen una información no despreciable, pero es con el análisis de la varianza y de la función de distribución cuando resulta posible desarrollar herramientas válidas de provisión de QoS. El desarrollo de dichas herramientas se ha ilustrado para un caso muy de actualidad, como es el uso de aplicaciones de VoIP. Adicionalmente, se ha desarrollado una plataforma de medición de resultados, sobre la que se ha definido una metodología experimental que permite la obtención de la configuración óptima basada en mediciones reales.

Por lo tanto, con esta tesis se proponen algoritmos concretos para la configuración de los mecanismos de EDCA para dos casos de uso bien diferenciados, WLAN abiertas y cerradas, lo que cubre gran parte de los escenarios posibles. Se ha demostrado la idoneidad de dichas configuraciones en los escenarios considerados, habiendo sido validadas mediante experimentación, mostrándose además la mejora frente a la configuración recomendada por el estándar.

Como principal trabajo futuro se pretende abordar el problema de la configuración óptima en un escenario genérico, con una serie de estaciones que generen diferentes modelos de tráfico en cada categoría de acceso, cada una con sus requisitos de calidad de servicio. En la presente tesis se han empleado las herramientas de modelado analítico para derivar la configuración óptima en un escenario VoIP; se plantea partir de tales herramientas, que modelan situaciones menos específicas que el caso VoIP, y obtener guías de configuración para esos entornos. Los resultados de dicha configuración serían validados mediante resultados experimentales, en la línea de lo realizado en la Sección 4.4.

Referencias

- [1] J. W. Robinson y T. S. Randhawa. Saturation Throughput Analysis of IEEE 802.11e Enhanced Distributed Coordination Function. *IEEE Journal on Selected Areas in Communications*, 22(5):917–928, Junio 2004.
- [2] A. Banchs y L. Vollero. A Delay Model for IEEE 802.11e EDCA. *IEEE Communications Letters*, 9(6):508–510, Junio 2005.
- [3] Z. Kong, D. H. K. Tsang, B. Bensaou, y D. Gao. Performance Analysis of IEEE 802.11e contention-based channel access. *IEEE Journal on Selected Areas in Communications*, 22(10):0–0, Diciembre 2004.
- [4] J. Hui y M. Devetsikiotis. A Unified Model for the Performance Analysis of IEEE 802.11 EDCA. *IEEE Transactions on Communications*, 53(9):1498–1510, Septiembre 2005.
- [5] A. Banchs y P. Serrano. Revisiting 802.11e EDCA Performance Analysis. (*enviado*).
- [6] B.-J. Kwak, N.-O. Song, y L. E. Miller. Analysis of the Stability and Performance of Exponential Backoff. En *Proceedings of WCNC'03*, Marzo 2003.
- [7] P. Serrano, A. Banchs, T. Melia, y L. Vollero. Performance Anomalies of Nonoptimally Configured Wireless LANs. En *Proceedings of IEEE WCNC '06*, Las Vegas, Nevada, EEUU, Abril 2006.
- [8] F. P. Kelly, A.K. Maulloo, y D.K.H. Tan. Rate Control in Communication Networks: Shadow Prices, Proportional Fairness and Stability. *Journal of the Operational Research Society*, 49:237–252, 1998.
- [9] A. Banchs, P. Serrano, y H. Oliver. Proportional Fair Throughput Allocation in Multirate 802.11e EDCA Wireless LANs. *Wireless Networks*, (pendiente de publicación) 2006.
- [10] P. Serrano, A. Banchs, y J. F. Kukielka. Detection of Malicious Parameter Configurations in 802.11e EDCA. En *Proceedings of IEEE Globecom 2005*, St. Louis, MI, EEUU, Noviembre 2005.
- [11] M. Raya, J.-P. Hubaux, y I. Aad. DOMINO: A System to Detect Greedy Behavior in IEEE 802.11 Hotspots. En *Proceedings of MobiSys '04*, pp. 84–97, Boston, MA, EEUU, 2004.

