Universidad de Zaragoza
Departamento de Ingeniería Electrónica y Comunicaciones

Provisión de Calidad de Servicio en Redes Móviles Ad Hoc Basada en el Diseño Cross-layer de Algoritmos de Encaminamiento

Tesis Doctoral

Autor:

María Canales Compés

Codirectores:

Dra. Ángela Hernández Solana

Dr. Antonio Valdovinos Bardají

Zaragoza, 2007

A mis padres y hermanas

"¿Por qué esta magnífica tecnología científica, que ahorra trabajo y nos hace la vida mas fácil, nos aporta tan poca felicidad? La repuesta es ésta, simplemente: porque aún no hemos aprendido a usarla con tino."

Albert Einstein (1879-1955), científico alemán.

Diseño gráfico de la portada principal: Esther Canales Compés

Agradecimientos

Quisiera agradecer a mis directores, Ángela Hernández y Antonio Valdovinos, su inestimable ayuda en la realización de esta tesis y su respaldo personal durante todos estos años.

A la gente de los grupos Wireless@KTH en Estocolmo y ANTLAB en Milán, su buena acogida y colaboración desinteresada.

A mis compañeros de GTC en la Universidad de Zaragoza, su cordialidad y compañerismo, que hacen de estas aulas y laboratorios un agradable lugar de trabajo.

A mis amigos, su compañía y su confianza y esos ratos irrepetibles en los que trabajar perdía parte de su importancia.

A mi familia, gracias por vuestro cariño y constante apoyo y, en especial, José Ramón, gracias por compartir conmigo algo más que largas horas de trabajo en la tesis.

Abstract

The fast growth that mobile communications have experienced in the last few years, with special interest in mobile ad hoc networks, has posed an important scientist and technical development. The self-organizing capacity of these networks, provided by the decentralization of the terminals functionalities, turns them into an attractive technological alternative for new scenarios of application, solving the deployment of networks where the difficulty to establish a net-infrastructure makes it unviable.

Nowadays applications heavily demand the fulfillment of their quality of service requirements. In this distributed environment also characterized by a great variability (terminals have free movement inside the network and they can dynamically appear and disappear) the solution motivates a great research effort, since the complexity to solve the tradeoff between an efficient resource utilization and the required reservation according to the application demands points to an efficient design of distributed admission mechanisms.

The role of routing in a cooperative environment, where the establishment of multi-hop routes is required, is especially relevant in order to define the admission criteria. A routing decision based on a metric that reflects the resources availability allows to identify the system ability to provide the demanded quality of service. However, this measurement highly depends on the resource management performed by the medium access control level, which suggests the collaboration between both layers (cross-layer) as a promising solution.

In this context, the main goal of this thesis is the proposal and evaluation of a quality of service routing based on a cross-layer design supported by the cooperation with the medium access layer. Therefore, it is intended to design a distributed admission control mechanism capable of guaranteeing end to end quality of service provision with an appropriate adaptation to the inherent dynamism of mobile ad hoc networks.

To this purpose, a measurement procedure for estimating the available end to end bandwidth has been integrated in the routing algorithm thus implementing a distributed admission mechanism. The evaluation of the proposal with a simplified propagation model has identified the problems stemmed from the cross-layer design and the distributed nature, leading to the development of specific strategies for intercommunication and resource management. The subsequent analysis has validated the proposal in a multiservice scenario.

The evaluation performed in more realistic conditions, determined by the interference among users and the terminals mobility has revealed new difficulties that demand the adaptation of the designed procedures. In this context, a full readmission mechanism has been proposed, which prioritizes the quality of service guarantee for the admitted connections trying to obtain a representative improvement in terms of grade of service. In addition, the implementation of an adaptive configuration according to the grade of mobility in the environment has proved the advantage of a flexible management capable of guaranteeing a reasonable performance in different scenarios.

Resumen

Dentro del amplio crecimiento que han experimentado en los últimos años las redes de comunicaciones móviles, las redes móviles ad hoc han despertado un especial interés que ha incentivado un importante desarrollo científico-técnico. La capacidad autosuficiente de este tipo de redes, gracias a la descentralización de las capacidades funcionales de los terminales, las convierte en una alternativa tecnológica atractiva para nuevos escenarios de aplicación, resolviendo además el despliegue de redes donde la dificultad en el establecimiento de una infraestructura de red lo hace inviable.

La creciente demanda de provisión de calidad de servicio en este entorno distribuido caracterizado además por su gran variabilidad (los terminales pueden moverse libremente dentro de la red así como aparecer y desaparecer dinámicamente) motiva la dedicación de un gran esfuerzo investigador para resolver el compromiso entre un uso eficiente de los recursos del sistema y su reserva conforme a la demanda de las aplicaciones, encaminado al diseño eficiente de mecanismos de admisión distribuidos.

El papel del encaminamiento en un entorno cooperativo caracterizado por el establecimiento de rutas multisalto es especialmente relevante a la hora de definir los criterios de admisión. Una decisión de encaminamiento basada en una métrica reflejo de la disponibilidad de recursos permite identificar la capacidad del sistema para proporcionar la calidad de servicio demandada. Sin embargo, la dependencia de esta medida de la gestión de recursos realizada en el nivel de acceso al medio sugiere como solución a esta necesidad la colaboración entre ambos niveles (*cross-layer*).

En este contexto, el objetivo fundamental de esta tesis es la propuesta y evaluación de un encaminamiento con calidad de servicio basado en un diseño *cross-layer* sustentado en la cooperación con el nivel de acceso al medio. Se pretende así diseñar un mecanismo de control de admisión distribuido que garantice la provisión de la calidad de servicio extremo a extremo y sea capaz de adaptarse al dinamismo inherente a las redes móviles ad hoc.

Partiendo de esta base, se ha desarrollado un procedimiento de medida del ancho de banda disponible extremo a extremo, integrado en el algoritmo de encaminamiento, que permite la implementación distribuida del mecanismo de admisión. La evaluación de la propuesta mediante un modelo de propagación simplificado ha permitido identificar los problemas derivados del diseño *cross-layer* y la naturaleza distribuida que han motivado el desarrollo de las estrategias específicas de intercomunicación y de gestión de recursos cuyo posterior análisis ha mostrado la validez de la propuesta en un escenario multiservicio.

La evaluación realizada en un escenario más realista de variabilidad condicionado por la interferencia entre usuarios y la movilidad de los terminales ha puesto de manifiesto nuevas dificultades que exigen la flexibilización de los procedimientos diseñados. En este contexto, se ha propuesto un mecanismo completo de readmisión que prioriza la garantía de calidad de servicio de las conexiones admitidas con el objetivo de obtener una mejora representativa en términos de grado de servicio. Además, la implementación de una configuración adaptativa de acuerdo al grado de movilidad del entorno ha evidenciado la ventaja de una gestión más flexible capaz de garantizar unas prestaciones razonables en diversos escenarios de aplicación.

Índice general

Lista de Figuras vi

Lista de Tablas xii

Lista de Acrónimos xv

1. Introducción 1

1.1. Motivación y objetivos de la tesis 1

1.2. Estructura de la tesis 5

2. Estado del Arte 7

2.1. QoS en redes móviles ad hoc 7

2.2. Encaminamiento en redes ad hoc 14

2.3. Encaminamiento con calidad de servicio 26

2.4. Diseño *cross-layer* 33

3. Control de Admisión Distribuido 39

3.1. Conectividad. Modelo de red 41

3.1.1. Modelos de Interferencia 41

3.1.2. Problemas del terminal oculto y el terminal expuesto 42

3.1.3. Modelo de red 45

3.2. Selección de un protocolo MAC eficiente 46

3.3. Encaminamiento con QoS 47

3.3.1. *Ad hoc On-demand Distance Vector* (AODV) 49

3.3.2. Métrica QoS. Cálculo del ancho de banda disponible en la ruta. 57

3.3.2.1. Algoritmo de cálculo del ancho de banda E2E 58

3.3.2.2. Integración en el protocolo de encaminamiento 60

3.3.2.3. Modificaciones y consideraciones adicionales 62
3.3.2.4. Soluciones a los problemas derivados de la reserva *backward* en la búsqueda bajo demanda 63
3.3.2.5. Funcionamiento multicamino 66
3.4. Gestión compartida de los recursos 68
3.5. Evaluación del mecanismo de admisión 78

4. Adaptación a la variabilidad del entorno 99
4.1. Implicaciones del *Physical Model*. Diseño de un MAC eficiente 101
4.1.1. Mantenimiento de reservas fiables 101
4.1.2. Estabilidad del sistema 108
4.2. Procedimiento de monitorización de la QoS y readmisión 113
4.2.1. Reacción ante la pérdida de QoS y readmisión 115
4.2.2. Expulsión de conexiones degradadas 121
4.2.3. Prioridades definidas en el nivel MAC. Criterios de admisión 126
4.3. Evaluación del efecto de la interferencia 128
4.4. Evaluación de la propuesta en condiciones de movilidad 157

5. Conclusiones y líneas futuras de investigación 207
5.1. Conclusiones 207
5.2. Líneas futuras de investigación 213

Apéndices 216

A. ADHOC MAC 217
A.1. Funcionamiento básico del protocolo 217
A.2. Reparto de recursos para datos de control y de usuario 219
A.3. Reserva de recursos PTP. Método de acceso BIAS 221
A.4. Disponibilidad de recursos. Información básica para el control de admisión. . . 224

B. Funciones para el cálculo del ancho de banda 231

C. Modelo de Movilidad 235

***Conclusions* 243**

Bibliografía 249

Lista de Figuras

1.1. Redes móviles ad hoc. Comparación con Wireless LAN basada en infraestructura. 2

2.1. Clasificación común de los algoritmos de encaminamiento [Zou y otros, 2002]. . 19

3.1. Estructura general de una propuesta de encaminamiento *cross-layer*. 40

3.2. Problemas del terminal oculto y el terminal expuesto. 44

3.3. Funcionamiento del protocolo AODV, *Request/Reply cycle*. 51

3.4. Detección de rotura en un enlace, envío de mensaje *RERR*. 55

3.5. Diagrama básico de estados para una entrada de la tabla de rutas. 56

3.6. Reutilización de slots. Terminal oculto. 58

3.7. Ejemplo del cálculo de ancho de banda integrado en el encaminamiento. Algoritmo *BWC-FA*. 63

3.8. Problemas derivados de la selección simultánea del mismo recurso (slot k) en diferentes caminos. 65

3.9. Propuesta *cross-layer* de encaminamiento basada en la integración del AODV y el protocolo ADHOC MAC. 68

3.10. Funcionamiento lógico del sistema de colas de un nodo, atendiendo a los enlaces activos, de acuerdo a los diferentes servicios. 70

3.11. Escenarios de muestra de las topologías simuladas (redes conexas). 30 y 50 nodos en un área de 2 Km^2, y 100 nodos en un área de 5 Km^2. 80

3.12. Tráfico correctamente cursado vs. tráfico ofrecido. AODV, QoSR-BN y QoSR-CL. T.O. CBR de 128 kbps. 30, 50 y 100 nodos. 82

3.13. Probabilidad de bloqueo vs. tráfico ofrecido. AODV, QoSR-BN y QoSR-CL. T.O. CBR de 128 kbps. 30, 50 y 100 nodos. 83

3.14. *Throughput* de conexiones admitidas vs. tráfico ofrecido. T.O. CBR de 128 kbps. 30, 50 y 100 nodos. 84

3.15. Número medio de saltos de las conexiones admitidas. T.O. CBR de 128 kbps. 30, 50 y 100 nodos. 85

3.16. Probabilidad de bloqueo vs. tráfico ofrecido. Comparativa versiones *cross-layer*. T.O. CBR de 64, 128 y 256 kbps. 50 nodos. 88

3.17. Probabilidad de bloqueo vs. tráfico ofrecido. AODV, QoSR-BN y QoSR-CL. T.O. CBR de 64, 128 y 256 kbps. 50 nodos. 89

3.18. *Throughput* de las conexiones admitidas vs. tráfico ofrecido. T.O. CBR de 64, 128 y 256 kbps. 50 nodos. 90

3.19. Tráfico correctamente cursado vs. tráfico ofrecido. AODV, QoSR-BN y QoSR-CL. T.O. CBR de 64, 128 y 256 kbps. 50 nodos. 91

3.20. Tráfico correctamente cursado vs. tráfico ofrecido. Comparativa de escenarios multiservicio. T.O. CBR de 128 kbps (100 % CBR vs. 25, 50 y 75 % BE). 50 nodos. 94

3.21. *Throughput* de las conexiones admitidas vs. tráfico ofrecido total. T.O. CBR de 128 kbps y sesiones BE (25, 50 y 75 %). 50 nodos. 95

3.22. Tráfico BE correctamente cursado vs. tráfico ofrecido. Escenarios multiservicio. T.O. CBR de 128 kbps y 25, 50 y 75 % BE. 50 nodos. 96

3.23. Tráfico cursado CBR y BE para los diversos escenarios multiservicio . T.O. CBR de 128 kbps (3200 kbps) y T.O. BE adicional (30, 56 y 80 % en kbps). 50 nodos. 97

4.1. *Physical Model*. Rango de cobertura variable. 102

4.2. *Physical Model*. Señalización básica del ADHOC MAC insuficiente. 103

4.3. Incremento de la interferencia sobre comunicaciones activas. 104

4.4. Transmisión de señal $BEACON$. 107

4.5. Margen frente a interferencias. 108

4.6. Margen de seguridad frente a oscilaciones de la SIR. Definición de $P_{rx,min}$. . . 112

4.7. Procedimientos de admisión. Conexiones nuevas en el sistema (I) y readmisión tras rotura (II) o pérdida de la QoS (III). 114

4.8. Actualización de los recursos asignados a un flujo *Quality of Service* (QoS) y señalización. 118

4.9. Pérdida del ancho de banda garantizado. Ruta QBE y actualización de números de secuencia. 121

4.10. Diagrama ejemplo. Procedimiento de expulsión. 123

4.11. Reservas de recursos y disponibilidad potencial. 127

4.12. Comparativa *Protocol Model* vs. *Physical Model*. AODV, QoSR-BN y QoSR-CL. T.O. CBR de 128 kbps. 50 nodos. 129

4.13. *Physical Model*. Uso de señales $BEACON$. Prestaciones AODV, QoSR-BN y QoSR-CL. T.O. CBR de 128 kbps. 50 nodos. 131

4.14. Comparativa $P_{rx,min}$. Escenarios de ejemplo. 134

4.15. Comparativa $P_{rx,min}$. Estrategia QoSR-CL. T.O. CBR de 128 kbps. 50 nodos. . 135
4.16. Probabilidad de bloqueo vs. tráfico ofrecido. Comparativa $P_{rx,min}$. AODV, QoSR-BN y QoSR-CL. T.O. CBR de 128 kbps. 50 nodos. 136
4.17. Probabilidad de boqueo *RREQ* - *RREP* en QoSR-CL. Comparativa $P_{rx,min}$. T.O. CBR de 128 kbps. 50 nodos. 137
4.18. *Throughput* de las conexiones admitidas. Comparativa $P_{rx,min}$. T.O. CBR de 128 kbps. 50 nodos. 138
4.19. Tráfico cursado vs. tráfico ofrecido. Comparativa $P_{rx,min}$. AODV, QoSR-BN y QoSR-CL. T.O. CBR de 128 kbps. 50 nodos. 139
4.20. Comparativa $P_{rx,min}$. Estrategia QoSR-CL. T.O. CBR de 128 kbps. 30 nodos. . 140
4.21. Comparativa $P_{rx,min}$. Estrategia QoSR-CL. T.O. CBR de 128 kbps. 100 nodos. 141
4.22. Probabilidad de bloqueo vs. tráfico ofrecido. Evaluación del control de potencia. AODV, QoSR-BN y QoSR-CL. T.O. CBR de 128 kbps. 50 nodos. 143
4.23. *Throughput* de las conexiones admitidas. Evaluación del control de potencia. T.O. CBR de 128 kbps. 50 nodos. 144
4.24. Tráfico cursado vs. tráfico ofrecido. Evaluación del control de potencia. AODV, QoSR-BN y QoSR-CL. T.O. CBR de 128 kbps. 50 nodos. 145
4.25. Rango de detección de señales $BEACON$ de acuerdo al valor del parámetro Δ: margen en dB sobre el umbral de sensado (CS_{th}). 146
4.26. Probabilidad de bloqueo vs. tráfico ofrecido. Señales $BEACON$. AODV, QoSR-BN y QoSR-CL. T.O. CBR de 128 kbps. 50 nodos. 147
4.27. *Throughput* de las conexiones admitidas. Señales $BEACON$. T.O. CBR de 128 kbps. 50 nodos. 148
4.28. Tráfico cursado vs. tráfico ofrecido. Señales $BEACON$. AODV, QoSR-BN y QoSR-CL. T.O. CBR de 128 kbps. 50 nodos. 150
4.29. Grado de servicio vs. tráfico ofrecido. Estrategia QoSR-CL. Señales $BEACON$. T.O. CBR de 128 kbps. 50 nodos. 151
4.30. Evaluación de las prestaciones AODV, QoSR-BN y QoSR-CL. Condiciones requeridas de diseño estable. T.O. CBR de 128 kbps. 30 nodos. 152
4.31. Evaluación de las prestaciones AODV, QoSR-BN y QoSR-CL. Condiciones requeridas de diseño estable. T.O. CBR de 128 kbps. 100 nodos. 153
4.32. Efecto de la demanda de ancho de banda individual en QoSR-CL. Escenario interferente. T.O. CBR de 64, 128 y 256 kbps. 50 nodos. 154
4.33. Evaluación de la propuesta en escenario multiservicio. Conexiones QoS. T.O. 50 % CBR de 128 kbps y 50 % BE. 50 nodos. 155
4.34. Trayectoria de un terminal móvil. Escenarios de ejemplo. 160
4.35. Número de readmisiones vs. tráfico ofrecido. AODV, QoSR-BN y QoSR-CL. T.O. CBR de 128 kbps. 50 nodos. 2, 9 y 20 Km/h. 162

4.36. *Throughput* de conexiones cursadas vs. tráfico ofrecido. Comparativa entre AODV, QoSR-BN y QoSR-CL. Escenarios de 2, 9 y 20 Km/h. T.O. CBR de 128 kbps. 50 nodos. 163

4.37. Tráfico correctamente cursado vs. tráfico ofrecido. AODV, QoSR-BN y QoSR-CL. T.O. CBR de 128 kbps. 50 nodos. 2, 9 y 20 Km/h. 164

4.38. Efecto del parámetro RT sobre las prestaciones del sistema. T.O. CBR de 128 kbps. 50 nodos. 9 Km/h. 165

4.39. Probabilidad de bloqueo vs. tráfico ofrecido. Evaluación del mecanismo de reasignación, QoSR-CL. T.O. CBR de 128 kbps. 50 nodos. 2, 9 y 20 Km/h. 168

4.40. Tráfico correctamente cursado vs. tráfico ofrecido. Evaluación del mecanismo de reasignación, QoSR-CL. T.O. CBR de 128 kbps. 50 nodos. 2, 9 y 20 Km/h. 169

4.41. Pérdidas de la QoS identificadas mediante monitorización vs. tráfico ofrecido. QoSR-CL. T.O. CBR de 128 kbps. 50 nodos. 2, 9 y 20 Km/h. 170

4.42. Prestaciones de la estrategia QoSR-CL. Mecanismo completo de readmisión. T.O. CBR de 128 kbps. 50 nodos. 2 Km/h. 174

4.43. Prestaciones de la estrategia QoSR-CL. Mecanismo completo de readmisión. T.O. CBR de 128 kbps. 50 nodos. 9 Km/h. 175

4.44. Prestaciones de la estrategia QoSR-CL. Mecanismo completo de readmisión. T.O. CBR de 128 kbps. 50 nodos. 20 Km/h. 176

4.45. Análisis del mecanismo de expulsión sobre la estrategia QoSR-CL. T.O. CBR de 128 kbps. Escenarios de 2, 9 y 20 Km/h. 50 nodos. 177

4.46. Comparativa entre monitorización aislada y readmisión completa en QoSR-CL. T.O. CBR de 128 kbps. 50 nodos. 20 Km/h. 178

4.47. Tráfico correctamente cursado vs. tráfico ofrecido. Estrategia QoSR-CL. T.O. CBR de 128 kbps y 200 seg. de duración. 50 nodos. 9 Km/h. 180

4.48. Tráfico CBR correctamente cursado vs. tráfico ofrecido en escenario multiservicio. T.O. 50 % CBR de 128 kbps y 50 % BE. 50 nodos. 2, 9 y 20 Km/h. . . . 182

4.49. Grado de servicio (CBR) vs. tráfico ofrecido en escenario multiservicio. T.O. 50 % CBR de 128 kbps y 50 % BE. 50 nodos. 2, 9 y 20 Km/h. 183

4.50. Tráfico BE cursado vs. tráfico ofrecido en escenario multiservicio. T.O. 50 % CBR de 128 kbps y 50 % de sesiones BE. 50 nodos. 2, 9 y 20 Km/h. 184

4.51. Prestaciones del sistema, QoSR-CL frente a ΔBW. T.O. CBR de 128 kbps. 50 nodos. 2 Km/h. 186

4.52. Tráfico correctamente cursado vs. tráfico ofrecido frente a ΔBW. T.O. CBR de 128 kbps. 50 nodos. 9 y 20 Km/h. 187

4.53. Prestaciones del sistema, QoSR-CL frente a ΔBW. T.O. CBR de 128 kbps. 50 nodos. 9 Km/h. 188

4.54. Prestaciones del sistema, QoSR-CL frente a ΔBW. T.O. CBR de 128 kbps. 50 nodos. 20 Km/h. 189

4.55. Evaluación del parámetro ΔBW. T.O. CBR de 64 kbps. 50 nodos. 9 Km/h. . . 190

4.56. Evaluación del parámetro ΔBW. T.O. CBR de 128 kbps. 50 nodos. 9 Km/h. . 191

4.57. Evaluación del parámetro ΔBW. T.O. CBR de 192 kbps. 50 nodos. 9 Km/h. . 192

4.58. Evaluación del parámetro ΔBW. T.O. CBR de 128 kbps. 50 nodos. 20 Km/h. 193

4.59. Efecto del margen ΔBW sobre la capacidad efectiva: número medio de saltos. T.O. CBR de 64 y 192 kbps. 50 nodos. 20 Km/h. 193

4.60. Evolución temporal del número de vecinos estables de un terminal concreto. Escenarios de 2, 5, 9 y 20 Km/h. 195

4.61. Evolución del número de vecinos y métrica $var_{NB,i}^{filt}$ para un nodo i. Escenario de 2 Km/h. 196

4.62. Variación filtrada del número de vecinos. Valor medio entre nodos. Escenarios de 2, 5, 9 y 20 Km/h. 197

4.63. Variación filtrada del número de vecinos. Comparativa respecto a la constante del filtro, α. Escenario de 9 Km/h 198

4.64. Variación filtrada del número de vecinos. Filtro de promediado: $\alpha = 0{,}99$. 50 nodos. Escenarios de 2, 5, 9 y 20 Km/h. 199

4.65. Ratio de tiempo en modo de funcionamiento "movilidad" vs. tráfico ofrecido. Mecanismo adaptativo de readmisión. T.O. CBR de 128 kbps. 50 nodos. 2, 5, 9 y 20 Km/h. 201

4.66. Tráfico cursado vs. tráfico ofrecido. Mecanismo adaptativo de readmisión en QoSR-CL. T.O. CBR de 128 kbps. 50 nodos. 2, 5, 9 y 20 Km/h. 202

4.67. Grado de servicio vs. tráfico ofrecido. Mecanismo adaptativo de readmisión en QoSR-CL. T.O. CBR de 128 kbps. 50 nodos. 2, 5, 9 y 20 Km/h. 203

4.68. Variación filtrada del número de vecinos. Filtro de promediado: $\alpha = 0{,}99$. Escenario de movilidad variable. 204

4.69. Evaluación del mecanismo de adaptación al entorno. T.O. CBR de 128 kbps. 50 nodos. Escenario de movilidad variable. 205

A.1. ADHOC MAC, estructura básica de trama. 218

A.2. ADHOC MAC, división estática de la trama. 220

A.3. Reserva previa de recursos PTP. Acceso BIAS. 222

C.1. Modelo de movilidad RWP modificado con parámetros R_{max} y α_{max}. 237

C.2. Ejemplo de la trayectoria descrita por un terminal. Comparación entre RWP y RWPM. Velocidad media de 9 Km/h. 238

C.3. Evolución de la duración media de los enlaces para distintas velocidades de desplazamiento con el modelo RWPM. 238

C.4. Distancia media al centro geométrico. Número medio de vecinos. RWP y RWPM (2 y 9 Km/h) y escenario estático uniforme. 50 nodos. 239

C.5. Efecto del modelo de movilidad (RWP y RWPM) sobre el funcionamiento básico del protocolo AODV. T.O. CBR de 128 kbps. 50 nodos. 2 y 9 Km/h. 240

Lista de Tablas

2.1. Provisión de QoS desde una perspectiva de capas 10
2.2. Protocolos de encaminamiento para MANETs 18
2.3. Características más representativas de los protocolos de encaminamiento. . . . 25
2.4. Propuestas de diseño *cross-layer* 35

3.1. Escenarios estáticos de análisis 79

4.1. Procedimiento de Admisión - Readmisión 128
4.2. Escenarios de movilidad. Modelo RWP modificado (RWPM). 159
4.3. Parámetros de simulación. 159
4.4. Mecanismo de monitorización y reasignación. 167
4.5. Mecanismo completo de readmisión 172

A.1. Estado del slot k visto por el nodo i, FI_i^k 225
A.2. $Status - Channel$ del slot k en el terminal i 227
A.3. Disponibilidad de recursos a partir del $Status - Channel$ (SC) 228

Lista de Acrónimos

ABR *Associativity Based Routing*

ADSL *Asymmetric Digital Subscriber Line*

AODV *Ad hoc On-demand Distance Vector*

AOMDV *Ad hoc On-demand Multipath Distance Vector*

AQOR *Ad hoc QoS On-demand Routing*

ART ACTIVE_ROUTE_TIMEOUT

BCH *Broadcast CHannel*

BE *Best Effort*

BIAS *Book In Advance Scheme*

BLER *BLock Error Rate*

BM *Bandwidth Management*

BSC *Base Station Controller*

BSS *Basic Service Set (Bluetooth mode)*

BWC-FA *BandWidth Calculation – Forward Algorithm*

CAC *Call Admission Control*

CBR *Constant Bit Rate*

CBRP *Cluster Based Routing Protocol*

CD *Collision Detection*

CDMA *Code Division Multiple Access*

CEDAR *Core-Extraction Distributed Ad hoc Routing*

CGSR *Clusterhead Gateway Switch Routing*

CRDF	*Cross-layer Route Discovery Framework*
CS	*Carrier Sense*
CSMA/CA	*Carrier Sense Multiple Access/Collision Avoidance*
CTS	*Clear To Send*
DCF	*Distributed Coordination Function*
DiffServ	*Differential Services*
DP	DELETE_PERIOD
DREAM	*Distance Routing Effect Algorithm for Mobility*
dRSVP	*Dynamic Resource Reservation Setup Protocol*
DSDV	*Destination-Sequenced Distance Vector*
DSR	*Dynamic Source Routing*
DST	*Distributed Spanning Tree*
DV	*Distance Vector*
DVR	*Distance Vector Routing*
E2E	*End to End*
ETP	*Exposed Terminal Problem*
FI	*Frame Information*
FORP	*Flow Oriented Routing Protocol*
FQMM	*Flexible QoS Model for MANETs*
FR	*Forward Route*
FSR	*Fisheye State Routing*
GEDIR	*GEographic DIstance Routing*
GoS	*Grade of Service*
GPRS	*General Packet Radio Service*
GPS	*Global Positioning System*
GPSR	*Greedy Perimeter Stateless Routing*
GSM	*Global System for Mobile communications*
GSR	*Global State Routing*

HSR	*Hierarchical State Routing*
HTP	*Hidden Terminal Problem*
IEEE	*Institute of Electrical and Electronics Engineers*
IETF	*Internet Engineering Task Force*
IIR	*Infinite Impulse Response*
INSIGNIA	*In-band SIGNalling QoS framework for mobile ad hoc networks*
IntServ	*Integrated Services Model*
IP	*Internet Protocol*
ITU-T	*International Telecommunication Union-Telecommunication*
LANMAR	*LANdMArk Routing*
LAR	*Location Aided Routing*
LB	*Link Bandwidth*
LMR	*Lightweight Mobile Routing*
LS	*Link State*
LSP	*Link State Packets*
LSR	*Link State Routing*
MAC	*Medium Access Control*
MANET	*Mobile Ad hoc NETwork*
MAODV	*Multicast Ad hoc On-demand Distance Vector*
MPR	*MultiPoint Relay*
NP	*Non-Polynomial*
ns-2	*Network Simulator version 2*
OLSR	*Optimized Link State Routing*
OSI	*Open Systems Interconnection*
PAN	*Personal Area Network*
PB	*Path Bandwidth*
PDT	PATH_DISCOVERY_TIME
PGB	*PiGgy-Backing*

PRDS	*Priority-based Route Discovery Strategy*
PTP	*Point-To-Point*
PTP_F	*Point-To-Point Flag*
$QLOST$	*QoS LOST*
QOLSR	*QoS for OLSR*
QoS	*Quality of Service*
QoSR-BN	*QoS Routing – BottleNeck*
QoSR-CL	*QoS Routing – Cross-Layer*
RA	$ROUTE_ACK$
$RERR$	*Route ERRor*
RFC	*Request For Comments*
RNC	*Radio Network Controller*
RR	*Reverse Route*
$RREP$	*Route REPly*
$RREQ$	*Route REQuest*
RS	*Reception Schedule*
RSVP	*Resource ReSerVation Protocol*
RT	RESERVATION_TIMEOUT
RTC	Red Telefónica Conmutada
RTS	*Request To Send*
RWP	*Random WayPoint mobility model*
RWPM	RWP modificado
SAT	*SATisfiability problem*
SC	$Status - Channel$
SEQM	*SRL-Enabled QoS Model*
SIR	*Signal to Interference Ratio*
SLA	*Service Level Agreement*
SNR	*Signal to Noise Ratio*

SRL	*Supernode-based Reverse Labeling*
SRR	*Slots Ready to Receive*
SRT	*Slots Ready to Transmit*
SSA	*Signal Stability-based Adaptive routing*
STAR	*Source Tree Adaptive Routing*
SWAN	*Stateless Wireless Ad Hoc Networks*
TBRPF	*Topology Broadcast based on Reverse Path Forwarding*
TCP	*Transport Control Protocol*
TDMA	*Time Division Multiple Access*
TETRA	*TErrestrial TRunked RAdio*
TIMELY	*The Illinois Mobile Environments LaboratorY*
T.O.	Tráfico Ofrecido
TORA	*Temporally Ordered Routing Algorithm*
TS	*Transmission Schedule*
TTL	*Time To Live*
UMTS	*Universal Mobile Telecomunications System*
VDIM	*Virtual Device Information Manager*
VF	*Virtual Frame*
WG	*Working Group*
WRP	*Wireless Routing Protocol*
WSN	*Wireless Sensor Networks*
ZHLS	*Zone-based Hierarchical Link State protocol*
ZRP	*Zone Routing Protocol*

Capítulo 1

Introducción

1.1. Motivación y objetivos de la tesis

El enorme desarrollo que han experimentado en los últimos años las redes de comunicaciones móviles y la creciente demanda de nuevas aplicaciones más versátiles cuyos servicios han de ser igualmente proporcionados en dicho contexto ha dado lugar a un gran esfuerzo científico-técnico para determinar los mecanismos más adecuados para el cumplimiento de la calidad de servicio – *Quality of Service* (QoS) – requerida por dichas aplicaciones.

A diferencia de las clásicas redes celulares (UMTS, GPRS, GSM, etc.) y las redes inalámbricas de área local (IEEE 802.11 - modo BSS) donde una infraestructura fija soporta el control y la gestión de la red (estaciones base o puntos de acceso), una red ad hoc – *Mobile Ad hoc NETwork* (MANET) – es una red auto-organizativa, distribuida y capaz de adaptarse a cambios en su topología (términos ingleses *self-creating, self-organizing, self-administering*) (figura 1.1).

La capacidad autosuficiente de las MANETs se consigue gracias a que la inteligencia de la red reside en los nodos o terminales y no está centralizada en un único elemento. Dadas su versatilidad, flexibilidad y facilidad de establecimiento, este tipo de redes representa por una parte una alternativa tecnológica atractiva para nuevos escenarios de aplicación, y por otra, resuelve el despliegue de redes allí donde la dificultad en el establecimiento de una infraestructura de red lo hacía inviable. Su importancia reside, por lo tanto, en su capacidad para proporcionar soporte a nuevos servicios en multitud de escenarios:

- Provisión de servicios de banda ancha en entornos específicos (centros comerciales, aeropuertos).
- Despliegue de redes en entornos de emergencias médicas.
- Aplicación en entorno militar, como por ejemplo, redes desplegadas en un campo de batalla.
- Redes de sensores inalámbricas: *Wireless Sensor Networks* (WSN).
- Redes de área personal: *Personal Area Network* (PAN).
- Acceso inalámbrico a redes de área extensa tanto fijas (RTC, ADSL, etc.) como móviles (TETRA, GPRS, UMTS).
- Aumento de la cobertura y la capacidad de redes celulares UMTS.
- Redes intervehiculares.

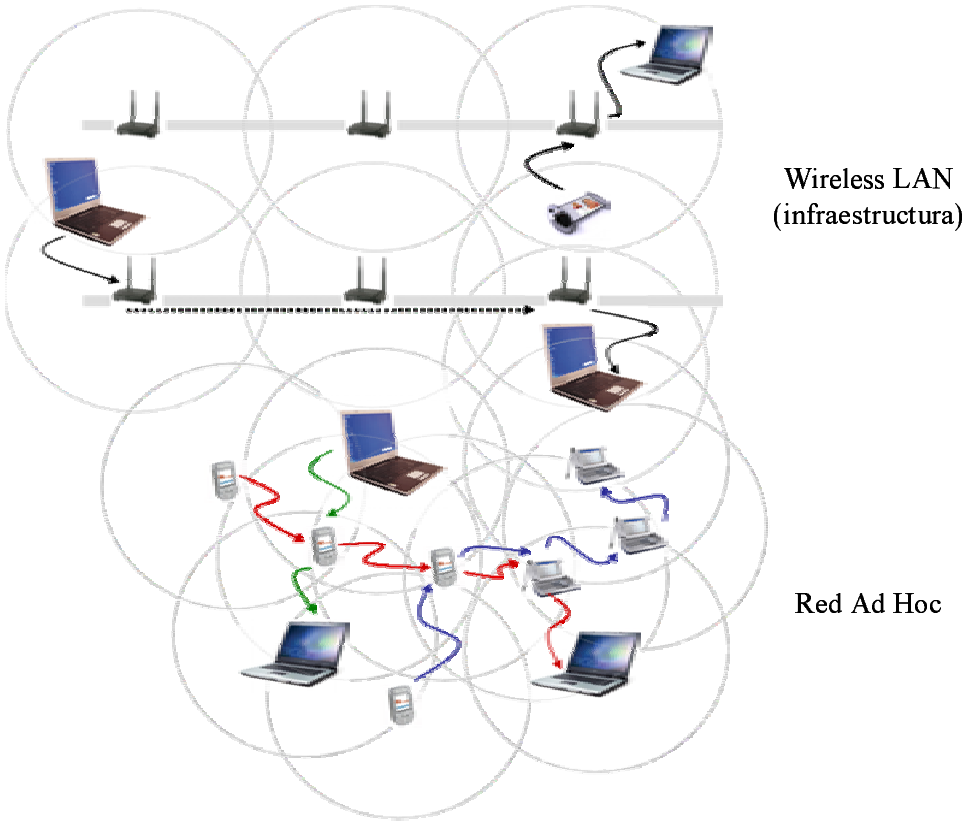


Figura 1.1. Redes móviles ad hoc. Comparación con Wireless LAN basada en infraestructura.

La falta de una infraestructura fija que soporte la gestión de la red exige la cooperación activa entre los diferentes nodos para lograr un funcionamiento distribuido. El alcance limitado de las transmisiones radio implica la necesidad de establecer rutas multisalto para conectar dos nodos cualesquiera de la red. Así pues, la funcionalidad de *host* de los terminales se complementa con la de *router*, retransmitiendo también el tráfico de comunicaciones ajenas, facilitando así un encaminamiento distribuido. De igual modo, el reparto de los recursos de la red entre las diversas comunicaciones no recae en un gestor específico, sino que debe realizarse mediante la colaboración de todos los terminales. La naturaleza compartida del medio de transmisión inalámbrico presenta una dificultad añadida. La eficiencia en su utilización depende en gran medida de la capacidad de los mecanismos de control de acceso al medio – *Medium Access Control* (MAC) – para proporcionar un acceso distribuido al sistema garantizando un elevado reúso de los recursos. Además, la provisión de QoS requiere una asignación diferenciada de los recursos entre las diversas aplicaciones de acuerdo a sus demandas. La solución al compromiso entre eficiencia y provisión de QoS debe basarse en mecanismos de admisión – *Call Admission Control* (CAC) – capaces de adecuarse a la naturaleza distribuida de las redes ad hoc.

Por otra parte, las MANETs se caracterizan por su dinamismo, puesto que los terminales pueden moverse libremente dentro de la red así como aparecer y desaparecer de la misma forma dinámica generando topologías muy variantes, lo que dificulta el establecimiento y mantenimiento de las rutas. Asimismo, garantizar las prestaciones de las comunicaciones activas en un escenario donde tanto la interferencia entre usuarios como la conectividad en sí misma resultan tan variantes todavía complica más el proceso de admisión y la provisión de QoS.

El mecanismo de admisión debe valorar el impacto real de una nueva activación sobre la eficiencia de la red y las prestaciones de las conexiones activas. Además, debe evaluar la posibilidad real de mantener la conexión activa en su duración completa garantizando el cumplimiento de sus demandas. La admisión de una aplicación en un contexto ad hoc no sería posible sin la confirmación previa de la existencia de un camino de conexión viable entre los dos extremos de la comunicación. Cobra así especial relevancia el problema del encaminamiento multisalto. Los protocolos de encaminamiento desarrollados en las redes fijas, donde los nodos permanecen estables, no son adecuados en redes ad hoc, por lo que han surgido multitud de propuestas para implementar algoritmos distribuidos capaces tanto de encontrar rutas como de reaccionar ante la variabilidad inherente a la red.

La mayoría de estos algoritmos se han diseñado sin tener en cuenta explícitamente la calidad de servicio de las rutas encontradas. Sin embargo, la demanda real de las aplicaciones exige el cumplimiento de ciertas restricciones (ancho de banda, retardo mínimo o su variación – *jitter*) a lo largo de la ruta completa, extremo a extremo – *End to End* (E2E). Se demanda, por lo tanto, la inclusión de nuevas métricas de encaminamiento, más allá del número de saltos, capaces de identificar la calidad existente en cada camino facilitando así las decisiones adecuadas por parte de los algoritmos. Por otra parte, puesto que la provisión de la QoS depende fundamentalmente de la gestión de los recursos del sistema, una nueva definición de la métrica de encaminamiento debe facilitar un conocimiento lo más exacto posible de la disponibilidad y viabilidad de estos recursos. Una correcta estimación de estos parámetros depende directamente de la tecnología de nivel físico empleada y la capa MAC implementada, de manera que una operación conjunta (*cross-layer*) se presenta como una de las respuestas más prometedoras a la necesidad planteada.

En este contexto, el objetivo fundamental de esta tesis es el diseño y evaluación de una propuesta de encaminamiento con QoS basada en un diseño *cross-layer* apoyado en la cooperación con el nivel MAC. Se pretende así diseñar un mecanismo de control de admisión distribuido que garantice la provisión de QoS extremo a extremo y sea capaz de adaptarse al entorno variante característico de las MANETs. Los problemas concretos que se plantean en esta tesis son: la caracterización de las principales necesidades tanto a nivel de encaminamiento como a nivel MAC finalmente representadas en la selección específica de los algoritmos; la adecuada definición de las métricas de encaminamiento que permiten identificar correctamente la disponibilidad real de los recursos; el diseño de los mecanismos de cooperación necesarios; el desarrollo de técnicas de adaptación al entorno orientadas a la flexibilización de las estrategias propuestas.

Con objeto de evaluar la arquitectura propuesta, se ha desarrollado una herramienta de simulación capaz de emular el comportamiento de los algoritmos involucrados y los diversos procedimientos desarrollados. Teniendo en cuenta la necesidad de realizar una adaptación consecuente con la naturaleza variante de las redes ad hoc se han tenido en cuenta aspectos como la interferencia entre usuarios, el modelo de propagación o la movilidad de los terminales. Dada la particularización de algoritmos considerada y el necesario diseño de procedimientos de cooperación, se ha buscado un desarrollo más flexible y autónomo respecto a los simuladores de red existentes, por lo que dicha herramienta ha sido implementada como parte del trabajo realizado en la tesis.

1.2. Estructura de la tesis

El contenido de la tesis se estructura del modo siguiente:

El capítulo 2 presenta la revisión del estado del arte especialmente centrada en las dificultades derivadas de la naturaleza de las redes ad hoc, que condicionan la capacidad de proporcionar QoS E2E. Se ha realizado un análisis crítico de los algoritmos de encaminamiento existentes en este entorno con especial énfasis en las propuestas que consideran la QoS y las soluciones *cross-layer* presentadas por otros autores. El estudio comparativo realizado, centrado en la viabilidad de las diferentes propuestas, trata de contextualizar las aportaciones de este trabajo de tesis.

En el capítulo 3 se identifican los requerimientos necesarios para conseguir al objetivo perseguido, definiendo así la selección de los protocolos base de la arquitectura propuesta: *Ad hoc On-demand Distance Vector* (AODV), del que se realiza una descripción detallada, y ADHOC MAC, referenciado al anexo A. En este capítulo se desarrollan la integración entre ambos protocolos y la definición de una métrica de encaminamiento basada en la medida del ancho de banda extremo a extremo que hacen posible un control de admisión distribuido. Una evaluación preliminar en condiciones estáticas sin interferencia entre usuarios permite verificar la viabilidad de la propuesta.

A continuación, el capítulo 4 presenta las limitaciones y dificultades surgidas ante la consideración de un escenario más realista caracterizado por la interferencia entre usuarios y la movilidad de los terminales. Se describe el impacto sobre la eficiencia del ADHOC MAC y las soluciones propuestas para seguir garantizando el soporte fiable de la arquitectura *cross-layer*. Se describen y evalúan las estrategias de adaptación del sistema orientadas a una mayor flexibilización en la utilización de los recursos gracias a la incorporación de procedimientos de monitorización de la QoS y reasignación de recursos. Asimismo, se proponen y analizan mecanismos de readmisión consecuentes con la necesidad de garantizar el mantenimiento de la QoS ante la variación de las condiciones iniciales.

Finalmente, en el capítulo 5 se exponen las principales conclusiones de este trabajo, así como las líneas de investigación futuras que se plantean.

Capítulo 2

Estado del Arte

El objetivo final del presente trabajo consiste en garantizar las prestaciones extremo a extremo demandadas por las aplicaciones gracias a un protocolo eficiente de encaminamiento basado en un diseño *cross-layer* que aproveche la interrelación con las capas inferiores del sistema para realizar una gestión eficiente de los recursos y adaptarse a las condiciones variantes del entorno. Con objeto de situar correctamente el contexto en que se ha desarrollado este estudio, se ha realizado una revisión bibliográfica general del entorno de las redes móviles ad hoc y las dificultades para proporcionar estas prestaciones derivadas de la especial naturaleza de las MANETs. Destaca, dentro de dicha revisión, el análisis crítico de los algoritmos de encaminamiento ya existentes y, especialmente, las propuestas de encaminamiento con QoS. El estudio detallado de las soluciones *cross-layer* presentadas por otros autores, y el análisis de la viabilidad de sus propuestas, pretenden focalizar las aportaciones de este trabajo de tesis.

2.1. QoS en redes móviles ad hoc

El término calidad de servicio – *Quality of Service* (QoS) – hace referencia, según la RFC[1] 2386 [Crawley y otros, 1998], al conjunto de requerimientos del servicio que debe garantizar la red cuando transporta el flujo de paquetes entre la fuente y el destino de la comunicación. Del mismo modo, la recomendación E.800 del *International Telecommunication Union-Telecommunication* (ITU-T) [ITU-T, 0894] define QoS como el efecto colectivo en las prestaciones del servicio que determina el grado de satisfac-

[1] *Request For Comments* (RFC)

ción del usuario final. En definitiva, la calidad de servicio debe ser entendida como la garantía de cumplimiento por parte de la red de ciertos requerimientos cuantitativos (medibles) con objeto de conseguir el efecto cualitativo deseado a nivel de usuario. Esta definición excluye la referencia explícita a los parámetros concretos que caracterizan dicha satisfacción y su traducción en términos cuantificables desde la perspectiva extremo a extremo – E2E – (retardo mínimo, ancho de banda garantizado, etc.), así como los mecanismos necesarios a nivel de red para lograr su cumplimiento (negociado, control de admisión y reserva de recursos, señalización, priorización de paquetes, etc.).

La provisión de QoS depende directamente de los servicios concretos, las topologías específicas de red, etc., por lo que la implementación de protocolos QoS requiere tener en mente el escenario concreto de aplicación. Si bien en redes fijas se han desarrollado muchos trabajos al respecto, éstos no son directamente aplicables en el entorno de las redes móviles ad hoc, especialmente limitado por un lado, por el propio medio de transmisión inalámbrico, y por otro, por el carácter móvil de los terminales.

Así pues, proporcionar un servicio satisfactorio a la demanda actual de aplicaciones con requerimientos cada vez más restrictivos en cuanto a la calidad de servicio y su implantación en el contexto de las MANETs plantea la necesidad de desarrollar propuestas específicas adaptadas a la problemática presentada en este nuevo escenario cuyas principales dificultades pueden resumirse en [Chakrabarti y Mishr, 2001; Mohapatra y otros, 2003; Reddy y otros, 2006]:

- **Limitaciones técnicas en los nodos de la red:** al tratarse de nodos móviles, cobra especial importancia la limitación impuesta por las baterías de los terminales así como las capacidades de almacenamiento y la complejidad asociada al procesado tanto a nivel de *host* como de *router*.

- **Topología variante:** la naturaleza dinámica de las redes móviles ad hoc implica potenciales roturas en conexiones ya establecidas, o incluso la posibilidad de particiones de la red, así como la pérdida de la QoS (si ésta fue garantizada en la admisión del servicio). Se aprecia la gran importancia del diseño de algoritmos de encaminamiento eficaces capaces de reaccionar con rapidez a dicha variabilidad y a su vez adecuados para encontrar caminos suficientemente fiables y con garantías de provisión de la QoS demandada.

- **Información imprecisa en los nodos:** fenómenos como los desvanecimientos de señal, el multicamino o la interferencia, hacen impredecible el estado de los

enlaces y dificultan la medida de la disponibilidad de éstos para proporcionar QoS (ancho de banda, retardo extremo a extremo, *BLock Error Rate (BLER)*).