- [12] P. Serrano, A. Banchs, y A. Azcorra. A Throughput and Delay Model for IEEE 802.11e EDCA Under Non Saturation. (*enviado*).
- [13] A. Banchs, P. Serrano, y A. Azcorra. End-to-end Delay Analysis and Admission Control in 802.11 DCF WLANs. *Computer Communications*, 29(7), Abril 2006.
- [14] A. Banchs, G. Iannello, P. Serrano, y L. Vollero. A CAPWAP Architecture for Dynamic Configuration of QoS MAC Parameters. En *Proceedings of 1st Workshop on multiMedia Applications over Wireless Networks (MediaWiN)*, Atenas, Grecia, Abril 2006.
- [15] V. J. García. Estudio experimental de VoIP en un escenario 802.11e EDCA. Proyecto Fin de Carrera, Ing. Telecom., Univ. Carlos III de Madrid, Noviembre 2005.
- [16] A. Cuevas, P. Serrano, C. J. Bernardos, J. I. Moreno, J. Jänhert, H.-W. Kim, J. Zhou, D. Gomes, P. Gonçalves, y R. L. Aguiar. Field Evaluation of a 4G "True-IP" Network. En *Proceedings of 13th IST Mobile and Wireless Communications Summit 2004.*, Lyon, Francia, Junio 2004.
- [17] A. Cuevas, P. Serrano, J. I. Moreno, C. J. Bernardos, J. Jänhert, R. L. Aguiar, y V. Marques. Usability and Evaluation of a Deployed 4G Network Prototype. *Journal of Communications and Networks*, 7(2), Junio 2005.
- [18] T. Melia, M. Liebsh, P. Serrano, y A. Banchs. The Daidalos Architecture for Mobility and QoS. En *Proceedings of World Telecommunications Congress*, Budapest, Hungría, Mayo 2006.
- [19] P. Serrano, D. Larrabeiti, M. Urueña, y A. G. Marques. Error-Aware Scheduling and its Effects on Efficiency and Fairness. *IFIP International Federation for Information Processing*, 196, Enero 2006.
- [20] P. Serrano, D. Larrabeiti, y A. León. Tunable Least Served First. *IFIP International Federation for Information Processing*, 196, Enero 2006.
- [21] J. Larocca y R. LaRocca. *802.11 Demystified*. McGraw-Hill, Nueva York, NY, USA, 2002.
- [22] IEEE 802.11e/D6.0 Part 11. *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Medium Access Control (MAC) Enhancements for Quality of Service (QoS), Supplement to IEEE 802.11e Standard*, Noviembre 2005.
- [23] IEEE 802.11. *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications), Standard*, Julio 1997.
- [24] B. O'Hara y A. Petrick. *The IEEE 802.11 Handbook: a designer's companion*. IEEE Press, Nueva York, NY, USA, 1999.
- [25] P. Roshan y J. Leary. *802.11 Wireless LAN Fundamentals*. Cisco Press, Indianapolis, IN, USA, 2004.

- [26] G. Anastasi y L. Lenzi. QoS Provided by the IEEE 802.11 Wireless LAN to Advanced Data Applications: A Simulation Analysis. *Wireless Networks*, 6(2):99–108, 1999.
- [27] G. Bianchi. Performance Analysis of the IEEE 802.11 Distributed Coordination Function. *IEEE Journal on Selected Areas in Communications*, 18(3):535–547, Marzo 2000.
- [28] M. Heusse, F. Rousseau, G. Berger-Sabbatel, y A. Duda. Performance Anomaly of 802.11b. En *Proceedings of INFOCOM '03*, pp. 836–840, 2003.
- [29] Q. Ni. Performance Analysis and Enhancements for IEEE 802.11e Wireless Networks. *IEEE Network*, 19(4), Julio 2005.
- [30] Y. Xiao y J. Rosdahl. Performance Analysis and Enhancement for the Current and Future IEEE 802.11 MAC Protocols. *SIGMOBILE Mob. Comput. Commun. Rev.*, 7(2):6–19, 2003.
- [31] H. Zhu y I. Chlamtac. Performance Analysis for IEEE 802.11e EDCF Service Differentiation. *IEEE Transactions on Wireless Communications*, 4(4):1779–1788, Julio 2005.
- [32] A. Nafaa, A. Ksentini, A. Mehaoua, B. Ishibashi, Y. Iraqi, y R. Boutaba. Sliding Contention Window (SCW): Towards Backoff Range-Based Service Differentiation over IEEE 802.11 Wireless LAN Networks. *IEEE Network*, 19(4), Julio 2005.
- [33] B. Li y R. Battiti. Performance Analysis of An Enhanced IEEE 802.11 Distributed Coordination Function Supporting Service Differentiation. En *Proceedings of QoIS'03*, Estocolmo, Suecia, Octubre 2003.
- [34] A. Banchs, A. Azcorra, C. Garcia, y R. Cuevas. Applications and Challenges of the 802.11e EDCA mechanism: An Experimental Study. *IEEE Network*, 19(4), Julio 2005.
- [35] A. Grilo y M. Nunes. Performance evaluation of IEEE 802.11e. En *Proceedings of PIMRC'02*, Lisboa, Portugal, September 2002.
- [36] A. Lindgren, A. Almquist, y O. Scheln. Quality of Service Schemes for IEEE 802.11 Wireless LANs: An Evaluation. *Mobile Networks and Applications*, 8(3):223–235, 2003.
- [37] N. H. Vaidya, P. Bahl, y S. Gupta. Distributed fair scheduling in a Wireless LAN. En *Proceedings of MOBICOM '00*, pp. 66–75, Boston, MA, EEUU, Agosto 2000.
- [38] J. L. Sobrinho y A. S. Krishnakumar. Real-Time Traffic over the IEEE 802.11 Medium Access Control Layer. *Bell Labs Technical Journal*, 2(1):172–187, 1996.
- [39] J.L. Sobrinho y A.S. Krishnakumar. Quality-of-Service in Ad Hoc Carrier Sense Multiple Access Wireless Networks. *IEEE Journal on Selected Areas in Communications*, 17(8):1353–1368, Agosto 1999.

- [40] D. He y C.Q. Shen. Simulation Study of IEEE 802.11e EDCF. En *Proc. 57th IEEE Semiannual VTC (VTC-Spring)*, pp. 685–689, Jeju, Corea del Sur, Abril 2003.
- [41] S. Mangold, S. Choi, G.R. Hiertz, O. Klein, y B. Walke. Analysis of IEEE 802.11e for QoS Support in Wireless LANs. *IEEE Wireless Communications*, 10(6):40–50, Diciembre 2003.
- [42] Y. Xiao. IEEE 802.11e: QoS provisioning at the MAC layer. *IEEE Wireless Communications*, 11(3):72–79, Junio 2004.
- [43] S. Choi, J. Prado, y S. Shankar. IEEE 802.11e Contention-Based Channel Access (EDCF) Performance Evaluation. En *Proceedings of IEEE ICC'03*, 2003.
- [44] L. Romdhani, Q. Ni, y T. Turletti. Adaptive EDCF: Enhanced Service Differentiation for IEEE 802.11 Wireless Ad Hoc Networks. En *Proceedings of IEEE WCNC '03*, New Orleans, LA, Marzo 2003.
- [45] H. S. Chhaya y S. Gupta. Performance modeling of asynchronous data transfer methods of IEEE 802.11 MAC protocol. *Wireless Networks*, 3(3):217–234, 1997.
- [46] G. Bianchi. IEEE 802.11 saturation throughput analysis. *IEEE Communication Letters*, 3(12):318–320, Diciembre 1998.
- [47] P. Chatzimisios, A. C. Boucouvalas, y V. Vitsas. Packet Delay Analysis of IEEE802.11 MAC Protocol. *IEEE Electronics Letters*, 18(39):1358–1359, 2003.
- [48] F. Cali, M. Conti, y E. Gregori. IEEE 802.11 wireless LAN: Capacity analysis and protocol enhancement. En *Proceedings of INFOCOM '98*, pp. 142–149, 1998.
- [49] F. Cali, M. Conti, y E. Gregori. Dynamic Tuning of the IEEE 802.11 Protocol to Achieve a Theoretical Throughput Limit. *IEEE/ACM Transactions on Networking*, 6(8):785–799, 2000.
- [50] Y. C. Tay y K. C. Chua. A Capacity Analysis for the IEEE 802.11 MAC Protocol. *Wireless Networks*, 7(2):159–171, 2001.
- [51] H. Wu, Y. Peng, K. Long, S. Cheng, y J. Ma. Performance of Reliable Transport Protocol over IEEE 802.11 Wireless LAN: Analisis and Enhancement. En *Proceedings of IEEE INFOCOM'02*, Nueva York, EEUU, Junio 2002.
- [52] A. Banchs, X. Pérez, y D. Qiao. Providing Throughput Guarantees in IEEE 802.11e Wireless LANs. En *Proceedings of the 18th International Teletraffic Congress (ITC18)*, Berlin, Alemania, Septiembre 2003.
- [53] G. Bianchi y I. Tinnirello. Remarks on IEEE 802.11 DCF Performance Analysis. *IEEE Communication Letters*, 9(8):765–767, Agosto 2005.
- [54] E. Ziouva y T. Antonakopoulos. CSMA/CA Performance under High Traffic Conditions: Throughput and Delay Analysis. *Computer Communications Journal*, 1(25):313–32, 2002.