- **Medio de comunicación inalámbrico:** las transmisiones son propensas a errores debido a la naturaleza variante del canal de transmisión. A su vez, dicho canal es un medio de difusión por su propia naturaleza, lo que dificulta la seguridad de las comunicaciones. Por otro lado, los recursos (la capacidad del canal) son limitados en comparación con la demanda de los mismos por parte de los usuarios (por ejemplo, aplicaciones multimedia con grandes requerimientos de ancho de banda). En definitiva, la gestión eficiente de dicho medio compartido (Control de Admisión, Gestión de Recursos y Control de Acceso al Medio) será uno de los principales retos para garantizar la provisión de QoS.

- **Falta de infraestructura:** la naturaleza autoorganizativa de las redes ad hoc implica una definición más compleja del papel a desempeñar por cada nodo. En último término, todos los nodos pueden adquirir la misma funcionalidad. Así pues, la división usuario-red resulta mucho más difusa, siendo necesaria la coordinación entre todos los nodos a la hora de desarrollar tareas específicas como el control de admisión o el encaminamiento.

Proporcionar calidad de servicio en este nuevo contexto abarca un campo de aplicación muy amplio, desde la perspectiva de todos y cada uno de los niveles de la arquitectura de la red, Tabla 2.1 [Mohapatra y otros, 2003]. Considerando las implicaciones que tiene desarrollar un sistema completo capaz de garantizar la provisión de servicios para cada uno de los usuarios o aplicaciones (lo que se entiende por *QoS Framework*) podemos generalizar dichos estudios desde la visión de cuatro aspectos principales [Wu y Harms, 2001]: Los modelos QoS, la señalización y reserva de recursos QoS, el encaminamiento con QoS y la QoS en el nivel de enlace – MAC. Dicho de otra forma, podremos hablar de QoS si la red es capaz de garantizar una comunicación fiable extremo a extremo entre cualquier par de nodos de la red mediante una gestión eficiente de los recursos que permita una adecuada diferenciación de servicios de acuerdo a las características y demandas propias de cada aplicación.

Tabla 2.1. Provisión de QoS desde una perspectiva de capas

NIVEL	ESTRATEGIAS QOS
Aplicación	Diseño de técnicas adaptativas a nivel de aplicación: Interfaz de usuario flexible y simple (conocimiento de servicios QoS). Establecimiento de rangos dinámicos de QoS (permitir en los niveles inferiores una reserva adaptativa de recursos). Algoritmos adaptativos de compresión, adaptación de tasa, control adaptativo de errores, suavizado de ancho de banda. Codificación de fuente (conjunta red - usuario).
Transporte	Diseño específico de mejoras sobre TCP adaptado a un escenario inalámbrico teniendo en cuenta el entorno MANETs (no existencia de estaciones base). Mecanismos de realimentación para distinguir los errores producidos por el canal variante de la congestión. Gestión de recursos: mecanismos eficientes a nivel de transporte para mejorar la QoS extremo a extremo.
Encaminamiento	Búsqueda selectiva de rutas de acuerdo a restricciones de QoS (ancho de banda, retardo, BLER...). Técnicas multicamino (existencia de rutas alternativas para futuras roturas o balanceo de carga).
Nivel de enlace	Gestión distribuida para resolver los problemas del terminal oculto y el terminal expuesto. Consideraciones específicas de QoS: diferenciación de servicios (prioridades), reparto justo de recursos, etc.
Nivel físico	Estimación eficiente del canal variante de transmisión: requiere realimentación transmisor-receptor. Técnicas adaptativas (codificación de canal, compresión) para mayor protección frente a errores de acuerdo al estado del canal: compromiso ancho de banda ocupado – minimización de errores. Técnicas de control de potencia.

Un **modelo QoS** especifica la arquitectura con la que ciertos servicios pueden ser proporcionados por la red. Desarrollados para su aplicación en Internet, los dos modelos *Integrated Services Model* (*IntServ*) [Braden y otros, 1994] y *Differential Services* (*DiffServ*) [Blake y otros, 1998] no son apropiados para su aplicación directa en el entorno de las MANETs. El primero de ellos consiste en mantener en los *routers* de la red información de estado asociada a cada uno de los flujos (o servicios) activos. Cada nueva conexión es tratada independientemente y el módulo de control de admisión se encarga, mediante señalización basada en el protocolo *Resource ReSerVation Protocol* (RSVP) [Braden y otros, 1997] de verificar la disponibilidad de recursos en el sistema así como de facilitar su reserva. Mediante un sistema de prioridades e identificadores de flujo *(per-flow)*, los paquetes son distribuidos en colas y manejados por el *scheduler* de acuerdo a las mismas. Tanto el *overhead* derivado del RSVP como el volumen de información manejado y la complejidad del sistema a implementar en los terminales dificultan enormemente la escalabilidad del modelo y su aplicación en el entorno ad hoc. *DiffServ* trata de paliar este problema agregando diferentes flujos en un número limitado de clases de servicios *(per-class)*. Asimismo simplifica las tareas a realizar por los terminales, repartiendo las funcionalidades entre tres tipos diferenciados de nodos *(ingress, interior and egress nodes)*. Si bien este esquema simplificado podría aplicarse al entorno de las MANETs, su desarrollo inicial, en el entorno cableado, no considera los problemas intrínsecos del medio inalámbrico: la frontera difusa entre red y usuario dificulta la clasificación funcional de los nodos así como la definición del concepto clásico de negociado (*Service Level Agreement* (SLA)).

En [Xiao y otros, 2000] se propone *Flexible QoS Model for MANETs* (FQMM), un modelo intermedio diseñado explícitamente para adaptarse al entorno ad hoc. Los nodos se clasifican igualmente en los tres tipos definidos en *DiffServ*, pero su papel cambia adaptativamente de acuerdo a la variabilidad de la topología de red y del tráfico. El tratamiento diferenciado a las aplicaciones consiste en un modelo híbrido basado en dar un trato *per-flow* a una cantidad limitada de tráfico – el más prioritario, asumiendo un volumen inferior del mismo – considerando a su vez cierto nivel de agregación para el resto de servicios (*per-class*). Sin embargo, como modelo QoS, ofrece una visión general carente de resolución práctica ante la diversa problemática derivada de la naturaleza de las redes ad hoc.

La **señalización QoS** consiste en el intercambio de información de control entre los agentes del sistema y la correcta interpretación de la misma para permitir la reserva, liberación, mantenimiento o incluso renegociado de los recursos solicitados para

cada servicio. Según el método seguido para la transmisión de dicha información se puede hablar de señalización en banda o fuera de banda. La primera aprovecha los propios paquetes de datos de usuario para enviar los mensajes de control, mientras que la segunda utiliza paquetes específicos de control (que además, pueden "viajar" por caminos de transmisión diferentes a los de los datos de usuario). Uno de los métodos de señalización más comúnmente usados en Internet es el RSVP, pero su utilización en el entorno de las MANETs resulta totalmente inapropiado debido al excesivo *overhead* introducido, como ya se comentó anteriormente, lo que supone mayor competencia por los recursos para los propios datos de usuario, y por la poca capacidad de adaptación de dicho protocolo a entornos muy variantes, al carecer de mecanismos de respuesta rápida ante cambios en la topología de la red. Por el contrario, el modelo propuesto en *In-band SIGNalling QoS framework for mobile ad hoc networks (INSIGNIA)* [Lee y Campbell, 1998] se desarrolló explícitamente para MANETs, basando su funcionamiento en un sistema de señalización en banda que permite establecer, restaurar y adaptar la información de estado de cada flujo de datos determinado por una sesión establecida extremo a extremo en respuesta a cualquier cambio tanto topológico como de la QoS experimentada por el servicio. No obstante, INSIGNIA debe trabajar conjuntamente con los demás bloques del modelo QoS (encaminamiento, *scheduling*, control de admisión, etc.) para garantizar realmente la satisfacción de la QoS demandada. Por otra parte, el mantenimiento de la información de control necesaria en cada uno de los terminales del sistema dificulta la escalabilidad del sistema.

Los **protocolos de encaminamiento con QoS** se ocupan de buscar rutas capaces de satisfacer la QoS demandada por las aplicaciones garantizando la disponibilidad de recursos en las mismas. En este sentido, estos protocolos deben ser capaces de trabajar junto con los mecanismos de gestión de recursos para ser capaces de establecer caminos a través de la red que realmente cumplan con los requisitos de QoS, definidos en último término por nuevas métricas de encaminamiento (métricas QoS) basadas en las restricciones de ancho de banda, retardo, limitaciones energéticas, etc. La implementación de algoritmos eficientes de encaminamiento es la base para conseguir la adaptación del sistema a la variación del entorno (capacidad de reacción, redescubrimiento de rutas...) así como supone, en última instancia, la posibilidad de definir mecanismos de control de admisión distribuidos (la admisión de un nuevo servicio impondrá, en primer lugar, la necesidad fundamental de posibilitar la conexión entre los dos extremos de una comunicación, es decir el establecimiento de una ruta multisalto entre diferentes

nodos intermedios). De este modo, el esfuerzo tecnológico desarrollado en este contexto supone uno de los mayores retos en la provisión de QoS en redes móviles ad hoc.

Finalmente, la **QoS a nivel MAC** es imprescindible a la hora de garantizar realmente la satisfacción de los requerimientos de las aplicaciones. Toda propuesta QoS basada en soportar la funcionalidad necesaria en las capas superiores de la arquitectura (como la señalización o el encaminamiento antes mencionados) presupone la existencia de un protocolo MAC que realmente es el que resuelve el problema de acceder al medio compartido estableciendo las comunicaciones a nivel de enlace y proporcionando la reserva real de los recursos. En este contexto, se han desarrollado diversos protocolos específicos para el entorno inalámbrico, y más concretamente en MANETs. El objetivo principal de dichos protocolos es resolver el acceso compartido al medio de transmisión, ya sea mediante contienda (*Carrier Sense Multiple Access/Collision Avoidance* (CSMA/CA)) o asignación dinámica de recursos (*Time Division Multiple Access* (TDMA), por ejemplo) tratando de maximizar el *throughput* conseguido así como de resolver los problemas del terminal expuesto [Haas y Deng, 2002] y el terminal oculto [Tobagi y Kleinrock, 1975], característicos del entorno ad hoc. En el primer caso, comunicaciones punto a punto paralelas que podrían establecerse, debido al sensado del canal efectuado por ambos transmisores, no se realizan, reduciendo el posible reúso del medio compartido, y los terminales ocultos son aquéllos que no se oyen – no se sensan – y que pueden intentar establecer comunicaciones concurrentes sobre un mismo receptor, vecino común, sobre el que colisionan. Con este propósito se desarrollaron los primeros protocolos MAC adaptados al medio inalámbrico, no proporcionando, sin embargo, reserva de recursos ni soporte para la QoS demandada por servicios en tiempo real (retardo mínimo, ancho de banda garantizado).

Las propuestas iniciales de protocolos de acceso al medio para MANETs (y actualmente las más desarrolladas), se basan en mecanismos de contienda (estándar IEEE 802.11 [IEEE 802.11 Working Group, 1999]). Dichos protocolos emplean competición directa por los recursos y resuelven las colisiones aleatorizando las retransmisiones. La QoS proporcionada, sin embargo, es algo limitada, a pesar de la diferenciación de servicios proporcionada por una de las versiones avanzadas del estándar (IEEE 802.11e *QoS enhancement* [IEEE 802.11 Task Group e, 2005]) dada la ineficiente gestión de recursos debida al funcionamiento asíncrono del protocolo. Por lo tanto, este mecanismo de acceso no resulta adecuado para asegurar unos requerimientos de QoS garantizados para la demanda actual de las aplicaciones. Con objeto de resolver estas limitaciones, han surgido diversas propuestas síncronas [Ahn y otros, 2003; Borgonovo y otros, 2004;

Tang y García-Luna-Aceves, 1999; Zhu y Corson, 2001, 2002] basadas en protocolos de asignación dinámica de recursos (TDMA). Mediante la aplicación de mecanismos distribuidos de asignación de recursos (ranuras temporales: slots), se proporciona una gestión de los mismos acorde con los requerimientos demandados por los usuarios. Generalmente, esta asignación se realiza de manera descentralizada, bajo petición de los terminales. En general, los protocolos MAC TDMA propuestos para MANETs, permiten a un usuario concreto reservar ciertos recursos que le garanticen su QoS. Alguna de las propuestas, [Ahn y otros, 2003], pretende además, mediante la asignación de prioridades dinámicas, regular el acceso de los distintos terminales de acuerdo con sus requisitos de QoS, lo cual puede favorecer la capacidad del sistema final de garantizar realmente la QoS demandada por las aplicaciones.

2.2. Encaminamiento en redes ad hoc

Los protocolos de encaminamiento tradicionalmente usados en las redes fijas ya no son apropiados en el entorno de las MANETs, debido al elevado *overhead* que suponen (que implica un gran consumo de recursos como ancho de banda o energía) así como a los elevados tiempos de convergencia de dichos algoritmos, demasiado grandes en comparación con el dinamismo de este tipo de redes [Zou y otros, 2002]. Con objeto de resolver la problemática inherente a la naturaleza de las redes ad hoc, se ha desarrollado un gran número de algoritmos específicos de encaminamiento. Este desarrollo ha sido principalmente impulsado por grupos de trabajo como el *Internet Engineering Task Force* (IETF) *Working Group* (WG): MANET[2]. Los objetivos que tratan de alcanzarse gracias a los esfuerzos llevados a cabo en este contexto, se basan en cumplir en la mayor medida posible ciertas propiedades cualitativas deseadas para cualquier protocolo de encaminamiento para redes móviles ad hoc, de acuerdo a la RFC 2501 del IETF [Corson y Macker, 1999]:

- Funcionamiento distribuido: se trata de una propiedad esencial, dada la naturaleza cooperativa de este tipo de redes.
- *Loop freedom*: es decir, rutas libres de bucles o lazos indeseados. Dichos lazos pueden formarse como consecuencia de un encaminamiento ineficiente provocando que posteriormente los paquetes se queden indefinidamente (hasta agotar su tiempo de vida) recorriendo dichos lazos sin llegar a su destino final.

[2] http://www.ietf.org/html.charters/manet-charter.html

- Funcionamiento en base a la demanda: en lugar de asumir una distribución uniforme del tráfico en la red (lo que implicaría la necesidad de un mantenimiento de rutas hacia todos los nodos), puede resultar más eficiente la adaptación del algoritmo de encaminamiento a las necesidades reales de tráfico en cada momento, aprovechando mejor tanto los recursos de la red como la propia energía de los terminales. Obviamente, este comportamiento implica cierto retardo de establecimiento de la conexión por la búsqueda inicial de un camino disponible.

- Funcionamiento proactivo[3]: en contraposición con lo anterior, en aquellos casos donde la latencia inicial sea inaceptable, siempre que los recursos energéticos y de ancho de banda lo permitan, la existencia de rutas hacia todos los destinos posibles, antes incluso de que éstas se necesiten, puede ser deseable.

- Control distribuido: Los protocolos de encaminamiento centralizados exigen que un nodo almacene toda la información acerca de la topología de la red con objeto de tomar posteriormete decisiones de encaminamiento óptimas. Esta solución no resulta ni robusta ni escalable.

- Decisiones rápidas: cuanto más rápidamente se tomen las decisiones de encaminamiento, más se reducen tanto la latencia de descubrimiento de ruta como los periodos faltos de ruta debidos a roturas por movilidad (y consecuente redescubrimiento de ruta).

- Reacción local a la variabilidad de la red: decisiones de encaminamiento tomadas localmente en respuesta a variaciones en la red que no afecten a zonas lejanas al lugar de dicho fenómeno reducen el *overhead* introducido aumentando la escalabilidad del protocolo.

- Multiplicidad de rutas: la existencia de rutas alternativas permite minimizar la pérdida de paquetes ante roturas de rutas ya establecidas.

- Robustez frente a la falta de bidireccionalidad: La mayor parte de los algoritmos asume la existencia de enlaces bidireccionales a la hora de establecer una ruta entre dos extremos. Por el contrario, en un entorno inalámbrico, dada la variabilidad del canal de transmisión, la conectividad en un sentido no implica necesariamente su recíproco. Sin embargo, el establecimiento de una ruta inversa

[3]La clasificación principal de los algoritmos de encaminamiento se realiza en función de la naturaleza proactiva o bajo demanda de los mismos. En la página 17 se explica concretamente en qué consiste dicho funcionamiento.

entre el destino y la fuente de la comunicación no es siempre necesario, de manera que limitar la posibilidad de encontrar una ruta unidireccional (fuente – destino) debido a la inexistencia de enlaces bidireccionales restringe la capacidad potencial del encaminamiento. Si bien normalmente existe un número suficiente de enlaces bidireccionales como para establecer el camino requerido, resulta un valor añadido considerar la existencia de enlaces unidireccionales en el desarrollo de algoritmos de encaminamiento eficientes en cualquier escenario de aplicación.

- Eficiencia en potencia: Dada la naturaleza particular de las redes ad hoc, donde la energía (baterías de los terminales) es un factor limitante, los algoritmos de encaminamiento deben ser eficientes en términos energéticos (distribución de la carga en la red, minimización de paquetes de control...) para evitar el apagado de los terminales y consecuentes particiones de la red. Asimismo, dada la necesidad de conservación de la energía, resulta conveniente que los algoritmos de encaminamiento sean capaces de acomodarse a periodos de inactividad de los terminales (estado "latente" – *sleep* – para no realizar gastos energéticos, ni siquiera permanecer en estado de recepción). Con este objetivo, resulta especialmente necesaria la operación conjunta entre el nivel de red y el nivel de enlace.

- Seguridad: debido a la propia inseguridad del canal de transmisión inalámbrico, y puesto que la principal característica de las redes ad hoc es su naturaleza distribuida y la cooperación entre nodos, resulta extremadamente crucial garantizar la seguridad de las comunicaciones, ya no sólo de datos, sino del propio encaminamiento, mediante mecanismos apropiados de autentificación, no-repudio y cifrado para evitar decisiones erróneas o "malintencionadas" de encaminamiento.

- Calidad de servicio o *QoS awareness*: un protocolo de encaminamiento eficiente debería tener en cuenta la calidad de servicio disponible en cada posible camino fuente-destino a la hora de determinar la ruta adecuada para cada flujo específico de datos de acuerdo a sus requerimientos particulares de QoS (ancho de banda, retardo, duración de la ruta estable...)

En base a las citadas propiedades, se pueden determinar las características que clasifican a cada una de las propuestas existentes en la literatura. Mediante dicha clasificación [Belding-Royer y Toh, 1999; Feeney, 1999; Lang, 2003; Perkins y Belding-Royer, 2001; Zou y otros, 2002] se pretende dar una visión global de todos los protocolos básicos

de encaminamiento existentes (tabla 2.2[4]), diferenciando así las propiedades principales de cada protocolo específico, con objeto de determinar las pautas de desarrollo y posibles contribuciones del presente trabajo.

Encaminamiento proactivo vs. encaminamiento bajo demanda.

Los protocolos proactivos, también llamados *table-driven*, se basan en el completo conocimiento de la topología de la red adquirido gracias a la propagación periódica de información relativa a la misma entre los terminales que cooperan en dicha red. Asimismo, ante repentinos cambios en la configuración de la red, nuevamente la información de estado es transmitida eventualmente. De este modo, cada nodo mantiene varias tablas de encaminamiento - estado (de ahí su nombre en inglés) que le permiten disponer en todo momento de información actualizada de las posibles rutas a cada potencial destino dentro de la red cuando éstas sean requeridas. La existencia de dicha información de encaminamiento, antes incluso de que ésta sea necesitada, hace despreciable la latencia de descubrimiento de rutas (reducción muy significativa cuando este retardo se produce una vez que la ruta ya ha sido activada ante una rotura por movilidad, por ejemplo). En contraposición, el hecho de mantener continuamente la información de encaminamiento actualizada supone el envío de mensajes de control en exceso sobre la red (compitiendo por los recursos con los datos útiles de usuarios), así como un almacenamiento elevado de información (incluso puede darse el caso de almacenar rutas hacia destinos que no son nunca realmente utilizadas). Todo ello dificulta la escalabilidad de este tipo de protocolos.

Los protocolos reactivos, se basan en una operación bajo demanda, de manera que únicamente cuando es necesario establecer una nueva comunicación multisalto, y no existe ya un camino establecido, la fuente inicia un proceso de búsqueda de ruta. La respuesta ante esa nueva conexión no es inmediata, debido a la latencia de búsqueda, pero el *overhead* de control se reduce al limitarse temporalmente a los periodos de búsqueda y reacción ante errores. Existen mecanismos de mantenimiento que conservan la ruta activa mientras ésta es necesaria reaccionando ante cambios en la topología de la red (detección de roturas y reconstrucción de las rutas).

[4]Existen multitud de protocolos de encaminamiento desarrollados para redes móviles ad hoc. En la revisión realizada tan sólo se muestran algunos de los más importantes, con objeto de dar una visión global del estado del arte en este contexto, prestando especial atención a las características diferenciadoras de todos ellos, sin tratar de describir concretamente el funcionamiento o las características particulares de cada uno de ellos. Para más detalles, pueden consultarse las referencias [Belding-Royer y Toh, 1999; Feeney, 1999; Lang, 2003; Perkins y Belding-Royer, 2001; Zou y otros, 2002].

Tabla 2.2. Protocolos de encaminamiento para MANETs

PROACTIVOS o TABLE - DRIVEN	
DSDV	*Destination-Sequenced Distance Vector*
WRP	*Wireless Routing Protocol*
CGSR	*Clusterhead Gateway Switch Routing*
GSR	*Global State Routing*
STAR	*Source Tree Adaptive Routing*
FSR	*Fisheye State Routing*
DREAM	*Distance Routing Effect Algorithm for Mobility*
OLSR	*Optimized Link State Routing* (**RFC 3626**)
TBRPF	*Topology Broadcast based on Reverse Path Forwarding* (**RFC 3684**)
LANMAR	*LANdMArk Routing*
GPSR	*Greedy Perimeter Stateless Routing*
REACTIVOS o ON - DEMAND	
AODV	*Ad hoc On-demand Distance Vector* (**RFC 3561**)
DSR	*Dynamic Source Routing* (**reemitido para RFC**)
TORA	*Temporally Ordered Routing Algorithm*
ABR	*Associativity Based Routing*
SSA	*Signal Stability-based Adaptive routing*
LAR	*Location Aided Routing*
FORP	*Flow Oriented Routing Protocol*
CEDAR	*Core-Extraction Distributed Ad hoc Routing*
GEDIR	*GEographic DIstance Routing*
CBRP	*Cluster Based Routing Protocol*
LMR	*Lightweight Mobile Routing*
HÍBRIDOS	
ZRP	*Zone Routing Protocol*
HSR	*Hierarchical State Routing*
ZHLS	*Zone-based Hierarchical Link State protocol*
DST	*Distributed Spanning Tree*

Las soluciones híbridas tratan de incrementar la escalabilidad de los protocolos mediante la combinación de cierta funcionalidad como protocolos reactivos así como proactivos. Los nodos cercanos comparten información de manera proactiva manteniendo continuamente las rutas entre ellos reduciendo el *overhead* que se produciría en el momento de buscarlas así como la latencia de este proceso. Por otro lado, en el caso de necesitar rutas hacia nodos más lejanos, tendría lugar el procedimiento de búsqueda bajo demanda.

La figura 2.1 muestra uno de los esquemas más comúnmente utilizados para clasificar los protocolos de encaminamiento, de acuerdo precisamente a su naturaleza reactiva, proactiva o híbrida. Las flechas continuas indican una descendencia directa y las discontinuas, una descendencia lógica. Las flechas punteadas relacionan protocolos que son resultado de la combinación de diversos antecedentes.

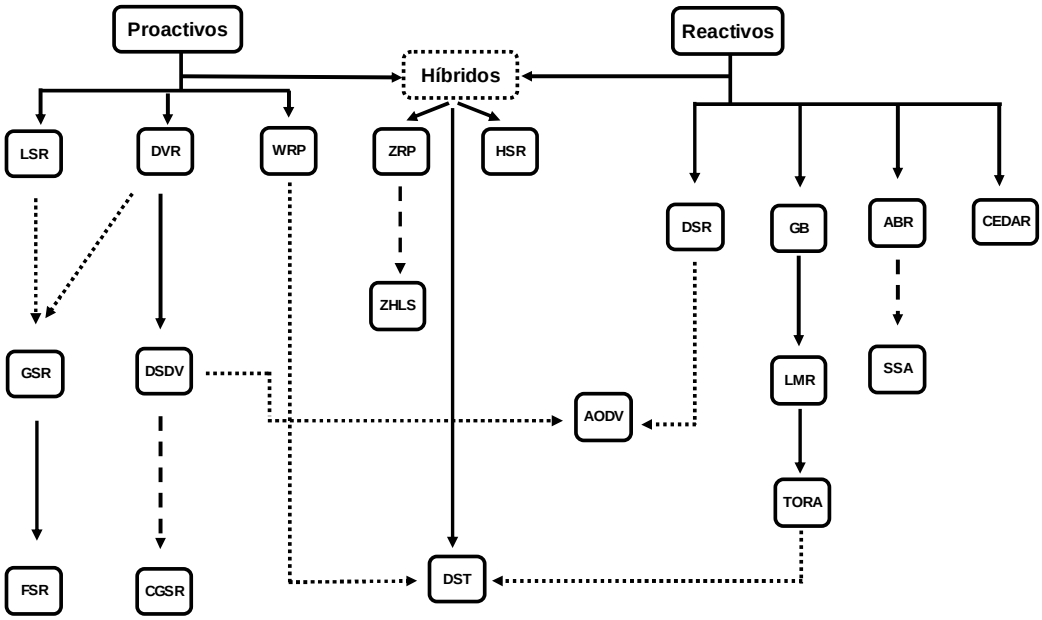


Figura 2.1. Clasificación común de los algoritmos de encaminamiento [Zou y otros, 2002].

Link State Routing (LSR)* vs. *Distance Vector Routing (DVR):

En el encaminamiento basado en la métrica *Link State* (LS) los nodos se intercambian información de estado relativa a sus vecinos en paquetes de control periódicos – *Link State Packets* (LSP) – y dichos paquetes son propagados por inundación (*flooding*) al resto de la red. De esta forma cada nodo puede construir y mantener la topología global de la red como base para sus decisiones de encaminamiento hacia cada uno de los

nodos. El problema que presenta esta estrategia es el excesivo *overhead* de control, especialmente en escenarios de gran movilidad, donde la información necesita propagarse con mayor periodicidad para permitir una rápida reacción a los posibles cambios. Los protocolos proactivos OLSR [Clausen y Jacquet, 2003; Jacquet y otros, 2005], TBRPF [Ogier y otros, 2004] o STAR [García-Luna-Aceves y Spohn, 1999a,b] se basan en este funcionamiento.

El uso de una métrica *Distance Vector* (DV), sin embargo, se basa en que cada nodo mantenga un denominado vector distancia con información del siguiente salto y la distancia (número de saltos) hacia cada posible destino (dicha distancia podría representar por otro lado cierto coste de la ruta, medida no únicamente en términos de saltos). Cada nodo actualiza la información en función de los vectores distancia recibidos por sus vecinos. La ruta se computa de manera distribuida, combinando el siguiente salto de cada nodo intermedio del camino. El problema de este tipo de protocolos es la lentitud de convergencia y la tendencia a crear lazos *(routing loops)*. DSDV [Perkins y Bhagwat, 1994] (protocolo proactivo) resuelve el problema de los posibles lazos introduciendo el concepto de números de secuencia. AODV [Perkins y otros, 2003] es una adaptación del DSDV a un funcionamiento reactivo.

Actualizaciones periódicas vs. actualizaciones generadas ante eventos específicos

Las actualizaciones periódicas utilizadas por los protocolos proactivos facilitan un conocimiento global de la topología de red por parte de todos los nodos, manteniendo así la estabilidad de la red de una manera simple. Sin embargo, si el periodo de actualización es considerablemente mayor que la variabilidad de la red (movilidad), la información no es suficientemente fiable a la hora de tomar decisiones a partir de ella y la capacidad de reacción se degrada. Por otro lado, si las actualizaciones son demasiado frecuentes, el elevado intercambio de paquetes entre los nodos provoca un consumo excesivo de los recursos de la red.

Las actualizaciones generadas únicamente como resultado de variaciones en la red (rotura de enlaces establecidos, aparición de nuevos vecinos...) permiten una reacción casi instantánea al dinamismo de la red minimizando el overhead introducido. No obstante, incluso en este caso, la generación de paquetes puede resultar excesiva cuando la movilidad de la red introduce demasiados eventos a los que reaccionar.

Los protocolos reactivos, por su operación bajo demanda, no necesitan del intercambio periódico de información. No obstante, algunos de ellos utilizan mensajes periódicos

transmitidos en modo *broadcast* a todos sus vecinos como mecanismo de gestión de la conectividad (por ejemplo, los mensajes HELLO del AODV). Dicho intercambio es innecesario si las capas inferiores (*link layer*) son capaces de realizar una indicación eficiente de la pérdida de los enlaces activos (por ejemplo, la falta de reconocimientos de paquetes transmitidos). Sin embargo, la existencia de ese intercambio periódico de información puede facilitar otro tipo de tareas, como la estimación de la calidad de los enlaces mediante medidas de relación señal a interferencia – *Signal to Interference Ratio (SIR)*, estabilidad, etc.

En cualquier caso, dadas las ventajas e inconvenientes de cada una de las estrategias, la mejor opción consistirá en una solución combinada, adaptada al escenario de aplicación.

Protocolos uniformes vs. no uniformes

Los protocolos de encaminamiento pueden diferenciarse de acuerdo al papel o funcionalidad de los nodos al involucrarse en las tareas de encaminamiento. Cuando cualquier nodo adquiere la misma responsabilidad en dichas tareas se considera una estructura plana (protocolos uniformes). Dicho funcionamiento, si bien es mucho más simple, puede resultar ineficiente en redes con un elevado número de nodos, ya que la información de encaminamiento a tratar, mantener y difundir puede resultar excesiva (con el consecuente consumo de recursos de la red).

Los protocolos no uniformes, sin embargo, diferencian la funcionalidad de cada uno de los nodos, incluso limitando sólo a ciertos terminales la posibilidad de tomar parte en el encaminamiento, de modo que se reduce considerablemente el *overhead* de control, así como se incrementa la escalabilidad. La gran mayoría de estos protocolos tiene una estructura jerárquica, basándose en *clusters* o particiones de la red, que emulan el comportamiento de las clásicas redes celulares, estableciendo como nodos responsables en las tareas de encaminamiento únicamente a un grupo reducido de nodos (los *cluster-heads*). La organización en *clusters* se realiza de manera dinámica, y de acuerdo a la variabilidad de la red, de manera que dicha organización permite mantener una topología de red relativamente estable limitando el dinamismo al interior de los *clusters*. Dada la limitada dimensión de los *clusters*, el encaminamiento entre los nodos internos (*intracluster*) puede realizarse de una manera proactiva y eficiente, mientras que únicamente entre diferentes *clusters* (*intercluster*) se realiza un encaminamiento bajo demanda (o híbrido). ZHLS [Joa-Ng y Lu, 1999], CEDAR [Sivakumar y otros, 1999], CBRP [Jiang y otros, 1999], CGSR [Chiang y otros, 1997] o HSR [Iwata y otros,

1999; O'Neil y Li, 2000] son ejemplos de protocolos jerárquicos. El protocolo LANMAR [Gerla y otros, 2001a; Pei y otros, 2000], resultado de la combinación de FSR y el encaminamiento Landmark [Tsuchiya, 1988], determina la formación de subgrupos de nodos de acuerdo a su patrón de movilidad, nodos que se mueven juntos forman un *cluster* encabezado por un nodo *landmark*.

Otro subconjunto de protocolos no uniformes lo componen aquéllos que realizan una selección de vecinos sin estructurarse jerárquicamente, de manera que la funcionalidad de los terminales de la red resulta diferenciada pero no existen particiones de la red establecidas como *clusters*. Cada nodo determina un subconjunto de vecinos que serán los que tomarán un papel diferenciado en la selección de la ruta y/o en el encaminamiento del tráfico. Dicha selección se realiza de manera individual, no coordinada entre los nodos, por lo que las prestaciones del protocolo son muy dependientes de cómo se haga realmente la misma. El funcionamiento del protocolo OLSR, por ejemplo, se basa en la selección de ciertos nodos específicos – *MultiPoint Relay (MPR)* – para desempeñar las tareas de encaminamiento. Un caso similar es el del protocolo ZRP [Haas y otros, 2002; Iwata y otros, 1999].

Computación centralizada vs. distribuida

De acuerdo al cálculo de la ruta los protocolos pueden clasificarse en aquéllos de funcionamiento centralizado y los de funcionamiento distribuido. Los primeros se basan en un mantenimiento de información global en los nodos, base para dicho cálculo centralizado (por ejemplo, mediante un algoritmo como Dijkstra). Los protocolos LSR se basan en este concepto.

En el caso de un funcionamiento distribuido, los nodos mantienen únicamente información parcial y local, por lo que se necesita la colaboración de todos para determinar la ruta completa entre la fuente y el destino. Los protocolos reactivos basados en una métrica DV son un claro ejemplo de este funcionamiento.

Encaminamiento fuente o salto por salto

Algunos protocolos de encaminamiento computan la ruta extremo a extremo en el nodo fuente, quien se encarga de propagar dicha información en la propia cabecera de los paquetes, de manera que cada nodo intermedio la utilice para realizar el encaminamiento (*source-routing*). En este caso, dichos nodos no mantienen ninguna información actualizada de encaminamiento, puesto que el reenvío de paquetes se realiza gracias a la información explícita en los mismos. De hecho, esta estrategia, junto con una operación bajo demanda, elimina la necesidad de actualizaciones periódicas y el almacenamiento

en los nodos. Sin embargo, la escalabilidad se reduce, dado el excesivo *overhead* introducido (rutas cada vez más largas acumulan mayor información en las cabeceras de los paquetes). Es el caso, por ejemplo, del algoritmo reactivo DSR, el protocolo CEDAR o el ABR [Toh, 1997].

Por otro lado, en los algoritmos basados en un encaminamiento salto por salto (*hop-by-hop routing*), comúnmente basados en un procedimiento distribuido (DVR bajo demanda), la ruta se distribuye entre todos los nodos implicados. El reenvío de paquetes se realiza de acuerdo a la información local de cada nodo (siguiente salto al destino). La problemática de este tipo de protocolos reside en la necesidad de implementar algoritmos eficientes capaces de detectar y evitar la formación de lazos. AODV, SSA [Dube y otros, 1997], TORA [Park y Corson, 2001, 1997], DSDV o CGSR son ejemplos de algoritmos distribuidos basados en métrica DV.

Un único camino vs. multicamino

La mayoría de los algoritmos de encaminamiento se basan en la búsqueda de una única ruta entre la fuente y el destino de la comunicación. El mantenimiento continuo de la misma se realiza en función de los mecanismos específicos del protocolo para detectar y responder a posibles roturas. Sin embargo, existen diversas propuestas en las que la búsqueda de la ruta no se limita a un único camino, lo que potencialmente ofrece numerosas ventajas: permite la disponibilidad inmediata de rutas alternativas ante roturas, posibilita el balanceo de carga o incluso la transmisión redundante (aumento de la fiabilidad) y permite la selección de "mejores" rutas en base a los criterios de calidad seleccionados. Sin embargo, la complejidad de dichos algoritmos reside en el mantenimiento fiable de todas las rutas conjuntamente, puesto que la variabilidad de la red afecta a mayor número de caminos que, en caso de no ser efectivamente usados (caso de rutas "en reserva") pueden quedar totalmente obsoletos en el momento de ser seleccionados. Protocolos como el DSR [Johnson y otros, 2004; Johnson y Maltz, 1996], GSR [Chen y Gerla, 1998], FSR [Gerla y otros, 2001b; Iwata y otros, 1999], la versión multicamino del AODV (*Ad hoc On-demand Multipath Distance Vector* (AOMDV) [Marina y Das, 2001, 2002, 2006]), o el TORA, proporcionan múltiples rutas para un mismo destino.

Protocolos basados en posición

Existen, por otra parte, los llamados algoritmos de encaminamiento basados en posición, cuyo funcionamiento es completamente diferente. La clave es el mantenimiento constante de información actualizada de posición. Los nodos deben conocer su propia

posición y la del destino de la comunicación (gracias a un GPS o extrayéndola de otro servicio independiente de localización). Una manera de mantener información actualizada de las diversas posiciones de los nodos es mediante el intercambio periódico de dicha información entre los diferentes nodos de la red (mediante *broadcast* a los vecinos, o mediante *piggybacking* en los paquetes de datos). Estos esquemas eliminan el establecimiento y mantenimiento de rutas puesto que la información es enviada automáticamente, al tiempo que se selecciona el camino a seguir, en cada nodo, de acuerdo a la información de la posición, incluida en los propios paquetes de datos. Estas estrategias no se definen realmente como de encaminamiento, sino como de "reenvío" (*packet-forwarding*). Ejemplos de este planteamiento son los protocolos DREAM [Basagni y otros, 1998], GEDIR [Stojmenovic y Lin, 1999] o GPSR [Karp y Kung, 2000]. Sin embargo, existen algunos algoritmos de encaminamiento propiamente dichos que si bien no se basan en *packet-forwarding* sino en el mantenimiento real de tablas de encaminamiento, mejoran el descubrimiento de la ruta, utilizando la información de la posición de una manera análoga a la descrita para los algoritmos mencionados. El protocolo LAR [Ko y Vaidya, 1998] o el protocolo FORP [Su y Gerla, 1999] son dos ejemplos de protocolos de encaminamiento reactivos que aprovechan cierta información de posición.

Como resumen de los aspectos más importantes descritos a lo largo de este apartado, la tabla 2.3 muestra las características más representativas de los algoritmos de encaminamiento para redes ad hoc, clasificadas según su funcionamiento proactivo, reactivo o híbrido.

Tabla 2.3. Características más representativas de los protocolos de encaminamiento.

Característica	**PROACTIVOS**	**REACTIVOS**	**HÍBRIDOS**
Latencia	ninguna	considerable	depende de la posición
Estructura	plana / jerárquica	mayoritariamente plana	mayoritariamente jerárquica
Actualizaciones	eventuales	no necesariamente	en las zonas (proactivo) y entre gateways
Movilidad	depende de la periodicidad de las actualizaciones	puede haber *local repair*, ABR y SSA utilizan estabilidad	disminuyen los fallos al trabajar como grupo, normalmente hay múltiples caminos
Señalización	elevada, existen optimizaciones (OLSR)	reducida (optimizado con GPS, como el LAR)	reducida, aunque incluye el mantenimiento de la jerarquía
Almacenaje	elevado	generalmente menor	depende del tamaño de los *clusters*
Escalabilidad	hasta 100 nodos (más en OLSR y TBRPF)	menor en los que usan *source routing* (DSR)	hasta 1000 nodos o más
Soporte QoS	*link-state* problema: información local	potencialmente más aprovechable (propagar información de estado de la red, bajo demanda)	$\approx$ reactivos

2.3. Encaminamiento con calidad de servicio

A pesar de los diferentes diseños, todos estos algoritmos de encaminamiento basan su funcionamiento en los mismos objetivos. Por un lado, responden a la necesidad de encontrar rápidamente un camino con el menor número de saltos posible entre los dos extremos de la comunicación (descubrimiento de ruta – métrica distancia). Por otro, ya sea gracias a un intercambio periódico de información (*regular updates*), o ante señalización explícita (*triggered updates*), todas las propuestas tratan de reaccionar a la variabilidad del entorno (mantenimiento de ruta – reencaminamiento, etc.), introduciendo el menor *overhead* posible, considerando la especial limitación de los recursos de una red inalámbrica (ancho de banda compartido y limitado).

Sin embargo, determinar el camino más corto, de acuerdo a la utilización de la distancia como métrica de encaminamiento, no tiene por qué corresponderse con la solución óptima a la hora de proporcionar QoS. Un encaminamiento cuyo objetivo principal sea la QoS, por el contrario, debe buscar una ruta capaz de satisfacer realmente la QoS demandada. Uno de los problemas más críticos a la hora de proporcionar QoS en una MANET es la capacidad de garantizar el ancho de banda requerido por las aplicaciones, puesto que éste es el recurso más escaso en comparación con la demanda real por parte de los usuarios. Satisfacer dicha demanda está directamente relacionado con la habilidad de las capas inferiores del sistema de proveer a los servicios de una asignación eficiente de recursos de transmisión. El doble objetivo de un encaminamiento eficiente es, por lo tanto, la satisfacción de la QoS para cada una de las aplicaciones admitidas en el sistema, consiguiendo a su vez la mayor eficiencia en la utilización de los recursos [Chen y Nahrstedt, 1998]. De este modo, un camino se considera viable para una aplicación cuando existen suficientes recursos disponibles para satisfacer las restricciones de QoS impuestas por el servicio. La función básica del encaminamiento con QoS es encontrar dicho camino, prestando atención, paralelamente, a la utilización óptima de los recursos, medida a través de cierta métrica abstracta o coste. El consiguiente problema de optimización a resolver consiste en encontrar el camino de mínimo coste de entre todos aquéllos viables. Sin embargo, resolver este problema, en el dinámico entorno de las redes móviles ad hoc, implica enfrentarse a nuevas dificultades [Chen y Nahrstedt, 1998]:

- La diferente naturaleza de las múltiples aplicaciones y en consecuencia, la diversa parametrización de la QoS (múltiples restricciones: ancho de banda, retardo, *jitter*, BLER...) dificulta la selección del camino óptimo capaz de cubrir las nece-

sidades, dado el compromiso que puede suponer el cumplimiento simultáneo de los diferentes criterios. De hecho, resolver conjuntamente múltiples restricciones suele hacer el problema de encaminamiento intratable. En [Wang y Crowcroft, 1996] se demuestra que encontrar un camino viable para satisfacer dos restricciones independientes es NP-completo.

- La integración de servicios con demandas de QoS y tráfico *best-effort*, ambos de naturaleza muy diferente, dificulta encontrar el punto de operación óptimo para los dos, complicando el proceso de optimización.

- El dinamismo del entorno (variaciones de carga, activación - desactivación de conexiones, pérdidas de conectividad...) implica continuos cambios de estado en la red, cuyo conocimiento continuado y actualizado es el que determina la posible resolución de dicho problema de optimización.

Como alternativa flexible a la complejidad que supone el desarrollo eficiente de algoritmos de encaminamiento con QoS y la integración de una reserva efectiva de recursos, máxime teniendo en cuenta la necesidad de adaptación al entorno cambiante, algunos trabajos defienden la diferenciación de servicios (clasificación - priorización de paquetes - gestión de colas) como mecanismo flexible de provisión de QoS. El modelo FQMM es una muestra del tratamiento diferenciado de servicios, adaptado además a la variabilidad tanto del tráfico como de la topología. Gerharz y de Waal [2003] proponen la completa separación del encaminamiento y la gestión de recursos mediante dicha diferenciación operativa, además, sobre una capa MAC sin capacidades adicionales de QoS. Un planteamiento similar ofrece la propuesta *cross-layer Stateless Wireless Ad Hoc Networks (SWAN)* [Ahn y otros, 2002] incluyendo, además, un mecanismo de control de admisión y limitación del tráfico *best-effort*. Si bien la flexibilidad proporcionada por estos esquemas parece una buena alternativa al dinamismo de la red, la falta de una reserva garantizada de recursos impide una satisfacción real de la QoS demandada, especialmente cuando los mecanismos de control de admisión se realizan según estimaciones de la capacidad de la red que no tienen en cuenta el reparto real de los recursos (como pueda ser, por ejemplo, la asignación de slots en una red basada en acceso TDMA).

Sin embargo, una reserva real de los recursos demandados, apoyada en la medida de su disponibilidad como funcionalidad propia de los algoritmos de encaminamiento es el mecanismo más eficiente para proporcionar la deseada QoS. Así pues, la mayor parte de los esfuerzos tecnológicos responden a dicho reto tratando de desarrollar algoritmos de

encaminamiento con QoS, capaces de resolver adecuadamente el mencionado problema de optimización.

La búsqueda del camino de menor coste con garantías de QoS tiene su aplicación directa en la adaptación de los protocolos de encaminamiento proactivos. La modificación del estado del enlace *(link-state)* de acuerdo a medidas adecuadas de la QoS (como el ancho de banda disponible en cada nodo) permite definir nuevos costes asociados a los caminos que permiten la selección final de aquél realmente adecuado para satisfacer la QoS extremo a extremo. El problema se transforma en la clásica búsqueda del "camino más corto", resoluble mediante los algoritmos de Dijkstra o Bellman-Ford. Éste es el caso, por ejemplo, del protocolo *QoS for OLSR* (QOLSR) [Badis y Agha, 2005], en el cual se añaden extensiones a la funcionalidad básica del OLSR y a los mensajes de encaminamiento ya existentes en el mismo para especificar nuevas métricas como el ancho de banda disponible, el retardo, la variación de retardo *(jitter)*, probabilidad de pérdidas, consumo de potencia, etc. La utilización de dichas métricas afectaría directamente a la selección de los MPR (únicos nodos seleccionados para transmitir la información específica de control) y al cálculo de las tablas de encaminamiento. La dificultad para escalar este tipo de propuestas dado el elevado *overhead* consumido por el intercambio periódico de la información de estado se acrecienta al considerar las medidas asociadas a la QoS (imprecisas, al ser locales y dependientes del período de actualización) suponiendo finalmente una clara desventaja.

Los protocolos reactivos, por el contrario, ofrecen mayores posibilidades, al permitir una mayor escalabilidad, reducir el consumo de recursos e incluso facilitar la búsqueda paralela de múltiples caminos, por lo que la mayor parte de las soluciones apuntan a un enfoque distribuido. El funcionamiento básico supone incluir dichas medidas como nuevas métricas, añadiendo a su vez las restricciones a cumplir en las propias peticiones de ruta. Las modificaciones genéricas necesarias sobre el protocolo AODV se presentan en [Perkins y Belding-Royer, 2000]. Las políticas de rechazo o propagación de los paquetes de encaminamiento, que finalmente definen las posibles rutas, se modifican de acuerdo al cálculo y evaluación de las nuevas métricas de QoS. En [Perkins y Belding-Royer, 2000] se asume, sin embargo, la disponibilidad de las medidas necesarias, sin ofrecer una solución real a la realización de las mismas. El trabajo de [Chen y Nahrstedt, 1999] propone una solución distribuida que incluye la utilización de *"tickets"* o permisos para propagar la información de búsqueda con objeto de limitar la información de *broadcast* difundida en la red, reduciendo el *overhead* de control. El protocolo funciona con medidas imprecisas del estado de la red: ancho de banda o retardo, evaluadas a partir

de los mensajes de control intercambiados ($HELLOs$), sin considerar ninguna capa de nivel de enlace en concreto (que sería la que realmente podría dar una información más acorde con la disponibilidad real de los recursos, al determinar la asignación de los mismos). La utilización de multicamino como redundancia ante posibles fallos mejora, por otra parte, las prestaciones en movilidad.

Lo que se describe en [Chen y Nahrstedt, 1999] es la existencia de un camino viable como criterio de admisión para el establecimiento de circuitos virtuales con ciertas garantías de QoS, en base a la reserva real de los recursos. Una reserva independiente del algoritmo de encaminamiento, aun siendo realizada sobre un camino considerado a priori con suficientes recursos, no puede realmente garantizarle al servicio las prestaciones demandadas. En este sentido, la viabilidad únicamente proporciona cierta probabilidad de disponer de los recursos necesarios. Es el caso del protocolo *Core-Extraction Distributed Ad hoc Routing* (CEDAR) [Sivakumar y otros, 1999]. Su funcionamiento se basa en establecer el encaminamiento bajo demanda, a través de un conjunto de nodos específicos ("corazón" en la red – *core*) encargados de propagar la información de estado necesaria (*link-state* – ancho de banda) con la que determinar la viabilidad del camino.

Por el contrario, en [Xue y Ganz, 2003] se introduce un mecanismo conjunto de encaminamiento y señalización basado en la reserva de recursos con objeto de proporcionar garantías de ancho de banda extremo a extremo. El algoritmo propuesto, *Ad hoc QoS On-demand Routing (AQOR)*, desarrollado sobre la capa de enlace IEEE 802.11 *Distributed Coordination Function (DCF)*, tiene en cuenta la característica compartida del medio inalámbrico para determinar el ancho de banda disponible en cada enlace de acuerdo al total (agregado) que consumiría el nuevo flujo considerando la interferencia generada sobre el entorno (vecinos). El cálculo se realiza extremo a extremo, determinando en cada nodo, de acuerdo a la medida de ancho de banda, la admisión y reserva de recursos asociada (reserva fuente→destino: *forward*). Dicha reserva es controlada mediante un *timeout* para permitir la liberación automática de los recursos en caso de admisión fallida (*soft-rsv*[5]).

La propuesta de [Dong y otros, 2002] ofrece un modelo genérico de QoS basado en la definición de una estructura jerárquica (*SRL-Enabled QoS Model* (SEQM)[6]). Los *supernode* actúan como los controladores en las redes celulares (BSC o RNC), determinando

[5]Una liberación explícita de los recursos implica el envío de información de señalización: *RELEASE packet*, con el *overhead* asociado, por lo que la estrategia de reserva a priori, controlada por tiempo – *timeout* – es la más utilizada.

[6]*Supernode-based Reverse Labeling* (SRL)

la admisión, renegociado y asignación de recursos para las nuevas conexiones. Esta funcionalidad se apoya en la suposición de la provisión, por parte de la capa MAC, de las medidas apropiadas de la QoS (ancho de banda disponible y retardo). El encaminamiento, de concepción análoga al DSR, permite encontrar múltiples caminos virtuales (sucesión de *supernodes*) con recursos suficientes en base a la admisión realizada por dichos *supernodes*. No obstante, aunque la admisión de las conexiones se determina a partir de cierta medida de disponibilidad de ancho de banda, la provisión final de recursos se realiza gracias a la diferenciación de servicios proporcionada por la función DCF del protocolo MAC IEEE 802.11e de manera que la QoS resulta más cualitativa que cuantitativa.

Un control de admisión eficiente resulta clave para garantizar el nivel de QoS demandado por las aplicaciones. Además de garantizar un reparto eficiente de los recursos, un objetivo clave para optimizar el uso de la red es minimizar en lo posible el bloqueo de conexiones de manera que los recursos sean explotados al máximo proporcionando la QoS para el máximo número de conexiones admitidas en el sistema. En el contexto de las redes ad hoc multisalto, donde resulta necesario garantizar primero la existencia de un camino de conexión entre los dos extremos de la comunicación, puede suceder que el reparto del ancho de banda de cada aplicación en un único camino no garantice un uso optimizado de los recursos de la red. Por ejemplo, la ocupación elevada de los enlaces de un camino puede bloquear la admisión de una nueva conexión que requiera de los mismos enlaces, cuando en la red, analizada en global, existen recursos suficientes. Para aliviar dicho problema se han planteado diferentes propuestas cuyo objetivo es la búsqueda de rutas multicamino en las que repartir el ancho de banda demandado [Liao y otros, 2002b; Lin y Huang, 2004; Shu y otros, 2004; Wang y otros, 2004]. Todas estas soluciones plantean la reserva parcial de los recursos en los subcaminos encontrados (reserva *forward*) de manera que finalmente se consiga garantizar la demanda total. Sin embargo, este planteamiento supone resolver el compromiso entre la reducción esperada de la probabilidad de bloqueo y el coste adicional que implica el reparto del tráfico en flujos de datos independientes y el posterior agrupamiento de la información en el destino (*buffers*, números de secuencia...). Más compleja se presenta, incluso, la verdadera gestión distribuida de todos los recursos reservados en diferentes caminos para una misma conexión (detección de roturas, reencaminamiento - nuevas reservas, etc).

Determinar la viabilidad de un camino para proporcionar ancho de banda depende directamente de la habilidad para medir la disponibilidad real. Teniendo en cuenta, además, la formación de un camino multisalto, dicha medida debe reflejar la capacidad

total extremo a extremo. Todas las soluciones hasta ahora comentadas definen como ancho de banda disponible el mínimo, determinado por uno de los enlaces, de un modo análogo al concepto de "cuello de botella" (*bottleneck link*) de una red cableada. La identificación de dicho enlace se realiza considerando total independencia entre los diferentes enlaces de la red, asumiendo que cada nodo es capaz de determinar aisladamente el ancho de banda del que dispone, que gestiona individualmente. La gran demanda de aplicaciones con requerimientos de QoS ha impulsado el desarrollo de redes basadas en asignación dinámica de recursos, donde es posible proporcionar un ancho de banda garantizado gracias a la reserva de los mismos. Gran parte del trabajo desarrollado en este contexto se ha centrado en el estudio de redes basadas en acceso TDMA, que también se ha considerado la base de esta tesis. El método de acceso empleado condiciona la compartición de los recursos, por lo que resulta necesario redefinir el cálculo del ancho de banda de acuerdo a dicha estructura.

El cálculo del ancho de banda extremo a extremo en una red TDMA no es una tarea trivial, dada la interdependencia entre los diferentes enlaces de un camino multisalto, especialmente teniendo en cuenta la necesidad de realizar una asignación de recursos (slots) que resuelva los problemas del terminal oculto y el terminal expuesto. En definitiva, el cómputo del ancho de banda en un camino, en una red TDMA, no sólo necesita determinar el ancho de banda de cada uno de los enlaces que conforman el camino, sino que requiere además determinar la asignación real de los slots libres en cada uno de ellos [Chen y otros, 1997]. El ancho de banda igualmente estará determinado por el mínimo disponible, pero el cálculo de dicho mínimo se relacionará directamente con la asignación real a lo largo de todo el camino, imposible de determinar a partir de medidas individuales de cada uno de los nodos.

En [Chen y otros, 1997] se propone la necesidad de establecer correctamente un control de admisión de acuerdo a la medida real del ancho de banda (slots) proporcionado por una red TDMA. No obstante, el problema se simplifica en este caso, al considerar una estructura basada en *clusters* (CDMA sobre TDMA) que independiza la medida entre diferentes enlaces. Sobre esta base, Lin y Liu [1999] plantean un método heurístico para el cálculo del ancho de banda, dado que la resolución óptima del problema – *SATisfiability problem (SAT)* – resulta NP-completo [Garey y Johnson, 1979]. Considerando el intercambio periódico de las tablas de rutas (encaminamiento DSDV), incluida la información relativa a slots, la aplicación de dicho algoritmo permite conocer la disponibilidad real del ancho de banda. Con esta medida, el control de admisión es invocado ante el establecimiento de una nueva aplicación en función de la ruta asociada.

Una vez admitida la conexión, es necesario realizar la asignación real de los recursos salto a salto, modificando adecuadamente la información necesaria (slots) de las tablas de rutas, de acuerdo a la señalización explícita llevada a cabo para determinar dicha reserva (*RESERVE packet*). En [Liao y otros, 2002a] se resuelven los problemas del terminal expuesto y el terminal oculto en el cálculo y asignación del ancho de banda para una red TDMA pura (sin *clusters* CDMA sobre TDMA) integrando el procedimiento con el algoritmo de encaminamiento bajo demanda DSR, incluyendo toda la información necesaria en los paquetes de encaminamiento (*source routing*). La aceptación por parte del destino, y posterior formación del camino, invoca la reserva real de los recursos (reserva destino→fuente: *backward*). Un procedimiento similar sobre el AODV se plantea en [Gerasimov y Simon, 2002], aunque la definición del ancho de banda disponible realizada por cada nodo se basa en la información de asignación de slots intercambiada entre vecinos (a través de paquetes $HELLO$) y la reserva de recursos se realiza antes de la confirmación por parte del destino (*forward*). En la propuesta de Kim y otros [2004] se incorporan restricciones en el nivel de la SIR deseada, seleccionando los slots disponibles de acuerdo con la interferencia medida en cada uno de ellos dada la recepción de señal estimada a partir de los propios mensajes de encaminamiento. En función de dicha selección se realiza la reserva *forward* de los recursos demandados. Con objeto de determinar el camino con mejores prestaciones, esta solución incluye la capacidad de búsqueda multicamino. El nodo destino espera la recepción de diversas peticiones antes de contestar favorablemente a una de ellas.

Todas estas soluciones, sin embargo, no contemplan la problemática adicional de la asignación simultánea de recursos consecuencia de la búsqueda de diversas rutas en paralelo (*Race condition*), lo que sí resuelve el trabajo de Jawhar y Wu [2005b] gracias a la inclusión de estados adicionales (asignado, libre, reservado) que "anuncian" la existencia de asignaciones antes de su reserva efectiva (*backward*). Tanto nodos vecinos realizando asignaciones paralelas, como los propios nodos, manejando rutas para diferentes conexiones simultáneamente, pueden evitar la selección repetida de recursos y el consecuente conflicto en la reserva.

La reserva efectiva de los recursos, sin embargo, depende de la capa MAC TDMA concreta sobre la que funciona el algoritmo de encaminamiento, lo que no es tenido en cuenta en ninguna de las propuestas anteriores. En [Zhu y Corson, 2002] se desarrolla un algoritmo de encaminamiento con QoS sobre la base de una capa MAC TDMA concreta, implementada por los mismos autores, en cooperación con el protocolo de encaminamiento AODV. El algoritmo de cálculo del ancho de banda y la gestión de los

slots asignados se integran completamente en el funcionamiento del protocolo. La asignación real y consecuente reserva se realizan únicamente durante la fase de respuesta del nodo destino (*backward*), momento en el que se determinan los recursos estrictamente necesarios de acuerdo a la demanda, ya que durante la evaluación de la disponibilidad se determinan todos los slots disponibles en cada uno de los enlaces, sin limitación, flexibilizando el encuentro de caminos posibles y la posterior asignación de recursos.

El alcance de la mayor parte de todas estas propuestas cubre por lo general únicamente soluciones teóricas o aproximaciones heurísticas, sin tener realmente en cuenta los potenciales problemas derivados de una implementación real de las soluciones planteadas. Considerar un escenario inalámbrico ad hoc supone importantes implicaciones físicas que no pueden ignorarse si se pretende asegurar la capacidad de la red de garantizar la QoS demandada. La naturaleza interferente del medio inalámbrico y la pérdida de conectividad debida a la movilidad afectarán a cualquier implementación práctica, puesto que las transmisiones teóricamente libres de colisiones (problemas del terminal oculto y el terminal expuesto solventados) ya no pueden asumirse. El efecto es incluso más crítico cuando se produce sobre las comunicaciones propias del control a nivel MAC, como puede ser el mantenimiento de la conectividad o la gestión de recursos. Dicho efecto tendrá un impacto directo sobre las decisiones de encaminamiento tomadas en función de parámetros determinados por el nivel de enlace. Dada la gran interdependencia entre el encaminamiento y las capas inferiores, una consideración más detenida de dicha relación parece apuntar en la dirección de un diseño más complejo y novedoso, basado en una estructura *cross-layer* que pueda explotar al máximo la cooperación real entre las diferentes capas.

2.4. Diseño *cross-layer*

La simplicidad y robustez de un diseño de protocolos basado en capas funcionales independientes (actual generación de protocolos OSI o TCP/IP) ha resultado exitosa en el desarrollo de Internet [Raisinghani y Iyer, 2004]. Sin embargo, esta estructura no resulta suficientemente flexible para afrontar el dinamismo y la limitación de recursos propios de las redes ad hoc [Xylomenos y Polyzos, 1999]. En consecuencia, ha surgido una nueva tendencia en el desarrollo de protocolos específicos para este tipo de redes denominada diseño *cross-layer*. Dicho concepto no es completamente nuevo en el entorno de las redes de comunicaciones, puesto que ya ha sido propuesto como optimización de la propia arquitectura de niveles en redes cableadas [Clark, 1985; Clark y Tennenhouse,

1990], lo que se ha denominado como realimentación entre capas (*cross-layer feedback*). En [Raisinghani y Iyer, 2004], los autores presentan las múltiples posibilidades que ofrece dicha intercomunicación entre niveles en el entorno de las comunicaciones inalámbricas con el objetivo de mejorar la calidad percibida por el usuario final gracias a la adaptación de toda la pila de protocolos en su conjunto a las condiciones variantes del medio.

Sin embargo, desde un punto de vista genérico, mucho más amplio, el concepto *cross-layer* hace referencia a un diseño de protocolos que, en lugar de desarrollar de manera completamente independiente cada uno de los niveles de la arquitectura, trata de explotar activamente la dependencia entre las diferentes capas para obtener cierta ganancia global en las prestaciones del sistema [Srivastava y Motani, 2005]. De este modo, existen muchas interpretaciones de lo que realmente supone un diseño de estas características. Srivastava y Motani ofrecen una completa revisión del estado del arte así como una clara identificación de las claves para entender este nuevo concepto de diseño. La motivación del mismo en el contexto de las redes móviles ad hoc se presenta de una manera clara: el nuevo entorno, dinámico e inalámbrico, presenta una problemática específica que no puede resolverse mediante la estricta separación de funcionalidades propia de la arquitectura de niveles tradicional. Por otra parte, la naturaleza variante del canal inalámbrico presenta potencialmente cierto beneficio en el uso oportunista del mismo (mediante ajuste de parámetros de transmisión se puede hacer un mejor uso del canal en aquellos momentos en que éste presenta mejores características). Y por último, el medio de transmisión ofrece nuevas posibilidades de comunicación, como puede ser la cooperación gracias al uso de un medio compartido de difusión, o ganancias adicionales mediante el uso de técnicas capaces de explotar al máximo la diversidad ofrecida por el canal de transmisión. Atendiendo a la definición dada por Srivastava y Motani, el diseño sujeto a una arquitectura de niveles implica el desarrollo independiente de cada uno de los protocolos involucrados, teniendo éstos en cuenta el uso de los servicios de capas inferiores, pero despreocupándose de cómo dichos servicios son realmente proporcionados. Del mismo modo, las interfaces de comunicación entre las capas siguen el modelo de referencia intercambiando únicamente las primitivas especificadas en dicho modelo. Por el contrario, una implementación *cross-layer* implica el incumplimiento de dichas reglas, consistiendo por lo tanto en una nueva metodología de diseño. Esta metodología puede entenderse de diversas formas, en función del grado de interacción que se supone entre las diferentes capas, implicando por ello diferentes grados de incumplimiento de la estructura por niveles (Tabla 2.4).

Tabla 2.4. Propuestas de diseño cross-layer

Interacción entre diferentes capas de la arquitectura	**Incumplimiento de la arquitectura por niveles**
Comunicación directa entre capas	Creación de nuevas interfaces (intercambio de información – *cross-layer feedback* [Raisinghani y Iyer, 2004])
Base de datos compartida [Conti y otros, 2004]	Nuevas interfaces entre los protocolos existentes y la nueva base de datos. Calibración vertical: ajuste de parámetros específicos de cada capa considerando la consecuente adaptación en el resto de niveles. No configuración independiente.
Nuevas abstracciones	Nuevos protocolos resultado de la combinación de capas adyacentes sin necesidad de añadir nuevas interfaces. Diseño acoplado de capas fijas sin añadir nuevas interfaces (problema: la dificultad de asegurar la no necesidad de propagar los cambios necesarios al modificar sólo una de las capas). Cambio completo en la organización de los protocolos. Un nuevo planteamiento requiere rediseñar completamente del sistema, con la inherente complejidad.

Hoy en día existen diversas propuestas *cross-layer*, cada una de ellas respondiendo de manera diferente a la problemática ad hoc, centrándose en diferentes relaciones de cooperación entre capas, y en resumen, desarrolladas cada una de ellas de manera independiente. Así pues, se presenta claramente la necesidad de analizar la importancia de cada una de ellas como base para identificar, en último término, cuál es el enfoque más prometedor, analizando cuidadosamente las ventajas e inconvenientes desde el compromiso coste-beneficio en términos de complejidad de implementación frente a la mejora de las prestaciones. En términos generales, sí existe una motivación real detrás de la aplicación de un diseño *cross-layer* en redes móviles ad hoc. Se presenta como una solución adecuada al dinamismo de las mismas, al hacer posible que todos los niveles ne-

cesarios de la arquitectura reaccionen conjuntamente adaptándose a las condiciones del entorno. Como contrapartida, sin embargo, se presentan nuevos retos a resolver. Por un lado, la gran versatilidad de propuestas lleva a determinar la necesidad de cuestionarse la interoperatividad de todas ellas. Por ejemplo, ¿puede una propuesta explícitamente diseñada para hacer cooperar ciertos protocolos específicos extrapolarse o apoyarse sobre otros mecanismos igualmente optimizados en base a un propósito diferente? Por otro lado, el hecho de diseñar arquitecturas cuyo objetivo final es la adaptación a la variabilidad de las condiciones implica analizar cuidadosamente las condiciones reales en que dicha solución es aplicable, ya que en determinadas circunstancias éstas podrían ser incluso contraproducentes. De hecho, la complejidad de un diseño *cross-layer* puede llegar a producir un intrincado código (*spaghetti-like code* [Conti y otros, 2004]) de difícil mantenimiento en el que cualquier modificación debe propagarse a través de todos los protocolos, e incluso pueden llegar a producirse interacciones no esperadas que resulten en una degradación de las prestaciones. Igualmente, la necesidad de tener un conocimiento actualizado del estado de la red, para responder adecuadamente al mismo, implica un intercambio necesario de información de estado (entre capas y/o entre diferentes terminales) que puede llegar a suponer una carga tanto computacional como de consumo de recursos muy elevada, con lo que ésta debe tenerse especialmente en cuenta. Finalmente, la pérdida de la definición tradicional de protocolos e interfaces de la arquitectura de niveles, dificulta el proceso de estandarización de los mismos y el establecimiento de ciertas referencias de diseño, convirtiéndose en todo un reto el diseño eficiente de sistemas completos de operación en el entorno de las redes móviles ad hoc.

Un planteamiento *cross-layer* abre un abanico de posibilidades de diseño en el entorno de las redes móviles ad hoc, y dada la dificultad de definir el método más apropiado, el esfuerzo tecnológico actual se ha desarrollado acorde con aquello que suscita mayor interés: la provisión de QoS. Con tal propósito, la definición de protocolos de encaminamiento eficientes, capaces de garantizar el establecimiento de caminos con recursos suficientes desempeña un papel importante, y asimismo, cobra especial relevancia la gestión eficiente de los recursos del sistema.

En este contexto, se han desarrollado diferentes propuestas. El modelo de servicio adaptativo *utility-fair* [Bianchi y otros, 1998], se basa en la adaptación de los recursos de acuerdo a la satisfacción de las aplicaciones según las denominadas curvas de utilidad (*utility-curves*). *The Illinois Mobile Environments LaboratorY* (TIMELY) [Bharghavan y otros, 1998] constituye un grupo de trabajo centrado en el diseño de una arquitectura

completa para la gestión de los recursos, gracias a la cooperación de diferentes capas del sistema para realizar adaptativamente el control de admisión, reserva de recursos y *handoff* entre diferentes puntos de acceso. Por otra parte, un protocolo de transporte específico se encarga de adaptar las aplicaciones al servicio proporcionado. Ejemplos más aplicados a un entorno verdaderamente ad hoc (sin puntos de acceso) los encontramos en la arquitectura *INSIGNIA*, un sistema completo que coordina control de admisión, encaminamiento y señalización (reserva y liberación de recursos); el protocolo *Dynamic Resource Reservation Setup Protocol (dRSVP)* [Mirhakkak y otros, 2001] encargado de la reserva dinámica de recursos en redes ad hoc; o en el modelo *SWAN*, un método flexible que, apoyado sobre una capa MAC sin prestaciones de QoS, mediante la gestión apropiada de los recursos e intercomunicación entre niveles, garantiza cierta QoS cualitativa gracias a la priorización de paquetes según el servicio, ajuste de tasa *best-effort* y un mecanismo de control de admisión. Por último, en [Shah y otros, 2005] se propone un método de gestión del ancho de banda en el entorno de un sólo salto – *Bandwidth Management (BM)* – en el que un gestor centralizado, en función de las medidas tomadas del enlace, se encarga de realizar adaptivamente la admisión y adaptación de los recursos para los servicios, basándose en el método de compartición justa por máximo - mínimo con garantía de tasa mínima.

En [Zhou y otros, 2005] se propone un claro ejemplo de la necesidad de cooperación entre capas para realizar una selección apropiada de la ruta de encaminamiento de acuerdo a múltiples métricas de QoS (estabilidad, ancho de banda, retardo y número de saltos). El objetivo principal de los autores es reducir el problema de inundación de mensajes de control (*broadcast storm*[7]), los mecanismos deficientes de *backoff* y la falta de detección de colisiones – *Collision Detection (CD)* – [Tseng y otros, 2002] característicos de las redes basadas en CSMA/CA, como es el caso del IEEE 802.11. En redes con una gestión de los recursos más eficiente, como puede ser el caso de redes TDMA basadas en reserva [Fasolo y otros, 2005], dicho problema puede verse aliviado. La arquitectura presentada identifica la interrelación entre un encaminamiento eficiente y la definición apropiada de métricas acorde con las medidas de niveles inferiores. En primer lugar, es necesario una definición clara de las prestaciones demandadas por las aplicaciones (determinadas gracias al bloque funcional *Virtual Device Information Manager* (VDIM) y, posteriormente, una búsqueda eficiente del mejor camino – *Priority-based Route Dis-*

[7]La transmisión *broadcast* por simple inundación (*flooding*) provoca, en redes CSMA/CA, una elevada redundancia de mensajes, un aumento de la contienda y el consecuente incremento de las colisiones, especialmente teniendo en cuenta la inexistencia de diálogo *Request To Send* (RTS)-*Clear To Send* (CTS) para *broadcast*.

covery Strategy (PRDS). La priorización en la propagación de los mensajes de encaminamiento, en función de una métrica combinada de QoS, limita la difusión de paquetes *broadcast* y, al mismo tiempo, permite encontrar un camino capaz de satisfacer la QoS demandada. La cooperación de todos los módulos implicados en el proceso constituye la arquitectura *cross-layer*: *Cross-layer Route Discovery Framework* (CRDF).

¿Por qué un encaminamiento basado en diseño cross-layer?

En definitiva, la estrecha relación entre la provisión de QoS y la gestión de los recursos del sistema, y la necesidad de aportar soluciones eficientes al tiempo que adaptables al dinamismo propio de las redes ad hoc evidencian la especial adecuación de una solución *cross-layer*. Desarrollar mecanismos distribuidos de control de admisión se presenta como uno de los principales retos a resolver. El primer paso necesario para garantizar un servicio satisfactorio es, naturalmente, el establecimiento de la ruta apropiada entre los dos extremos de la comunicación. La definición de la idoneidad de dicho camino para garantizar las prestaciones demandadas no es más que la representación cuantificable, a través de determinada métrica de encaminamiento, de la capacidad del sistema para proporcionar los recursos necesarios (ancho de banda principalmente). Un diseño conjunto del protocolo de encaminamiento y la capa MAC permitirá determinar la métrica adecuada así como gestionar de una manera distribuida los recursos del sistema realizando una admisión eficiente de los servicios de acuerdo a sus demandas.

Capítulo 3

Control de Admisión Distribuido

Como ya se comentó en §2, la clara necesidad de establecer mecanismos de control de admisión distribuidos tiene una respuesta potencialmente atractiva en el diseño *cross-layer* de algoritmos de encaminamiento capaces de proporcionar caminos que garanticen la provisión de la QoS. Dicha garantía se fundamenta en una reserva eficiente de los recursos demandados acorde con nuevas métricas de QoS que reflejen adecuadamente la disponibilidad real de los recursos, tal y como éstos son gestionados por el nivel de enlace.

Un diseño *cross-layer* aplicado sobre una propuesta de encaminamiento con QoS implica la necesidad de evaluar cuidadosamente el grado de colaboración necesario entre las capas implicadas (encaminamiento y nivel de enlace) prestando especial atención al objetivo real de dicha cooperación, que consiste en resolver el compromiso entre la provisión de la QoS y la gestión eficiente de los recursos de la red:

- El algoritmo de encaminamiento debe ser capaz de extraer una medida fiable de la disponibilidad de recursos de acuerdo a la gestión realizada por el nivel MAC. Dicha medida debe reflejarse de manera apropiada en la métrica de acuerdo a la cual se tomen las decisiones.

- El nivel MAC debe responder eficientemente a la petición de asignación y reserva de recursos demandada por el nivel de encaminamiento.

- Tanto el nivel de encaminamiento como el MAC deben reaccionar conjuntamente ante la variabilidad del entorno actualizando eficientemente (sin redundancias o complejidades innecesarias) tanto las rutas como la asignación de recursos.

Dicha comunicación es totalmente dependiente de los protocolos específicos implementados como base. Teniendo en cuenta que diferentes capas MAC tienen formas diversas de gestionar los recursos, un algoritmo de encaminamiento con QoS basado en una única capa MAC no será fácilmente generalizable a otro tipo de capas MAC. De igual modo, los diferentes mecanismos empleados por diversos protocolos de encaminamiento, tanto en la manera de reaccionar a la variabilidad del entorno como en las decisiones tomadas en el encaminamiento propiamente dicho, suponen una comunicación diferente con los niveles inferiores. No obstante, el trabajo de investigación desarrollado pretende analizar la viabilidad de esta nueva metodología de diseño y su especial adecuación a la naturaleza dinámica de las redes móviles ad hoc, para lo cual se propone la evaluación bajo simulación de una implementación práctica concreta. Así pues, se trata de localizar los problemas a los que hacer frente a la hora de llevar a cabo una implementación real, presentando soluciones específicas que, si bien adaptadas a las selecciones particulares aquí realizadas, pueden permitir establecer las pautas de diseño de una manera generalizada.

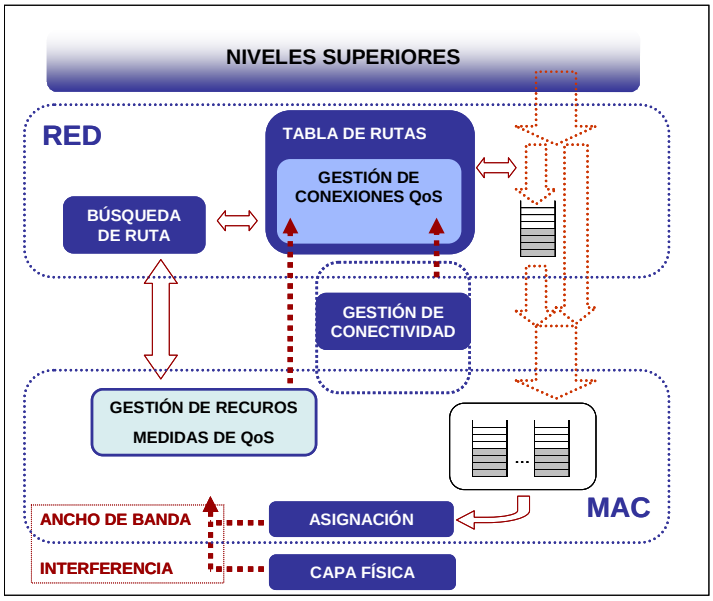


Figura 3.1. Estructura general de una propuesta de encaminamiento *cross-layer*.

El diseño finalmente propuesto se resume conceptualmente en el esquema de la figura 3.1. Éste, se corresponde con un desarrollo independiente de los diferentes niveles (encaminamiento y enlace) a los que se incorporan las interfaces adecuadas para realizar una comunicación directa entre ellos [Canales y otros, 2005a]. Se añade, por otra parte,

el acceso común a cierta información compartida relativa a la gestión de los recursos. En este capítulo se describe la selección de protocolos realizada para cada nivel y los bloques funcionales que constituyen la arquitectura encaminamiento – MAC conjunta, con especial atención a la problemática de diseño asociada, los procedimientos de gestión de recursos llevados a cabo y las soluciones específicas propuestas.

3.1. Conectividad en las redes multisalto. Modelo de red

3.1.1. Modelos de Interferencia

En una red ad hoc la potencia de transmisión es limitada, lo que reduce la cobertura de los terminales de manera que la comunicación entre cualquier par de usuarios dentro de la red se basa en el soporte de encaminamiento (colaboración) de cierto número de nodos, ajenos a la comunicación. Dicha conexión multisalto se establece entre aquellos nodos alcanzables entre sí. Para comprender y por lo tanto ser capaces de modelar y analizar el comportamiento de una red ad hoc resulta imprescindible definir correctamente en qué consiste realmente la conectividad o alcance entre dos nodos cualesquiera.

Consideramos una red inalámbrica genérica, compuesta por un conjunto T de nodos arbitrariamente distribuidos en el plano espacial. Dados dos nodos $n_i, n_j \in T$, de coordenadas (x_i, y_i) y (x_j, y_j) respectivamente, se define su distancia euclídea $d_{ij} = |n_i - n_j| = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$. El equipamiento radio de cada terminal i determina su radio de cobertura R_i y el rango de interferencia R'_i, potencialmente mayor. En este escenario, es posible establecer un enlace entre dos nodos (es decir, son alcanzables) cuando una transmisión entre ellos se recibe correctamente. A la hora de identificar cuándo dicha transmisión es realmente correcta, existen dos modelos de interferencia comúnmente utilizados [Gupta y Kumar, 2000]:

Protocol Model: En este modelo, una transmisión entre n_i y n_j se considera correcta (libre de colisión) siempre y cuando se cumpla:

1. $d_{ij} < R_i$
2. Ningún nodo $n_k \in T, d_{kj} \leq R'_k$ está transmitiendo

En la mayoría de las ocasiones, se define un modelo simplificado en el que $R_i = R'_i = R$, de manera que todo aquel terminal fuera del rango de cobertura R del receptor n_j no contribuye a la interferencia sufrida por dicho terminal.

Este modelo de interferencia es el más utilizado, ya que permite realizar una descripción de la red como un grafo conectado. De este modo, la teoría de grafos es aplicable al análisis de multitud de aspectos relacionados, tales como la asignación óptima de slots en una red TDMA (problema del coloreado), el encaminamiento (Dijkstra, Bellman-Ford) o la conectividad propiamente dicha (matriz de adyacencias). Por otra parte, la utilización de un modelo simplificado permite focalizar los problemas básicos a resolver a la hora de diseñar protocolos específicos en el contexto ad hoc. Su posterior análisis mediante modelos físicos realistas permite plantear las mejoras necesarias para su correcto funcionamiento en este entorno.

Physical Model: En este caso, la recepción correcta de una comunicación depende de la relación señal a interferencia (SIR) con que la información llega al receptor. Una transmisión se considera correcta si la SIR con la que se recibe la señal es mayor o igual a la mínima necesaria para decodificar correctamente la información. En este caso, todas las transmisiones activas en la red contribuyen al nivel de interferencia sufrido por el receptor. Por lo tanto, la regla que debe cumplirse en este caso es:

$$SIR_{ij} = \frac{P_{tx}^{i} \cdot L_{ij}}{P_{n,j} + \sum\limits_{\substack{n_k \in TX, \\ n_k \neq n_i}} P_{tx}^{k} \cdot L_{kj}} \geq SIR_{th} \tag{3.1}$$

donde SIR_{ij} es la relación señal a interferencia en recepción, en el enlace (i, j), $TX \in T$ es el conjunto de nodos transmisores en el mismo canal[1], P_{tx}^{k} es la potencia transmitida por el nodo n_k, $P_{n,j}$ el ruido térmico en el receptor n_j y L_{ij} las pérdidas de propagación entre los nodos n_i y n_j.

3.1.2. Problemas del terminal oculto y el terminal expuesto

La naturaleza compartida del medio inalámbrico y la comunicación multisalto de las redes ad hoc tienen como consecuencia problemas específicos de este entorno, concre-

[1]Modelo generalizado tanto para un único canal de transmisión como para la existencia de un medio de transmisión multicanal, como por ejemplo una estructura de trama TDMA, donde cada slot temporal define un canal de transmisión.

tamente los comúnmente conocidos problemas del terminal oculto – *Hidden Terminal Problem* (HTP) – y el terminal expuesto – *Exposed Terminal Problem* (ETP), brevemente comentados en §2.1 y descritos a continuación. A la hora de entender la problemática real, dichos fenómenos suelen explicarse basándose en un modelo de interferencia básico (*Protocol Model*) que permite un acercamiento más directo a la raíz del problema. La aplicación de un modelo basado en el cómputo de la interferencia real entre diferentes usuarios, sin embargo, no hace perder generalidad a dicha descripción, como se destacará en el capítulo 4.

Los terminales considerados "ocultos" son aquéllos que, por encontrarse fuera del rango de cobertura, no son capaces de "escucharse mutuamente" pudiendo así transmitir concurrentemente al mismo terminal (receptor) que sufriría una colisión. En el caso de la figura 3.2(a), existe una comunicación activa entre los nodos A y B. Puesto que el nodo A se encuentra fuera del rango de cobertura del nodo C, éste no es consciente de la potencial interferencia que puede causar sobre el terminal B si intenta establecer una comunicación paralela (con D, por ejemplo). Una señalización apropiada entre los diferentes nodos permitiría resolver sencillamente dicho problema. Es el caso del diálogo RTS/CTS[2] añadido a CSMA/CA [IEEE 802.11 Working Group, 1999] o una señalización más explícita de la ocupación de cada slot en una estructura de trama TDMA[3].

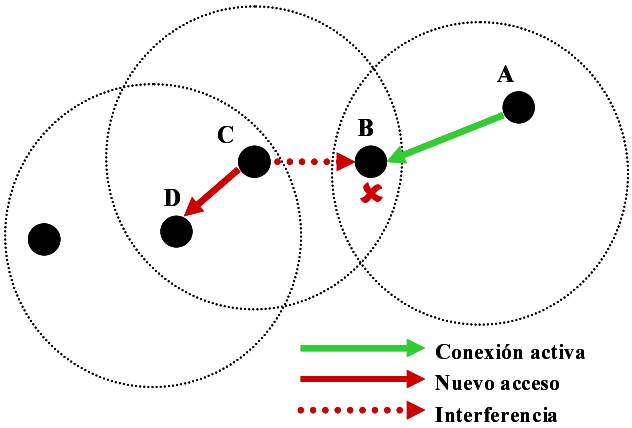
El problema del terminal expuesto hace referencia, por el contrario, a la reducción en la capacidad de la red como resultado de transmisiones no establecidas al considerar ocupado el canal de comunicación (sensado de otras transmisiones dentro del rango de cobertura) cuando realmente, dado que las mismas no afectarían al potencial receptor (éste no vería colisión, por estar fuera de su rango de cobertura), podrían coexistir en paralelo. La figura 3.2(b) muestra esta situación. Existiendo una comunicación activa entre los nodos C y B, el nodo D, que es capaz de escuchar dicha comunicación, presupone que la ocupación del canal de transmisión le impide establecer nuevas comunicaciones. En el caso descrito en la figura, una comunicación con el terminal E, sin embargo, no produciría colisión alguna sobre el receptor B, por lo que realmente sí podría establecerse. El diálogo RTS/CTS no resuelve dicho problema. Se necesita, por el contrario, una señalización explícita por parte de los receptores, que indique la existencia de éstos

[2] *Request To Send / Clear To Send*

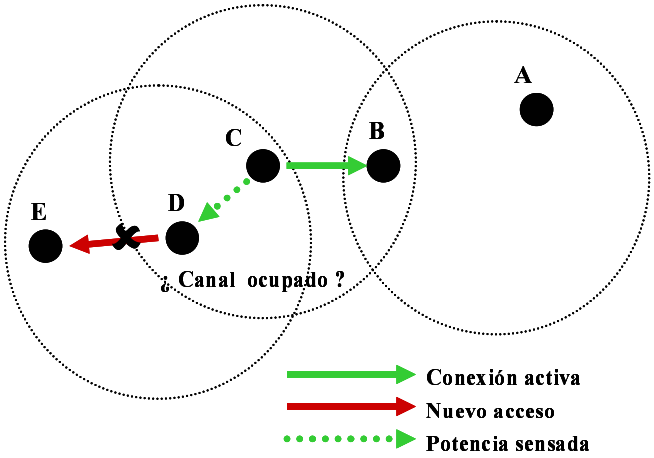
[3] Señalización como la utilizada en el protocolo ADHOC MAC, base del diseño propuesto, descrito en el anexo A.

en las proximidades del transmisor. De esta manera, la ausencia de dicha señalización permite darse cuenta de que éstos no se encuentran en el rango de cobertura.

En definitiva, una adecuada señalización en el nivel MAC puede solventar completamente el problema del terminal oculto así como maximizar la reutilización del medio de transmisión eliminando el problema del terminal expuesto. Los mecanismos particularmente empleados a tal efecto varían entre las diferentes propuestas de protocolos MAC desarrollados en el entorno de las redes ad hoc [Jurdak y otros, 2006; Kumar y otros, 2006]. En el anexo A se detallan las soluciones concretas desarrolladas en el protocolo MAC seleccionado en la arquitectura propuesta.



(a) Terminal oculto



(b) Terminal expuesto

Figura 3.2. Problemas del terminal oculto y el terminal expuesto.

3.1.3. Modelo de red

Una red ad hoc formada por N_t terminales puede modelarse mediante un grafo[4] $G = (T, E)$, donde T es el conjunto finito de terminales (nodos o vértices del grafo) y E el conjunto de enlaces, de manera que $N_t = |T|$. Para un determinado nodo de la red, n_i, se define el conjunto de sus vecinos a un salto como $NB_i = \{n_j \in T : (n_i, n_j) \in E\}$. Recordando el modelo de interferencia *Protocol Model* anteriormente explicado (página 41), se definen como vecinos de un nodo aquellos terminales dentro del rango de cobertura, de manera que sería posible establecer con ellos una transmisión correcta. Si se considera una red basada en tecnología de acceso TDMA, el ancho de banda se divide en ranuras temporales (slots) $S = \{s_1, s_2, ..., s_N\}$, que conforman la trama de transmisión.[5] La asignación de slots dentro de dicha trama para cada uno de los nodos se denomina TS_i – *Transmission Schedule* (TS), siendo R_i^k el conjunto de receptores para cada uno de los slots $s_k, s_k \in TS_i, R_i^k \in NB_i$. Por simplicidad en la notación, se usará únicamente la referencia a TS_i. Por otra parte, $RS_i = \left\{s_k \in S : n_i \in R_j^k, n_j \in NB_i\right\}$ define el conjunto de slots en los que el nodo n_i recibe de cada uno de sus vecinos. $TN^k = \{n_i \in T : s_k \in TS_i\}$ hace referencia al conjunto de nodos que transmiten en el slot s_k. Una transmisión activa en determinado enlace se denota por $(n_i \rightarrow n_j)$, o $(n_i \rightarrow n_j)^k$ si se quiere enfatizar que dicha transmisión tiene lugar en el slot s_k. La asignación de recursos de la red completa TS consistirá por tanto en la colección de todas las asignaciones individuales de cada nodo $\{TS_i : n_i \in T\}$. Esta asignación estará sujeta a las características específicas del nivel MAC empleado. Un protocolo eficiente diseñado para eliminar completamente el problema del terminal oculto permite asumir que dicha asignación TS cumplirá la propiedad de no-conflicto (*conflict-free*)[6]: Si un nodo n_i transmite en el slot $s_k \left(n_i \in TN^k\right)$, para todos aquellos nodos receptores $n_j \in R_i^k, NB_j \cap TN^k = \{n_i\}$ y $n_j \notin TN^k$. En otras palabras, cuando un nodo n_i transmite a uno de sus vecinos n_j en el slot s_k, dicho vecino no transmite, y para él, n_i es el único de sus vecinos que transmite en ese slot. Así pues, dada una asignación de recursos TS determinada, se definen los siguientes conjuntos para el nodo n_i:

$$SRT_i = \left\{ s_k \in S : s_k \notin TS_i, s_k \notin RS_i, s_k \notin \bigcup_{n_j \in NB_i} RS_j \right\} \tag{3.2}$$

[4]Notación utilizada en [Zhu y Corson, 2002].

[5]A lo largo de esta memoria se utilizará indistintamente la notación s_k o directamente slot k para hacer referencia a un slot en concreto. Del mismo nodo, se utilizará la notación nodo i o n_i.

[6]Es el caso del protocolo seleccionado, el ADHOC MAC, descrito en el anexo A.

$$SRR_i = \left\{ s_k \in S : s_k \notin TS_i, s_k \notin RS_i, s_k \notin \bigcup_{n_j \in NB_i} TS_j \right\} \tag{3.3}$$

que determinan el conjunto de slots disponibles para transmitir sin causar interferencia sobre los vecinos receptores – *Slots Ready to Transmit* (SRT) – y el conjunto de slots disponibles para recibir sin sufrir interferencia de los vecinos transmisores – *Slots Ready to Receive* (SRR), de manera que $TS = \{SRT_i, SRR_i, n_i \in N\}$. Por último, el conjunto de slots que realmente pueden usarse en una comunicación a nivel de enlace entre dos nodos n_i y n_j viene determinado por la intersección de los correspondientes SRT_i y SRR_j: *Link Bandwidth* (LB).

3.2. Selección de un protocolo MAC eficiente

Diversos protocolos de nivel de enlace desarrollados en el contexto de redes móviles ad-hoc responden de manera diferente a las necesidades específicas de este tipo de redes. Sin embargo, a la hora de realizar la selección de una capa de acceso al medio apropiada para formar parte del diseño propuesto se han tenido en cuenta diversos factores que, si bien no limitan la solución a una única implementación óptima, permiten definir las características necesarias para cumplir con los objetivos de dicho diseño, facilitando la provisión de QoS e interactuando adecuadamente con el nivel de encaminamiento:

Toda capa de acceso al medio desarrollada en el entorno de las redes móviles ad hoc debe resolver los problemas específicos de dicho entorno, cumpliendo un doble objetivo: **resolver** completamente los **problemas del terminal oculto y del terminal expuesto** derivados de la naturaleza multisalto de este tipo de redes y regular el acceso al medio compartido **optimizando la utilización de los recursos** (medio de transmisión inalámbrico limitado). En definitiva, se trata de dar respuesta al compromiso entre eficiencia y provisión de QoS.

En relación con la esperada colaboración con el protocolo de encaminamiento de nivel superior, teniendo en cuenta el papel crucial que éste desempeña en una red multisalto, el nivel de enlace debe garantizar una transmisión eficiente de los mensajes de control asociados. En este sentido, cobra especial interés **garantizar transmisiones *broadcast* fiables**. Para ello, un protocolo MAC adecuado debe resolver apropiadamente el problema del terminal oculto para este tipo de transmisiones, y no únicamente en comunicaciones punto a punto (RTS-CTS *handshake*).

Como ya se introdujo en §2, la asignación dinámica y reserva de recursos proporcionada por las redes basadas en **acceso TDMA** se presenta como una de las soluciones prometedoras para la provisión de QoS en redes móviles ad hoc, habiendo centrado el esfuerzo de numerosos autores.

La provisión de la QoS extremo a extremo requiere una gestión eficiente de los recursos capaz de diferenciar apropiadamente servicios de acuerdo a sus requerimientos particulares. El **uso de prioridades** en el nivel MAC permite realizar una asignación de recursos dependiente de las demandas específicas así como resolver adecuadamente la competencia por los mismos. Igualmente, garantizar realmente dichas demandas proporcionando un ancho de banda asegurado se fundamenta en la **reserva efectiva** del mismo, por lo que el nivel de enlace debe posibilitar el acceso exclusivo a los recursos necesarios, una vez reservados, limitando la competencia al momento de resolver nuevas reservas simultáneas.

Por último, dado que la propuesta final se basa en aprovechar la cooperación entre niveles de modo que el encaminamiento, último encargado de establecer el camino apropiado en el que la provisión de QoS esté asegurada, pueda tener una visión real del entorno, el nivel MAC debe ser capaz de proporcionarle dicha visión, gracias a **medidas fiables de la disponibilidad de recursos** (slots TDMA disponibles) y una rápida detección del dinamismo de la red (**gestión de la conectividad local**).

Teniendo en cuenta las características descritas, necesarias para servir como soporte de la arquitectura *cross-layer* propuesta, se ha tomado como base de diseño el protocolo ADHOC MAC [Borgonovo y otros, 2004]. Sin embargo, con objeto de garantizar la provisión de la QoS, el esquema de acceso finalmente implementado es resultado de la evolución de dicho protocolo hacia un diseño multiservicio capaz de garantizar una reserva diferenciada de recursos de acuerdo a la demanda diversa de las aplicaciones [Gállego, 2007]. La descripción detallada de esta propuesta se muestra en el anexo A.

3.3. Protocolo de encaminamiento con calidad de servicio

Con objeto de seleccionar un algoritmo de encaminamiento lo más genérico posible, respondiendo a su vez a las pretensiones de proporcionar QoS, se han tenido en cuenta las siguientes consideraciones:

Dado que el objetivo final es desarrollar un mecanismo distribuido de control de admisión, el protocolo seleccionado debe determinar, en el mismo momento del establecimiento de una nueva conexión, la capacidad de la red de proporcionarle los recursos demandados garantizando la no degradación de las conexiones previamente admitidas. La cooperación de los diferentes nodos en la determinación de un camino fiable entre dos extremos, potencialmente puede permitir un conocimiento mucho más global del estado de la red, tanto a la hora de evaluar la disponibilidad de recursos como para analizar el efecto de una nueva admisión sin la necesidad de un intercambio periódico de información (con el *overhead* asociado). Esta naturaleza bajo demanda resulta determinante para considerar como más apropiado un **protocolo reactivo**.

Si bien desarrollar un ejemplo práctico implica tomar decisiones específicas adaptadas a los escenarios considerados, se ha pretendido generalizar el diseño en la medida de lo posible sin asociarlo a un contexto de aplicación determinado. En consecuencia, se ha tomado como partida una red completamente ad hoc, en la que todos los terminales de la red puedan desempeñar el mismo papel dentro del encaminamiento, sin la existencia de nodos con capacidades superiores o funcionalidades específicas (como los *clusterheads* en estructuras jerárquicas). Teniendo esto en cuenta, la elección de un **algoritmo de encaminamiento de estructura plana**, proporciona mayor flexibilidad al diseño final.

Un algoritmo basado en una **métrica vector-distancia** (DV) exige mayor complejidad de desarrollo para evitar ciertos problemas potenciales (creación de lazos). Sin embargo, resulta más apropiado para un cómputo distribuido de la ruta. Por otra parte, puesto que cada nodo puede mantener información parcial de la ruta hacia el destino (siguiente salto, slots TDMA seleccionados para transmitir...) no se requiere un direccionamiento desde la fuente (en la cabecera de los paquetes) limitando así el *overhead* asociado y dando mayor versatilidad a la composición de múltiples rutas.

Un protocolo de encaminamiento capaz de encontrar **múltiples rutas** hacia un mismo destino presenta una clara ventaja a la hora de permitir la selección de aquél de los caminos que mayores garantías ofrezca de proporcionar la esperada QoS. Los principales objetivos perseguidos con una estrategia multicamino son el mantenimiento de rutas alternativas en previsión de roturas, el balanceo de carga o la diversidad de caminos de transmisión. Sin embargo, ya que asegurar las prestaciones deseadas implica realmente la reserva efectiva de los recursos asociados, el mantenimiento de varios caminos no resulta eficiente. Por otra parte, la complejidad asociada a la actualización de cada uno de ellos (máxime de aquéllos que no se mantienen realmente activos) se

incrementa al tener en cuenta consideraciones de QoS. De cualquier modo, la ventaja potencial de una búsqueda multicamino se presenta como una opción atractiva para el diseño propuesto.