- [55] Y. Xiao. A Simple and Effective Priority Scheme for IEEE 802.11. *IEEE Communications Letters*, 7(2):70–72, Febrero 2003.
- [56] C. H. Foh y J. W. Tantra. Comments on IEEE 802.11 Saturation Throughput Analysis with Freezing of Backoff Counters. *IEEE Communication Letters*, 2(9):130–132, Febrero 2005.
- [57] J. W. Tantra, C. H. Foh, I. Tinnirello, y G. Bianchi. Analysis of the IEEE 802.11e EDCA Under Statistical Traffic. En *Proceedings of ICC'06*, Estambul, Turquía, Junio 2006.
- [58] A. Kumar, E. Altman, D. Miorandi, y M. Goyal. New Insights from a Fixed Point Analysis of Single Cell IEEE 802.11 Wireless LANs. En *Proceedings of INFOCOM '05*, 2005.
- [59] G.-S. Ahn, A. T. Campbell, A. Veres, y L.-H. Sun. Supporting Service Differentiation for Real-Time and Best-Effort Traffic in Stateless Wireless Ad Hoc Networks (SWAN). *IEEE Transactions on Mobile Computing*, 1(3):192–207, Julio-Septiembre 2002.
- [60] A. Zaki y M. El-Hadidi. Modeling the 802.11 Distributed Coordination Function with Heterogenous Finite Load. En *Proceedings of ISCCSP '04*, pp. 535–538, 2004.
- [61] M. Ergen y P. Varaiya. Throughput Analysis and Admission Control for IEEE 802.11a. *MONET*, 10(5):705–716, 2005.
- [62] D. Malone, K. Duffy, y D. J. Leith. Modeling the 802.11 Distributed Coordination Function with Heterogenous Finite Load. En *Proceedings of RAWNET'05*, Riva Del Garda, Italia, Abril 2005.
- [63] C. H. Foh y M. Zukerman. Performance Analysis of the IEEE 802.11 MAC Protocol. En *Proceedings of European Wireless 2002*, Florencia, Italia, Febrero 2002.
- [64] P. Patil y V. Apte. Sizing of IEEE 802.11 Wireless LANs. En *Proc. ACM WMASH '05*, Colonia, Alemania, Septiembre 2005.
- [65] C. H. Foh y M. Zukerman. Performance Evaluation of IEEE 802.11. En *Proceedings of 57th Semiannual VTC (VTC Spring)*, vol. 2, Mayo 2001.
- [66] Y. Ge y J. Hou. An Analytical Model for Service Differentiation in IEEE 802.11. En *Proceedings of ICC'03*, vol. 2, pp. 1157–1162, Anchorage, Alaska, EEUU, Mayo 2003.
- [67] D. Qiao y K. Shin. Achieving Efficient Channel Utilization and Weighted Fairness for Data Communications in IEEE 802.11 WLAN under the DCF. En *Proceedings of IEEE Int'l Workshop on QoS*, pp. 227–236, Miami Beach, Florida, EEUU, Mayo 2002.
- [68] Y. Xiao. Performance Analysis of Priority Schemes for IEEE 802.11 and IEEE 802.11e Wireless LANs. *IEEE Transactions on Wireless Communications*, 4(4):1506–1515, Julio 2005.