Sobre estas bases, el protocolo utilizado como punto de partida en el desarrollo de la propuesta final es el AODV, incluyendo ciertas consideraciones de su versión multicamino, el AOMDV. A continuación se muestran las modificaciones y consideraciones prácticas llevadas a cabo durante la realización final, para establecer la comunicación del protocolo con el nivel de enlace y definir adecuadamente la nueva métrica de encaminamiento.

3.3.1. *Ad hoc On-demand Distance Vector* (AODV)

El algoritmo *Ad hoc On-demand Distance Vector* (AODV)[7] es capaz de establecer y mantener dinámicamente rutas multisalto entre diferentes nodos que requieren de una comunicación extremo a extremo en una red ad hoc, permitiendo a su vez una rápida reacción ante variaciones en la topología de la red. El protocolo se basa en la obtención bajo demanda de rutas hacia los destinos deseados no siendo necesario mantener almacenados caminos en los que no existen comunicaciones activas. De hecho, dichas rutas únicamente se mantiene activas mientras son necesarias por las fuentes de las comunicaciones implicadas. Una de las características principales del protocolo es la utilización de los números de secuencia, concepto heredado del protocolo DSDV, para identificar en cualquier momento la información de encaminamiento más reciente (número de secuencia mayor). Cada nodo mantiene un número de secuencia monotónicamente creciente que le identifica, y a cada destino de la tabla de rutas se le asocia igualmente el número de secuencia conocido más reciente. Gracias precisamente a la utilización de dichos números de secuencia, el AODV permite establecer rutas libres de lazos (*loop-free*), ofreciendo a su vez un tiempo de convergencia razonable evitando el problema de la "cuenta al infinito"[8] de Bellmand-Ford. La utilización de una métrica DV, y el consecuente establecimiento de rutas mediante cooperación de nodos intermedios (cada uno almacenando el siguiente salto en la ruta), reduce el *overhead* que se introduciría

[7] `http://moment.cs.ucsb.edu/AODV/aodv.html`

[8] Otros autores han demostrado la inexactitud de esta afirmación, teniendo en cuenta su dependencia con una correcta configuración de los parámetros propios del protocolo, ante lo que presentan nuevas alternativas [Garcia-Luna-Aceves y Rangarajan, 2004; Mosko y Garcia-Luna-Aceves, 2006; Rangarajan y Garcia-Luna-Aceves, 2005a,b]. Dado que no es el objetivo final del trabajo descrito en esta memoria, no se ha tenido en cuenta dicha limitación, asumiendo la correcta implementación del protocolo AODV seleccionado.

en caso de utilizar una estrategia *source-routing* (como en el DSR), proporcionado por lo tanto gran escalabilidad.

AODV permite realizar un encaminamiento tanto *unicast*, como *multicast* (*Multicast Ad hoc On-demand Distance Vector* (MAODV) [Belding-Royer y Perkins, 2000]), si bien únicamente nos centraremos en el primero de ellos. Por otra parte, existe una versión multicamino del protocolo (AOMDV) de la que se ha considerado la capacidad para buscar múltiples rutas, como ya se ha comentado, por lo que se describirá más adelante el empleo particular de esta capacidad. No obstante, el diseño final no se basa en una implementación completa del algoritmo multicamino sino en el aprovechamiento de parte de su funcionalidad. A continuación se muestra, de un modo simplificado, el funcionamiento básico del protocolo. Una descripción detallada del mismo puede encontrarse en [Perkins y otros, 2003]:

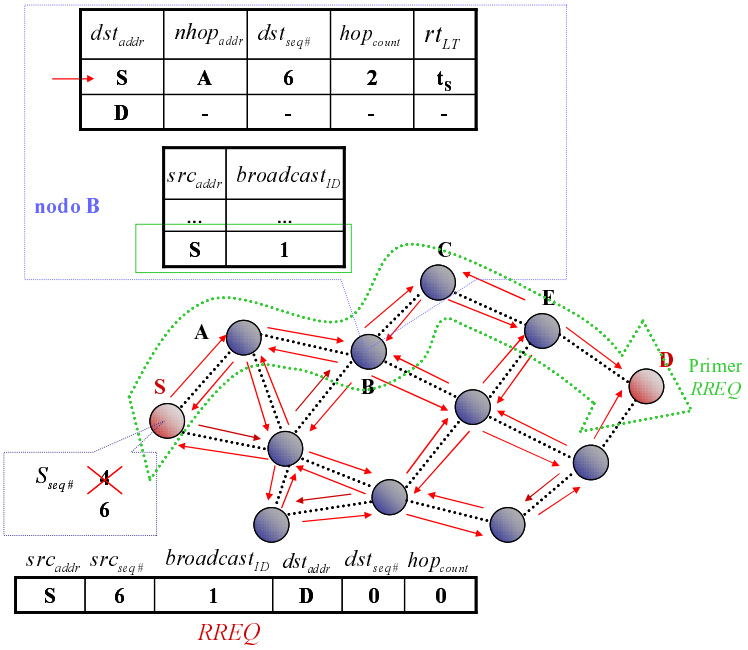
Descubrimiento de ruta

AODV construye las rutas basándose en un mecanismo de interrogación mediante petición / contestación (*route request / route reply query*, figura 3.3). Cuando un nodo fuente desea establecer una comunicación con un destinatario concreto, y no existe ninguna ruta activa, genera un mensaje de petición *Route REQuest* (*RREQ*) que será retransmitido en modo *broadcast* a través de toda la red (inundación - *flooding*). La información contenida en dicho mensaje especifica los campos mostrados en (3.4).

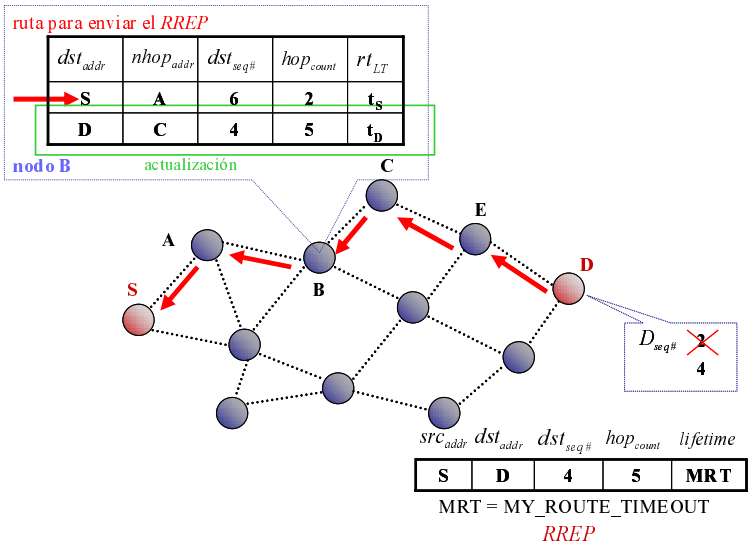
$$RREQ = \langle src_{addr}, src_{seq\#}, broadcast_{ID}, dst_{addr}, dst_{seq\#}, hop_{count} \rangle \tag{3.4}$$

El par $\langle src_{addr}, broadcast_{ID} \rangle$ identifica la búsqueda en curso, puesto que hace referencia a la fuente de la misma y el contador $broadcast_{ID}$ es incrementado para cada nuevo $RREQ$ generado por dicho nodo. Esta identificación permite detectar y eliminar copias recibidas del mismo mensaje para evitar una inundación innecesaria de la red, así como la formación de lazos. Cuando un nodo recibe el $RREQ$ y no tiene en su tabla de rutas información reciente de la ruta demandada (de acuerdo a los números de secuencia), reenvía dicho mensaje actualizando la información relativa al número de saltos hop_{count}. Si dicho nodo es el propio destinatario, o un nodo intermedio con una ruta actualizada hacia el mismo, responde con un mensaje *Route REPly* ($RREP$), cuyos campos se describen en (3.5).

$$RREP = \langle src_{addr}, dst_{addr}, dst_{seq\#}, hop_{count}, lifetime \rangle \tag{3.5}$$



(a) Descubrimiento de ruta (*broadcast*)



(b) Confirmación del establecimiento de la ruta (*unicast*)

Figura 3.3. Funcionamiento del protocolo AODV, *Request/Reply cycle*. MRT $= 2 \cdot$ ART, ACTIVE_ROUTE_TIMEOUT (ART). t_S y t_D tiempos de vida de acuerdo a lo indicado en el *RREQ* y *RREP* respectivamente.

Cada nodo almacena una tabla de rutas con información de los posibles destinos hacia los que mantiene comunicaciones activas (ya sea nodo fuente o intermedio). Puesto que el protocolo funciona bajo demanda, únicamente se crean y almacenan aquellas rutas que realmente están en uso. Dado que el AODV es un encaminamiento basado en una métrica vector-distancia la ruta completa se constituye por concatenación de la información del siguiente salto almacenada en cada nodo intermedio. Una entrada de la tabla de rutas almacena, por lo tanto, los siguientes campos:

- Dirección IP del destinatario (dst_{addr}).
- Dirección IP del siguiente salto en la ruta ($nhop_{addr}$).
- Número de saltos (hop_{count}) (**métrica**).
- Número de secuencia del destinatario ($dst_{seq\#}$).
- Vecinos activos que utilizan dicha ruta.
- Tiempo de vida (rt_{LT}).

Durante la transmisión del $RREQ$, en cada nodo intermedio se crea (o actualiza, si existe) la entrada de la tabla correspondiente a la ruta inversa hacia la fuente – ***Reverse Route*** **(**RR**)**. El número de secuencia $src_{seq\#}$ se utiliza para mantener la información de esta ruta actualizada. Es necesario mantenerla activa al menos el tiempo suficiente para que el $RREQ$ atraviese la red y el $RREP$ vuelva hacia la fuente (utilizando dicha ruta). En la fase de transmisión del $RREP$, se actualiza la verdadera ruta demandada – ***Forward Route*** **(**FR**)**. Al igual que en la RR, el número de secuencia $dst_{seq\#}$ permite mantener la información de dicha ruta actualizada.

El número de secuencia del destinatario se corresponde con el último conocido de dicha ruta. Los números de secuencia asociados a los nodos fuente ($src_{seq\#}$) y destino ($dst_{seq\#}$), recibidos en los paquetes $RREQ$ y $RREP$ respectivamente, permiten actualizar correctamente la información asociada a las rutas hacia cada uno de dichos nodos (*Reverse Route* y *Forward Route* respectivamente). Cuando se recibe nueva información relativa a una entrada de la tabla, se considera información "nueva" si se corresponde con un número de secuencia del destinatario mayor o, en caso de igualdad, una métrica menor. De acuerdo con este criterio, a continuación se enumeran las reglas de actualización de los números de secuencia y de las entradas a la tabla.

1) **Mantenimiento de números de secuencia:** El número de secuencia asociado al destinatario de una entrada de la tabla de rutas debe actualizarse cada vez que se recibe información nueva acerca de dicho número contenida en un mensaje de

encaminamiento. Los propios nodos de la red mantienen actualizado el número de secuencia asociado a ellos mismos, que es incrementado en dos situaciones[9]:

- En el momento de iniciar una nueva búsqueda de ruta.
- Antes de enviar un $RREP$ en respuesta a un $RREQ$. El número de secuencia se incrementa igualándolo al valor máximo entre el número almacenado por el nodo y el relativo a él mismo que ha sido recibido en el $RREQ$.

Por otro lado, la actualización de los números de secuencia asociados a entradas de la tabla de rutas se realiza:

- En respuesta a la pérdida del siguiente salto (rotura del enlace) o cuando el tiempo de vida ha expirado.
- Cuando se recibe un mensaje AODV con nueva información relativa al número de secuencia del destino de la entrada (por ejemplo, al recibir un $RERR$ – mantenimiento de ruta, o el $src_{seq\#}$ recibido en un $RREQ$...)

2) **Actualización de las entradas de la tabla de rutas:** Cuando se recibe un nuevo mensaje AODV que afecta a una entrada de la tabla, la ruta asociada es actualizada únicamente si la información recibida es nueva de acuerdo al número de secuencia almacenado, es decir:

- Si la entrada a la tabla no es válida (se rompió anteriormente, o expiró el tiempo de vida y todavía no se ha eliminado).
- Si el número de secuencia recibido es mayor que el almacenado.
- Si el número de secuencia recibido es igual, pero la métrica[10] es menor.
- Si se desconoce el número de secuencia (al añadir una entrada completamente nueva, no existe información previa del nuevo destino).

Mantenimiento de ruta

Debido a la movilidad de los terminales, algunos enlaces pertenecientes a rutas activas pueden desconectarse, provocando la rotura de dichas rutas y el consecuente

[9] Al igual que en DSDV se mantiene el criterio de incrementar los números de secuencia de modo par en los nodos y de modo impar en las entradas de la tabla para las que se detecten roturas.

[10] En caso de actualizar la información en un nodo y considerar la nueva ruta como válida, la métrica número de saltos se actualizaría de acuerdo a los saltos recorridos hasta el momento más el último hasta dicho nodo, es decir $RREQ(hop_{count}) + 1$.

corte de la comunicación. El protocolo debe reaccionar ante dicha eventualidad, para lo cual, existen dos posibilidades:

1) **Gestión autónoma** del protocolo de la conectividad local en su entorno. Se basa en la transmisión periódica de mensajes *broadcast* entre vecinos (*Time To Live* (TTL) = 1 salto), denominados mensajes de $HELLO$. La pérdida de conectividad puede ser detectada cuando mensajes esperados de $HELLO$ no son recibidos. El envío de mensajes de control periódicos entre los vecinos implica un incremento en el *overhead* introducido por el protocolo, máxime en entornos de elevada movilidad que exigen una frecuencia mayor para reaccionar apropiadamente a los cambios en la topología. No obstante, si los niveles inferiores no permiten realizar una detección eficiente de la pérdida del enlace, o no es posible su notificación (interfaz con el AODV), el propio protocolo puede realizar, de manera autónoma, el mantenimiento de la conectividad.

2) Existencia de un **mecanismo de niveles inferiores** que alerte de dicha desconexión[11]. Puesto que el diseño propuesto se basa en la cooperación entre capas, únicamente se ha considerado esta posibilidad, aprovechando la capacidad del protocolo MAC considerado para gestionar la conectividad local (anexo A). De hecho, dicho protocolo incluye un intercambio periódico de información (slot *Broadcast CHannel* (BCH)) que desempeña un papel análogo a los paquetes de $HELLO$ del AODV.

En cualquier caso, la pérdida de un vecino detectada por un nodo perteneciente a varias rutas activas es comunicada a las correspondientes fuentes que, en caso de necesitarlo, realizarán una nueva búsqueda de ruta (figura 3.4). La inactivación de las rutas en los nodos intermedios implicados supone actualizar su número de secuencia e invalidar la métrica (número de saltos = ∞). La notificación de pérdida de ruta se realiza mediante el envío de un mensaje de error *Route ERRor* ($RERR$) que contiene las direcciones IP de los n destinos inalcanzables y sus correspondientes números de secuencia[12]:

$$\left\langle undst^1_{addr}, undst^1_{seq\#}, ..., undst^n_{addr}, undst^n_{seq\#} \right\rangle \tag{3.6}$$

[11]Detección de pérdida de enlace: fallo persistente en la recepción del CTS tras el envío de un RTS – IEEE 802.11 – o fallo en la recepción de ACKs a los paquetes transmitidos al siguiente salto.

[12]Puesto que el $RERR$ debe llegar a todas las fuentes de rutas que incluyeran el enlace roto, el envío puede hacerse bien en modo *broadcast* o *unicast* iterativo.

El tiempo de vida de la ruta permite mantener un estado *soft-state* que identifica la "actualidad" de la misma. Mientras ésta permanece activa, su tiempo de vida es incrementado. En el caso de que cese el flujo de paquetes a través de ella se agota el tiempo de vida, momento en que se inactiva. Sin embargo, la ruta permanece todavía en la tabla, al menos durante un tiempo DELETE_PERIOD (DP). Esta permanencia de rutas inactivas permite mantener el último número de secuencia conocido hacia el destino al menos durante el tiempo suficiente para que todos aquellos nodos que pudieran estar usándola la inactiven. Es el caso, por ejemplo, de la rotura de algún enlace. Es necesario que todos los nodos puedan darse cuenta de dicho efecto y que, en caso de reiniciarse una nueva búsqueda (que irá asociada a un número de secuencia mayor) las actualizaciones de acuerdo a la nueva información se realicen correctamente. Esto es debido a la suposición de que las rutas libres de lazos y la eliminación de la "cuenta al infinito" se consiguen exclusivamente gracias al mantenimiento de los números de secuencia, siempre y cuando éstos "no se olviden". Dicha suposición, al borrar las entradas de la tabla, o producirse reinicios del sistema en los terminales, ya no sería cierta, pero mantener dicha información (número de secuencia) durante un "tiempo extra" suficiente, permite emular dicho efecto de "memoria"[13].

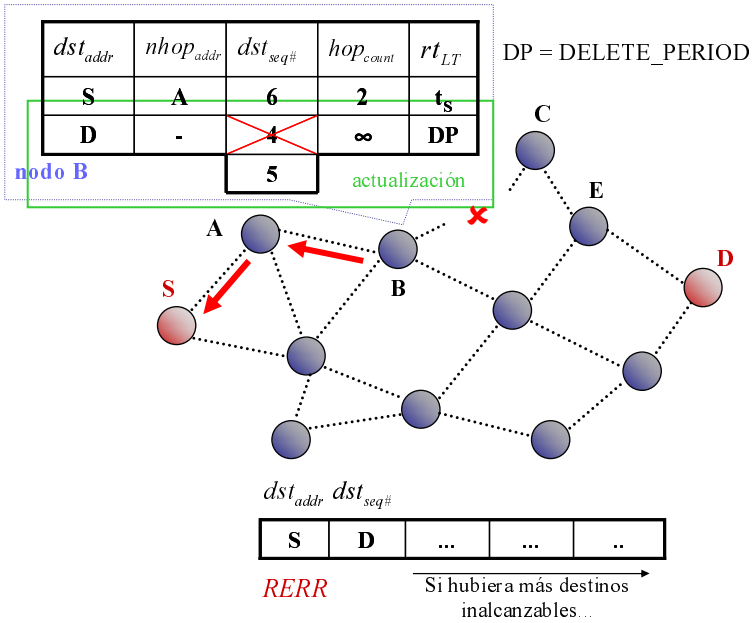


Figura 3.4. Detección de rotura en un enlace, envío de mensaje *RERR*.

[13]La necesidad de dicho efecto de memoria es la principal crítica a los algoritmos de encaminamiento teóricamente libres de bucles gracias a los números de secuencia, tal y como se describe en [Rangarajan y Garcia-Luna-Aceves, 2005a]

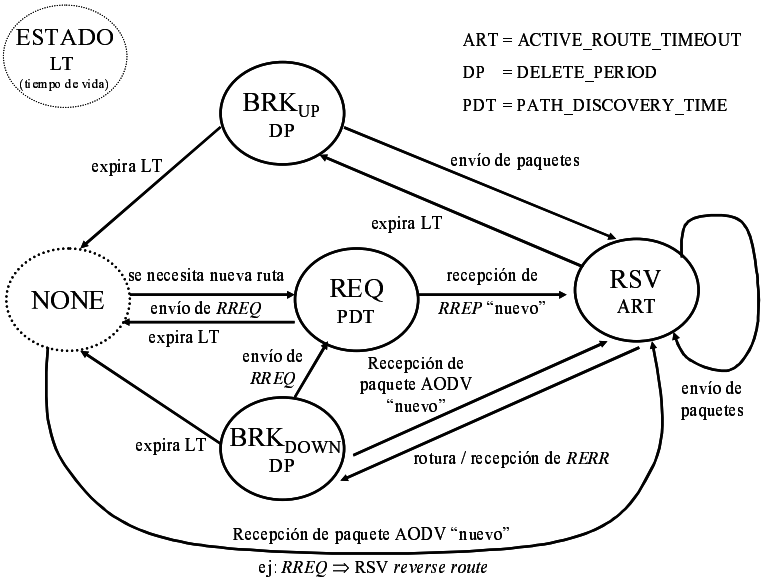


Figura 3.5. Diagrama básico de estados para una entrada de la tabla de rutas.

La figura 3.5 muestra un diagrama de los diferentes estados en los que se encuentra una entrada de la tabla (con la duración relativa de cada uno de ellos, desde su activación), de acuerdo a las fases de establecimiento y mantenimiento de la ruta. La notación empleada en dicha figura se corresponde con la siguiente definición: NONE = ruta inactiva; REQ = ruta en fase de búsqueda; $\mathrm{BRK_{UP}}$ = pérdida del enlace con el salto previo (no se han recibido paquetes); $\mathrm{BRK_{DOWN}}$ = pérdida del enlace con el siguiente salto (detección de rotura). ACTIVE_ROUTE_TIMEOUT (ART) es el tiempo que una entrada de la tabla permanece activa. Durante el tiempo DP una ruta inactiva (por rotura, o por cese del flujo de paquetes) permanece en la tabla de rutas, de manera que durante todo ese tiempo se mantiene la información del último número de secuencia conocido hacia el destino correspondiente. El PATH_DISCOVERY_TIME (PDT) hace referencia al tiempo que debe mantenerse un nodo en espera de recibir una respuesta del destino ante una petición de ruta (*RREQ*) antes de asumir que ésta no puede establecerse. Todos estos parámetros son configurables, siendo las prestaciones del protocolo altamente sensibles a los valores elegidos, dependientes en gran medida de las características de los protocolos inferiores, las tecnologías de radio, etc. [Perkins y otros, 2003]. Las consideraciones generales, sin embargo, a la hora de darles el valor apropiado, deben ser: El valor PDT ha de ser suficiente para que el paquete *RREQ* atraviese la red al completo, siendo así capaces de encontrar cualquier nodo destino dentro de la misma, y que el *RREP* alcance de nuevo al nodo fuente. Mantener la ruta activa un tiempo de vida

ART debe conservar las entradas de la tabla el tiempo suficiente para ser capaces de transmitir el flujo de paquetes asociado a la conexión (absorbiendo periodos de inactividad debidos a congestiones, o a la tasa variable de generación de paquetes), y el DP, como ya se comentó anteriormente, es un requisito indispensable para mantener el funcionamiento eficiente del protocolo, en cuanto al manejo de los números de secuencia de manera que se respete su propósito general (evitar lazos y "cuenta al infinito").

3.3.2. Métrica QoS. Cálculo del ancho de banda disponible en la ruta.

Con objeto de definir apropiadamente una métrica QoS que identifique la disponibilidad real de los recursos del sistema se ha integrado en el funcionamiento del protocolo de encaminamiento un procedimiento de cálculo distribuido del ancho de banda a lo largo de la ruta completa, basándose en la propuesta inicial de Zhu y Corson [2002]. Dicha propuesta se apoya en una capa MAC TDMA desarrollada por los mismos autores. Su funcionamiento no difiere conceptualmente de la capa MAC TDMA que se ha seleccionado finalmente, por lo que el procedimiento responde a las necesidades aquí presentadas. La implementación final, no obstante, es el resultado de ciertas modificaciones llevadas a cabo en respuesta a la diversa problemática estudiada, así como al propio desarrollo *cross-layer* considerado. Dichas modificaciones se explicarán detalladamente a lo largo del proceso de análisis descrito en este capítulo.

La medición del ancho de banda se basa en encontrar el número de slots TDMA disponibles para transmitir en cada uno de los enlaces que conforman el camino de transmisión entre la fuente y el destino de la comunicación. La complejidad en la definición de "disponible" en un contexto TDMA reside en la interdependencia de todos los enlaces activos de la red, dada la naturaleza interferente de un escenario inalámbrico. Así pues, el problema de determinar el ancho de banda disponible extremo a extremo no se limita en este caso a encontrar el enlace "cuello de botella". La idea básica del algoritmo de Zhu y Corson consiste en encontrar en los diferentes enlaces slots libres de colisiones, en el caso de que finalmente sean reservados. Para ello, debe poder resolverse completamente el problema del terminal oculto. Resolviendo además el problema del terminal expuesto puede maximizarse el reúso, en la medida que lo permite el funcionamiento distribuido del protocolo.

3.3.2.1. Algoritmo de cálculo del ancho de banda E2E

El cálculo del ancho de banda de una ruta completa, basado en el algoritmo denominado *BandWidth Calculation – Forward Algorithm* (*BWC-FA*) consiste en la determinación de los slots disponibles en cada enlace de manera iterativa, salto a salto, desde el nodo fuente hacia el destino. El problema del terminal oculto limita a una distancia de tres saltos la reutilización del mismo slot (figura 3.6), lo que determina la necesidad de definir conjuntos disjuntos de slots cada tres saltos. Sobre dicha base, el algoritmo propuesto realiza el cálculo "hacia delante" (*forward*) a partir de la limitación impuesta por los saltos previos. El número mínimo de slots en cada enlace determina el ancho de banda máximo que puede garantizarse en el camino completo, lo que limita la dimensión máxima con la que se define cada conjunto de slots disponibles (un ancho de banda de R slots extremo a extremo implica la existencia de al menos ese número de slots en todos los enlaces del camino).

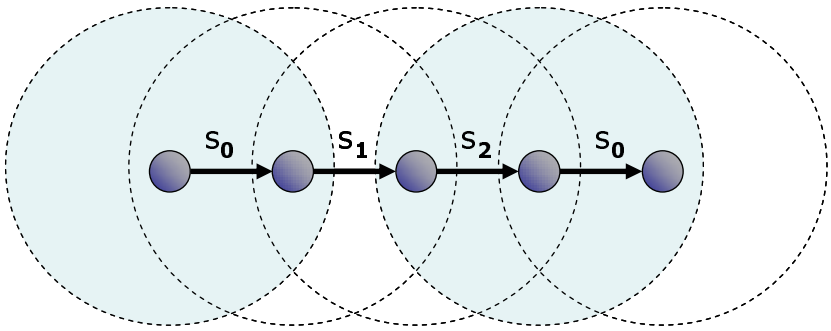


Figura 3.6. Reutilización de slots. Terminal oculto.

Teóricamente, una capa de acceso al medio capaz de garantizar la eliminación de los problemas del terminal oculto y el terminal expuesto gracias a la señalización adecuada (como es el caso del ADHOC MAC) permite realizar una reserva de los recursos totalmente libre de colisión una vez que estos se seleccionan adecuadamente (*BWC-FA*) maximizando el reúso. Esto, como se comprobará en §4, no es cierto en un escenario realista que considera la interferencia real de los usuarios y los efectos derivados de la movilidad. Sin embargo, antes de tener en cuenta estas limitaciones, este capítulo se centra en el estudio de la propuesta *cross-layer* como mecanismo de control de admisión distribuido, a evaluar en un contexto próximo al ideal (**escenario estático, *Protocol Model***). Esto permitirá una descripción detallada de la estructura genérica, que deberá mejorarse posteriormente cuando se introduzcan los factores no tenidos en cuenta.

Se pretende así caracterizar la interacción entre el AODV y el ADHOC MAC así como introducir los procedimientos básicos de la gestión de los recursos.

La información de disponibilidad proporcionada por el nivel MAC se basa en la definición de los conjuntos SRT y SRR descritos anteriormente – (3.2) y (3.3) respectivamente. No obstante, dicha definición se ha modificado de acuerdo a las reglas específicas del protocolo ADHOC MAC teniendo en cuenta la capacidad del protocolo de reasignar recursos de servicios menos prioritarios para nuevas conexiones con demandas de QoS más prioritarias. La descripción detallada de los mecanismos de reasignación y la definición particular de "disponible" dada por el protocolo se describen en el anexo A, donde se redefinen el SRT y el SRR siguiendo la notación utilizada en la descripción del protocolo – (A.9) y (A.10).

Notación:

- Se define $P = \{n_m, n_{m-1}, ..., n_0\}$ como un camino donde los nodos n_m y n_0 son respectivamente la fuente y el destino de la comunicación.
- $FP^k = \{n_m \rightarrow n_{m-1} \rightarrow ... \rightarrow n_{k+1} \rightarrow n_k\}$ es el camino parcial de P comenzando en el nodo fuente y hasta el nodo n_k, de manera que $FP^0 = P$.
- Se define PB_i^k como el conjunto de slots usados en el enlace $(n_i \rightarrow n_{i-1})$ en el camino parcial FP^k, con $m \leq i \leq k+1$.
- La disponibilidad de recursos en en el enlace $(n_{k+1} \rightarrow n_k)$ se define como $LB_{k+1} = SRT_{k+1} \cap SRR_k$, donde SRT_{k+1} y SRR_k se corresponden con $SRT_{k+1}^{p_0}$ y $SRR_k^{p_0}$, de acuerdo a las definiciones de disponibilidad dadas en (A.9) y (A.10), siendo p_0 la máxima prioridad establecida en el nivel MAC.[14].

El cálculo del ancho de banda extremo a extremo en el camino P se realiza iterativamente de acuerdo al algoritmo propuesto por Zhu y Corson y descrito a continuación[15]:

1. Si $m = 1$

$$PB_1^0 = LB_1 \tag{3.7}$$

[14]Como primera aproximación, los servicios QoS a admitir se consideran prioritarios, permitiendo reasignación de recursos de menor prioridad para darles servicio. No obstante, en §4.2.3 se analiza la posibilidad de flexibilizar esta priorización con objeto de realizar una gestión más apropiada para un entorno dinámico.

[15]Las funciones BW_1,BW_2 y BW_3 se detallan en el apéndice B.

2. Si $m = 2$

$$\left(PB_2^0, PB_1^0\right) = BW_2\left(LB_2, LB_1\right) \tag{3.8}$$

3. Si $m \geq 3$

$$\left(PB_m^{m-2}, PB_{m-1}^{m-2}\right) = BW_2\left(LB_m, LB_{m-1}\right) \tag{3.9}$$

desde $k = m - 3$ hasta 0:

$$\left(PB_{k+3}^k, PB_{k+2}^k, PB_{k+1}^k\right) = BW_3\left(PB_{k+3}^{k+1}, PB_{k+2}^{k+1}, LB_{k+1}\right) \tag{3.10}$$

De manera que el ancho de banda disponible en el camino FP^k viene determinado por:

$$BW\left(FP^k\right) = \left|PB_{k+1}^k\right| \tag{3.11}$$

y el ancho de banda extremo a extremo del camino completo $P = FP^0$:

$$BW\left(P\right) = BW\left(FP^0\right) = \left|PB_1^0\right| \tag{3.12}$$

3.3.2.2. Integración en el protocolo de encaminamiento

El algoritmo descrito permite calcular el ancho de banda para un camino P dado. Sin embargo, necesita de un protocolo de encaminamiento con el que funcionar conjuntamente para constituir un verdadero encaminamiento con QoS.

La propuesta de Zhu y Corson aprovecha la característica bajo demanda del AODV para integrar el algoritmo dentro del procedimiento de búsqueda de la ruta. El funcionamiento, brevemente descrito a continuación, se basa en propagar la información necesaria para computar iterativamente el ancho de banda extremo a extremo desde la fuente n_m hasta el destino n_0, a lo largo del camino P, durante la fase de *route discovery*. Cada vez que un nodo intermedio recibe un *RREQ* el algoritmo *BWC-FA* permite calcular el ancho de banda en el camino parcial desde la fuente al nodo:

1. Cuando un nodo n_{k+1} propaga un *RREQ* añade en el mismo la siguiente información: $\left\langle PB_{k+3}^{k+1}, PB_{k+2}^{k+1}, SRT_{k+1}\right\rangle$

2. Si un nodo n_k recibe dicha información, calcula:

$$LB_{k+1} = SRT_{k+1} \cap SRR_k \tag{3.13}$$

$$\left(PB_{k+3}^{k}, PB_{k+2}^{k}, PB_{k+1}^{k}\right) = BW_3\left(PB_{k+3}^{k+1}, PB_{k+2}^{k+1}, LB_{k+1}\right) \tag{3.14}$$

Para $k = m-1$ la ecuación (3.14) se sustituye por $PB_m^{m-1} = LB_m$, y en el caso de $k = m-2$ por $\left(PB_m^{m-2}, PB_{m-1}^{m-2}\right) = BW_2\left(LB_m, LB_{m-1}\right)$ El ancho de banda del camino parcial FP^k es $BW\left(FP^k\right) = \left|PB_{k+1}^{k}\right|$. Puesto que el objetivo del protocolo de encaminamiento es encontrar caminos con el número de slots suficientes R para satisfacer las demandas, si $BW\left(FP^k\right) < R$ el $RREQ$ ya no se propaga, reduciendo así el *overhead* en la red. Si existe suficiente ancho de banda, se propaga con la información $\left\langle PB_{k+2}^{k}, PB_{k+1}^{k}, SRT_k\right\rangle$. En definitiva, el cálculo parcial $BW\left(FP^k\right)$ constituye la **métrica de encaminamiento QoS** utilizada por los nodos intermedios como parámetro decisorio a la hora de propagar la información de control.

3. Cuando el $RREQ$ llega al destino, el cálculo del $BW\left(FP^0\right)$ es el ancho de banda extremo a extremo del camino completo P desde n_m hasta n_0. Si dicho ancho de banda es $BW\left(FP^0\right) \geq R$, el destino envía un mensaje de $RREP$. Nuevos $RREQ$s son ignorados para reducir la latencia de descubrimiento de ruta, si bien la solución obtenida puede no resultar óptima ni en número de saltos ni en ancho de banda. Dado que el cálculo definitivo del ancho de banda en el camino completo se obtiene en el extremo final, es únicamente el nodo destino el que puede contestar ($RREP$) estableciendo finalmente el camino, a diferencia del funcionamiento básico del AODV, donde un nodo intermedio con una ruta actualizada hacia el destino puede contestar sin reenviar la petición hacia dicho nodo.

4. Los slots específicos que debe usar cada nodo transmisor son seleccionados por su receptor, mientras el $RREP$ atraviesa la ruta inversa hacia el nodo fuente:

   Un nodo n_k determina los slots TS_{k+1}^{P} en que debe transmitir el vecino n_{k+1} en el enlace $(n_{k+1} \rightarrow n_k)$ dentro del camino P. Para ello, utiliza la información obtenida del $RREP$: $\left\langle TS_k^P, TS_{k-1}^P\right\rangle$:

$$TS_{k+1}^{P} = BW_1\left(PB_{k+1}^{k} \cap \overline{TS_k^P} \cap \overline{TS_{k-1}^P}, R\right) \tag{3.15}$$

   El nodo n_k reenvía el $RREP$ sustituyendo: $\left\langle TS_{k+1}^P, TS_k^P\right\rangle$.

3.3.2.3. Modificaciones y consideraciones adicionales

A la hora de asegurar la admisión de una nueva conexión, dada la disponibilidad del ancho de banda demandado, deben reservarse enlace a enlace los slots determinados durante el encaminamiento. Así pues, la interacción con el nivel MAC debe promover la reserva de los TS_i seleccionados en cada uno de los enlaces conforme el mensaje de $RREP$ alcanza cada nodo intermedio. No obstante, la variabilidad de una red ad hoc puede cambiar las condiciones de interferencia sobre las que se asumió la disponibilidad de determinados slots (§4), lo que puede provocar que en alguno de los enlaces no resulte posible reservar el número suficiente de slots para el TS_i, bloqueando consecuentemente la conexión. La determinación que el receptor hace de los R slots en concreto para transmitir al definir el TS_i, limita las posibilidades de selección cuando los conjuntos disponibles PB_i tienen en realidad una dimensión suficientemente mayor. De hecho, un conflicto puntual en uno de los R slots, percibido por parte del transmisor, podría solventarse con la elección de otro de los disponibles dentro del PB_i, pero dicha selección ya fue realizada por el receptor, por lo que cualquier cambio implicaría una comunicación adicional entre los dos extremos del enlace para buscar una solución al conflicto.

Con objeto de incrementar la probabilidad de acierto en la selección del conjunto TS_i, se ha optado por la propagación de la información completa, el PB_i. Puesto que la versión más actualizada de dicho conjunto se encuentra tres saltos más allá en la dirección del destino, se ha modificado completamente la información propagada en el $RREP$ para que cada nodo n_k pueda disponer del conjunto PB_k^{k-1} más actualizado posible para elegir convenientemente los R slots dentro del mismo, una vez actualizado de acuerdo a su SRT_k (modificable dadas las condiciones variantes del entorno). De igual modo, a la hora de propagar hacia la fuente dicha información actualizada, cada nodo recalcula el PB_{k+1}^k de acuerdo a las variaciones experimentadas en su SRR_k. Así pues, un nodo n_k que recibe de su siguiente salto n_{k-1} $\left\langle PB_{k+2}^{k-1}, PB_{k+1}^{k-1}, PB_k^{k-1} \right\rangle$, actualiza los conjuntos que originariamente calculó, propagando así el $RREP$ $\left\langle PB_{k+3}^k, PB_{k+2}^k, PB_{k+1}^k \right\rangle$:

$$PB_k^{k-1} = PB_k^{k-1} \cap SRT_k \tag{3.16}$$

$$PB_{k+1}^k = PB_{k+1}^{k-1} \cap SRR_k \tag{3.17}$$

$$PB_{k+2}^k = PB_{k+2}^{k-1} \tag{3.18}$$

Por otra parte, n_k selecciona los slots en los que transmitir, TS_k^P:

$$TS_k^P = BW_1\left(PB_k^{k-1}, R\right) \tag{3.19}$$

La figura 3.7 muestra un ejemplo del procedimiento completo de cálculo del ancho de banda y la selección de los slots de transmisión por enlace para un camino P de 5 saltos [Canales y otros, 2005a].

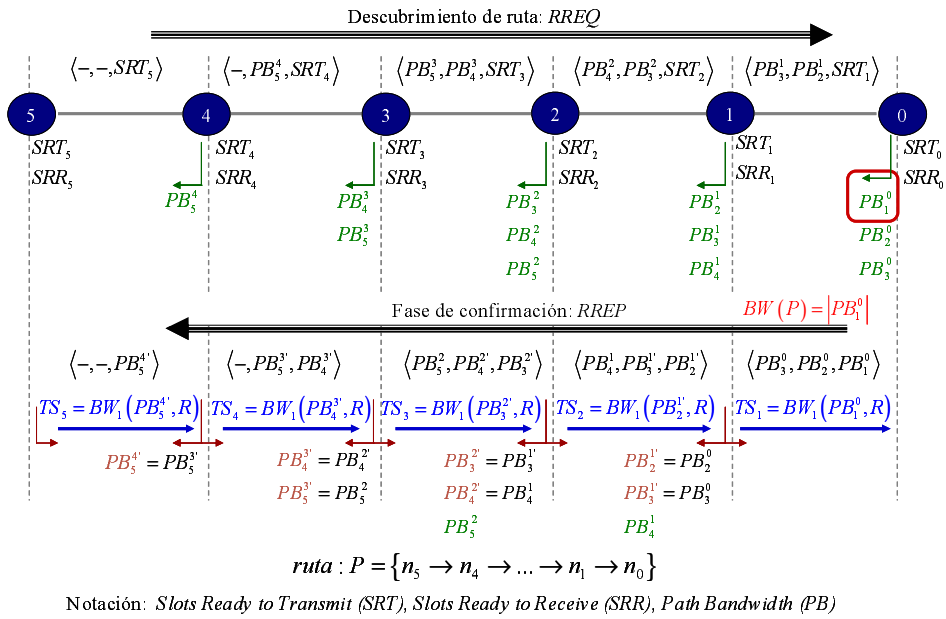


Figura 3.7. Ejemplo del cálculo de ancho de banda integrado en el encaminamiento. Algoritmo *BWC-FA*.

3.3.2.4. Soluciones a los problemas derivados de la reserva *backward* en la búsqueda bajo demanda

Cuando se realiza una selección bajo demanda de los recursos a reservar durante la fase de búsqueda de ruta, pero la reserva efectiva se realiza únicamente cuando se recibe el *RREP*, pueden darse ciertas situaciones de conflicto, lo que en [Jawhar y Wu, 2005a] se enuncia como *race condition* y reservas paralelas.

***Race condition*:** Cuando un nodo intermedio toma parte en la búsqueda de dos caminos en paralelo, si la asignación de un slot como disponible para el primero de ellos fuera únicamente propagada en el mensaje de *RREQ*, sin almacenarla en dicho nodo, el mismo recurso podría volver a asignarse a una nueva conexión en proceso de búsqueda (ya que la primera todavía no se habría hecho efectiva) – figura 3.8(a).

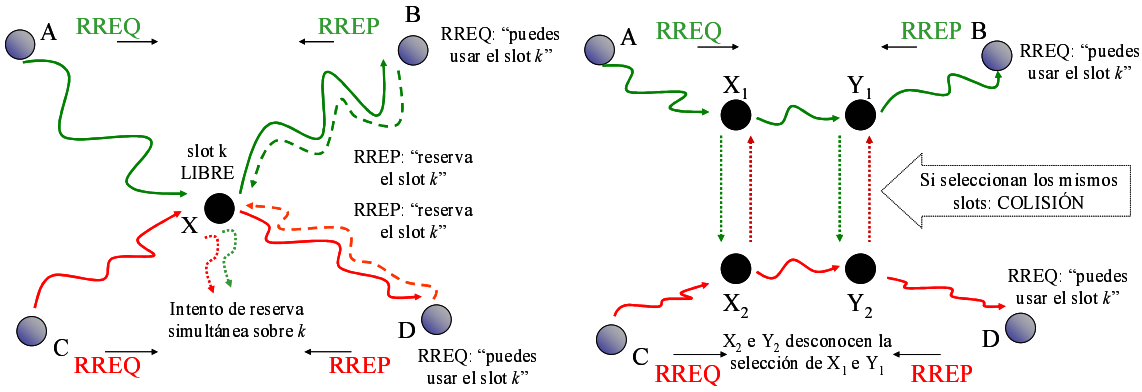
Tras recibir el primer *RREP* la reserva se confirmaría para la conexión asociada. Sin embargo, el slot seleccionado ya no estaría disponible para la segunda, por lo que su reserva fallaría bloqueando la conexión (no admisión por falta de ancho de banda). En caso de que hubiera más slots disponibles se habría producido un bloqueo innecesario. La degradación de las prestaciones debida al intento de reservas simultáneas en un nodo intermedio es lo que se denomina *race condition*. En [Jawhar y Wu, 2005a] se propone como solución tener en cuenta la "futura ocupación" (estado *ALLOCATED*) de un slot seleccionado como disponible para una conexión, considerándolo inaccesible ante nuevas peticiones de recursos. Tal como proponen los autores, se resuelve completamente la *race condition*. Sin embargo, esto es posible considerando que la elección de los slots se realiza en la misma fase de búsqueda, limitándose exactamente al número de recursos demandados para la conexión, sin esperar el *RREP*, a diferencia de lo propuesto en el algoritmo *BWC-FA* modificado[16].

En nuestra propuesta, por el contrario, se ha considerado que una selección ajustada a la demanda específica, durante la fase de búsqueda, resulta ser demasiado restrictiva y poco flexible, pudiendo limitar la aceptación de conexiones debido a una mala selección de los slots, especialmente en situaciones de baja disponibilidad. Por ejemplo, un nodo con 3 slots disponibles, elige de entre ellos uno que en uno o dos saltos posteriores resulta ser el único disponible, impidiendo la reserva que, de elegir apropiadamente otro de sus slots disponibles, habría sido posible. Con ese propósito se ha implementado el algoritmo *BWC-FA*, para medir adecuadamente la disponibilidad máxima de recursos. De hecho, incluso la selección de los slots se realiza con la mayor información posible, a diferencia de la propuesta original de Zhu y Corson, como se ha descrito anteriormente. Teniendo en cuenta dichas consideraciones, resulta impracticable una solución basada en un estado de asignación para los slots antes de su reserva ya que, al determinarse todos los disponibles, una única conexión bloquearía cualquier otra que se buscara en paralelo. No obstante, el hecho de realizar la asignación real cuando el mensaje *RREP* es recibido por el nodo intermedio, flexibiliza la selección de los slots reduciendo la probabilidad de conflicto dada por la *race condition* tal como ésta se ha descrito. Por otra parte, si bien la solución de Jawhar y Wu no es directamente aplicable, la gestión de la base de datos compartida entre el encaminamiento y el nivel de acceso supone la implementación indirecta de una solución parcial: una vez que un nodo intermedio actualiza como válida la entrada asociada a la conexión en la tabla de rutas, indicando

[16]En [Jawhar y Wu, 2005a], el concepto *backward* es únicamente aplicado a la materialización de la reserva y no a la selección de los slots específicos.

en ella los slots asignados, dichos slots pueden ser identificados por el MAC como no disponibles[17] [Canales y otros, 2005a].

Problema de reservas paralelas: Cuando dos caminos se buscan en paralelo, sin nodos intermedios comunes, la selección particular de slots realizada para uno de ellos es desconocida por los nodos implicados en el otro camino. En estos escenarios, si algunos de los nodos intermedios pertenecientes a caminos diferentes son vecinos a un salto, en caso de producirse la selección simultánea del mismo recurso, la posterior reserva produciría un conflicto, tal como se muestra en la figura 3.8(b). En este caso, si la reserva fuera efectiva en uno de los caminos (siendo detectada por el vecino a un salto del camino paralelo), la segunda conexión no podría realizar la reserva, bloqueándose.



(a) *Race Condition* (b) Reservas paralelas

Figura 3.8. Problemas derivados de la selección simultánea del mismo recurso (slot k) en diferentes caminos.

La solución propuesta por Jawhar y Wu se basa precisamente en aprovechar la vecindad de los nodos que entrarían en conflicto y que, consecuentemente, con capaces de comunicarse entre sí correctamente. Cada nodo transmite periódicamente, y ante cualquier cambio, la información de estado de sus slots (incluyendo el estado *ALLOCATED*), previniendo así la aparición de futuros conflictos. Nuevamente, puesto que la selección real de los recursos propuesta en nuestro diseño no se realiza en la fase de búsqueda, la solución vuelve a no ser aplicable. Sin embargo, la transmisión de información de estado descrita es análoga al empleo del esquema de acceso BIAS del ADHOC MAC propuesto en este trabajo[18]. Una vez que existe la entrada a la tabla de rutas que asocia la información de slots asignados a una conexión específica, que es compartida

[17]Explicado con detalle en §3.4.

[18]El funcionamiento del esquema de acceso *Book In Advance Scheme* (BIAS) se describe en el anexo A.

por el nivel de enlace, se invoca la reserva de los recursos. Dicha reserva consiste en marcar los slots en el FI como ACCEDIENDO (información de estado, ver anexo A), indicando el enlace específico, hasta que la reserva sea efectiva [Canales y otros, 2005a]. Si bien la transmisión comienza una vez recibido el *RREP*, la flexibilización en la selección de slots conseguida al utilizar el conjunto total de disponibles, actualizado nuevamente al recibir el *RREP*, permite que la segunda conexión en intentar reservar descarte los slots señalizados en el BCH de su vecino perteneciente al camino paralelo. De este modo, el bloqueo se reduce a aquellas situaciones de elevada congestión donde encontrar slots adicionales resulte imposible. En caso de bloquearse una conexión, con objeto de facilitar la admisión de una posible comunicación coincidente, se genera un paquete específico de liberación de recursos en dirección al nodo destino (*RELEASE packet*). De este modo se garantiza que aquellos recursos reservados una vez que el *RREP* activó el mecanismo BIAS en el nodo correspondiente puedan asignarse a nuevos flujos QoS.