- [69] A. Banchs y L. Vulliamy. Throughput Analysis and Optimal Configuration of 802.11e EDCA. *Elsevier Computer Networks*, 50(11):1749–1768, Agosto 2006.
- [70] A. K. Parekh y R. G. Gallager. A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks: the Multiple Node Case. *IEEE/ACM Transactions on Networking*, 2(2):137–150, Abril 1994.
- [71] G. Bianchi y I. Tinnirello. Analysis of Priority Mechanisms Based on Differentiated Inter Frame Spacing in CSMA-CA. En *Proceedings of VTC'03*, vol. 3, pp. 1401–1405, Orlando, Florida, EEUU, Octubre 2003.
- [72] J. Zhao, Z. Guo, Q. Zhang, y W. Zhu. Performance Study of MAC for Service Differentiation in IEEE 802.11. En *Proceedings of IEEE GLOBECOM '02*, Taiwan, Noviembre 2002.
- [73] R. Battiti y B. Li. Supporting Service Differentiation With Enhancements Of The IEEE 802.11 MAC Protocol: Models And Analysis. Technical report, Universidad de Trento, Mayo 2003.
- [74] P. E. Engelstad y O. Østerbø. Non-saturation and Saturation Analysis of IEEE 802.11e EDCA with Starvation Prediction. En *Proceedings of ACM MSWiM '05*, pp. 224–233, Montreal, Quebec, Canadá, 2005. ACM Press.
- [75] D. Gao, J. Cai, y K. N. Ngan. Admission Control in IEEE 802.11e Wireless LANs. *IEEE Network*, 19(4):6–13, Julio/Agosto 2005.
- [76] Y. Xiao y H. Li. Evaluation of Distributed Admission Control for the IEEE 802.11e EDCA. *IEEE Communications Magazine*, 42(9):20–24, 2004.
- [77] Y. Xiao, H. Li, y S. Choi. Protection and Guarantee for Voice and Video Traffic in IEEE 802.11e Wireless LANs. En *Proceedings of INFOCOM '04*, pp. 2152–2162, Hong Kong, Marzo 2003.
- [78] M. Barry, A. T. Campbell, y A. Veres. Distributed Control Algorithms for Service Differentiation in Wireless Packet Networks. En *Proceedings of IEEE INFOCOM '01*, pp. 582–590, Anchorage, Alaska, EEUU, 2001.
- [79] A. Veres, A.T. Campbell, M. Barry, y L.-H. Sun. Service Differentiation in Wireless Packet Networks Using Distributed Control. *IEEE Journal on Selected Areas in Communications*, 19(10), Octubre 2001.
- [80] A. Iera, G. Ruggeri, y D. Tripodi. An Algorithm for Dynamic Priority Assignment in 802.11e WLAN MAC Protocols. En *Proc. of 11th Int'l Conf. Telecommun., LNCS*, vol. 3124/2004, pp. 1267–73, 2004.
- [81] N. Ramos, D. Panigrahi, y S. Dey. Quality of Service Provisioning in 802.11e Networks: Challenges, Approaches, and Future Directions. *IEEE Network*, 19(4), Julio 2005.