3.3.2.5. Funcionamiento multicamino

En el funcionamiento básico del AODV, el nodo destino únicamente contesta con un *RREP* al primero de los *RREQ* que recibe, lo que no determina el "mejor" de los caminos posibles desde la fuente, aunque minimiza la latencia de descubrimiento de ruta. Sin embargo, puesto que el objetivo real de un control de admisión eficiente consiste en realizar (en la medida de lo posible) un reparto más apropiado de los recursos para las conexiones a lo largo de las diferentes rutas, debería realizarse una selección más cuidadosa de dichos caminos. Con esta finalidad se ha seguido el planteamiento multicamino del AOMDV, de manera que el destino cuente con la información del ancho de banda de diversas rutas pudiendo así seleccionar una de ellas según la métrica apropiada. Para ello, una vez recibido el primer *RREQ* asociado a una nueva conexión, el destino espera a recibir nuevos mensajes durante un tiempo configurable t_{wait}.

El protocolo AOMDV determina diversos caminos alternativos entre la fuente y el destino aceptando tanto diversos *RREQ*s para conformar posibles rutas inversas hacia la fuente, como diversos *RREP*s que serán los que determinen las múltiples rutas hacia el destino. La aceptación de diferentes copias de los mensajes de control podría dar lugar a la formación de bucles de encaminamiento, por lo que en dicho protocolo las reglas básicas de actualización de rutas del AODV se han modificado ligeramente para evitar esa posibilidad. Al igual que en el AODV, se propaga el primer *RREQ* recibido, para evitar la inundación de la red, puesto que las diferentes copias

únicamente sirven para determinar los caminos hacia el nodo fuente y, en la dirección destino, sólo aportan la información del nodo que los ha propagado (lo que ya consta en la primera copia enviada). En el caso que nos ocupa, dado que un *RREQ* incluye información relativa al ancho de banda de diferentes enlaces previos, dicha única copia no permitiría disponer de la información necesaria para los diferentes caminos. Sin embargo, si se pretendiera difundir toda esa información, los diversos *RREQ*s asociados introducirían un *overhead* excesivo. Por lo tanto, al igual que ambos protocolos, en la propuesta realizada únicamente se propaga el primer *RREQ*. Así pues, considerando que, desde el punto de vista del nodo intermedio, el destino final sólo tendrá información relativa al primero de los caminos establecidos a través de dicho nodo, no tendría sentido acumular en el mismo toda la información (conjuntos de slots) relativa a los diferentes caminos hacia la fuente. Por lo tanto, cualquier nodo intermedio no tendrá necesidad real de considerar las diferentes copias de *RREQ*s. Sólo el nodo destino deberá acumular dicha información, puesto que en su caso permite realizar la selección del mejor de los caminos. Dicha selección se realiza en función del valor final $BW\,(P)$, el ancho de banda disponible extremo a extremo en el camino P, lo que constituye la **métrica de encaminamiento con QoS** con la que efectivamente se selecciona el camino.

Dado que la ocupación real de recursos en el sistema debida a la admisión de una nueva conexión está directamente relacionada con el número de enlaces ocupados por ésta (ya que el entorno inalámbrico no permite un reúso completo entre enlaces, estando éste limitado por el problema del terminal oculto), la decisión final realizada por el destino, a la hora de seleccionar el camino "óptimo" de acuerdo a la métrica QoS, tiene en cuenta tanto el ancho de banda, como el número de saltos. De este modo, la QoS es medida en función de una métrica doble, utilizada como base para la ordenación de los caminos que cumplen con el requisito mínimo de ancho de banda (los demás no se incluyen en la decisión):

- El camino a escoger cumplirá primeramente el requisito de mínimo número de saltos.
- Entre caminos de igual número de saltos, se considera "óptimo" aquél con un mayor ancho de banda disponible.

El primero de los *RREQ*s recibidos para una conexión nueva (identificado por el par $\langle source_address, broadcast_id \rangle$) conlleva la creación de una entrada en la lista de *RREQ*s recibidos. Las sucesivas copias son detectadas gracias a dicha lista. En el caso de que un nodo intermedio descarte un *RREQ* por incumplimiento de la QoS (ancho

de banda insuficiente en el camino parcial desde la fuente), la entrada correspondiente, o bien es eliminada, o se admiten las sucesivas copias para no limitar la posibilidad de encontrar un camino. Para dotar de mayor flexibilidad a la propuesta (potencial uso futuro del multicamino como tal) se ha determinado la última solución como única excepción en la que un nodo intermedio no descarta los subsiguientes *RREQ*s.

3.4. Gestión compartida de los recursos

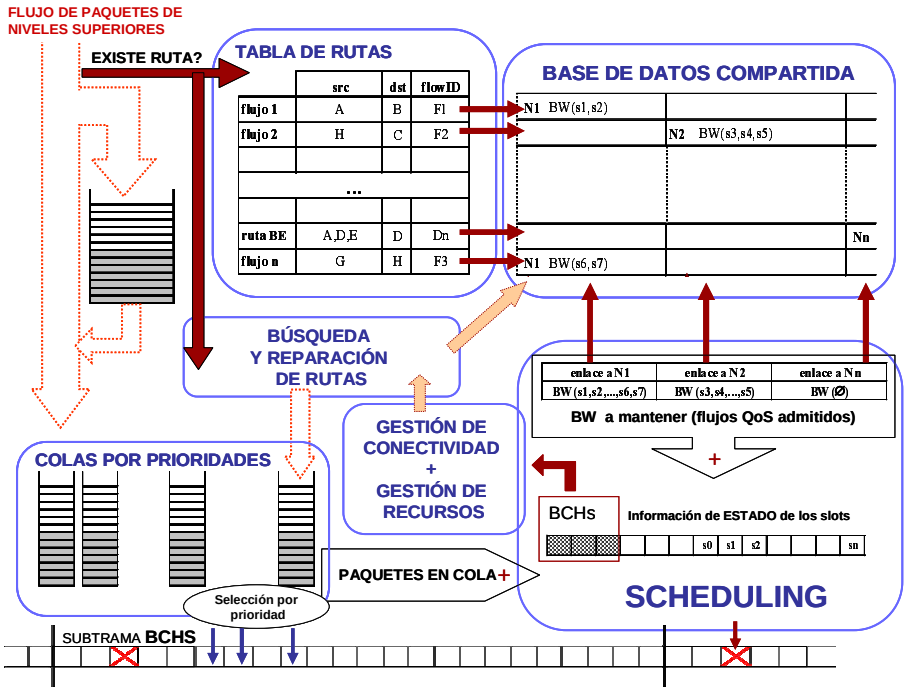


Figura 3.9. Propuesta *cross-layer* de encaminamiento basada en la integración del AODV y el protocolo ADHOC MAC.

El cálculo del ancho de banda integrado en el procedimiento del encaminamiento permite llevar a cabo un control de admisión distribuido. Su funcionamiento es, tanto determinar la capacidad de la red para servir a la nueva conexión, como garantizar la reserva de los recursos asociados dentro del camino seleccionado. Una vez la conexión es admitida, cobra vital importancia la gestión real que se realiza de dichos recursos, es decir, cómo se materializan las reservas, cómo se mantienen, y qué procedimientos permiten mantener la conexión admitida, con las prestaciones prometidas, cuando la variabilidad de la red modifica los parámetros con que inicialmente se produjo la admisión[19]. Gracias a la intercomunicación entre el encaminamiento y el nivel MAC,

[19] Análisis realizado en §4.

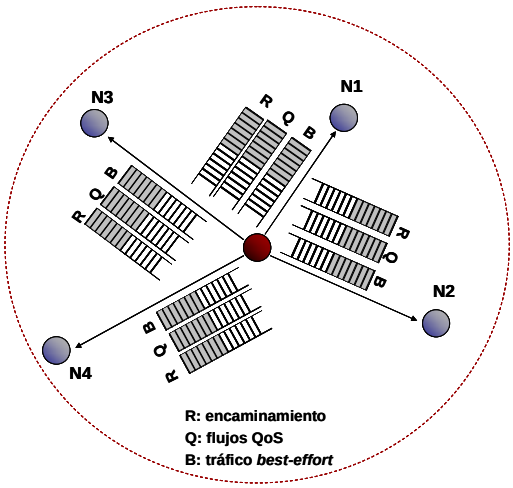
puede establecerse una base de datos compartida asociada a las conexiones (flujos de datos) y los enlaces correspondientes, especificando además el ancho de banda requerido (parámetro de la admisión) – Figura 3.9.

En el caso de rutas para tráfico *best-effort* no es necesario identificar cada uno de los flujos que atraviesan un nodo hacia el mismo destino (a través del $flow_{ID}$) como en el caso de las conexiones con QoS. Dicha simplificación es posible considerando que no es necesario establecer el ancho de banda para dichas conexiones, sino tan sólo saber el siguiente salto en la ruta para permitir que la información llegue a su destino. Esta información es válida para cualquier posible ruta *best-effort* hacia el mismo destino que pase por ese nodo, independientemente del nodo fuente correspondiente. Los recursos a utilizar por dichas conexiones consistirán en el ancho de banda remanente una vez hecha la asignación para las conexiones con restricciones de QoS. Desde este punto de vista, el nivel MAC únicamente necesitará saber la cantidad de tráfico asociado (los paquetes *best-effort* en cola) para establecer las necesidades de recursos. Mientras las rutas de encaminamiento se mantengan activas, el enlace correspondiente en el nivel MAC y las reservas asociadas permanecerán activados. En caso de desaparecer dicha ruta (rotura, o simplemente, aplicación completada y consecuente finalización del flujo de paquetes) el mismo temporizador que controla la eliminación de la información asociada al encaminamiento, activará la liberación de los recursos asociados. La diferenciación de servicios en el nivel de acceso al medio, necesaria para proporcionar QoS, debe realizarse en dos contextos:

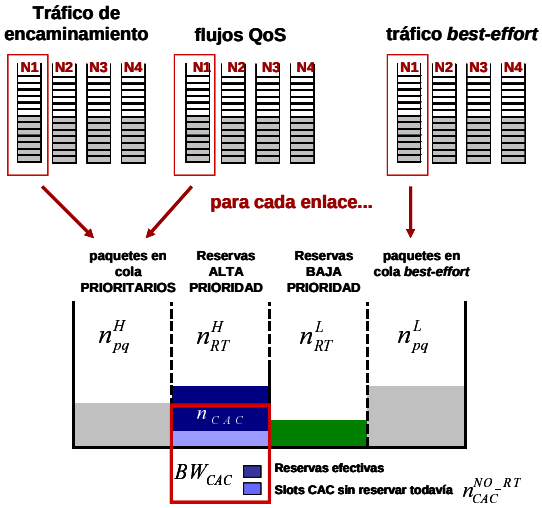
1. Un terminal debe gestionar adecuadamente los diversos servicios que maneja, asignando apropiadamente los recursos y transmitiendo la información de acuerdo a ciertas prioridades teniendo en cuenta, además, la existencia de conexiones con diversos vecinos (múltiples enlaces). Las tareas asociadas a dicha gestión se corresponden con:

   - Cálculo de las nuevas asignaciones de acuerdo al tráfico en cola, las prioridades y las reservas existentes (Figura 3.10).
   - Asignación de slots libres, por prioridad.
   - Una vez reservados los recursos dentro de un enlace, multiplexación de todos los paquetes (independientemente del servicio) asociados a dicho enlace. La priorización se realiza seleccionando el siguiente paquete a transmitir de la cola correspondiente, por orden de prioridad (*Priority Queueing*).

2. Las reglas de competencia básicas del protocolo deben tratar adecuadamente las reservas asociadas a cada servicio para resolver de un modo diferenciado la competición entre diferentes terminales por el canal de transmisión. En este sentido, a lo largo de este trabajo se emplearán las reglas diseñadas en [Gállego, 2007].



(a) Estructura lógica del sistema de colas de un nodo



(b) Colas equivalentes para la determinación de las necesidades de asignación

Figura 3.10. Funcionamiento lógico del sistema de colas de un nodo, atendiendo a los enlaces activos, de acuerdo a los diferentes servicios.

Las características propias del protocolo ADHOC MAC considerado determinan los mecanismos de asignación de recursos. Un nodo únicamente podrá reservar recursos y enviar datos de usuario del nivel superior cuando disponga de un slot BCH en el que transmitir la información de control correspondiente: estado percibido de cada uno de los slots de la trama, reservas activas y nuevas reservas demandadas de recursos incluyendo la prioridad asociada. El slot BCH se transmite al inicio de la trama virtual de cada nodo (todos en la primera subtrama de slots de control). Puesto que precisamente en este momento se dispone de la información completa de todos los vecinos que nos han transmitido su BCH en la trama anterior, así como de nuestra propia percepción del entorno, será entonces, una vez cada trama, cuando el terminal realice la gestión de los recursos, la cual consistirá en lo siguiente:

Fase 1 *Análisis del estado de las actuales reservas:*

Por un lado, ha de verificarse el resultado de las peticiones de reservas nuevas realizadas en la trama anterior (pre-reservas). Para ello, el destino debe haber aceptado la comunicación (puede haber resuelto un conflicto en favor de dichas reservas) y ningún otro vecino debe indicar que va a recibir en ese slot – verificación de (A.7). En aplicación de las reglas de competencia por los recursos del ADHOC MAC, alguna de las reservas puede no haberse hecho efectiva. Por otra parte, debe comprobarse la pérdida de reservas anteriores, por ejemplo, como resultado de una reasignación de recursos para un servicio prioritario.

Fase 2 *Determinación inicial del estado de los slots:*

Con objeto de determinar qué recursos pueden utilizarse para realizar nuevas asignaciones, es necesario disponer de una información actualizada del estado de cada uno de los slots. Con dicha información se puede definir correctamente el $Status - Channel$ (tabla A.2) y en consecuencia la disponibilidad (tabla A.3), como se describe en §A.4.

Fase 3 *Cálculo de las necesidades de nuevos recursos para cada enlace:*

Considerando el tráfico que es necesario transmitir (correspondiente a flujos de datos originados en el terminal) o reenviar (nodo intermedio en otros caminos, información de encaminamiento), debe gestionarse adecuadamente cuántos slots será necesario reservar en la siguiente trama. Dado que el procedimiento BIAS realmente supone una petición

de reserva, será en la trama siguiente cuando, si la reserva es efectiva, dichos recursos puedan ser utilizados. El procedimiento detallado se describe a continuación.

Dado un nodo i que mantiene k enlaces activos (receptores $R_i \subset NB_i$, $|R_i| = k$), el acceso a la base de datos compartida con el encaminamiento permite determinar el ancho de banda asignado en cada enlace $(i \rightarrow n), n \in R_i$, para las NC^n_{QoS} conexiones con QoS admitidas: BW^n_{CAC}. Cada conexión se corresponde con una ruta extremo a extremo $P(m), m \in NC^n_{QoS}$ asociada al flujo QoS identificado por el algoritmo de encaminamiento. Para dicha ruta, se ha asignado un conjunto de slots específico en cada enlace del camino, determinado por el algoritmo *BWC-FA* (3.20). En el caso del nodo i, dicho conjunto se identifica como $TS_i^{P(m)}$, tal como se definió en (3.19).

$$BW^n_{CAC} = \bigcup_{m \in NC^n_{QoS}} TS^{P(m)}_{i,n} \tag{3.20a}$$

$$P(m) = \text{camino}\left(fuente \rightarrow destino\right), m \in NC^n_{QoS} \tag{3.20b}$$

La actualización del $Status{-}Channel$ y la verificación de las reservas, por otra parte, proporcionan la información de los slots reservados para cada prioridad $n^{p_j}_{RT,n} = \left|TS^{p_j}_{i,n}\right|$, $p_j \in PR$[20]:

$$TS^{p_j}_{i,n} = \left\{k \in S : FI^k_i\left[st, id_{TX}, pr, id_{RX}\right] = \left[\text{OCUPADO}, i, p_j, n\right]\right\} \tag{3.21}$$

La gestión del BW_{CAC} se realiza independientemente del resto de recursos. Teniendo en cuenta que una vez establecido el procedimiento de admisión dichos slots son teóricamente reservados y la reserva se mantiene libre de colisiones[21], el gestor puede asumir que va a tener disponibles dichos recursos y no va a necesitar realizar nuevas asignaciones mientras los paquetes en cola asociados "quepan" en dichos recursos. De hecho, puesto que los recursos pertenecientes al BW_{CAC} son reservados para garantizar la QoS de las conexiones para los que fueron asignados, éstos son considerados los recursos más prioritarios ($pr = p_0$), de manera que no existe reasignación posible que implique la pérdida de los mismos. En este sentido, el BW_{CAC} puede considerarse un recurso permanentemente garantizado para la duración completa de la conexión. Por este motivo, no se tiene en cuenta como volumen de tráfico en cola demandante de re-

[20]Conjunto de prioridades: $PR,\ N_p = |PR|\,,\ 0 \leq i \leq N_p,\ p_i > p_j\ \forall i < j$.

[21]Como se explica en el anexo A, teóricamente, en un modelo estático y libre de interferencias (*Protocol Model*), una vez efectivas la reservas, gracias al intercambio de señalización del BCH, éstas serán permanentes hasta la finalización de las aplicaciones correspondientes.

cursos de alta prioridad aquél que está dentro de la capacidad asegurada (CAC), siendo el "exceso" sobre ésta el que realmente requerirá nuevas asignaciones. La suposición de reservas "permanentes", lógicamente, no será cierta en un contexto dinámico (interferencia y topología cambiantes), pero su efecto y las soluciones adoptadas se mostrarán en el capítulo siguiente.

Así pues, en función del estado de las reservas y del volumen de tráfico existente (paquetes en cola, sin incluir aquellos a transmitir en el BW_{CAC} ya reservado), debe calcularse cuántos recursos nuevos son necesarios para cada prioridad. Como se observa en la figura 3.10, la priorización de los servicios se realiza de acuerdo a la cola en la que se encuentran los paquetes. En dicha figura se ejemplifica un caso particular, en la que únicamente se han utilizado dos prioridades para diferenciar los paquetes pertenecientes a flujos QoS y encaminamiento ($pr = p_0$, HIGH) y al tráfico *best-effort* ($pr = p_{N_p}$, LOW). Por simplificación en la notación de los algoritmos de gestión de recursos mostrados a continuación, se han considerado únicamente las prioridades menor y mayor ($p_{N_p} < p_0$) de un conjunto de prioridades potencialmente mayor, dada la capacidad del ADHOC MAC para realizar una diferenciación de recursos mucho más flexible (anexo A.3). No obstante, se ha realizado una definición clara de la máxima prioridad como aquélla asignada a los recursos seleccionados por el control de admisión (*BWC-FA* integrado en el encaminamiento) para los servicios con requerimientos de QoS . La protección de dichos recursos gracias a su prioridad evita su reasignación para servicios menos prioritarios.

El algoritmo 3.1 muestra la implementación real del cálculo de los nuevos recursos necesarios. En caso de tener reservas en exceso en algún enlace, se liberan inmediatamente los recursos oportunos, aumentando así la disponibilidad para aquellos enlaces que lo necesiten. Igualmente, en un enlace específico, ante la necesidad de nuevas asignaciones de mayor prioridad, si existen recursos por liberar de prioridad menor, puede realizarse directamente un cambio de asignación, aprovechando la ventaja de tener una reserva efectiva, reduciendo el retardo introducido por la petición de los recursos (BIAS).

Una vez establecida la necesidad de nuevas reservas, puesto que también han podido liberarse recursos, o simplemente cambiado la prioridad de ciertas reservas activas, es necesario actualizar con dicha información el $Status - Channel$ con el que, como paso previo a la asignación, se defina correctamente la disponibilidad actual de recursos.

Algoritmo 3.1 Evaluación de las necesidades de asignación

for all $n \in R_i$ **do**
  BW^n_{CAC} = slots seleccionados durante los procedimientos de admisión
  $n^{p_i}_{pq,n}$ = número de paquetes en cola de prioridad $p_i \in PR$ (ver figura 3.10).
  $n^{p_i}_{RT,n}$ = número de slots reservados de prioridad p_i
  for $p_i, 0 \leq i \leq N_p$ **do**
    if $i == 0$ **then** ▷ *Caso particular* $(*)$
      $n^{p_0}_{pq,n} = n^{rout}_{pq,n} + excess^n_{CAC}$
      donde
      $n^{rout}_{pq,n}$ = paquetes en cola de alta prioridad, correspondientes únicamente a datos de encaminamiento.
      $excess^n_{CAC}$ = paquetes de flujos QoS que no "caben" en el correspondiente BW^n_{CAC}.
      $n^{p_0}_{RT,n} = n^{p_0}_{RT,n} - |CAC^n_{RT}|$
      $CAC^n_{RT} = BW^n_{CAC} \cap TS^{n,p_0}_i$
    end if
    $(news^{p_i}_n, frees^{p_i}_n) = calcula\left(n^{p_i}_{pq,n}, n^{p_i}_{RT,n}\right)$ ▷ *Ver algoritmo 3.2*
  end for
  Una vez determinados $news^{p_i}_n$ y $frees^{p_i}_n$ se pueden determinar los cambios directos de prioridad.
  for $p_i, 0 \leq i \leq N_p$ **do**
    for $p_j, 0 \leq j \leq N_p$, $pr_j \neq p_i$ **do**
      if $news^{p_i}_n == 0$ **then**
        break ▷ *Parar si no necesito más*
      end if
      if $frees^{p_j}_n > 0$ **then**
        $value = min\left(frees^{p_j}_n, news^{p_i}_n\right)$
        $news^{p_i}_n = news^{p_i}_n - value$
        $frees^{p_j}_n = frees^{p_j}_n - value$
        $changes^{p_i,p_j}_n = value$
      end if
    end for
  end for
  Utilizamos el valor de $changes^{p_i,p_j}_n$ para realizar el cambio de asignación que posteriormente habrá que indicar en el FI a transmitir en el BCH correspondiente.
  Igualmente, de acuerdo a los valores $frees^{p_i}_n$ finales realizamos la liberación de recursos (con objeto de disponer, posteriormente, de dichos slots para nuevas asignaciones).
end for

Algoritmo 3.2 Cálculo de nuevas asignaciones, liberación, etc.

function $(news, frees) = calcula(n_{pq}, n_{RT})$
  Caso particular: $pr == p_0$, gestión de los recursos asociados a conexiones admitidas $(*)$.
  $news = n_{pq} - n_{RT}$
  if $news < 0$ **then**
    $frees = news$
    $news = 0$
  end if
  return
end function

Fase 4 *Definición de la disponibilidad actual de recursos para nuevas demandas (recursos adicionales a BW^n_{CAC}):*

En el caso de la evaluación de la disponibilidad hay que tener en cuenta la capacidad de reasignación que proporciona el ADHOC MAC. Los slots disponibles para transmitir incluyen aquellos slots aparentemente ocupados de menor prioridad – ecuación (A.9). A la hora de decidir qué recursos ya ocupados pueden reasignarse, pueden definirse dos criterios:

- Proteger las reservas ya confirmadas, reasignando únicamente slots en los que exista un acceso (estado ACCEDIENDO).
- Decidir que cualquier recurso ocupado (OCUPADO/ACCEDIENDO) puede ser reasignado.

En cualquier caso, incluso si la decisión tomada es ocupar cualquier recurso de menor prioridad, aun cuando éste ya está efectivamente reservado, la utilización de los slots ya ocupados puede priorizarse de acuerdo a la existencia o no de reserva en firme. Es decir, de todo el conjunto de slots disponibles, es preferible utilizar primeramente recursos completamente libres y, ante la falta de éstos, primero aquéllos en estado ACCEDIENDO y por último los verdaderamente OCUPADOS. De acuerdo a este criterio, pueden definirse como subconjuntos de $SRT_i^{p_j}$ $(p_j \in PR)$:

$$SRT_i^{p_j,LIBRES} = \left\{k \in SRT_i^{p_j} : SC_i^k\,[st] = LIBRE\right\} \tag{3.22}$$

$$SRT_i^{p_j,ACC} = \left\{k \in SRT_i^{p_j} : SC_i^k\,[st] = ACCEDIENDO\right\} \tag{3.23}$$

$$SRT_i^{p_j,OCU} = \left\{k \in SRT_i^{p_j} : SC_i^k\,[st] = OCUPADO\right\} \tag{3.24}$$

Los SRT_i descritos definen de un modo genérico la disponibilidad, para cualquier potencial destino. Sin embargo, a la hora de asignar los recursos necesarios ordenadamente, enlace a enlace, se requiere la disponibilidad particular de cada enlace. Dicha particularidad afecta únicamente a $SRT_i^{p_j}\big|_{ACC}$ y $SRT_i^{p_j}\big|_{OCU}$, ya que entre los potenciales terminales que pueden perder su recurso tras una reasignación puede encontrarse precisamente el destino del enlace para el que se pretende asignar, lo que puede limitar dichos conjuntos (ver tabla A.3).

$$SRT_{i,n}^{p_j,OCU} = \{k \in SRT_i^{pr,OCU} : FI_n^k[st] = LIBRE \vee FI_n^k[pr] > p_j\} \tag{3.25}$$

$$SRT_{i,n}^{p_j,ACC} = \{k \in SRT_i^{pr,ACC} : FI_n^k[st] = LIBRE \vee FI_n^k[pr] > p_j\} \tag{3.26}$$

Fase 5 *Asignación ordenada (por prioridades) de los recursos:*

Una vez conocidas las necesidades finales de nuevos recursos y las disponibilidades de cada enlace, se seleccionan aleatoriamente los slots que sea posible de cada conjunto disponible. Dichas asignaciones consistirán en reservas nuevas (pre-reservas) de las que informar a continuación en el BCH (estado ACCEDIENDO).

Considerando todos los parámetros de entrada de que se dispone se realiza, enlace a enlace, la asignación de recursos, por orden de prioridad. El tráfico de encaminamiento se considera prioritario sobre todo lo demás. Sin embargo, a nivel MAC, tiene la misma categoría que los servicios de QoS en cuanto a prioridad del recurso, por lo que las necesidades reales asociadas al encaminamiento están englobadas en las aquéllas de alta prioridad ($news_n^{p_0}$). El mecanismo para realizar una asignación más prioritaria se ha simplificado, realizando únicamente una ordenación de la información relativa a los enlaces, determinando primero las asignaciones hacia destinos para los que existan paquetes de encaminamiento en cola. Puesto que una vez reservados los recursos de un enlace los paquetes son multiplexados por orden de prioridad de las colas (máxima para el encaminamiento) independientemente de la prioridad con que se reservaran los recursos, siempre que logremos asignar recursos en enlaces donde exista encaminamiento, éste se transmitirá con mayor urgencia.

Con el algoritmo 3.3 se determina, para cada enlace ($i \rightarrow n$), las nuevas asignaciones necesarias S_n^{ass,p_m}. A la hora de transmitir el *Frame Information* (FI) en el BCH, dichas asignaciones se marcan como pre-reservas hacia los correspondientes destinos n, con la prioridad p_m adecuada.

Algoritmo 3.3 Asignación prioritaria de recursos

for $p_m, 0 \leq m \leq N_p$ **do**
  Ordenar n, $n \in R_i$ de acuerdo a la existencia de $n_{pq}^{n,rout}$
  for all $n \in R_i$ **do**
    $SRT_{i,n}^{p_m,0} = SRT_i^{p_m,LIBRES}$
    $SRT_{i,n}^{p_m,1} = SRT_{i,n}^{p_m,ACC}$
    $SRT_{i,n}^{p_m,2} = SRT_{i,n}^{p_m,OCU}$
    $S_n^{ass,p_i} = \emptyset$
    for $k = 0$ to $k = 2$ **do**
      if $\left[(news_n^{p_m} == 0) \vee \left(SRT_{i,n}^{p_m,k} == \emptyset\right)\right]$ **then**
        break ▷ *Parar si ya no tengo disponibles o no necesito más*
      end if
      if $[(m == 0) \wedge (k > 0)]$ **then**
        break ▷ *Parar si no puedo reasignar*
      end if
      $S = BW_1\left(SRT_{i,n}^{p_m,k}, new_n^{p_m}\right)$
      $S_n^{ass,p_m} = S_n^{ass,p_m} \cup S$
      $new_n^{p_m} = new_n^{p_m} - |S|$
    end for
  end for
end for

Fase 6 *Actualización del Status − Channel*

Una vez realizadas las asignaciones oportunas, deben eliminarse aquellos recursos seleccionados, puesto que dejarán de estar disponibles si las reservas concluyen con éxito. Dicha información de estado se mantendrá durante la trama siguiente (pese a recibir nuevos BCHs o detectar transmisiones nuevas en los slots de datos) con objeto de servir de base estable para las decisiones de encaminamiento. Gracias al *Status − Channel* puede extraerse la información necesaria (SRT_k y SRR_k) para el cálculo del ancho de banda. El hecho de no actualizar esta información hasta la trama siguiente tiene como objetivo considerar el estado real una vez resueltos los posibles conflictos entre reservas evitando los accesos fallidos. Desde el punto de vista de un terminal, únicamente tras recopilar la información de toda la trama virtual, en aplicación de las reglas del protocolo (prioridades, resolución de colisiones, etc.), la información de estado será consecuente con todos sus vecinos.

Fase 7 *Decisión acerca de la información a transmitir en el BCH (FI):*

El FI a transmitir en el BCH debe incluir las reservas que sigue manteniendo y las pre-reservas resultado de la asignación realizada indicando la prioridad adecuada. A la hora de mantener el BW_{CAC}, que no debe ser liberado mientras la conexión esté activa, hay que tener en cuenta los transitorios en los que todavía no se han conseguido reservar completamente todos los recursos. En definitiva, considerando $CAC^n_{noRT} = BW^n_{CAC} \cap \overline{TS^{n,p_0}_i}$, como el conjunto de slots pendientes de reservar, los $s_k \in S, s_k \in CAC^n_{noRT}$ se marcan como pre-reservados de alta prioridad en el FI mientras que el resto se marcan reservados. Una vez todas las reservas son efectivas, los slots pertenecientes al BW_{CAC} permanecerán continuamente marcados en el FI como reservados.

3.5. Evaluación del mecanismo de admisión

La descripción realizada en este capítulo presenta la arquitectura básica que permite realizar un control de admisión distribuido acorde con la demanda de las aplicaciones. Con objeto de evaluar la conveniencia de las soluciones adoptadas, se ha analizado la propuesta frente a diferentes condiciones de entorno, asumiendo un escenario simplificado: **escenario estático, *Protocol Model***[22].

Para implementar el diseño *cross-layer* propuesto se ha desarrollado un simulador orientado a eventos, programado en C++, que incluye los protocolos implicados, AD-HOC MAC y AODV[23], así como las interfaces de comunicación adecuadas y el algoritmo *BWC-FA* modificado integrado en el propio encaminamiento.

Se ha considerado un promedio de diversos escenarios correspondientes a topologías de 30, 50 y 100 nodos uniformemente distribuidos en una cuadrícula de 2 Km^2 en los dos primeros casos, y de 5 Km^2 en el caso de 100 nodos. La generación de los distintos escenarios se corresponde con el análisis de la propuesta tanto en diferentes condiciones de densidad de terminales posicionados en una misma región espacial (30 y 50 nodos) como en la influencia de la distribución de rutas (aumento de saltos) para una misma densidad de terminales (50 y 100 nodos). La tabla 3.1 muestra la densidad media de vecinos obtenida con los escenarios propuestos. Una muestra representativa de los

[22]Definido en §3.1.1.

[23]Se ha realizado la implementación de ciertos modelos basándose en la adaptación y modificación de funciones existentes en el simulador de redes ns-2. No obstante, la arquitectura implementada no se ha integrado en el citado simulador.

mismos se ha representado en la figura 3.11, donde se refleja a su vez la conectividad de todos ellos, considerando el rango de cobertura especificado $R = 515$ m. Éste se ha definido de acuerdo a una *SIR* mínima de decodificación (SIR_{th}) de 5 dB, con una potencia de transmisión de 20 dBm, un ruido térmico de -103 dBm y el modelo de propagación de Kammerman:

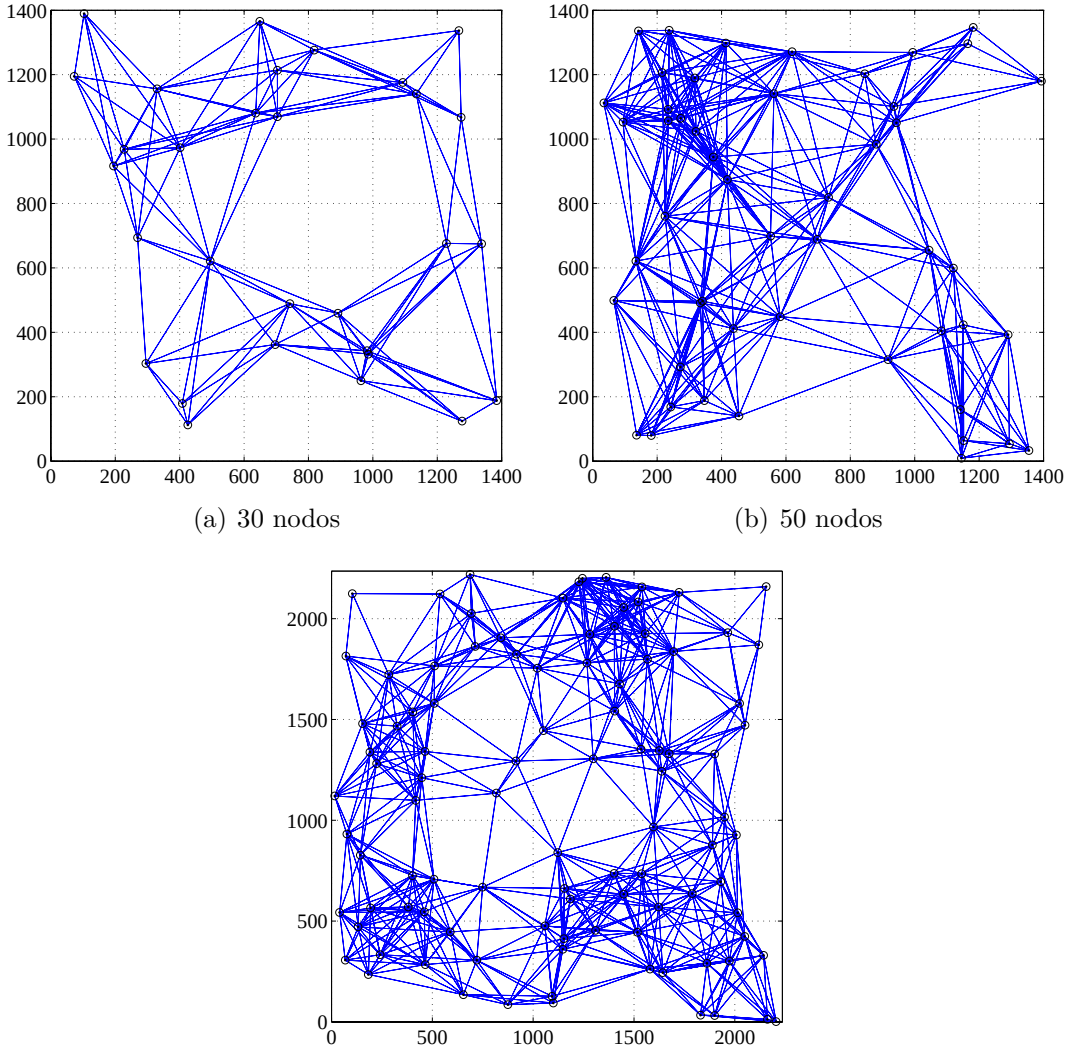
$$L_{i,j} = \begin{cases} -20 \cdot \log_{10}\left(\frac{4 \cdot \pi \cdot d}{\lambda}\right) & \text{if } d < 8, \\ -58{,}3 - 33 \cdot \log_{10}(\frac{\mathrm{d}}{8}) & \text{if } d \geq 8. \end{cases} \tag{3.27}$$

Las conexiones entre diferentes pares de nodos se generan de acuerdo a un proceso de Poisson, de tasa λ [conexiones/seg.] calculada en función del tráfico ofrecido (T.O. en Erlangs) y la duración constante por conexión (50 seg.). Los diferentes flujos QoS se corresponden con fuentes *Constant Bit Rate* (CBR) con diversas demandas de ancho de banda (64, 128 y 256 kbps). El T.O. representado en kbps se corresponde con el T.O. en Erlangs por la tasa demandada por conexión en kbps. Con objeto de limitar el retardo de los paquetes recibidos correctamente se aplica una política de descarte anticipado en cada nodo. La estructura de trama definida en el protocolo ADHOC MAC consiste en 25 slots reservados para datos y 50 destinados a la transmisión del BCH[24]. Una trama dura 50 ms., estableciendo una tasa de transmisión por slot equivalente a 64 kbps, de manera que la demanda individual por conexión se corresponde con 1, 2 ó 4 slots, de acuerdo a las tasas de 64, 128 y 256 kbps.

Tabla 3.1. Escenarios estáticos de análisis

nodos	superficie	densidad media
30	2 Km^2	8,21 vecinos
50	2 Km^2	14,75 vecinos
100	5 Km^2	13,41 vecinos

[24]El número mínimo de slots necesarios, teniendo en cuenta la capacidad de reutilización espacial, es mucho menor de los 50 propuestos, especialmente en este escenario simplificado [Gállego, 2007]. Sin embargo, se ha sobrestimado su dimensionado con objeto de asegurar, de manera genérica, incluso en situaciones de elevada movilidad e interferencia, una probabilidad de acceso igual a 1. De este modo, se pretende que la problemática asociada al mantenimiento de los slots BCH y el acceso de terminales al sistema no condicione las conclusiones respecto a la gestión de los recursos verdaderamente dedicados a tráfico de usuario. El análisis detallado del dimensionado adecuado se ha desarrollado en profundidad en [Gállego, 2007].



(a) 30 nodos

(b) 50 nodos

(c) 100 nodos

Figura 3.11. Escenarios de muestra de las topologías simuladas (redes conexas). 30 y 50 nodos en un área de 2 Km^2, y 100 nodos en un área de 5 Km^2.

A la hora de valorar la eficacia de un control de admisión adecuado se han considerado las siguientes métricas de calidad:

- **Probabilidad de bloqueo:** ratio de conexiones no admitidas en el sistema entre el total de conexiones ofrecidas. Una conexión es rechazada cuando no se encuentra un camino de conexión viable, una vez superados el máximo número de intentos posibles ($n_{disc,max}$), establecido a 3 en las simulaciones realizadas.

- ***Throughput*:** ratio entre paquetes correctamente recibidos y paquetes generados. La referencia al *throughput* de las conexiones admitidas se corresponde con el ratio entre paquetes recibidos y paquetes transmitidos, al no incluir aquellas aplicaciones bloqueadas.

- **Tráfico correctamente cursado:** tráfico correspondiente a las conexiones admitidas que experimentan la QoS demandada. A la hora de medir la satisfacción a nivel de usuario se ha considerado una pérdida de paquetes máxima tolerable (P_{loss}) del 3 %.

A continuación se evalúa la propuesta presentada, de acuerdo a las citadas métricas, comparando diversas estrategias:

- **AODV básico**: correspondiente a la aplicación del algoritmo de encaminamiento AODV sin consideraciones de QoS ni control de admisión explícito en el nivel de acceso.

- ***QoS Routing – BottleNeck* (QoSR-BN):** Encaminamiento con QoS basado en la medida del ancho de banda en el enlace "cuello de botella" (*bottleneck*). Dicha estimación únicamente evalúa la potencial disponibilidad de recursos en cada enlace (LB_k), sin tener en cuenta la selección particular de slots que sería necesaria considerando el mecanismo de acceso definido en el ADHOC MAC. De igual modo, la reserva de recursos se realiza dinámicamente, en función del tráfico en las colas una vez son activadas las aplicaciones admitidas (modelando por tanto una admisión y una reserva independientes).

- ***QoS Routing – Cross-Layer* (QoSR-CL):** Arquitectura *cross-layer* propuesta que incluye la reserva de recursos asociada, de acuerdo al criterio de admisión (algoritmo *BWC-FA*). La funcionalidad de espera en el nodo destino permite la recepción de diversos *RREQ* correspondientes a caminos alternativos de manera que se puede seleccionar el "mejor" de acuerdo a la métrica de QoS. El parámetro

configurable correspondiente, t_{wait} (§3.3.2.5), se ha establecido a un valor de 0,3 seg. a lo largo de todas las simulaciones[25].

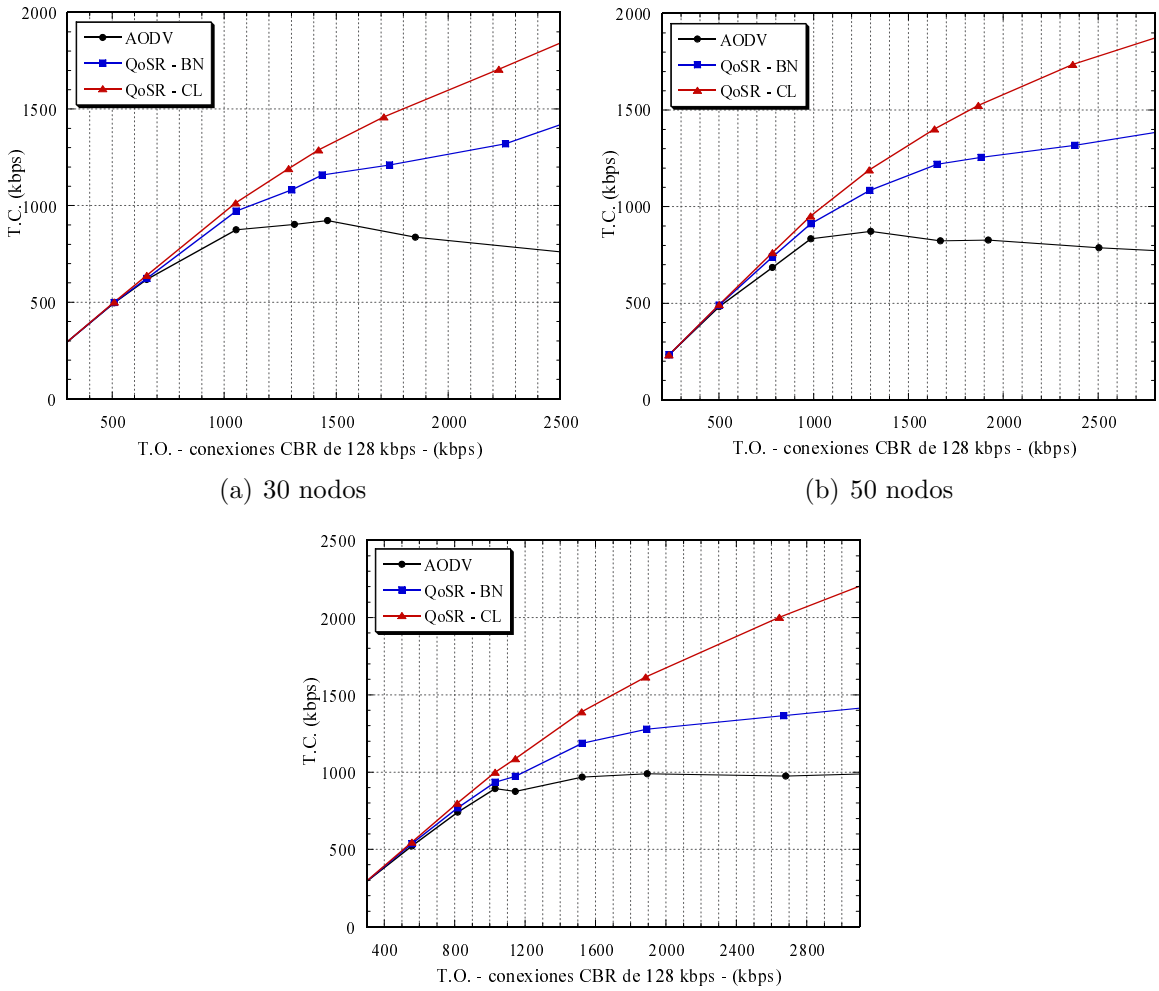
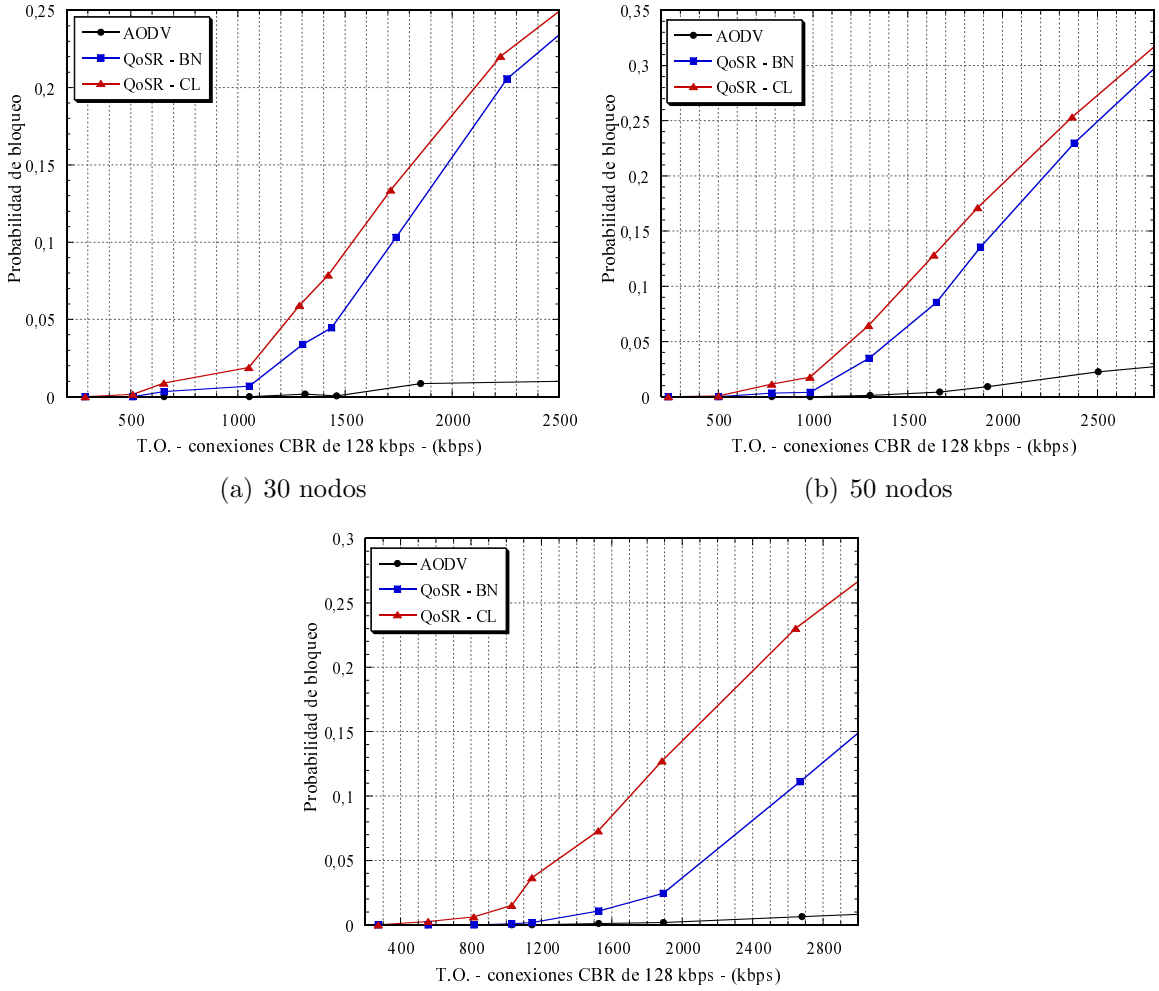


Figura 3.12. Tráfico correctamente cursado vs. tráfico ofrecido. Comparativa de estrategias de encaminamiento y admisión. Escenarios de 30, 50 y 100 nodos.

Sin la existencia de un control de admisión, dado que únicamente puede garantizarse una ocupación de recursos igual a la capacidad del sistema, el exceso de tráfico ofrecido provoca una continua competencia por los recursos existentes que tiene como resultado una gestión ineficiente de los mismos. El tráfico en exceso, por otra parte, se bloquea

[25]El estudio detallado de las diversas topologías utilizadas ha permitido establecer el valor necesario para recoger un porcentaje significativo de mensajes *RREQ*, y por lo tanto caminos alternativos, sin aumentar excesivamente la latencia de búsqueda.

en las colas provocando finalmente el descarte de los paquetes que superan el límite tolerable de retardo. De este modo, las aplicaciones activas no experimentan la QoS deseada. Por otra parte, si la ocupación de recursos no se distribuye en los caminos más apropiados, se infrautiliza la capacidad realmente disponible. En todos los escenarios analizados, cualquier estrategia capaz de controlar la admisión de conexiones al sistema (QoSR-BN o QoSR-CL) mejora considerablemente el tráfico correctamente cursado respecto al máximo obtenido en condiciones de continua congestión (AODV sin control de admisión), como puede observarse en la figura 3.12.



(a) 30 nodos

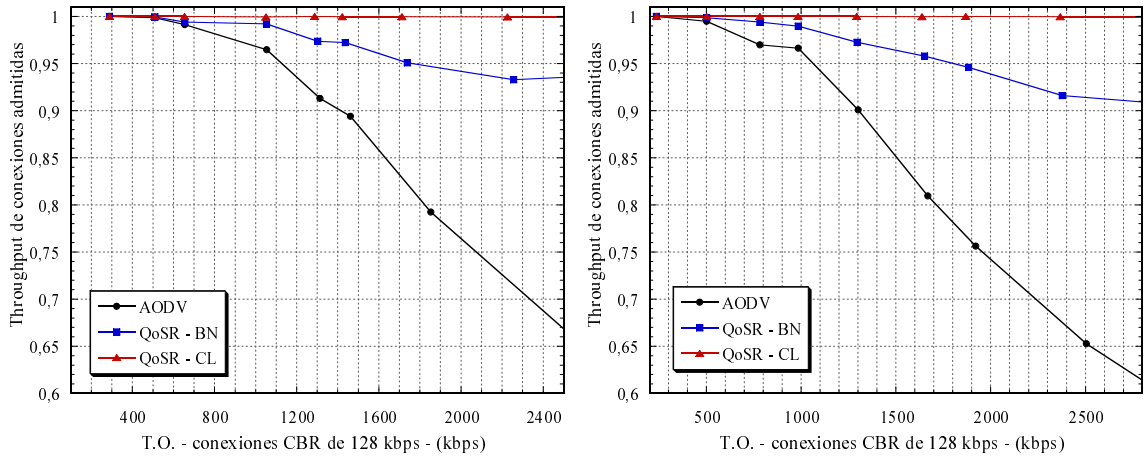
(b) 50 nodos

(c) 100 nodos

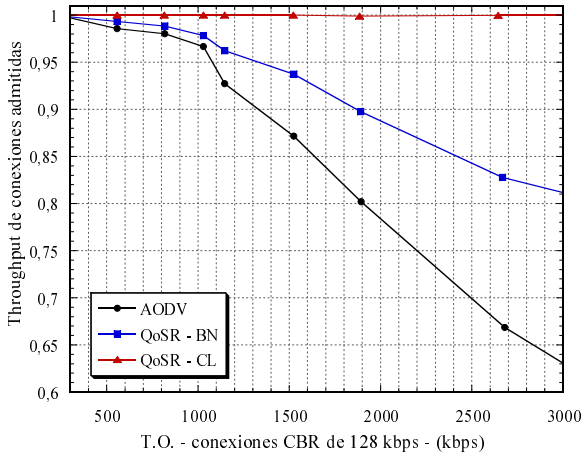
Figura 3.13. Probabilidad de bloqueo vs. tráfico ofrecido. Escenarios de 30, 50 y 100 nodos. Comparativa de estrategias de encaminamiento y admisión.

El beneficio observado se debe a la aplicación de un encaminamiento con QoS (*QoS Routing*) basado en una métrica de ancho de banda. La funcionalidad básica del en-

caminamiento de encontrar una ruta viable limita el acceso al sistema al bloquear las conexiones para las que no es posible garantizar un camino capaz de proporcionar la QoS demandada. Una sobreestimación de la disponibilidad de recursos puede provocar, sin embargo, una admisión en exceso que limita la capacidad del sistema de asegurar la QoS demandada. Tal es el caso, por ejemplo, de la estimación basada en el "cuello de botella", cuando se aplica en una red basada en acceso TDMA, sin tener en cuenta las dependencias reales entre los diferentes enlaces. La figura 3.13 muestra que dicha estrategia supone un bloqueo menor que en caso de realizar una asignación específica de los slots TDMA en los que transmitir (QoSR-CL).



(a) 30 nodos (b) 50 nodos



(c) 100 nodos

Figura 3.14. *Throughput* de conexiones admitidas vs. tráfico ofrecido. Escenarios de 30, 50 y 100 nodos. Comparativa de estrategias de encaminamiento y admisión.

Como resultado, las conexiones admitidas sufren cierta degradación, no cumpliendo con la restricción de descarte especificada ($P_{loss} < 3\,\%$), como se observa en la figura 3.14. Por el contrario, la asignación de recursos realizada durante el mismo procedimiento de admisión (QoSR-CL) identifica de una manera mucho más realista la capacidad de la red de asegurar el ancho de banda demandado. Además, la reserva realizada en fase $RREP$ garantiza el mantenimiento de los slots especificados, obteniendo el *throughput* idealmente esperado, tal y como se muestra en dicha figura.

Los resultados obtenidos en diversos escenarios muestran que un aumento de densidad en una misma región espacial (comparativa de 30 y 50 nodos en 2 Km^2) no supone una diferencia apreciable en las prestaciones siempre que la red se mantenga suficientemente conexa. El mayor número de vecinos existente en el escenario de 50 nodos disminuye ligeramente el número medio de saltos por conexión admitida (figura 3.15), dado que aumentan las probabilidades de encontrar caminos más cortos que en el escenario de 30 nodos. Éste constituye una red igualmente conexa, pero más dispersa. Puesto que la reducción en el número de saltos no supone variaciones en el orden de magnitud, la ocupación espacial de los recursos es análoga, al situarse en la misma superficie. Así pues, la congestión en ambos escenarios (30 y 50 usuarios) es comparable. En consecuencia, el tráfico cursado correctamente no varía considerablemente.

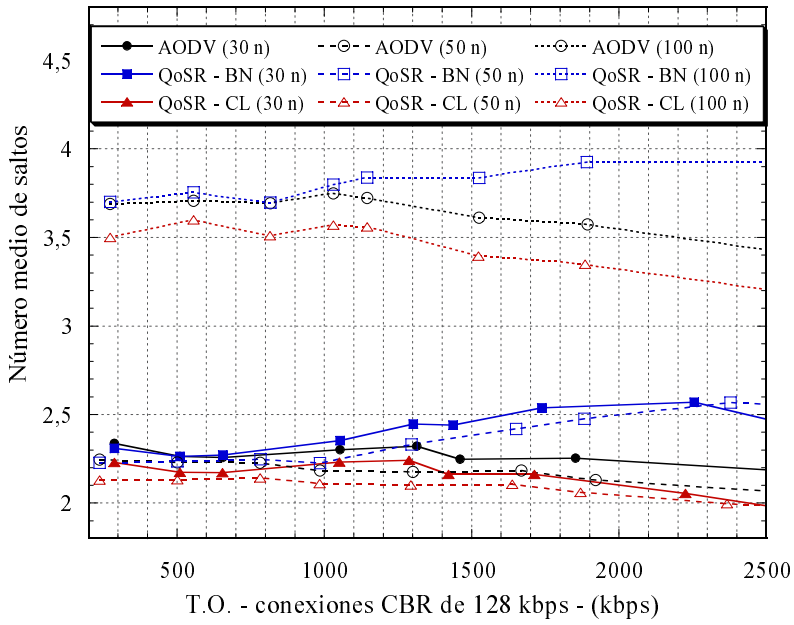


Figura 3.15. Número medio de saltos de las conexiones admitidas. Comparativa entre escenarios de 30, 50 y 100 nodos.

Una densidad similar en un área mayor (100 nodos en 5 Km^2) supone por un lado, un aumento en el número de saltos necesarios para establecer un camino, y por otro, una mayor distribución espacial de los recursos ocupados (mayor superficie), lo que permite aumentar la reutilización de recursos. El aumento de saltos necesarios que se observa en la figura 3.15 supone un tráfico ofrecido efectivo mayor sobre el sistema (mayor número de enlaces con recursos ocupados para una misma conexión). La estrategia QoSR-CL permite aprovechar dicha capacidad gracias a una distribución eficiente de los recursos, consiguiendo por lo tanto un tráfico correctamente cursado mayor (menor probabilidad de bloqueo). Por otra parte, la diferencia entre las estrategias QoSR-BN y QoSR-CL se hace mucho más pronunciada en esta topología, debido a la mayor probabilidad de estimar incorrectamente el ancho de banda disponible considerando únicamente el enlace de mínimo ancho de banda. Cuantos más saltos conforman la ruta, mucho más apropiada resulta una estrategia capaz de asignar y reutilizar adecuadamente los recursos específicos en cada uno de los enlaces siendo por el contrario mucho más probable equivocarse al considerar las medidas aisladas de cada uno de ellos (QoSR-BN). En situaciones de alta carga, por ejemplo, la existencia de pocos recursos disponibles puede llevar erróneamente a interpretar un pequeño número de slots, únicos disponibles en toda la red, como el valor de ancho de banda disponible (mínimo, "cuello de botella") cuando la dependencia entre diversos enlaces haría imposible utilizarlos simultáneamente en enlaces consecutivos de un mismo camino.

La estimación errónea del ancho de banda disponible extremo a extremo realizada en la estrategia QoSR-BN provoca, por otro lado, un aumento innecesario del número medio de saltos, como se observa en la figura 3.15. La dificultad de encontrar un camino viable (descarte de *RREQs* por incumplimiento de ancho de banda mínimo) aumenta el bloqueo por parte de la estrategia QoSR-CL – figura 3.13(c), de manera que únicamente es posible que las rutas de pocos saltos, y por tanto con menos recursos efectivos ocupados, se establezcan correctamente. En el caso de no realizar ninguna admisión (AODV) se produce un efecto similar a consecuencia de la congestión en la red. Como resultado, el número medio de saltos de las conexiones admitidas, a elevadas tasas de tráfico ofrecido, decrece ligeramente. Sin embargo, cuando la admisión se basa en la medida del enlace de ancho de banda mínimo (QoSR-BN), sin tener en cuenta el efecto real del número de saltos que componen el camino, es posible encontrar rutas supuestamente viables, de mayor número de saltos, que llevan al aumento observado de dicha métrica. Una admisión en exceso basada en rutas que añaden una carga superior al sistema, en cuanto a tráfico ofrecido efectivo, en las que además no es posible cursar

realmente dicho tráfico, provoca el aumento en la degradación del *throughput* obtenido, tal y como se observa en la figura 3.14(c), lo que reduce drásticamente el tráfico correctamente cursado – figura 3.12(c).

A la hora de proponer la estrategia analizada QoSR-CL, se ha tenido en cuenta la implementación finalmente realizada del algoritmo conjunto de cálculo de ancho de banda y asignación de recursos. Dicha propuesta se ha basado en el planteamiento inicial mostrado en [Zhu y Corson, 2002] con ligeras modificaciones. La diferencia sustancial entre ambas propuestas reside en la asignación definitiva de los recursos en un enlace, de acuerdo a (3.19) (selección de TS^k, en $(n_k \rightarrow n_{k-1})$) por parte del extremo transmisor (n_k). En la propuesta original de Zhu y Corson, el nodo receptor, que dispone de la información más actualizada del conjunto de recursos disponibles en ese enlace (PB_k^{k-1}), determina los recursos específicos en los que transmitir indicándoselos en el $RREP$ al transmisor. En la propuesta considerada, sin embargo, se indica el conjunto completo de manera que sea el transmisor el que realice la asignación. La principal ventaja de esta medida reside en disponer de un rango de recursos mayor, a la hora de realizar la reserva, que permita evitar la posible pérdida de alguno de los recursos disponibles. Dicha pérdida puede deberse a múltiples efectos en un entorno realmente dinámico (movilidad, interferencia...). Sin embargo, en el escenario analizado, aparentemente estable, la problemática de reservas paralelas y *race-condition*[26], generalizables como la asignación simultánea del mismo recurso para diversas aplicaciones, puede igualmente originar la pérdida de los recursos. Si bien en algún caso esta pérdida es inevitable (cuando las aplicaciones originalmente admitidas no pueden cursarse simultáneamente, en situaciones de alta carga), una asignación más flexible de los recursos puede evitar un bloqueo excesivo en situaciones menos congestionadas. El análisis de las simulaciones llevadas a cabo en los diversos escenarios ha permitido constatar dicha ventaja potencial, lo que ha conducido a la selección final de la propuesta descrita.

En la figura 3.16, tomada como muestra representativa del análisis realizado, se observa la comparativa de ambas estrategias, denominadas QoSR-CL v1 (versión basada en la asignación propuesta por Zhu y Corson) y QoSR-CL v2 (versión finalmente seleccionada, basada en una asignación más flexible). Puesto que ambas estrategias realizan una reserva específica de recursos, cuando éstos son efectivamente reservados se proporciona un ancho de banda garantizado, lo que permite obtener las prestaciones de QoS demandadas. Así pues, el parámetro diferenciador entre ambos, que influye directamente en el tráfico cursado, es la probabilidad de bloqueo, por lo que se ha repre-

[26]Descritas en §3.3.2.4

sentado dicha métrica, analizada en un escenario de 50 nodos, con diversas demandas individuales de recursos por conexión (64, 128 y 256 kbps).

En cualquiera de los escenarios, la reducción en la probabilidad de bloqueo conseguida con la versión 2 es apreciable. No obstante, ésta es mucho más significativa cuanto menor es la demanda de recursos por conexión. Esto es debido al mayor impacto que supone decidir erróneamente un único recurso, que finalmente no puede ser reservado, cuando existen muchos otros posibles, especialmente cuando el tráfico ofrecido supone la existencia de suficientes recursos disponibles (baja carga en el sistema). Conforme aumenta el tráfico ofrecido, aumenta las probabilidad de aparición de situaciones de reservas paralelas o *race-condition*, aumentando la diferencia en el bloqueo. Sin embargo, cuando la carga en el sistema supone una reducción considerable de la capacidad disponible, la determinación de recursos realizada por ambas estrategias resulta muy similar, de manera que ya no es posible obtener una ganancia considerable. El mismo efecto puede apreciarse cuando la demanda por conexión aumenta (256 kbps), puesto que la ocupación por enlace crece reduciendo, por la misma razón, la diferencia entre ambas estrategias. En cualquier caso, la mayor flexibilidad conseguida siempre supone una mejora sobre la propuesta original (versión 1), lo que se ha considerado, además, beneficioso para adaptarse posteriormente a escenarios mucho más dinámicos.

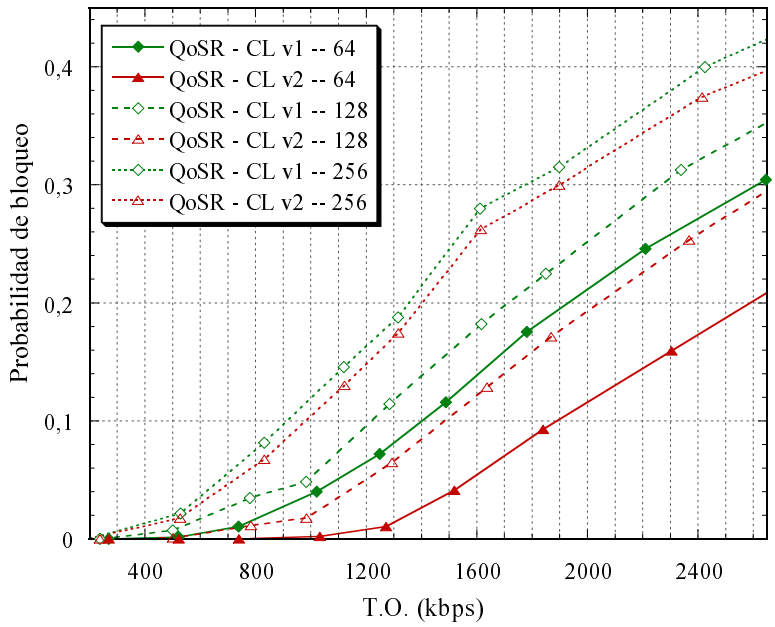


Figura 3.16. Probabilidad de bloqueo vs. tráfico ofrecido. Demandas individuales de recursos por conexión de 64, 128 y 256 kbps. Escenario de 50 nodos. Comparativa de versiones *cross-layer*.

El encaminamiento propuesto se basa en encontrar y reservar una única ruta multisalto entre los dos extremos de la comunicación. Dado que el protocolo considerado no proporciona la facilidad de dividir el tráfico generado entre diversos caminos (multicamino con división por ancho de banda: *BW splitting*) la flexibilidad para distribuir el tráfico adecuadamente está condicionada por la demanda de ancho de banda realizada sobre un único camino. Cuanto mayor número de recursos demanda una aplicación, mayor es la probabilidad de que ésta sea bloqueada, aún existiendo aparentemente recursos suficientes en la red. Para analizar este efecto, se han evaluado las prestaciones de la propuesta frente al mismo tráfico ofrecido total pero con diferentes requerimientos individuales. Dicho de otro modo, se han simulado escenarios con el mismo tráfico ofrecido, medido en kbps, pero generado a partir de conexiones CBR de diversas tasas (64, 128 y 256 kbps).

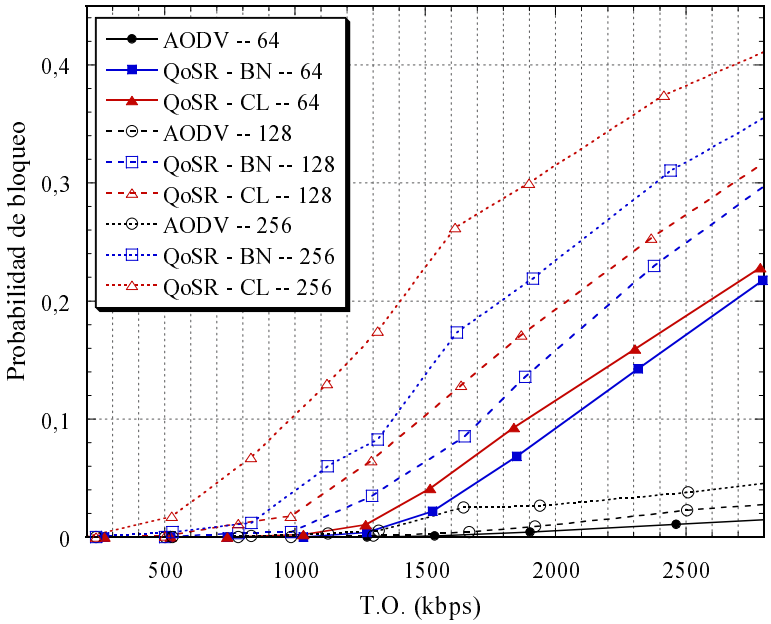


Figura 3.17. Probabilidad de bloqueo vs. tráfico ofrecido. Demandas individuales de recursos por conexión de 64, 128 y 256 kbps. Escenario de 50 nodos. Comparativa de estrategias de encaminamiento y admisión.

Una mayor demanda individual por conexión supone la necesidad de encontrar caminos con más recursos disponibles en todos sus enlaces. El establecimiento de las conexiones admitidas supone una ocupación más elevada por enlace, dificultando así la admisión de nuevas aplicaciones sobre los mismos enlaces. En consecuencia, el bloqueo

experimentado (ratio de conexiones no admitidas) crece, como se observa en la figura 3.17. En el caso de la estrategia QoSR-BN, la búsqueda de ruta es menos restrictiva, al no buscar slots no interferentes en enlaces consecutivos. Puesto que el ancho de banda mínimo determina la admisión, pese al aumento de ocupación resulta más probable encontrar un camino posible que en el caso de la estrategia QoSR-CL. Esta diferencia se hace mucho más patente conforme aumenta la demanda individual por conexión. Por otra parte, dado que la sobreestimación de capacidad disponible realizada por QoSR-BN es mucho más significativa para una demanda superior (256 kbps), las conexiones admitidas experimentan una mayor degradación en el *throughput*, como se refleja en la figura 3.18. En consecuencia, disminuye el tráfico correctamente cursado (figura 3.19).

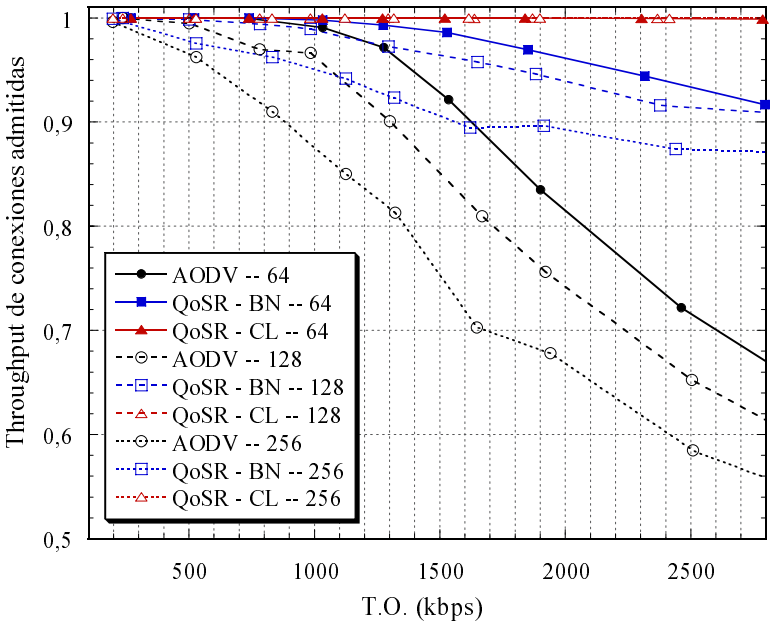


Figura 3.18. *Throughput* de las conexiones admitidas vs. tráfico ofrecido. Demandas individuales de recursos por conexión de 64, 128 y 256 kbps. Escenario de 50 nodos. Comparativa de estrategias de encaminamiento y admisión.

La estrategia QoSR-CL, sin embargo, consigue proporcionar un ancho de banda garantizado que asegura la QoS demandada. La disminución de tráfico cursado observada demuestra la infrautilización de la capacidad del sistema. Una mejor distribución de los recursos a lo largo de diferentes caminos permitiría aproximarse a la capacidad obtenida con una demanda por enlace de un único slot (64 kbps).

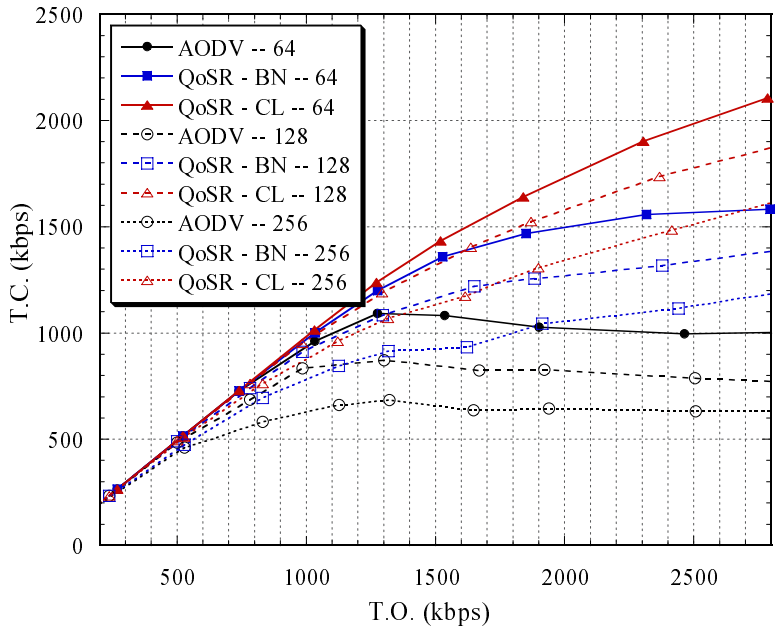


Figura 3.19. Tráfico correctamente cursado vs. tráfico ofrecido. Demandas individuales de recursos por conexión de 64, 128 y 256 kbps. Escenario de 50 nodos. Comparativa de estrategias de encaminamiento y admisión.

En caso de disponer de una facilidad que permitiera dividir el tráfico en peticiones de menor ancho de banda, cada uno de los subcaminos tendría una probabilidad de admisión mucho mayor, incrementando la ocupación de recursos y aprovechando al máximo la capacidad de la red. No obstante, la implementación de dicha alternativa no es trivial. El principal requisito a resolver sería desarrollar una algoritmo de encaminamiento multicamino con la posibilidad de reservar recursos en diferentes rutas. En este contexto, hay que tener en cuenta la dificultad de realizar asignaciones de recursos que garanticen la total independencia (no interferencia) de diversos caminos en paralelo (problemática de la *race condition* y reservas paralelas, §3.3.2.4). El planteamiento de búsqueda de disponibilidad máxima que se ha considerado en la propuesta descrita no permite determinar, paralelamente, recursos disponibles no coincidentes. La selección de recursos debería realizarse, por lo tanto, durante la misma fase de búsqueda, resolviendo de manera eficiente tanto el problema de la división del ancho de banda como las reservas paralelas. Dichas consideraciones suponen un planteamiento completamente diferente al desarrollado en este estudio. Si bien existen diversos trabajos en este sentido [Liao y otros, 2002b; Lin y Huang, 2004; Shu y otros, 2004], todas las soluciones suponen ciertas limitaciones. El ancho de banda disponible se estima en función del enlace "cuello de botella" [Shu y otros, 2004] o asumiendo independencia entre los diversos

enlaces [Liao y otros, 2002b], siendo su aplicación errónea en un contexto TDMA. Por otra parte, la selección y reserva de recursos durante la misma fase de búsqueda, limita la posibilidad de encontrar el "mejor" camino y la flexibilidad necesaria para adaptarse a un escenario variante (tal y como se mostrará en §4). Por esta razón, se ha optado por un análisis centrado en un único camino tratando de optimizar, en la medida de lo posible, la selección de éste y la gestión de recursos asociada una vez admitida la conexión.

La aplicación del control de admisión se basa en el establecimiento de **circuitos virtuales**, es decir, caminos con un ancho de banda garantizado constante, reservado mientras las correspondientes sesiones permanecen activas. Sin embargo, dicho comportamiento no se adapta a la naturaleza de todas las aplicaciones. En el caso de aquéllas no orientadas a conexión, sin restricciones temporales, puede no ser necesario el mantenimiento de un ancho de banda garantizado, especialmente cuando dicho tráfico sea racheado. La arquitectura propuesta, por lo tanto, debe ser capaz de realizar una diferenciación acorde con las características propias de las aplicaciones [Canales y otros, 2006a]. Con objeto de evaluar esta capacidad diferenciadora se ha analizado el comportamiento del sistema en un escenario multiservicio, en el que el tráfico ofrecido equivale tanto a aplicaciones CBR (128 kbps) como a tráfico *best-effort*. Este último se ha modelado mediante sesiones constituidas por llamadas con un número medio de 50 paquetes, de acuerdo a una distribución geométrica, generados a una tasa de 64 kbps. El tiempo medio entre paquetes y el intervalo entre llamadas están exponencialmente distribuidos, con medias 62.5 ms y 5 seg. respectivamente. Se han simulado tres escenarios multiservicio, que se corresponden con el mismo tráfico ofrecido total (medido en kbps) pero con diferentes porcentajes de tráfico *best-effort* (correspondencia del 25, 50 y 75 % de las conexiones generadas con sesiones BE). El análisis realizado incluye igualmente la comparativa con las estrategias analizadas del AODV y QoSR-BN, con objeto de validar si la demostrada superioridad de la arquitectura propuesta se mantiene en este escenario.

La aplicación básica del AODV, sin procedimiento de admisión, permite dar un tratamiento diferenciado a los distintos servicios gracias a la priorización realizada en el propio nivel de acceso (ADHOC MAC), es decir, la provisión de QoS se realiza únicamente desde la perspectiva del nivel MAC. De un modo análogo, la estrategia QoSR-BN se apoya en la capacidad diferenciadora del nivel de enlace. Sin embargo, el acceso de las conexiones CBR se limita gracias al procedimiento de admisión, con objeto de evitar la saturación de la red, puesto que en esas condiciones sería imposible

proporcionar la QoS demandada. En cualquiera de los casos, el ADHOC MAC se encarga de asignar recursos bajo demanda a las aplicaciones admitidas, CBR o BE, por orden de prioridad, permitiendo además potenciales reasignaciones para los servicios prioritarios.

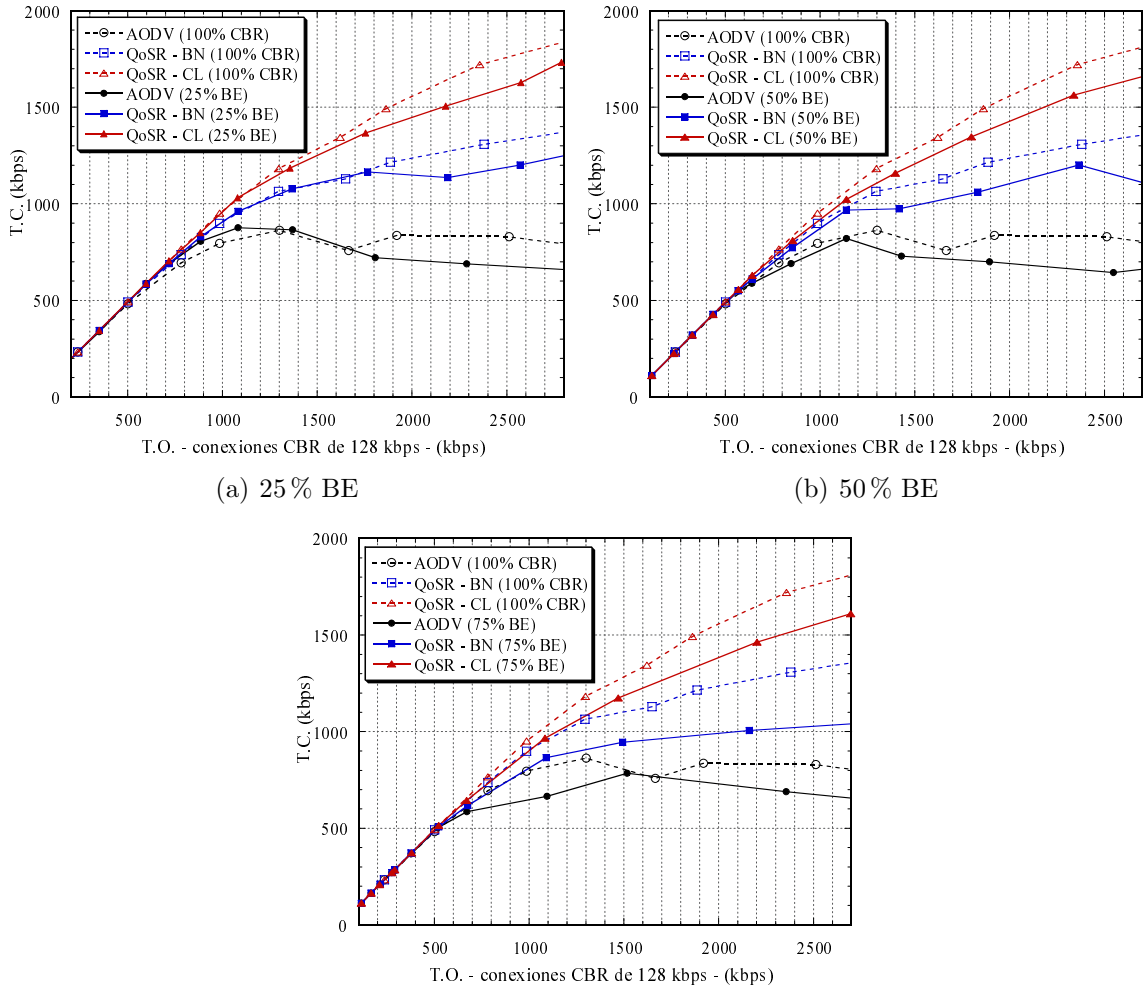
En el caso de la propuesta *cross-layer* desarrollada (QoSR-CL), por el contrario, la interacción entre los diferentes niveles permite explotar la capacidad diferenciadora del ADHOC MAC teniendo además en cuenta la asignación de recursos específica realizada desde el nivel de encaminamiento. Los circuitos virtuales se asocian a flujos QoS prioritarios para los que se mantiene de manera garantizada (prioridad máxima, p_0) el ancho de banda seleccionado (BW^n_{CAC} definido en (3.20)). Las rutas establecidas sin criterios de QoS (sin identificación de $flow_{ID}$) no son explícitamente tenidas en cuenta desde el punto de vista del nivel MAC. Los recursos necesarios para transmitir el tráfico acumulado en las colas de la prioridad asociada (la mínima, p_{N_p}) se asignan bajo demanda, cuando existe dicho tráfico en cola. Esta asignación, además, únicamente tiene lugar cuando existen recursos sobrantes disponibles, una vez que todo el tráfico prioritario ha sido asignado (BW_{CAC} y recursos necesarios para transmitir la información propia del encaminamiento).

La figura 3.20 muestra el tráfico correctamente cursado frente al tráfico ofrecido, evaluado en el caso de las conexiones CBR, para cada uno de los escenarios multiservicio analizados, de acuerdo a la estrategia de admisión empleada. En cada figura se incluye la curva obtenida cuando el tráfico ofrecido corresponde exclusivamente a tráfico CBR.

Como se observa, conforme aumenta el porcentaje de tráfico *best-effort* disminuyen ligeramente las prestaciones obtenidas por las conexiones con QoS, en cualquiera de las estrategias utilizadas (AODV, QoSR-BN y QoSR-CL). La continua competencia por los recursos debida a la falta de un control de admisión explícito en el AODV disminuye el tráfico correctamente cursado.

La ligera degradación observada en aplicación de la propuesta QoSR-CL no se debe, por el contrario, a la pérdida de la QoS garantizada. Sin embargo, se produce cierto incremento de la probabilidad de bloqueo que reduce el tráfico correctamente cursado. A pesar de la posibilidad de reasignación existente, un escenario de admisión donde existen ciertos recursos ocupados de baja prioridad no se corresponde exactamente con un escenario totalmente libre. Esto es debido a que la disponibilidad medida durante la fase de descubrimiento de ruta incluye todos los recursos *potencialmente* reasignables. Sin embargo, durante la fase de $RREP$, la verificación de la disponibilidad real puede determinar ciertos recursos como no disponibles, llegado incluso a impedir la

admisión de la conexión. Si bien existe una mayor probabilidad de bloquear conexiones, la diferenciación permite flexibilizar considerablemente la asignación de recursos dando prioridad a las conexiones CBR admitidas, que pueden mantener la QoS demandada a pesar de la coexistencia de tráfico no prioritario en la red (figura 3.21).



(a) 25 % BE

(b) 50 % BE

(c) 75 % BE

Figura 3.20. Tráfico correctamente cursado vs. tráfico ofrecido. Comparativa de escenarios multiservicio (100 % CBR vs. 25, 50 y 75 % BE). Escenario de 50 nodos.

La ineficacia ya demostrada de la estrategia QoSR-BN, que supone una admisión de aplicaciones CBR en exceso, es todavía más patente en un escenario multiservicio. Del mismo modo que en la estrategia QoSR-CL, la no idealidad del procedimiento de reasignación conduce a un ligero incremento en el bloqueo de conexiones. El hecho de estimar como recursos disponibles todos aquéllos que podrían ser reasignados implica en la estrategia QoSR-BN un error de estimación adicional al ya existente por la propia

medida del "cuello de botella". Como resultado, las conexiones admitidas experimentan una degradación mayor que en ausencia de sesiones BE. La diferenciación realizada por el protocolo ADHOC MAC consigue tratar de un modo preferente a las conexiones CBR, por lo que éstas compiten por los recursos con el tráfico *best-effort* con una clara ventaja. Sin embargo, dado que se admiten más aplicaciones CBR de las que el sistema es capaz de cursar correctamente, es la competencia mutua entre todas ellas la que provoca la degradación observada en la figura 3.20.

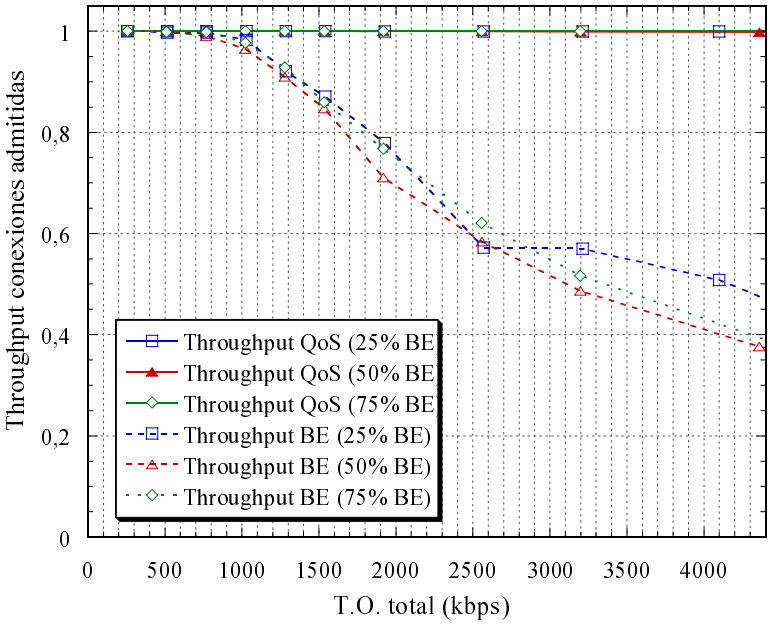
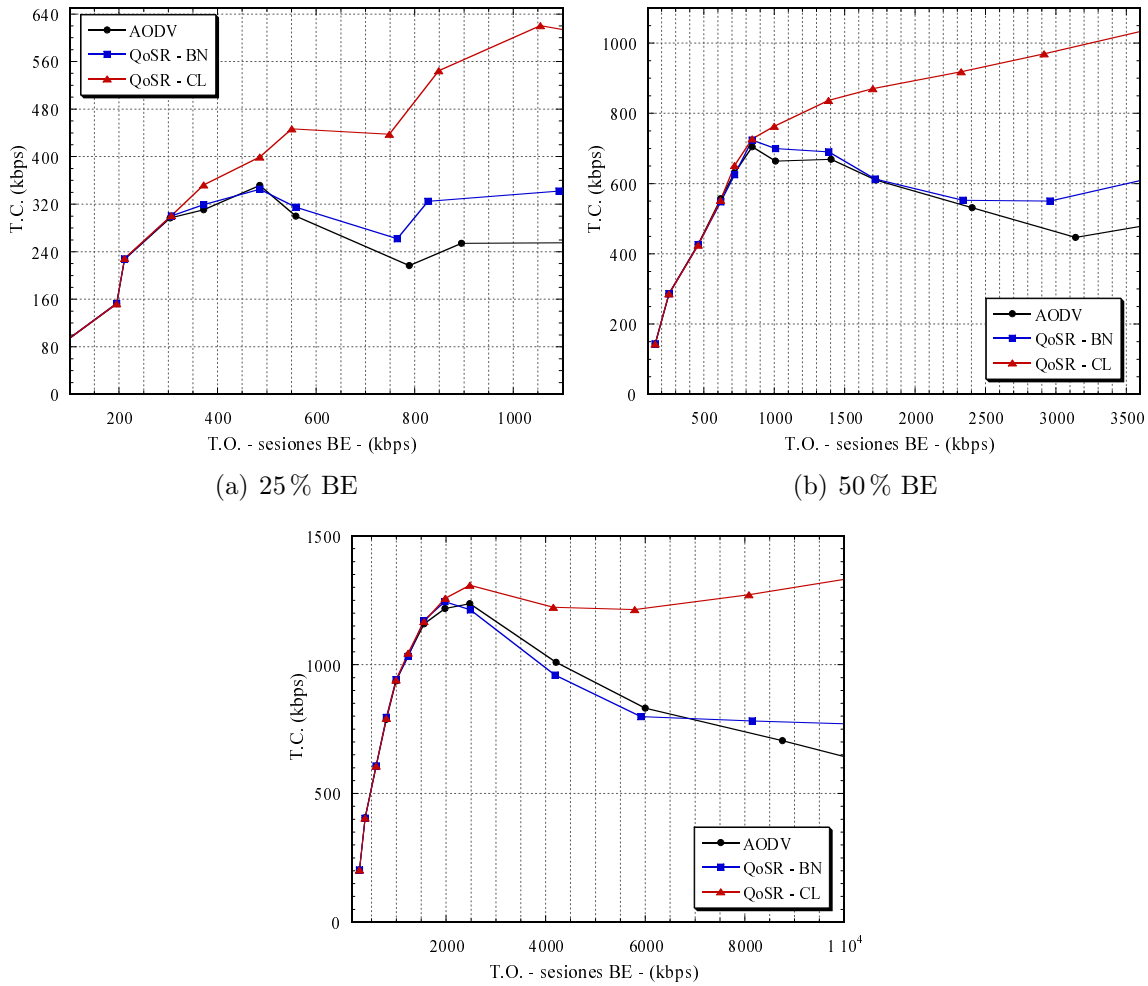


Figura 3.21. *Throughput* de las conexiones admitidas vs. tráfico ofrecido total (CBR y BE). Porcentaje de sesiones BE de 25, 50 y 75 %. Escenario de 50 nodos.

Gracias a la distribución eficiente de los recursos disponibles para las conexiones CBR, los recursos sobrantes pueden ser utilizados por las conexiones BE. Éstas, sin embargo, compiten sin restricciones de admisión, por lo que el tráfico BE efectivamente cursado disminuye conforme aumenta el correspondiente tráfico ofrecido, tal y como se observa en la figura 3.22. Sin embargo, puesto que las aplicaciones CBR ocupan exclusivamente los recursos especificados en el procedimiento de admisión, gracias a la competencia exclusiva entre aplicaciones de igual prioridad, las sesiones BE consiguen mediante la estrategia QoSR-CL unas prestaciones superiores a los casos de aplicación de las estrategias AODV o QoSR-BN.

La figura 3.23 muestra un análisis más detallado de la eficiencia de la diferenciación de servicios llevada a cabo. Se han analizado los diversos escenarios multiservicio considerados, para un caso particular de tráfico ofrecido CBR (3200 kbps). En dicha figura,

cada escenario se ha representado de acuerdo al % de tráfico ofrecido (en kbps) que representan los diversos porcentajes de conexiones BE generados, correspondiéndose de esta manera con un 30, un 56 y un 80 % del tráfico ofrecido.



(a) 25 % BE

(b) 50 % BE

(c) 75 % BE

Figura 3.22. Tráfico correctamente cursado vs. tráfico ofrecido (sesiones BE). Comparativa de escenarios multiservicio (25, 50 y 75 % BE). Escenario de 50 nodos.

La capacidad máxima del sistema para cursar dicho tráfico viene limitada por la habilidad del control de admisión de distribuir adecuadamente las conexiones en las rutas multisalto más apropiadas (circuitos virtuales). Esta medida se observa en la barra correspondiente al escenario 100 % CBR. Los recursos no asignados a circuitos virtuales pueden ser aprovechados por tráfico *best-effort*, siempre y cuando pueda garantizarse un trato preferente a las conexiones con demandas de QoS. Sin embargo, como se comprueba en la figura 3.23, conforme aumenta el porcentaje de tráfico *best-effort*, las

prestaciones para tráfico CBR se degradan ligeramente. Como se ha comentado a lo largo de la descripción de la arquitectura propuesta, dada la naturaleza distribuida del control de admisión, la capacidad de reasignar recursos menos prioritarios para nuevas conexiones CBR no es completamente ideal. A pesar de esta limitación, el tráfico efectivamente cursado se encuentra muy próximo al ideal. Adicionalmente, cuanto mayor es el porcentaje de tráfico *best-effort* ofrecido, se cursa un volumen de tráfico mayor, lo que constata el aprovechamiento de recursos conseguido.

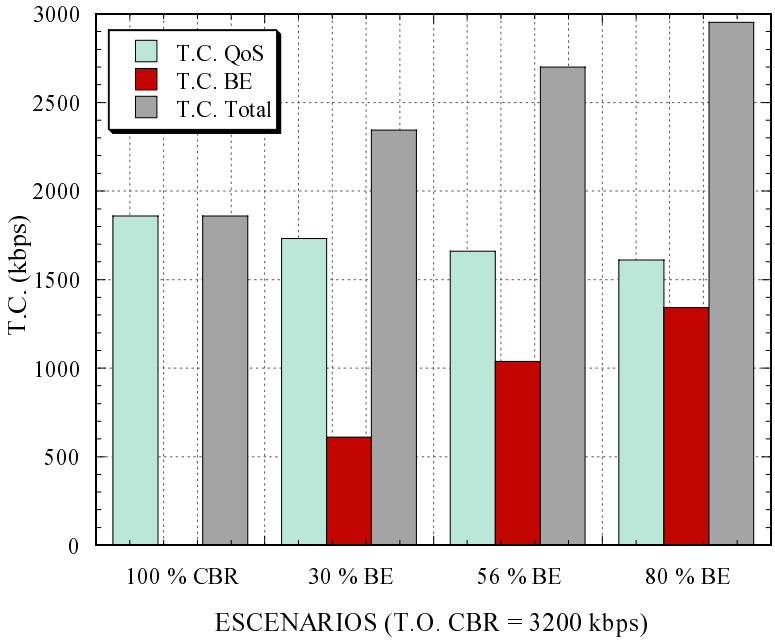


Figura 3.23. Tráfico cursado CBR y BE para los diversos escenarios (100 % del tráfico ofrecido CBR; tráfico adicional BE (30, 56 y 80 % del tráfico, en kbps); tráfico ofrecido CBR = 3200 kbps; escenario de 50 nodos.

Capítulo 4

Adaptación a la variabilidad del entorno

El control de admisión descrito en §3 asume una estimación fiable del ancho de banda extremo a extremo (algoritmo *BWC-FA*) y la consiguiente selección y reserva permanente de recursos libres de colisiones (*conflict-free*). Dicha fiabilidad se apoya en la señalización proporcionada por el protocolo ADHOC MAC, capaz de eliminar el problema del terminal oculto y garantizar el reúso evitando terminales expuestos. El análisis del esquema propuesto mediante un modelo de interferencia simplificado (condiciones estáticas, *Protocol Model*) demuestra la certeza de dicha afirmación. Sin embargo, el contexto de aplicación real – MANETs: redes inalámbricas, móviles y distribuidas – implica un escenario variante que requiere la adaptación de las soluciones inicialmente adoptadas.