- [82] A. Iera, A. Molinaro, G. Ruggeri, y D. Tripodi. Improving QoS and Throughput in Single- and Multihop WLANs through Dynamic Traffic Prioritization. *IEEE Network*, 19(4), Julio 2005.
- [83] D. Gu y J. Zhang. A New Measurement-Based Admission Control Method for IEEE802.11 Wireless Local Area Networks. En *Proceedings of IEEE PIMRC '03*, pp. 2009—2013, Septiembre 2003.
- [84] M. Li, B. Prabhakaran, y S. Sathyamurthy. On Flow Reservation and Admission Control for Distributed Scheduling Strategies in IEEE 802.11 Wireless LAN. En *Proceedings of ACM MSWIM '03*, pp. 108–115, San Diego, CA, USA, 2003.
- [85] L. Zhang y S. Zeadally. HARMONICA: Enhanced QoS Support with Admission Control for IEEE 802.11 Contention-based Access. En *Proceedings of RTAS*, pp. 64–71, Mayo 2004.
- [86] D. Pong y T. Moors. Call Admission Control for IEEE 802.11 Contention Access Mechanism. En *Proc. of IEEE GLOBECOM'03*, pp. 174–178, San Francisco, California, EEUU, Diciembre 2003.
- [87] A. Di Stefano, A. Scaglione, G. Terrazzino, I. Tinnirello, V. Ammirata, L. Scalia, G. Bianchi, y C. Giaconia. On the Fidelity of IEEE 802.11 Commercial Cards. En *Proceedings of First International Conference on Wireless Internet*, pp. 10–17, Julio 2005.
- [88] P. Kyasanur y N. Vaidya. Detection and Handling of MAC Layer Misbehavior in Wireless Networks. En *Proceedings of DSN '03*, pp. 173–182, Junio 2003.
- [89] B. Li y R. Battiti. Analysis of the IEEE 802.11 DCF with Service Differentiation Support in Non-Saturation Conditions. En *Proceedings of QofIS'04*, Septiembre 2004.
- [90] R. Vijayakumar, T. Javidi, y M. Liu. From Saturation to Nonsaturation: A Study on 802.11 Networks. Dept. of Electrical Eng., University of Michigan, Tech. Rep. CSPL-363, 2005.
- [91] H. Pan, S. Sato, y K. Kobayashi. On the Throughput of an IEEE 802.11a Wireless LAN System with Terminals under Heterogeneous Radio Conditions. En *Proceedings of the 18th International Teletraffic Congress (ITC18)*, Berlin, Alemania, Septiembre 2003.
- [92] F. Kelly. Charging and Rate Control for Elastic Traffic. *European Transactions on Telecommunications*, 8(1):33–37, Enero 1997.
- [93] J.W. Roberts y L. Massoulié. Bandwidth Sharing and Admission Control for Elastic Traffic. En *Proc. of ITC Specialist Seminar*, Yokohama, Japón, 1998.
- [94] T. Nandagopal, T. Kim, X. Gao, y V. Bharghavan. Achieving MAC Layer Fairness in Wireless Packet Networks. En *Proc. of ACM Mobicom'00*, Boston, MA, EEUU, Agosto 2000.

- [95] J.-Y. Le Boudec. Rate Adaptation, Congestion Control and Fairness: A Tutorial. Disponible en http://icalwww.epfl.ch/PS_files/LEB3132.pdf (última comprobación: Septiembre 2006).
- [96] G. Bianchi y I. Tinnirello. Kalman Filter Estimation of the Number of Competing Terminals in an IEEE 802.11 Network. En *Proceedings of IEEE INFOCOM'03*, San Francisco, CA, Marzo 2003.
- [97] A. Banchs y X. Pérez. Distributed Weighted Fair Queuing in 802.11 Wireless LAN. En *Proceedings of IEEE ICC'02*, New York City, NY, EEUU, Mayo 2002.
- [98] J. Zhao, Z. Guo, Q. Zhang, y W. Zhu. Distributed MAC Adaptation for WLAN QoS Differentiation. En *Proceedings of IEEE GLOBECOM'03*, San Francisco, CA, Diciembre 2003.
- [99] R. Litjens, F. Roijers, J.L. van den Berg, R.J. Boucherie, y M. Fleuren. Performance Analysis of WLANs: An Integrated Packet/Flow Level Approach. En *Proceedings of the 18th International Teletraffic Congress (ITC18)*, Berlin, Alemania, Septiembre 2003.
- [100] K. Claffy, G. Miller, y K. Thompson. The Nature of the Beast: Recent Traffic Measurements from an Internet Backbone. En *Proceedings of INET'98*, Ginebra, Suiza, Julio 1998.
- [101] H. Zhai y Y. Fang. Performance of Wireless LANs Based on IEEE 802.11 MAC Protocols. En *Proceedings of IEEE PIMRC'03*, Septiembre 2003.
- [102] O. Tickoo y B. Sikdar. Queueing Analysis and Delay Mitigation in IEEE 802.11 Random Access MAC based Wireless Networks. En *Proceedings of IEEE INFOCOM'04*, Hong Kong, China, Marzo 2004.
- [103] A. Gurtov. Technical Issues of Real-Time Network Simulation on Linux. Bachelor of Science thesis, University of Helsinki, Department of Computer Science, Mayo, 1999.
- [104] ITU-T. *One-way Transmission Time, Recommendation G.114*, Febrero 1996.
- [105] G. Bianchi, I. Tinnirello, y L. Scalia. Understanding 802.11e Contention-Based Prioritization Mechanisms and Their Coexistence with Legacy 802.11 Stations. *IEEE Network*, 19(4):28–34, Julio/Agosto 2005.