Dada la existencia de un medio de transmisión compartido e inalámbrico, toda transmisión activa genera niveles de interferencia sobre el resto de terminales receptores (*Physical Model*). Dichos niveles, además, varían de acuerdo al dinamismo de la red: actividad intermitente de los terminales, movilidad, desconexiones debidas a fallos de alimentación o batería insuficiente, ajustes derivados del control de potencia y desvanecimientos del canal. Como consecuencia, la relación señal a interferencia (SIR) en recepción puede llegar a ser muy variante. Una transmisión se considera correcta si la SIR con la que se recibe la señal es mayor o igual a la mínima necesaria para decodificar adecuadamente la información. De este modo, la fiabilidad de las reservas estará condicionada por la capacidad de mantener la SIR por encima del umbral requerido a pesar de las variaciones sufridas. Una vez reservado un recurso, el receptor

de la comunicación puede decodificar erróneamente la información, con cierta probabilidad, en función de la variación de la SIR experimentada. La pérdida total de la reserva estará determinada por la tolerancia a errores considerada. A diferencia de lo analizado mediante el *Protocol Model*, una transmisión errónea puede no ser resultado de la colisión debida a dos transmisiones concurrentes realizadas dentro del rango de cobertura. El problema del terminal oculto, si bien existe igualmente, no es la única causa de errores en la transmisión. Por lo tanto, la señalización básica del protocolo ADHOC MAC, desarrollada para permitir un acceso distribuido capaz de eliminar este problema, garantizando reservas totalmente fiables, resulta insuficiente en este nuevo escenario.

Como consecuencia, se presentan diversos efectos a los que resulta necesario adaptar la propuesta original:

- Comunicaciones establecidas asumiendo su fiabilidad no están exentas de sufrir cierta degradación ante las variaciones de SIR experimentadas. Así pues, el ancho de banda asignado por el control de admisión no está totalmente garantizado.

- La estimación del ancho de banda disponible depende de la definición de estado realizada por el ADHOC MAC. En un escenario de interferencia variable, la información transmitida en el BCH de los vecinos a un salto puede no ofrecer una visión realista de la probabilidad de mantener un recurso estable. Además, dado que alguno de los slots BCH puede no ser decodificado correctamente, se genera cierta imprecisión al no tener en cuenta toda la información. Como resultado, una sobreestimación de la capacidad de la red limita las bondades del control de admisión.

- Las variaciones de SIR limitan la posibilidad de establecer comunicaciones fiables entre nodos supuestamente vecinos, por lo que dicha vecindad resulta ser muy variante y no sólo dependiente de la distancia entre nodos. La conectividad de la red, por lo tanto, puede sufrir grandes oscilaciones incluso en condiciones de baja movilidad, generando gran inestabilidad.

Teniendo en cuenta todas estas limitaciones, en este capítulo se analizan los problemas específicos del control de admisión original desarrollando en consecuencia soluciones que permitan adaptar la arquitectura *cross-layer* propuesta a un escenario variante. Un primer apartado analiza la problemática derivada de la falta de fiabilidad en las reservas y la inestabilidad del sistema, centrando las soluciones en el nivel MAC [Canales

y otros, 2005b]. Mediante dicho análisis se pretende garantizar un nivel de enlace que sirva de soporte fiable para el encaminamiento, tratando de independizar, en la medida de lo posible, la funcionalidad específica de ambos niveles. Posteriormente se presentan los mecanismos de adaptación desarrollados a partir de la cooperación explícita del encaminamiento y la gestión de recursos del nivel MAC [Canales y otros, 2006a,b,c]. Se ha considerado un escenario donde los factores de variabilidad determinantes son la intermitencia en la activación de conexiones, que provoca variaciones en los niveles de interferencia, y la movilidad de los nodos, considerando pérdidas por propagación sin desvanecimientos del canal.

4.1. Implicaciones del *Physical Model*. Diseño de un MAC eficiente

4.1.1. Mantenimiento de reservas fiables

A pesar de la señalización proporcionada por el protocolo ADHOC MAC, las comunicaciones establecidas mediante el procedimiento BIAS no están libres de colisiones. La capacidad de decodificar correctamente la información no depende únicamente de la distancia entre transmisor y receptor, (*Protocol Model*: rango de cobertura) sino también de la interferencia que éste último sufre como consecuencia de otras comunicaciones activas (3.1) - *Physical Model*. En este escenario, no puede definirse un rango de cobertura constante, sino que la correcta recepción de la información dependerá de la interferencia sufrida por cada vecino, tal y como se muestra en la figura 4.1.

En estas condiciones se plantean dos limitaciones principales: la estimación errónea de los recursos disponibles y la posible degradación de las comunicaciones activas. Ambos efectos se relacionan con la pérdida de la propiedad ideal de no-conflicto. Slots aparentemente libres son susceptibles de sufrir futuras colisiones de manera que su definición de disponibilidad puede no ser exacta. De igual modo, al establecerse nuevas conexiones pueden generarse colisiones sobre comunicaciones lejanas. La definición de **colisión**, en el escenario de interferencia considerado, sin desvanecimientos de canal, supone una degradación no tolerable de la SIR debida a:

- La activación de una nueva comunicación en el mismo slot.
- La reducción de la potencia de señal recibida debida a la movilidad relativa de los nodos implicados.

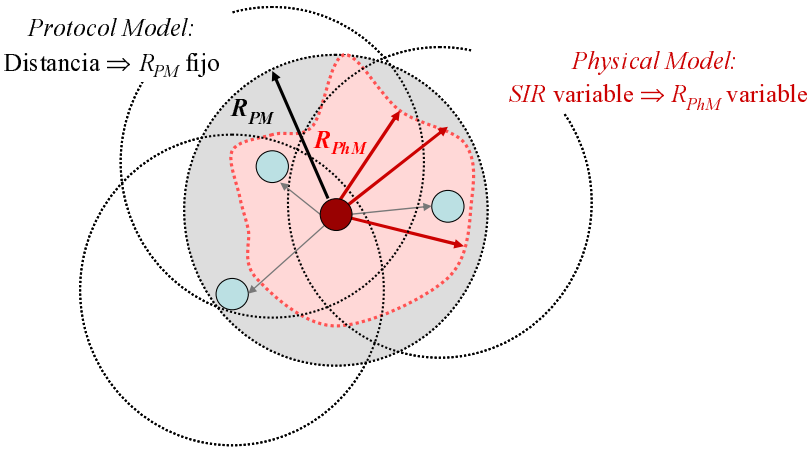


Figura 4.1. Rango de cobertura variable debido a la variabilidad de la SIR. (PM = *Protocol Model*, PhM = *Physical Model*)

En cualquiera de los casos, la interferencia sufrida impide la recepción correcta de la información mientras las comunicaciones concurrentes sobre el mismo slot permanecen activas. Esto supone, por lo tanto, la necesidad de asignar inmediatamente un nuevo recurso para la misma comunicación. Es importante hacer esta distinción dado que en un escenario que incluya oscilaciones en la potencia de señal debidas a desvanecimientos de canal o ajustes de la potencia de transmisión, una recepción errónea puede no ser siempre consecuencia de la activación de una comunicación coincidente en ese slot. En este contexto, la decisión de reasignación debería realizarse únicamente con la seguridad de estar sufriendo una degradación permanente no tolerable en base a sucesivas medidas tanto de la potencia recibida como de la SIR. De hecho, un canal de transmisión de baja calidad (desvanecimientos profundos) puede producir el mismo efecto que una colisión (alta probabilidad de error como resultado de una degradación continuada de la SIR) sin la existencia de comunicaciones coincidentes en el slot. Como resultado, será igualmente necesaria una reasignación de recursos para mantener la calidad deseada de la comunicación [Gállego, 2007].

1. Estimación errónea de recursos disponibles:

La señalización básica del protocolo ADHOC MAC resulta insuficiente para determinar de una forma realista la disponibilidad de recursos libres de colisiones, como puede observarse en la figura 4.2.

Asumiendo que todos los terminales implicados disponen de un slot BCH donde transmitir la información de control, completamente libre de colisiones, la señalización intercambiada genera el siguiente escenario a la hora de gestionar los recursos: Dado que

todos los vecinos del terminal A transmiten en un slot libre de conflictos, dicho terminal es capaz de decodificar el BCH de todos ellos, lo que define para ese nodo su conjunto de vecinos a un salto (rango de cobertura del nodo A equivalente al definido de acuerdo al *Protocol Model*). De la información recibida de los terminales F,B y C acerca del slot k se concluye que dicho recurso está ocupado por una comunicación que no "afecta" a los nodos B y F, por lo que podría reutilizarse en una nueva conexión con cualquiera de ellos. Sin embargo, en el caso de activarse dicha comunicación, la interferencia real generada por la comunicación entre D y E influiría en ambos receptores. Mientras el terminal B experimentaría una SIR por debajo del umbral de decodificación, el terminal F recibiría correctamente. Dado que tanto F como B señalizan el slot como LIBRE, A puede intentar realizar la reserva con cualquiera de ellos. La aplicación de las reglas del protocolo resolvería de manera favorable una de las reservas (dependiendo del criterio aplicado, [Gállego, 2007]). Sin embargo, a la hora de realizar la transmisión efectiva en el slot k, la comunicación en el enlace (A,B) sufriría una colisión.

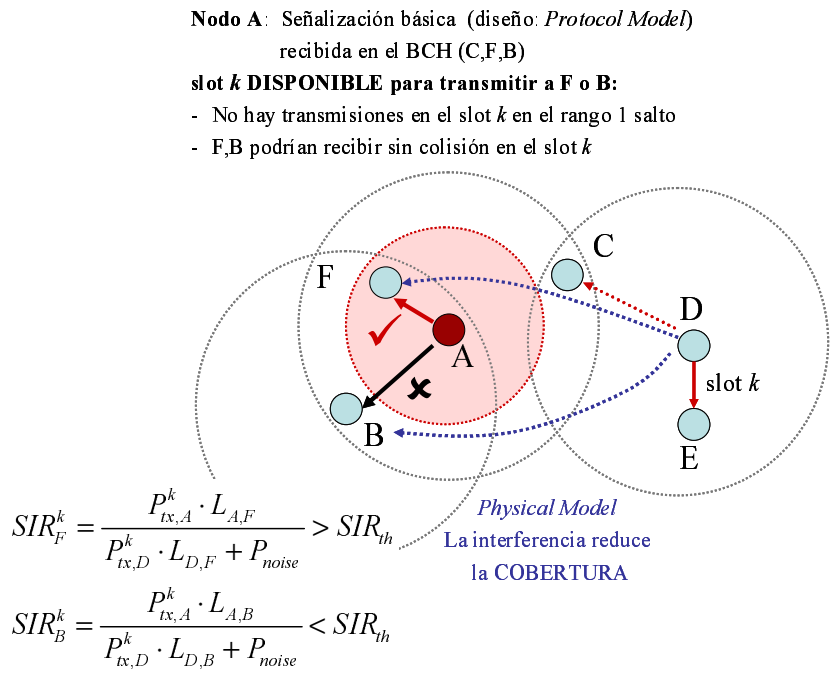


Figura 4.2. *Physical Model*. Disponibilidad dependiente de la interferencia de transmisiones activas lejanas. Señalización básica del ADHOC MAC insuficiente.

2. Degradación de comunicaciones activas:

La activación de una reserva supuestamente fiable ($SIR > SIR_{th}$) no está exenta de sufrir futuras colisiones debido a la posible activación de conexiones lejanas que idealmente no afectarían por encontrarse fuera del rango de cobertura teórico del receptor.

Considerando la definición de colisión utilizada, y siguiendo el ejemplo anterior, si se observa la figura 4.3, una vez activa la comunicación entre A y F (posible según el nivel de SIR en F), ésta influye sobre la comunicación previamente establecida entre D y E, de manera que la SIR en el receptor E decae en función de la nueva interferencia sufrida. Si dicha disminución supusiera un descenso de la SIR por debajo del umbral, la conexión experimentaría una colisión, por lo que la reserva, establecida inicialmente como fiable, se perdería.

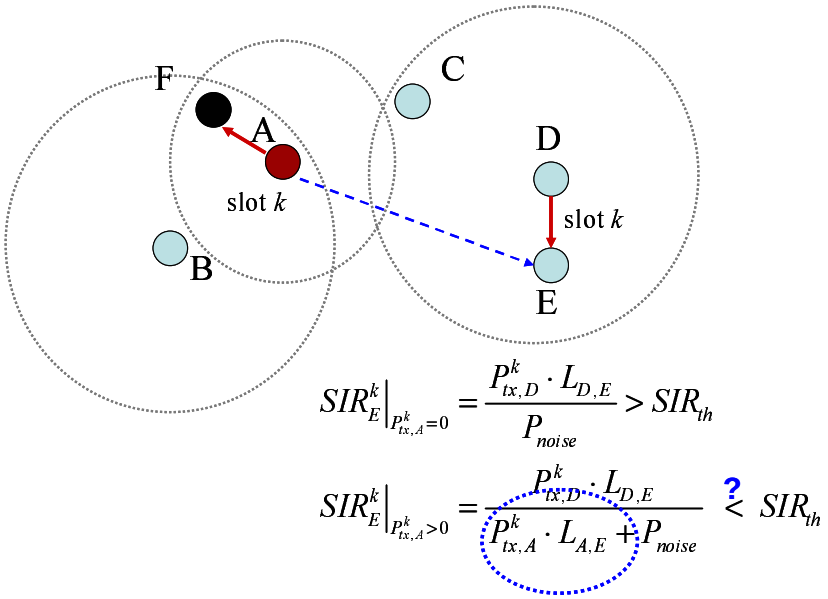


Figura 4.3. *Physical Model*. Incremento de la interferencia sobre comunicaciones activas.

Con objeto de hacer frente a estas limitaciones, es necesario incrementar la señalización proporcionada por el protocolo ADHOC MAC garantizando:

- Una definición de ocupación acorde con la naturaleza interferente del entorno capaz de identificar apropiadamente los recursos disponibles.
- Un acceso compartido a los recursos que permita el establecimiento de reservas fiables con el menor impacto posible sobre los recursos ya ocupados.

Como primera aproximación, en [Gállego y otros, 2005b] se plantea la definición de un nuevo estado de slot: DIRTY, además de los ya existentes de LIBRE, OCUPADO y ACCEDIENDO. Se identifican, de esta manera, los slots donde se detecta cierto nivel de potencia pese a no decodificar la información (lo que a priori sería señalizado como

slots LIBREs). Sería el caso, por ejemplo, de los nodos F y B en la figura 4.3. La señalización del slot k como DIRTY en sus respectivos BCH permite que el nodo A conozca la posibilidad de que los potenciales receptores F y B sufran interferencia en dicho slot. De esta manera, se evita el intento de reservar recursos que pueden fallar, resolviendo así la detección errónea de recursos disponibles. Esta solución, sin embargo, sigue presentando ciertas limitaciones:

1. Evitar la utilización de todos los slots señalizados como DIRTY puede producir una reducción excesiva de la capacidad de reúso de la red (en el ejemplo descrito se impedirían las comunicaciones paralelas de D a E y de A a F que, dependiendo del nivel de SIR, podrían establecerse correctamente). Un control de potencia eficiente puede paliar este efecto, como se describirá a continuación.

2. La imposibilidad de decodificar la información transmitida por receptores lejanos (BCH de terminales fuera del rango de cobertura) impide evitar completamente el impacto que una nueva activación puede tener sobre los mismos. La limitación en la potencia transmitida para una nueva comunicación (control de potencia) reduce el rango de influencia de la misma, aliviando en parte este problema.

3. Un slot señalizado como LIBRE no garantiza el éxito de la reserva, ya que la sensibilidad de los receptores limita la capacidad de sensado de los terminales (umbral de sensado CS_{th} – *Carrier Sense* (CS)). A pesar de no detectar potencia en los alrededores, un enlace con un margen muy limitado frente a interferencia puede verse afectado por niveles de potencia por debajo de dicho umbral. Dada la naturaleza impredecible de la potencia no sensada, no es posible asegurar una reserva estable en este tipo de enlaces, lo que requerirá una restricción adicional en el establecimiento de comunicaciones (margen de seguridad, que se explicará en el apartado §4.1.2).

Frente a la reducción excesiva de la capacidad de reúso provocada por la inclusión de los slots DIRTY, la implementación de mecanismos eficientes de control de potencia puede permitir la reutilización de recursos gracias al ajuste de la potencia de transmisión de los terminales al mínimo necesario. Dicho ajuste, por otra parte, reduce el consumo energético en la red, que es una de las limitaciones de diseño más importantes en el entorno de las MANETs. La reducción conseguida en los niveles de interferencia generados, permite disminuir el número de slots DIRTY, aumentando la disponibilidad. La transmisión explícita de información relativa a la interferencia soportada por

los receptores podría permitir, además, la activación de transmisiones en estos slots DIRTY. Gracias a dicha reutilización, la capacidad del sistema podría verse incrementada [Gállego, 2007].

El incremento de capacidad a nivel de enlace se ha analizado en detalle en [Gállego y otros, 2006a,b]. Sin embargo, con objeto de aislar la problemática asociada a la integración del control de potencia sobre el protocolo ADHOC MAC[1] de la evaluación del control de admisión desarrollado (propuesta *cross-layer*), se ha considerado únicamente un control de potencia básico. La potencia de transmisión se ajusta al nivel necesario para conseguir una SIR objetivo específica, una vez conocidas las pérdidas de propagación, sin considerar desvanecimientos del canal.

Aún optimizando en la medida de lo posible la potencia de transmisión necesaria, garantizando unos niveles de interferencia considerablemente menores, sigue siendo posible que una conexión ya establecida pueda verse degradada a consecuencia de una nueva activación. El efecto sobre comunicaciones muy lejanas se debe principalmente a la limitación en la sensibilidad de los receptores, lo que se resolverá mediante un margen de seguridad (§4.1.2). Sin embargo, los receptores fuera del rango de cobertura, suficientemente cerca como para verse afectados por la potencia de transmisión (aun limitando ésta por el control de potencia), tienen potencialmente un riesgo mucho mayor de ver degradadas sus comunicaciones. Dada la imposibilidad de decodificar correctamente el BCH, el estado de sus reservas es desconocido. Siguiendo el ejemplo de la figura 4.2, desde el punto de vista del potencial transmisor interferente (nodo A), la detección de una potencia de transmisión, sea o no decodificable (nodo D), no limita la disponibilidad, siempre y cuando no existan receptores cerca. En caso de que éstos se encuentren dentro del área de cobertura, su BCH informa a dicho transmisor evitando el acceso, pero cuando el BCH no es decodificado (nodo E), ningún mecanismo evita el acceso a los recursos, siendo así muy probable la colisión. Una posible solución para proteger estas reservas es la inclusión de una pequeña señal de alerta ($BEACON$) al inicio de cada slot, transmitida por el terminal receptor [Gállego y otros, 2005b]. Dicha alerta permite que todos aquellos nodos incapaces de decodificar el BCH en el que el receptor informa de su estado, puedan sensar cierta potencia e interpretar el slot como OCUPADO. En la figura 4.4 se muestra la utilización del $BEACON$ aplicada al ejemplo anteriormente descrito.

Puesto que el nodo E se encuentra fuera del rango de cobertura máximo del nodo A, éste no es capaz de decodificar el BCH que transmite E, donde informa de la transmisión

[1] Mecanismos de control de potencia desarrollados en [Gállego, 2007]

que recibe del nodo D. Como ya se comentó anteriormente, A podría establecer una comunicación paralela – enlace (A,F) – sin conocer el efecto que la misma tiene sobre E. La detección en el slot k del $BEACON$ transmitido por E, matiza la definición de DISPONIBLE que A ha establecido tras la recopilación de lo recibido en dicho slot (campo de datos) y de los FIs recibidos de sus vecinos.

Al igual que la reducción de capacidad de reúso que supone la inclusión del estado DIRTY, asumir que todos los slots en los que se detecta cierta potencia de $BEACON$ están ocupados puede resultar demasiado conservador (realmente existirá colisión únicamente cuando la SIR decrezca por debajo del umbral SIR_{th}).

A diferencia de la detección de colisiones realizada en función de la señalización explícita recibida en el BCH, que únicamente tiene efecto dentro del rango de cobertura, la transmisión de una señal de $BEACON$, sin información de control incorporada, permite identificar a los receptores cercanos pero no decodificables, evitando así una activación dañina de comunicaciones, protegiendo las reservas ya establecidas y garantizando en mayor medida la estabilidad de las mismas. Esta solución puede ser especialmente atractiva a la hora de proteger los recursos más prioritarios, como pueden ser los asignados específicamente durante el procedimiento de control de admisión (BW_{CAC}).

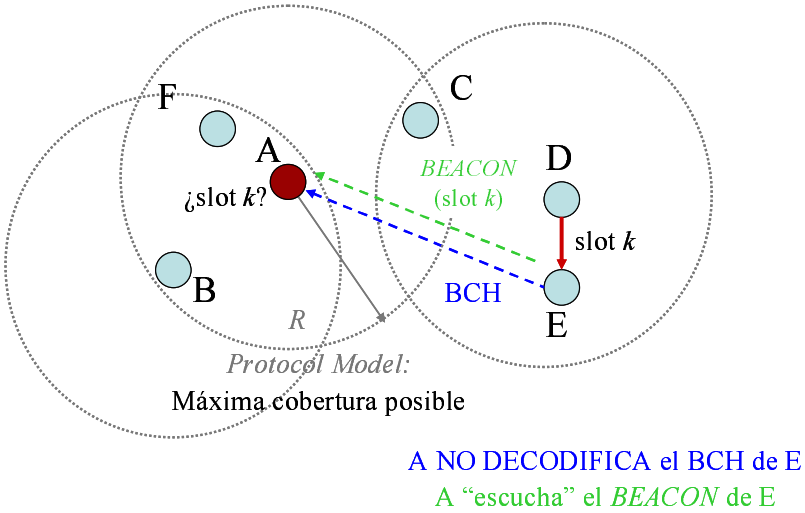


Figura 4.4. Transmisión de señal $BEACON$.

4.1.2. Estabilidad del sistema

La problemática descrita tiene un impacto mayor en los enlaces entre nodos cuya distancia es próxima al límite de la cobertura teórica (rango del *Protocol Model*). Dichos enlaces disponen de un margen muy limitado frente a interferencias (figura 4.5), por lo que los slots reservados pueden potencialmente sufrir una colisión, ya sea en el mismo momento de establecerse, debido a una incorrecta definición de los mismos como disponibles, o posteriormente, por la activación futura de una transmisión interferente. Son lo que se ha denominado como **enlaces débiles**.

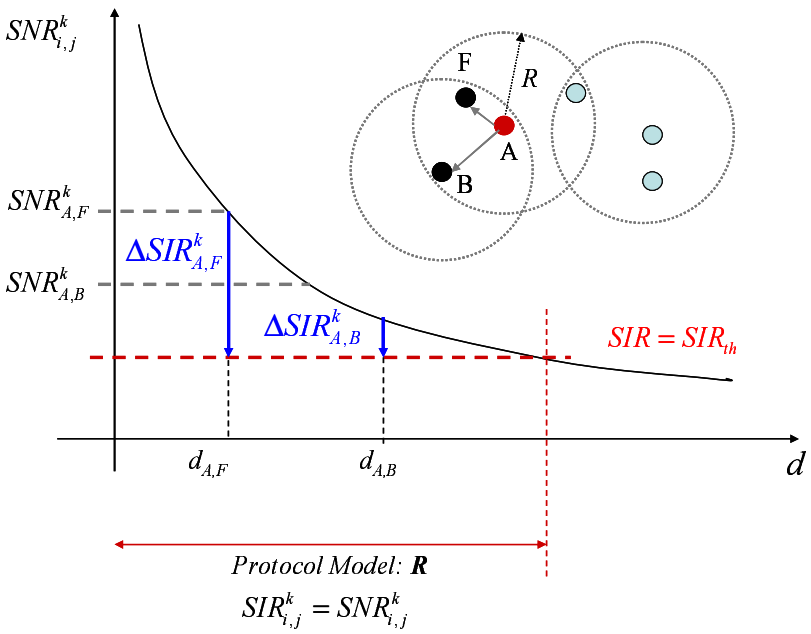


Figura 4.5. Margen frente a interferencias.

En la figura 4.5 se expresa el margen en función de la relación señal a ruido (*SNR – Signal to Noise Ratio*), interpretada ésta como la SIR teórica, en ausencia de interferencia, equivalente a la utilización del *Protocol Model*:

$$SIR^k_{i,j}\big|_{teórica} = SNR^k_{i,j} = \frac{P^k_{tx,i} \cdot L_{i,j}\left(d_{i,j}\right)}{P_{n,j}} \tag{4.1}$$

$$\Delta SIR^k_{i,j} = SNR^k_{i,j} - SIR_{th} \tag{4.2}$$

Los recursos reservados en enlaces débiles tienen una probabilidad mayor de sufrir una colisión. La pérdida de dichos recursos implica la necesidad de realizar nuevas asig-

naciones que, por el hecho de realizarse en los mismos enlaces, siguen teniendo una elevada probabilidad de fallar. Esta inestabilidad en las reservas repercute negativamente en las prestaciones de las conexiones de usuarios que demandan los recursos. Parece evidente, por lo tanto, que estos enlaces no son apropiados para formar parte de ninguna ruta con QoS.

Por otra parte, la inestabilidad en las reservas tiene un efecto mucho más contraproducente sobre el establecimiento de slots BCH:

- Un slot BCH fiable es condición necesaria para permanecer como usuario activo dentro del sistema. Todo terminal activo que transmite su BCH es detectado por sus vecinos. Un BCH inestable implica una densidad de vecinos muy variante y una definición de estado para los slots imprecisa, puesto que ésta depende de los FIs recibidos de los vecinos. El funcionamiento correcto del protocolo distribuido, tanto en la gestión de los recursos como en el propio mecanismo de acceso, es altamente sensible a dicha imprecisión.
- El algoritmo de encaminamiento con QoS se apoya en una medida fiable del ancho de banda disponible proporcionada por el protocolo ADHOC MAC. Dicha medida depende directamente de la precisión en la definición del estado de los slots. Por otra parte, el método de acceso BIAS, encargado de materializar las reservas necesarias, se sustenta en la transmisión de las peticiones sobre el slot BCH.

En conclusión, un recurso reservado, ya sea el propio slot BCH o un slot utilizado para transmitir datos de usuario, debe permanecer suficientemente estable dentro del entorno variante producido por las oscilaciones en los niveles de interferencia. Para garantizar esta estabilidad, se propone la utilización de un margen de protección frente a interferencias. Dicho margen, supone establecer enlaces entre nodos suficientemente cercanos para que la tolerancia a oscilaciones de la SIR sea mayor. Esta medida podría considerarse una métrica de QoS adicional a la del ancho de banda, que fuera tenida en cuenta como criterio de encaminamiento, de manera que se establecieran rutas con suficiente ancho de banda que incluyeran únicamente enlaces estables[2] [Canales y otros, 2005b]. No obstante, como ya se ha comentado anteriormente, la inestabilidad de las reservas resulta mucho más perjudicial en los slots BCH, propios del nivel de enlace, cuyo establecimiento no depende de la activación de ninguna comunicación extremo a extremo. El mantenimiento de dichos recursos es especialmente importante para garantizar la estabilidad desde un punto de vista de sistema, de manera que se gestione

[2]Selección de rutas con enlaces "cortos", a costa de incrementar el número de saltos.

eficientemente tanto de la conectividad como la utilización de los recursos suponiendo así un soporte fiable para la arquitectura *cross-layer* de manera generalizada. Teniendo esto en cuenta, se ha considerado más apropiado trasladar esta métrica al propio nivel MAC. En este sentido, en [Gállego y otros, 2006b] se propone la utilización de un margen de protección frente a interferencias definido de acuerdo a un requerimiento de potencia mínima en recepción.

Se define $P_{rx,min}$ (4.3) como el valor de potencia de señal en recepción necesario para tener una SIR teórica equivalente (*Signal to Noise Ratio* (SNR)) ΔSIR dBs superior a SIR_{th}.

$$P_{rx,\min} : SNR = \frac{P_{rx,\min}}{P_n} = SIR_{th} + \Delta SIR \tag{4.3}$$

Teniendo en cuenta dicha potencia, un nodo considera como vecinos estables a todos aquellos nodos de los que es capaz de decodificar su BCH (vecino j, slot s_{BCH_j}), y de los recibe una potencia igual o superior a $P_{rx,min}$. De acuerdo a este criterio, se define NB_i como el conjunto de vecinos "a un salto", de los que se recibe un BCH decodificable ($SIR > SIR_{th}$), y NB_i^{est} como el subconjunto de vecinos estables, con los que puede garantizarse además un margen de seguridad de ΔSIR dBs.

$$NB_i = \left\{ n_j \in T : SIR_{rx,j,i}^{s_{BCH_j}} \geq SIR_{th} \right\} \tag{4.4}$$

$$NB_i^{est} \subset NB_i = \left\{ n_j \in NB_i : P_{tx,j}^{s_{BCH_j}} \cdot L_{j,i} \geq P_{rx,\min} \right\} \tag{4.5}$$

La definición del conjunto de vecinos estables tiene vital importancia dentro de la arquitectura *cross-layer* propuesta. La calidad de las posibles transmisiones con cualquier vecino decodificable pero fuera del conjunto NB_i^{est} puede verse rápidamente degradada por interferencia, por lo que no se contempla la activación de comunicaciones de datos con estos nodos. De este modo, la selección de la potencia mínima $P_{rx,min}$ define el rango máximo de cobertura (R_{max}^{est}) determinando por lo tanto la conectividad de la red así como la capacidad de la misma [Gupta y Kumar, 2000]. La cobertura real se reduce respecto a la teórica determinada por el umbral de decodificación (*Protocol Model*) ($R_{max}^{est} < R$). Sin embargo, se proporciona un margen de seguridad frente a interferencias que pretende garantizar la estabilidad de las transmisiones realizadas con cualquiera de los vecinos de dicho rango.

Es importante hacer una distinción clara entre vecinos estables (NB_i^{est}) y vecinos decodificables fuera del rango R_{max}^{est} ($NB_i \cap \overline{NB_i^{est}}$). Los primeros determinan la con-

figuración real de la topología de red, en cuanto al establecimiento de enlaces y las consiguientes rutas multisalto. Por lo tanto, la información de control recibida de los mismos (FI en el BCH) es determinante a la hora de mantener una información actualizada de la conectividad de la red, con el consiguiente impacto sobre el encaminamiento. De los segundos, sin embargo, se puede decodificar la información de control con cierta probabilidad, dependiendo de la interferencia sufrida. Dada su inestabilidad, no son considerados dentro de las operaciones de mantenimiento de la conectividad. No obstante, en el caso de que su información sea correctamente decodificada, debe tenerse en cuenta a la hora de definir el estado de ocupación de los slots (actualización del $Status - Channel$[3]), con objeto de evitar una degradación de las reservas realizadas por dichos vecinos o una estimación errónea de los recursos disponibles.

De este modo, esta nueva definición de vecinos repercute directamente sobre las operaciones de gestión realizadas en el nivel de enlace:

- **Mantenimiento del slot BCH:** Un terminal i debe disponer permanentemente de un slot BCH asignado dentro de la trama para considerarse activo dentro el sistema. La reserva debe ser continuamente verificada mediante reconocimiento por parte de sus vecinos, es decir, siempre y cuando éstos identifiquen en su FI el slot ocupado por el terminal i. Una transmisión errónea de BCH tiene como consecuencia una reasignación de slot con objeto de establecer una nueva reserva libre de colisiones. Los vecinos no estables suponen transmisiones de BCH intermitentes. Si éstas fueran tenidas en cuenta, la verificación de la transmisión correcta del BCH sería frecuentemente errónea, lo que tendría como resultado continuas reasignaciones de slots, con la consiguiente pérdida de actividad en el sistema. Por el contrario, la recepción de un reconocimiento negativo de uno de los vecinos estables se considera suficientemente determinante como para decidir la reasignación del slot BCH. En caso de tener en cuenta desvanecimientos de canal, como se explicó anteriormente, la decisión debería ser más conservadora (reasignación tras un número finito de BCHs fallidos, por ejemplo) [Gállego, 2007].

- **Identificación de vecinos:** Gracias al funcionamiento intrínseco del protocolo ADHOC MAC, la necesidad de transmitir periódicamente un slot BCH permite identificar la actividad de los nodos vecinos. No recibir correctamente el BCH de un terminal que ha sido decodificado en tramas anteriores permite percibir la posible pérdida de éste como vecino. Este hecho, tiene mayor impacto en el

[3]Anexo A.

contexto *cross-layer* considerado, especialmente en un entorno verdaderamente dinámico. El mantenimiento de la rutas estables por parte del encaminamiento se basa en una pronta reacción ante roturas. Éstas se identifican gracias a la información recibida del ADHOC MAC. Cuando este protocolo detecta la pérdida de un vecino, lo notifica al nivel superior. En esta ocasión, una detección prematura de la pérdida de un vecino debida a una desconexión momentánea, puede tener consecuencias negativas más allá de la propia gestión del nivel de enlace: en un escenario muy variante, las roturas de las rutas tienen graves repercusiones sobre las prestaciones de las aplicaciones establecidas en ellas (pérdida de paquetes), incrementando además el *overhead* de control debido a las nuevas búsquedas de caminos. Por lo tanto, a la hora de determinar la pérdida de un vecino estable, la decisión es tomada una vez transcurrido un número finito de tramas ($n_{frnolisten}$) sin recibir el BCH de ese nodo con la potencia requerida ($P_{rx,min}$).

El valor adecuado de $P_{rx,min}$ depende del escenario concreto de aplicación. Un escenario variante debido a oscilaciones en el canal de propagación (desvanecimientos) supone una probabilidad de error mayor con niveles de interferencia menores. Así pues, la elección de la $P_{rx,min}$ debe realizarse con la intención de mantener una potencia media de recepción entre vecinos capaz de garantizar un margen adicional, sobre el umbral mínimo de detección, capaz de absorber las oscilaciones de SIR esperadas en dicho escenario, tal y como se muestra en la figura 4.6.

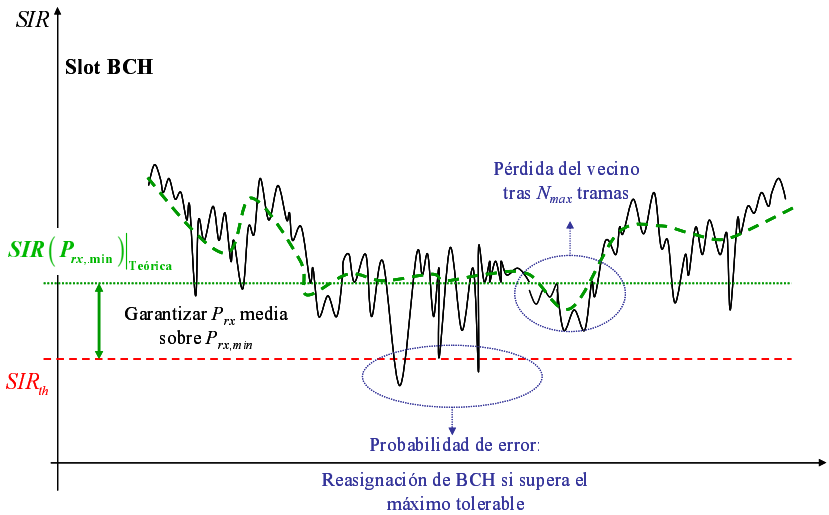


Figura 4.6. Margen de seguridad frente a oscilaciones de la SIR. Definición de $P_{rx,min}$.

Dado que en nuestro caso no se han considerado desvanecimientos del canal, las oscilaciones de la SIR se deben exclusivamente a la activación y desactivación de conexiones y a la movilidad, lo que reduce el margen de protección necesario. Si bien un análisis detallado de la configuración adecuada de este parámetro queda fuera del alcance del trabajo desarrollado en esta tesis[4], en §4.3 se analizará el impacto del mismo sobre las prestaciones del sistema.

4.2. Procedimiento de monitorización de la QoS y readmisión

El control de admisión distribuido permite realizar una asignación eficiente de recursos adaptada a las demandas de las aplicaciones evitando congestionar la red por encima de su capacidad. De esta forma se garantizan las prestaciones para los servicios admitidos a costa de limitar el acceso al sistema de aplicaciones a las que no se podría proporcionar un servicio satisfactorio. No obstante, la variabilidad de la red como consecuencia de la movilidad de los nodos y las oscilaciones en el tráfico implica una constante modificación de las condiciones iniciales de admisión. Una vez activo un determinado conjunto de aplicaciones, la evolución de dichas condiciones conduce a un escenario completamente distinto al de su admisión. Como resultado, la nueva topología puede implicar que los caminos elegidos durante dicho proceso dejen de ser adecuados para proporcionar la QoS esperada, provocando por lo tanto la degradación de las prestaciones de las aplicaciones correspondientes. Esta degradación, sin embargo, puede deberse a dos efectos diferenciados:

- **La pérdida de conectividad**: Un enlace roto supone la rotura simultánea de todas las conexiones activas que lo incluyan en su camino, implicando la pérdida de paquetes y retardos que pueden llegar a no ser asumibles por algunas de las aplicaciones.

- **La pérdida de ancho de banda**: Enlaces todavía activos experimentan un ancho de banda variable debido a las oscilaciones en la interferencia sufrida. El algoritmo *BWC-FA* asume que la selección de recursos realizada implica que las conexiones admitidas no se interfieran entre sí, lo cual es únicamente cierto en

[4]Diseño y evaluación realizados en [Gállego, 2007].

la topología inicial de admisión[5]. La variación real en la interferencia sufrida condiciona el mantenimiento de los slots reservados, pudiendo incluso provocar la pérdida de los mismos (colisiones).

Cualquiera de las situaciones descritas conlleva la necesidad de iniciar un procedimiento de readmisión que determine si el sistema todavía es capaz de proporcionar un servicio ajustado a las demandas de la aplicación. El criterio utilizado, sin embargo, depende de la naturaleza de la degradación sufrida. La figura 4.7 muestra el esquema de admisión propuesto [Canales y otros, 2006b], diferenciando los criterios aplicados sobre conexiones completamente nuevas y las ya activadas, señalando además la diferencia en el tratamiento de las roturas y las pérdidas de ancho de banda (criterios II y III)[6]. En dicha figura, P_b denota la probabilidad de bloquear una nueva conexión, mientras que P_d es la probabilidad de expulsar una conexión que fue originalmente admitida. A continuación se detallan los mecanismos desarrollados para implementar el procedimiento mostrado, resumiendo en §4.2.3 la correcta definición de los criterios establecidos.

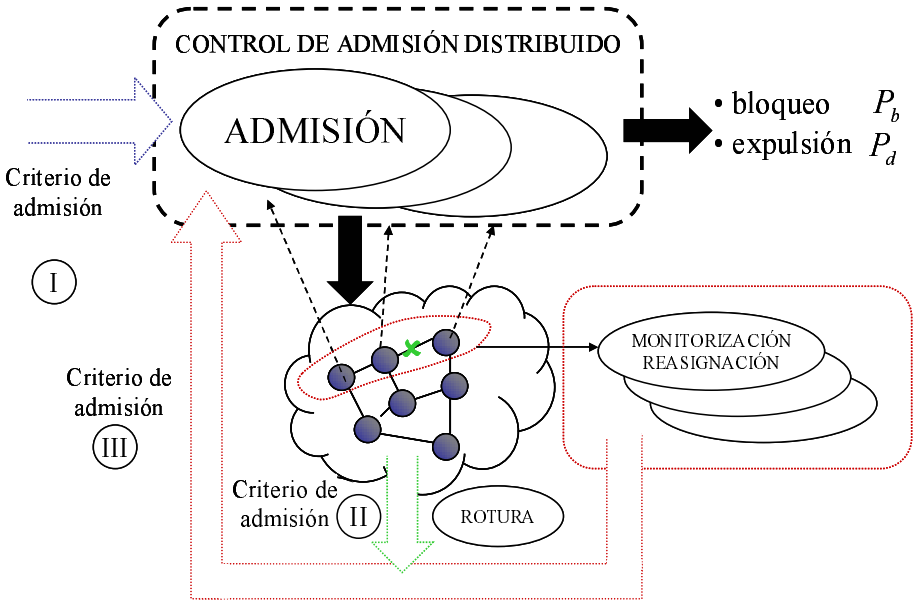


Figura 4.7. Procedimientos de admisión. Conexiones nuevas en el sistema (I) y readmisión tras rotura (II) o pérdida de la QoS (III).

[5]Esta suposición no es completamente cierta, como se ha descrito en el apartado anterior, cuando se considera la problemática de un escenario de interferencia realista.

[6]Las conexiones asociadas a tráfico *best-effort* son siempre admitidas, considerando que su ocupación de recursos se verá limitada a la utilización de los recursos no asignados a conexiones con QoS.

4.2.1. Reacción ante la pérdida de QoS y readmisión

1. Pérdida de conectividad, rotura de rutas:

El protocolo de encaminamiento seleccionado, AODV, incorpora un mecanismo de mantenimiento de rutas basado en la detección de roturas y consecuente alerta (mensaje de error $RERR$). Cuando un enlace ya no es viable, todas las conexiones que lo utilizan son informadas, los nodos intermedios en cada una de las rutas actualizan sus tablas de encaminamiento y las fuentes asociadas reactivan nuevas búsquedas. El incremento de los números de secuencia permite mantener una información fidedigna de la fiabilidad de las rutas evitando retomar enlaces inviables garantizando además que no se forman bucles.

Cuando una conexión con QoS sufre una rotura, se mantiene inactiva durante todo el periodo de búsqueda, por lo que el ancho de banda experimentado es nulo y la pérdida de paquetes asociada (posterior descarte) puede llegar a ser muy elevada. Para reducir la latencia en la búsqueda, el procedimiento de readmisión de un servicio prioritario "parado por rotura" no contempla la espera en el nodo destino ($t_{wait} = 0$). Por otra parte, teniendo en cuenta que se pretende encontrar rápidamente una ruta viable, lo cual, dadas las demandas de ancho de banda, puede resultar complicado, se han relajado las restricciones en la búsqueda de la ruta no siguiendo la política de descarte de $RREQ$s cuando el ancho de banda es insuficiente. Como resultado de estas modificaciones, la búsqueda es análoga a la de una conexión *best-effort* ($P_b \approx 0$), con la ventaja de incluir el procedimiento *BWC-FA* capaz de reservar el ancho de banda necesario en caso de que sea posible.

Si bien el procedimiento descrito no garantiza una admisión acorde con la demanda real de la aplicación, mantenerla en el sistema facilita la posterior actualización y reasignación de recursos. Gracias a la relajación en las restricciones de la búsqueda resulta mucho más probable encontrar un camino posible. De este modo, cuando la conexión compite con nuevas aplicaciones que demandan QoS por primera vez, tiene mayores probabilidades de ganar la competencia por los recursos. En definitiva, se trata de un modo preferente a las conexiones admitidas que están sufriendo roturas. Se pretende así minimizar la degradación sufrida ante una rotura permitiendo que el procedimiento de monitorización de la QoS seguido con posterioridad logre readmitir adecuadamente a la conexión (como se explicará a continuación).

Cuando una ruta se rompe a consecuencia de la movilidad, la reacción de los nodos responde a la recepción de mensajes $RERR$ dirigidos hacia la fuente, lo que supone

la liberación inmediata de los recursos asociados (BW^n_{CAC} – (3.20)) al tiempo que se invalida la ruta. Sin embargo, esto sucede únicamente en los nodos intermedios previos al enlace roto. Todos aquellos enlaces activos hacia el destino, marcan la ruta como inválida a partir del momento en que expira el tiempo de vida asociado a la misma (ART[7]), al cesar el envío de paquetes. La definición de dicho tiempo de vida, responde a un mantenimiento de la ruta capaz de absorber ceses temporales del tráfico que pudieran deberse a congestión en la red o tráfico racheado, tal y como se explicó en §3.3.1. Sin embargo, un dimensionado excesivo del ART puede suponer un retardo elevado en la reacción de los nodos posteriores, donde el BW^n_{CAC} permanece reservado durante todo ese tiempo. La reserva se mantiene por el hecho de estar asociada a diversas entradas en la tabla de rutas (base de datos compartida) que identifican dichos slots con conexiones con QoS (flujos QoS). No obstante, dado que las conexiones se han perdido, se produce un mantenimiento de reservas innecesario, lo que supone una infrautilización de la disponibilidad real de la red para aceptar nuevas conexiones, e incluso poder readmitir adecuadamente la conexiones fallidas. Como mecanismo de liberación de recursos podría plantearse el envío de señalización específica. Sin embargo, dado que las rutas se han roto, hacer llegar la información a los nodos implicados generaría nuevas búsquedas (o envíos *broadcast*) por lo que esta solución se considera inviable. Como alternativa, se propone una reserva temporal de los recursos (*soft-rsv*). De modo análogo al ART, se define un tiempo de vida asociado a una ruta con QoS, denominado RESERVATION_TIMEOUT (RT), menor que el ART. El mantenimiento de un ancho de banda garantizado para una conexión QoS permite definir un RT adecuado para el flujo de paquetes esperado. Cuando dicho flujo cesa, los recursos se liberan al expirar el RT. La entrada correspondiente en la tabla de rutas, sin embargo, permanece activa durante el tiempo restante hasta completar el ART. Dicha solución conservadora se ha adoptado para no eliminar precipitadamente entradas de la tabla de rutas aunque se liberen los recursos asociados. Este mantenimiento permite recuperar la actividad de la ruta cuando la pérdida de los recursos que hizo expirar el RT se deba a descensos temporales de ancho de banda provocados por un incremento en la interferencia del entorno, tal y como se explicará a continuación. El procedimiento de actualización de recursos que se describirá permite recuperar al ancho de banda sin una eliminación precipitada de la entrada asociada en la tabla de rutas. Por otra parte, cuando realmente se produzca una rotura, la pronta liberación de los recursos tras el RT garantiza una gestión eficiente de los mismos.

[7]ACTIVE_ROUTE_TIMEOUT (ART), ver §3.3.1

2. Incremento de la interferencia, pérdida de la QoS:

Monitorización de la QoS y reasignación de recursos:

Con objeto de flexibilizar la asignación de recursos de manera que ésta se adapte a la variabilidad del entorno, se ha planteado un mecanismo de monitorización y actualización de reservas en conexiones con QoS. El procedimiento propuesto se basa en la inclusión del procedimiento *BWC-FA* en el flujo de paquetes asociado a cada conexión (*piggy-backing*), tal y como se muestra en la figura 4.8. Con una periodicidad configurable (t_{update}), se transmite la misma información que durante el procedimiento de búsqueda de ruta (PB y SRT) añadida como control en dichos paquetes. Los paquetes de datos originados por la fuente ($DATA$) funcionan de manera análoga a un $RREQ$. Cuando éstos se reciben en el nodo destino, el propio algoritmo de encaminamiento genera reconocimientos específicos ($ROUTE_ACK$ (RA)) que actúan como un mensaje $RREP$. La información enviada a través de los $ROUTE_ACKs$ puede incluirse en los reconocimientos propios de la aplicación en caso de que éstos existan. No obstante, su generación desde el algoritmo de encaminamiento permite aplicar el mismo procedimiento en aplicaciones unidireccionales (como por ejemplo el *videostreaming*).

Recordando la notación seguida en §3, para determinada conexión admitida, la sucesión $P = \{n_m \rightarrow n_{m-1} \rightarrow ... \rightarrow n_0\}$ constituye el camino seleccionado entre el nodo fuente, n_m, y el destino, n_0. El cálculo *BWC-FA* realizado para dicha conexión durante el procedimiento de monitorización mide el ancho de banda disponible adicional al originalmente reservado (TS_k^P, para cada nodo intermedio k). Sin embargo, durante la fase $ROUTE_ACK$ se toman como nuevos TS_k^P los slots originariamente seleccionados que permanecen reservados ($TS_k^P \cap CAC_{RT}^{k-1}$, en el enlace($k \rightarrow k-1$)). Sólo en caso de que se hayan perdido reservas, se añaden los recursos necesarios del conjunto correspondiente de disponibles calculado en el proceso (PB_k^{k-1} indicado en el $ROUTE_ACK$ enviado desde el nodo $k-1$). El $ROUTE_ACK$ actúa como un $RREP$ invocando en el nivel MAC la reserva de los nuevos recursos (procedimiento BIAS). En definitiva, considerando R el número de slots demandado:

$$TS_k^P\big|_{nuevo} = TS_k^P \cup BW_1\left(PB_k^{k-1}, R - \left|TS_k^P\right|\right) \tag{4.6}$$

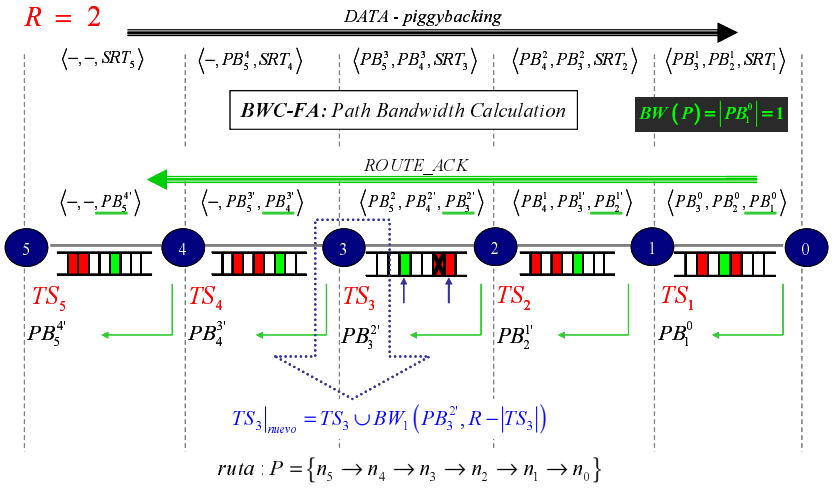


Figura 4.8. Actualización de los recursos asignados a un flujo QoS. Señalización sobre $DATA-ROUTE_ACK$

Cuando en un enlace determinado ($k \rightarrow k-1$) se pierden las reservas asociadas al TS_k^P se intenta continuamente recuperarlas[8] mientras el procedimiento de monitorización no proporcione un conjunto diferente de slots ($TS_k^P\big|_{nuevo}$). Sin embargo, para evitar un incremento en el descarte de paquetes por falta de recursos, se tratan de reservar temporalmente nuevos slots.

Teniendo en cuenta las múltiples conexiones admitidas sobre un enlace ($k \rightarrow k-1$), lo cual determina el BW_{CAC}^{k-1}, se define como ***tráfico QoS en exceso***[9] la diferencia entre los recursos necesarios para acomodar la transmisión del tráfico prioritario y los recursos efectivamente reservados dentro del BW_{CAC}^{k-1}:

$$excess_{CAC}^{k-1} = \sum_{m \in NC_{QoS}} n_{pq,m}^{k-1} - \left|CAC_{RT}^{k-1}\right| \tag{4.7}$$

donde NC_{QoS} es el número total de conexione admitidas, $n_{pq,m}^{k-1}$ se corresponde con los paquetes en cola para la conexión $m \in NC_{QoS}$ y CAC_{RT}^{k-1} son los recursos pertenecientes al BW_{CAC}^{k-1} efectivamente reservados, siendo este último la unión de todos los TS_k^P asociados a cada una de las conexiones (enlace $n_k \rightarrow n_{k-1} \in P(m)\ \forall m \in NC_{QoS}$).

[8]La pérdida puede haberse debido, por ejemplo, a una colisión momentánea con una conexión *best-effort* que, tras sufrir la colisión dejará de interferir sobre los slots reservados de mayor prioridad para la conexión con QoS e intentará reservar otros recursos.

[9]Algoritmo 3.1 en la página 74.

Es importante destacar que en un escenario completamente variante, la existencia de tráfico QoS en exceso no sólo se debe a la pérdida momentánea de recursos por la aparición de interferencia. Cada rotura en la ruta supone la acumulación de paquetes en espera del redescubrimiento de un camino posible. Dicha acumulación implica posteriormente ráfagas de paquetes con una demanda real de ancho de banda mayor del estimado como suficiente para la conexión admitida ($\left|TS_k^P\right| = R$), lo que puede generar momentáneamente situaciones de exceso aún garantizando el TS_k^P reservado.

En cualquier caso, la selección adecuada de la prioridad asociada al exceso debe responder a una asignación preferente frente al tráfico *best-effort*, dado que se trata de recursos necesarios para conexiones con QoS. Sin embargo, teniendo en cuenta el carácter temporal de las reservas, que no se realizan en respuesta a una selección específica por parte del control de admisión o del procedimiento de actualización, la prioridad no puede igualarse a la de TS_k^P. El uso de una prioridad menor permite, por otra parte, que el procedimiento de monitorización incluya estos recursos como reasignables. Gracias a dicha reasignación, cuando no es posible "renovar" el TS_k^P perdido con nuevos recursos, estas reservas pueden incorporarse directamente al $TS_k^P\big|_{nuevo}$, dado que son reservas firmes de las que ya dispone el propio nodo.

La diferenciación de servicios finalmente propuesta se detallará a continuación (4.2.3) una vez descritos los procedimientos en los que ésta será aplicada.

Procedimiento de readmisión:

La medida inicial de ancho de banda y la consecuente selección de slots sirve de criterio inicial de admisión. Dar el servicio apropiado a las conexiones admitidas, sin embargo, depende de la capacidad de la red de gestionar los recursos en un escenario variante. La actualización antes descrita permite redefinir dinámicamente los recursos asignados en un camino ya establecido, pero dicha redefinición puede verse limitada por la nueva topología. La pérdida de los recursos debida a las colisiones entre diversas conexiones con QoS cuyos caminos comienzan a interferirse puede suponer que, en la competencia simultánea por los nuevos recursos, no sea posible encontrar ancho de banda suficiente para todas ellas, degradando considerablemente las prestaciones de alguna o incluso todas ellas. Resulta imprescindible detectar apropiadamente la situación para tratar de redistribuir adecuadamente las conexiones en el nuevo escenario, incluso llegando a tomar la decisión de descartar alguna de ellas en beneficio de las otras. Dicha detección se consigue gracias a dos mecanismos:

- Las sucesivas actualizaciones $DATA - ROUTE_ACK$ y la medida correspondiente de ancho de banda (métrica QoS) permiten detectar una pérdida de QoS irrecuperable en el camino original. La estimación de "irrecuperable" se determina parametrizando tanto la periodicidad de las actualizaciones (t_{update}) como el número de reasignaciones fallidas ($n_{failupd}$). Cualquier nodo intermedio dentro del camino que detecta la imposibilidad de recuperar el ancho de banda genera un paquete *QoS LOST* ($QLOST$) para informar a la fuente de la comunicación.

- De igual modo, cuando el ancho de banda decae lo suficiente para ralentizar el flujo de paquetes (incrementando por lo tanto el descarte), la falta de las actualizaciones correspondientes (durante un tiempo también configurable t_{noupd}) alerta directamente al nodo fuente de la pérdida de QoS.

En cualquiera de los casos, el resultado es la activación de una nueva búsqueda de ruta que trata de readmitir adecuadamente la conexión. El camino original, por otra parte, aunque no dispone de recursos suficientes, sigue siendo un camino viable para transmitir tráfico *best-effort*. Teniendo en cuenta esta posibilidad, aun buscando un camino alternativo, el primero se mantiene activado mientras sea posible, evitando así una pérdida excesiva de paquetes (ruta QBE). Puesto que la búsqueda de recursos realizada mediante el algoritmo *BWC-FA* trata de identificar los recursos disponibles en la red, es imprescindible asegurar que la reserva parcial de recursos en el camino que todavía se mantiene no limite esta disponibilidad. Esto es posible identificando todo el tráfico asociado como $excess_{CAC}^{k-1}$, en cada uno de los enlaces ($k \rightarrow k-1$), considerando un TS_k^P nulo. La prioridad intermedia con la que los slots necesarios intentan ser reservados permite obtener unas prestaciones superiores al tráfico *best-effort* sin suponer una competencia directa con la nueva búsqueda en curso. Para asignar adecuadamente la prioridad al tráfico QoS activo en una ruta QBE es necesario informar a los nodos intermedios que puedan no haberse percatado de la pérdida de QoS (al no recibir el $QLOST$). Para ello, tanto si la pérdida de QoS se detecta en el nodo fuente o en un nodo intermedio, se envía desde éste un mensaje específico $QOS_{RELEASE}$ que cambia la prioridad de los slots que todavía permanezcan reservados dentro del TS_k^P. Asimismo, éste es anulado en la entrada a la base de datos compartida, de manera que la identificación del $excess_{CAC}^{k-1}$ resulte inmediata. La figura 4.9 muestra el intercambio de señalización necesario.

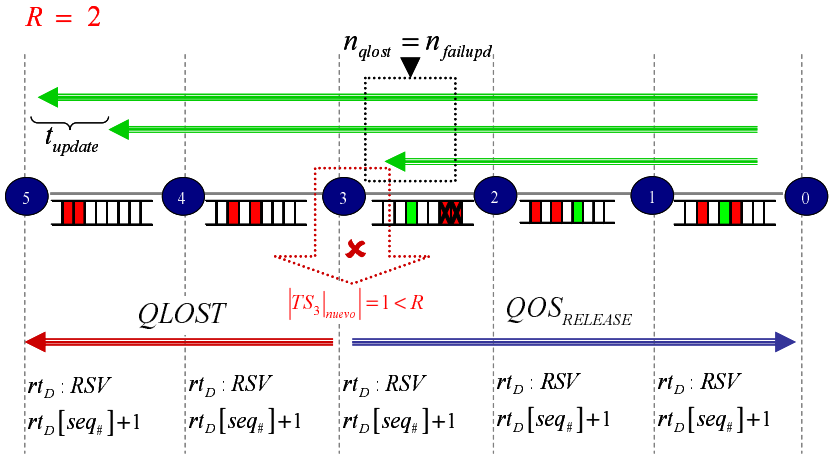


Figura 4.9. Pérdida del ancho de banda garantizado. Envío de $QLOST$ y $QOS_{RELEASE}$. Actualización de los números de secuencia y mantenimiento del camino reservado (ruta QBE).

A la hora de buscar la nueva ruta, el algoritmo de encaminamiento se comporta como si de una nueva admisión se tratara. Se mantienen tanto la política de descarte de $RREQ$s ($BW\left(FP^k\right) < R$) como la espera realizada por el nodo destino (t_{wait}) para seleccionar el "mejor" de los caminos. La priorización de los recursos, por el contrario, responde al criterio seguido para conexiones admitidas, al igual que tras una rotura de la ruta. El procedimiento conjunto se ha denominado criterio III de admisión.

Es importante destacar que, dado que es necesario realizar la búsqueda de una ruta nueva, aún manteniendo activo el camino inicial, los nodos de la red, incluidos los que toman parte en dicho camino, deben ser capaces de reconocer la búsqueda iniciada como nueva. Para ello, el número de secuencia conocido hacia el destino debe incrementarse en la entrada a la tabla de todos ellos, como si de una rotura se tratara, tal y como se muestra en la figura 4.9.

4.2.2. Expulsión de conexiones degradadas

La decisión de expulsar del sistema una conexión activada (*dropping*) realmente depende de la degradación sufrida por la aplicación correspondiente (nivel de usuario). Sin embargo, desde un punto de vista de la red, las prestaciones de la conexión se evalúan únicamente desde la perspectiva del ancho de banda garantizado. Resulta importante, por lo tanto, medir adecuadamente cuándo se considera la pérdida de ancho de banda equivalente a una degradación de la QoS irrecuperable. Dicha equivalencia se ha definido como un tiempo máximo sin el ancho de banda suficiente ($t_{qlost,max}$). De este modo, la

definición de ***soft-QoS*** se corresponde con un ancho de banda garantizado con una pérdida temporal asumible de $t_{qlost,max}$.

Tal y como se ha descrito, la pérdida de la QoS se debe tanto a roturas en las rutas como a incrementos en la interferencia en el entorno del camino seleccionado. El tiempo equivalente a la pérdida de ancho de banda para cada situación concreta se acumula sobre un tiempo t_{qlost} (4.8) que sirve de referencia a la hora de decidir la posible eliminación de la conexión:

$$t_{qlost} = \sum_{disc_{brk}} t^{i}_{discovery} + \sum_{disc^{failupd}_{qlost}} t_{update} \cdot n_{failupd} + \sum_{disc^{noupd}_{qlost}} t_{noupd} \tag{4.8}$$

- Cada rotura de la conexión implica una nueva búsqueda $disc_{brk}$ durante la cual el ancho de banda experimentado por la conexión es nulo. Así pues, la duración de cada descubrimiento i-ésimo de ruta asociado, $t^{i}_{discovery}$, se acumula sobre el t_{qlost}.

- Mientras la comunicación está activada, las sucesivas actualizaciones de la ruta permiten mantener el TS^{P}_{k} en cada enlace. Cuando se recibe un $QLOST$, se incrementa el t_{qlost} con el tiempo equivalente sin un BW_{CAC} garantizado, igual a $t_{update} \cdot n_{failupd}$, y en consecuencia se inicia una nueva búsqueda de ruta $disc^{failupd}_{qlost}$.

- Cuando la variabilidad de la topología conduce a una saturación en el entorno, incrementando excesivamente la interferencia, la falta de actualizaciones no permite asignar nuevos recursos ni estimar correctamente el ancho de banda que percibe el destino de la comunicación. Sin embargo, se ha considerado una medida equivalente de ancho de banda nulo, al considerar la situación como crítica. Por lo tanto, el t_{qlost} se incrementa en t_{noupd} y, nuevamente, se trata de readmitir la conexión en un camino diferente, $disc^{noupd}_{qlost}$.

En el caso de superar el límite establecido $t_{qlost} > t_{qlost,max}$, en el momento de intentar una nueva admisión de la conexión, ésta es eliminada del sistema. El diagrama mostrado en la figura 4.10 ejemplifica el procedimiento descrito.

Cuando la readmisión de una conexión es activada por rotura, una búsqueda poco restrictiva minimiza la probabilidad de bloqueo, como se comentó anteriormente. Sin embargo, la movilidad puede llegar al extremo de producir particiones en la topología (red inconexa), imposibilitando encontrar un camino. En ese caso, el tiempo que un nodo fuente espera la respuesta de su destino ($RREP$) se ha considerado excesivo

para cumplir con la condición de ***soft-QoS***, por lo que una vez agotado el PDT[10], la conexión es directamente expulsada ($\uparrow P_d$). Cuando este mismo efecto sucede durante el procedimiento de admisión de una conexión nueva, por el contrario, puede reintentarse la búsqueda más de una vez ($n_{disc,max}$), incrementando la latencia inicial, ya que ésta no se considera una degradación equivalente a la de una aplicación que ya se encuentra en el sistema.

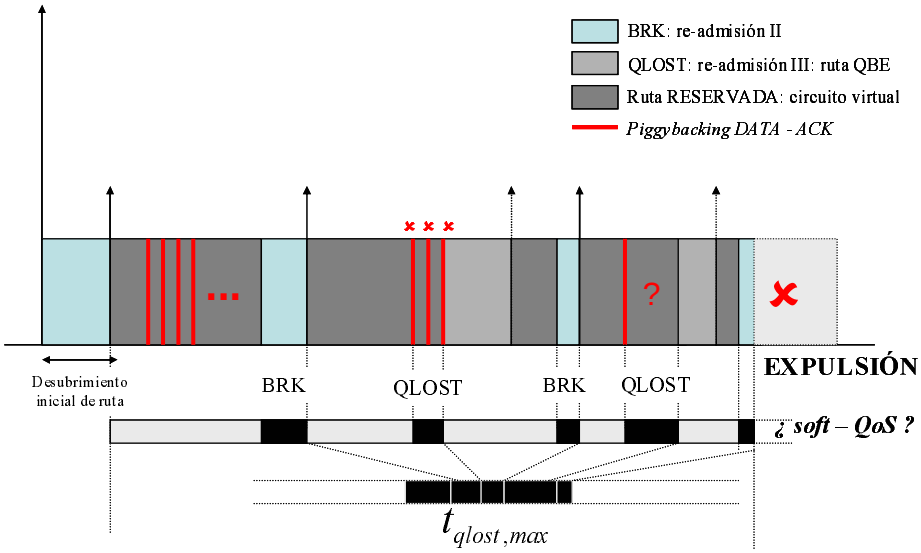


Figura 4.10. Diagrama ejemplo del procedimiento de decisión que conduce a la expulsión de una conexión activada. Medidas temporales de la pérdida de ancho de banda.

Expulsar selectivamente las conexiones que continuamente sufren pérdidas de QoS permite aprovechar mejor los recursos utilizados. Si una conexión que ya no puede cumplir el requerimiento de ***soft-QoS*** se mantiene dentro del sistema, dado que no es posible readmitirla en un camino apropiado, por un lado mantiene una reserva parcial de recursos en la ruta original (QBE) y por otro genera continuas búsquedas de ruta. Cuando la situación se repite para varias conexiones en paralelo, la competencia continua entre todas ellas para tratar de mantener reservados esos recursos (de prioridad intermedia), impide que ninguna de ellas resulte beneficiada, a su vez que bloquea la entrada de nuevas conexiones. Por otra parte, el tráfico *best-effort* utiliza los recursos sobrantes de la red y, puesto que demanda recursos de mínima prioridad, no puede competir con dichas aplicaciones. En consecuencia, ni los servicios prioritarios obtienen las prestaciones deseadas, ni se pueden aprovechar los recursos para nuevas aplicaciones o, en su defecto, para maximizar en lo posible el *throughput best-effort*.

[10]PATH_DISCOVERY_TIME (PDT), ver §3.3.1

Ancho de banda de guarda:

En un escenario muy dinámico, las constantes pérdidas de QoS provocan continuamente nuevas búsquedas. Si la admisión inicial trata de aprovechar al máximo la capacidad del sistema, ocupando en la medida de lo posible todos los recursos, la variación de las condiciones iniciales dificulta la readmisión, dado que resulta extremadamente complicado maximizar la ocupación de los recursos para todas las topologías posibles. Cuanto mayor es la variabilidad del entorno, mayor número de conexiones corren el riesgo de no resultar readmitidas incrementando considerablemente la probabilidad de expulsión. Para reducir dicha probabilidad, a costa de incrementar el bloqueo de conexiones, se propone restringir la admisión de conexiones nuevas dejando cierto ancho de banda desocupado. Esta infrautilización de los recursos pretende aumentar las posibilidades de una redistribución adecuada de las conexiones cuando las variaciones de la topología lo exijan. Permitiendo que las conexiones admitidas tengan garantizada la QoS esperada se mejora el grado de servicio (*Grade of Service* (GoS)) ofrecido por la red. El concepto GoS hace referencia a la probabilidad de no recibir las prestaciones deseadas. La representación típica en redes celulares, por ejemplo, define el GoS de acuerdo a (4.9), donde P_b es la probabilidad de bloqueo de nuevas conexiones y P_h la probabilidad de no admisión de una conexión en proceso de *handoff*.

$$GoS = \alpha \cdot P_b + \beta \cdot P_h \tag{4.9}$$

El perjuicio sufrido por una conexión expulsada del sistema supone una percepción de grado de servicio mucho mayor que no poder establecer la llamada, por lo que ambas probabilidades se ponderan dando una relevancia típica 10 veces mayor a P_h, de manera que $GoS = P_b + 10 \cdot P_h$ [Dimitriou y otros, 2000].

En el escenario de movilidad descrito para redes ad hoc, la readmisión en un camino diferente de una conexión ya establecida es equivalente al concepto de *handoff*. La probabilidad de expulsión P_d, por lo tanto, representaría lo equivalente a P_h. Al igual que en las redes celulares, no encontrar una nueva ruta en la que garantizar la QoS demandada supone un perjuicio mayor que el bloqueo de nuevas conexiones. Dada la naturaleza distribuida tanto de la actualización de recursos ($DATA - ROUTE_ACK$) como del procedimiento de admisión, no todas las aplicaciones que perciben una degradación de las prestaciones a nivel de usuario son correctamente identificadas a nivel de red y por lo tanto no son expulsadas. Así pues, existe cierto ratio de conexiones que no son gestionadas adecuadamente una vez admitidas. Si bien un diseño eficiente

del algoritmo distribuido trata de minimizar este porcentaje, es necesario ser consciente de su existencia a la hora de medir el GoS real ofrecido por la red. Dado que una implementación ideal del mecanismo de admisión, tal y como sería realizado desde un punto de vista centralizado, eliminaría del sistema dichas comunicaciones, éstas se han incluido dentro de las conexiones cuyo proceso de readmisión (*handoff*) ha fallado. De este modo, se redefine el concepto de GoS de acuerdo a:

$$GoS = \alpha \cdot P_b + \beta \cdot P_f \tag{4.10}$$

donde

- P_b, **Probabilidad de bloqueo**: representa el porcentaje de conexiones para las que no se encuentra un camino que cubra sus demandas de QoS en el momento en que realizan su petición.
- P_f, **Probabilidad de fallo**: denota el ratio de conexiones admitidas que sufren una degradación de las prestaciones por encima de su límite tolerable. Se incluyen en dicha definición tanto las conexiones no readmitidas como las gestionadas inadecuadamente ($P_f > P_d$).

La relación entre α y β depende de la importancia relativa de ambos efectos. Siguiendo el mismo criterio que en redes celulares se ha mantenido la definición de grado de servicio: $GoS = P_b + 10 \cdot P_f$[11].

Con objeto de limitar P_f a costa de aumentar P_b se ha definido un margen de recursos adicional al realmente demandado por las aplicaciones a la hora de admitir su entrada al sistema. Con este propósito, cuando se activa una nueva aplicación que demanda R slots, el algoritmo *BWC-FA* busca un ancho de banda disponible igual al demandado más un margen establecido ΔBW. Si el requisito puede satisfacerse, garantizando que $BW(P) = R + \Delta BW$, el destino envía el mensaje de confirmación $RREP$. No obstante, a la hora de seleccionar los recursos específicos para transmitir en cada uno de los enlaces, sólo se reserva exactamente el ancho de banda demandado, tal y como se describió en §3.3.2.3 – ecuación (3.19), de manera que $\left|TS_k^P\right| = R, \forall n_k \in P$. La capacidad sobrante se mantiene inicialmente desocupada, pero cuando las variaciones de la topología conllevan la readmisión de las aplicaciones, al existir recursos disponibles resulta mucho más probable que éstas no sean expulsadas, manteniendo la ***soft-QoS***

[11] En la representación realizada en la evaluación de resultados (§4.3 y §4.4) se ha utilizado la definición de grado de servicio normalizado: $GoS = (1/11) \cdot P_b + (10/11) \cdot P_f$.

deseada. Tanto en el procedimiento de actualización ($DATA - ROUTE_ACK$) como en la readmisión, la definición de ancho de banda disponible se limita a lo estrictamente demandado (R slots). El margen se utiliza únicamente como criterio conservador a la hora de decidir la entrada de una conexión nueva al sistema.

La definición del ΔBW es muy dependiente del escenario de movilidad. Un entorno poco dinámico posibilita que las conexiones se mantengan activas durante el desarrollo completo de la comunicación haciendo muy probable mantener la QoS deseada. Igualmente, al existir un número reducido de conexiones perjudicadas es posible encontrar recursos cuando se produce la readmisión. Por el contrario, en un escenario de elevada movilidad, se necesitan muchos más recursos desocupados para soportar la redistribución de las conexiones. La influencia de éste y otros parámetros se estudia con detalle en §4.3, donde se evalúa el comportamiento de la arquitectura propuesta en condiciones de movilidad.

4.2.3. Prioridades definidas en el nivel MAC. Criterios de admisión

Gracias al uso de diferentes prioridades en el $excess_{CAC}$ y el BW_{CAC} y a la definición de **criterios diferenciados de admisión**, mostrados en la figura 4.11, se trata de un modo preferente a las conexiones admitidas que han sufrido una pérdida de su QoS, tanto frente a tráfico *best-effort* como a nuevas conexiones QoS:

- **Prioridades**
  - Los recursos asignados a conexiones QoS (TS_k^P) se mantienen reservados con la máxima prioridad (p_0).
  - Las reservas temporales de recursos debidas a un BW_{CAC} insuficiente, lo que se ha denominado $excess_{CAC}$, se realizan con una prioridad intermedia ($p_1 < p_0$).
  - El tráfico *best-effort* utiliza recursos sobrantes asignados con la mínima prioridad ($p_2 < p_1 < p_0$).

- **Criterios de admisión**

  - Para una **conexión completamente nueva** (criterio I) únicamente se consideran REASIGNABLES[12] los recursos ya reservados para tráfico *best-effort*. Además, puede tomarse como ancho de banda requerido, el correspondiente al demandado más el margen de seguridad considerado ($R + \Delta BW$).

  - Una **conexión en proceso de readmisión** puede utilizar todos los recursos reservados de prioridades inferiores a la asignada a los posibles BW^n_{CAC}, $\forall n \in T$. Cuando la conexión ha sufrido rotura – recepción de $RERR$ – (criterio II), se relajan las restricciones en el encaminamiento. En el caso de responder a un $QLOST$, la búsqueda es análoga a una nueva admisión (criterio III). En cualquier caso, el ancho de banda buscado corresponde únicamente con los R slots demandados.

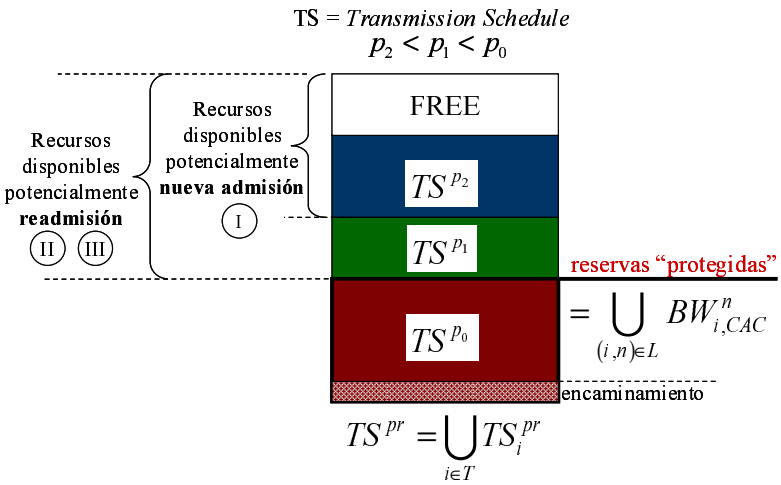


Figura 4.11. Reservas de recursos desde el punto de vista de la red $G = (T, L)$. Recursos potencialmente utilizables (en función de la situación real de los nodos: HTP y ETP).

A modo de resumen, la tabla 4.1 muestra la configuración apropiada de los parámetros característicos del algoritmo de encaminamiento correspondientes a los diferentes criterios de admisión establecidos.

[12]Anexo A.4, definición de disponibilidad.

Tabla 4.1. Procedimiento de Admisión - Readmisión

	Conexiones con QoS Encaminamiento QoS – *BWC-FA*			**Tráfico *best-effort***
Criterio de admisión[a]	I	II	III	siempre
Número de búsquedas	$n_{disc,max}$	1	1	si $n_{pq} > 0$[b]
descarte de *RREQ*s	✓	-	✓	-
Espera en destino	t_{wait}	0	t_{wait}	0
Margen de recursos	ΔBW	0	0	0
Prioridad disponible (recursos reasignables)	p_1 (p_2)	p_0 (p_2, p_1)	p_0 (p_2, p_1)	p_2 -

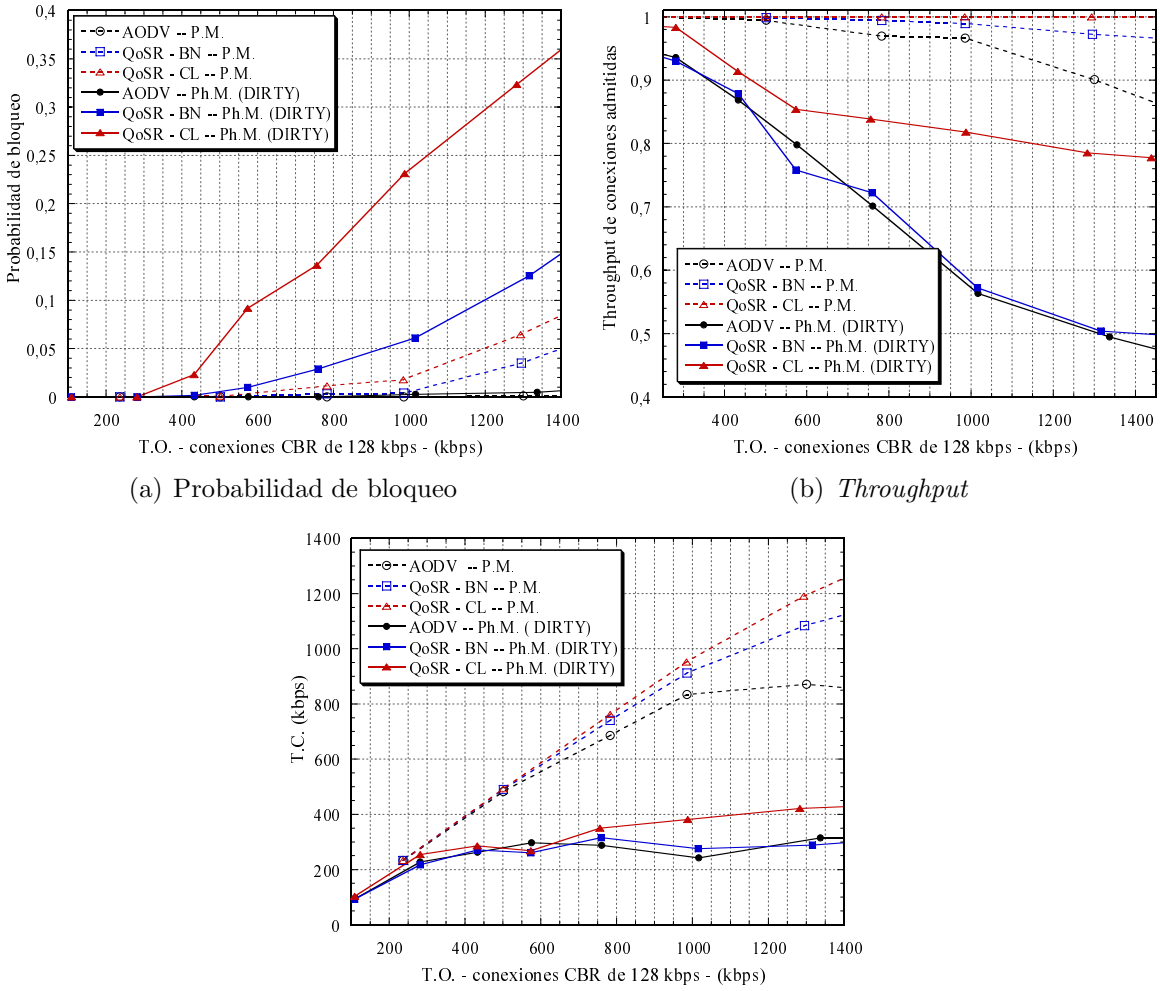
[a]I: conexiones nuevas, II: búsqueda tras rotura y III: readmisión tras $QLOST$.
[b]Número de paquetes en cola.

4.3. Evaluación del efecto de la interferencia sobre el mecanismo de admisión

Una vez estudiada la problemática existente en un escenario inalámbrico realista, se ha realizado en primer lugar un análisis comparativo considerando ambos modelos de escenario (*Protocol Model* y *Physical Model*). Los resultados mostrados se han obtenido en condiciones análogas a las utilizadas en la evaluación preliminar (§3.5). El análisis realizado en situación estática permite centrar el estudio en los efectos propios de la interferencia entre usuarios, independizando éstos de la influencia de la movilidad, que se evaluará posteriormente (§4.4). Se ha considerado el promedio de diversas topologías estáticas de 30 y 50 nodos distribuidos uniformemente en un área de 2 Km^2 y de 100 nodos en una superficie de 5 Km^2. Se ha considerado una sensibilidad del receptor que permite detectar la presencia de señal cuando la potencia media total recibida supera en 3,5 dB al nivel medio de ruido térmico. Esto implica que la potencia mínima de señal útil que puede ser detectada está 1 dB por encima de la potencia de ruido (CS_{th} = 1dB). La SIR mínima de decodificación (SIR_{th}) es de 5 dB, considerando una potencia de transmisión de 20 dBm, una potencia de ruido de -103 dBm y la utilización del modelo de propagación de Kammerman (3.27). La trama del ADHOC MAC se constituye de 50 slots BCH y 25 de datos. La medida del tráfico correctamente cursado, de acuerdo a la satisfacción de usuario, se define mediante una pérdida de paquetes tolerable P_{loss} del 3 %.

Aunque posteriormente se analizará cada una de las modificaciones ideadas como solución a las limitaciones del sistema en un escenario de interferencia real, la inclusión del estado DIRTY se ha considerado una propiedad intrínseca del protocolo ADHOC MAC en este nuevo escenario, cuya necesidad, explicada en §4.1, ha sido evaluada en [Gállego y otros, 2005b] y [Gállego, 2007]. Su utilización, sin embargo, produce una drástica reducción de la capacidad disponible. La definición de estado DIRTY ante la detección de cualquier nivel de potencia implica que incluso slots con un nivel de interferencia que sería tolerado por el receptor se consideran ocupados.

Evaluación del esquema básico en escenario de interferencia



(a) Probabilidad de bloqueo

(b) *Throughput*

(c) Tráfico cursado

Figura 4.12. Análisis comparativo de prestaciones. *Protocol Model* (P.M.) vs. *Physical Model* (Ph.M.). Estado DIRTY incluido en el protocolo ADHOC MAC. Estrategias AODV, QoSR-BN y QoSR-CL. Escenario de 50 nodos.

Los resultados obtenidos en condiciones de interferencia – figura 4.12 – muestran que todas las estrategias analizadas sufren una degradación considerable de las prestaciones en comparación con las obtenidas tras la evaluación realizada mediante el *Protocol Model*. La reducción de capacidad incrementa el bloqueo de conexiones durante los procedimientos de admisión, como se muestra en la figura 4.12(a). El resultado demuestra, además, la reducción de la ganancia teóricamente obtenida gracias al control de admisión (QoSR-BN y QoSR-CL). Aunque la inclusión del estado DIRTY permite realizar una medida de disponibilidad acorde con el escenario interferente, ésta depende de la posibilidad de decodificar correctamente la información transmitida en el BCH (donde efectivamente el receptor señaliza todos aquellos slots en los que recibe interferencia). Cuando los propios slots de control sufren colisiones debidas a interferencia no detectada, como la potencia no sensada en enlaces muy próximos al límite de la cobertura, la información de estado resultante no es totalmente precisa. En consecuencia, por una parte siguen produciéndose errores en la definición de la disponibilidad, y por otra, las reservas correctamente establecidas pueden llegar a perderse, resultando en la disminución de *throughput* observada en la figura 4.12(b).

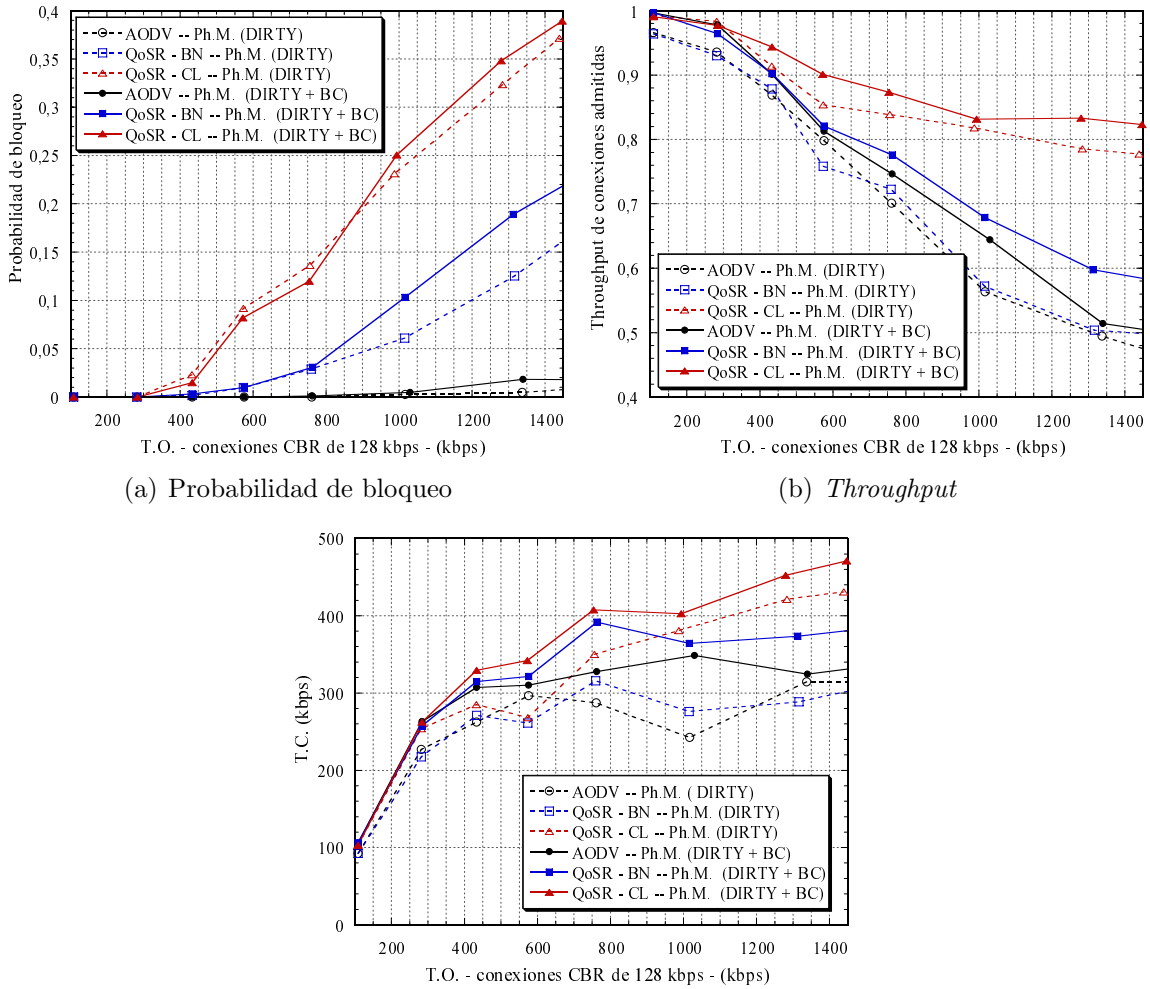
En el caso de la estrategia QoSR-BN, la sobreestimación de la disponibilidad basada en el enlace "cuello de botella" y el consecuente exceso de conexiones admitidas genera una situación de colisiones que degrada el *throughput* que éstas obtienen, como ya se explicó en §3.5. Este efecto es mucho más contraproducente en un escenario interferente debido al mayor número de colisiones que supone la interferencia real entre usuarios. Las prestaciones obtenidas son incluso equivalentes a las del AODV básico, como se muestra en la figura 4.12(b), a pesar del bloqueo realizado en la admisión que en el caso del AODV es inexistente.

Finalmente, el tráfico correctamente cursado sólo experimenta cierta ganancia en el caso de la estrategia QoSR-CL, como puede observarse en la figura 4.12(c). La definición de disponibilidad realizada se aproxima más a la capacidad real, lo que genera el elevado bloqueo representado en la figura 4.12(a). Sin embargo, aunque se limita el acceso de conexiones al sistema, la degradación de las conexiones admitidas no se corresponde con la QoS teóricamente garantizada.

La inclusión de señales $BEACON$[13] en los slots reservados, transmitidas por los terminales receptores, pretende proteger estas reservas de la posible colisión debida a transmisores que no han podido decodificar correctamente el BCH transmitido por dichos receptores. La potencia de señal necesaria para determinar la recepción de un

[13]Notación: BC en las gráficas de resultados.

$BEACON$, diferenciándola así de la interferencia que define el estado DIRTY, debe ser un parámetro de diseño. Éste se determina mediante un margen de potencia Δ (dB) sobre el umbral de sensado (CS_{th}). Puesto que todos aquellos slots definidos como OCUPADOS limitan la disponibilidad de recursos, se presenta el compromiso entre la reducción de capacidad resultante y la protección de reservas establecidas respecto a posibles transmisores interferentes.



(a) Probabilidad de bloqueo

(b) *Throughput*

(c) Tráfico cursado

Figura 4.13. *Physical Model*. Efecto de la inclusión de señales $BEACON$ (Protección frente a cualquier potencia detectada: margen (dB) sobre el umbral de sensado $\Delta = 0$ dB) Estrategias AODV, QoSR-BN y QoSR-CL. Escenario de 50 nodos.

Como análisis preliminar, se ha adoptado un criterio conservador, considerado que toda potencia detectada en un slot puede potencialmente corresponderse con una señal $BEACON$ transmitida por un receptor. Dado que perder reservas establecidas supone una degradación no tolerable en aquellos recursos utilizados por conexiones con ancho de banda garantizado y que realmente sólo estas conexiones deberían limitar la capacidad disponible para nuevas admisiones, únicamente los receptores con reservas de alta prioridad transmiten señales $BEACON$ (de igual modo si éstas se deben a transmisiones de control, como las del encaminamiento, dada la necesidad de garantizar una transmisión eficiente de dicha información).

La reducción de capacidad es ligeramente mayor que en el caso de utilizar únicamente el estado DIRTY debido a la identificación como OCUPADO de todos aquellos slots no identificados originariamente como DIRTY u OCUPADO por problemas en la decodificación del BCH. Como resultado, se produce cierto incremento en la probabilidad de bloqueo, tal y como se muestra en la figura 4.13(a). Por otra parte, el *throughput* de las conexiones admitidas mejora, como se muestra en la figura 4.13(b), consiguiendo el incremento observado del trafico cursado – figura 4.13(c).

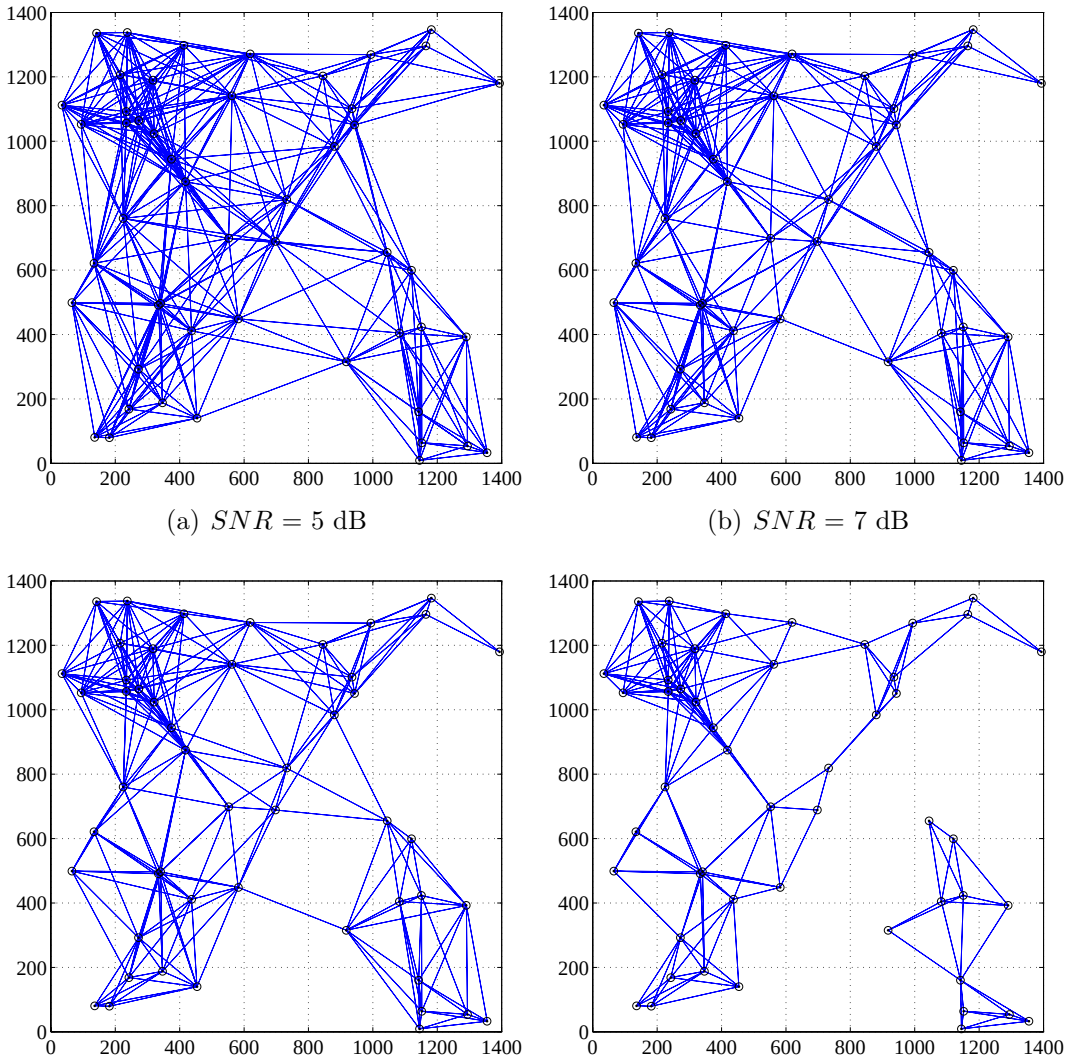
Configuración del nivel mínimo de potencia en recepción (estabilidad)

Como ha podido analizarse a través de los resultados obtenidos, a pesar de la mejora observada gracias a la utilización del estado DIRTY o las señales $BEACON$, las prestaciones distan mucho de las conseguidas mediante el *Protocol Model*. Las conexiones admitidas siguen sufriendo una degradación considerable dificultando que la QoS se corresponda con la deseada, incluso con un procedimiento de admisión adecuado (QoSR-CL) teóricamente capaz de gestionar eficientemente los recursos. Como se describió en §4.1.2, aun limitando el acceso a recursos susceptibles de sufrir interferencia y protegiendo con señalización las reservas más prioritarias, el principal problema reside en la inestabilidad de la red. Los enlaces denominados débiles, por su escaso margen para soportar interferencias (distancias próximas al rango máximo de cobertura) suponen constantes fallos en las reservas establecidas, provocando así continuas reasignaciones. Dicho problema es especialmente crítico en cuanto a las reservas del slot BCH necesario para garantizar la actividad de los usuarios en el sistema.

La solución propuesta consiste en limitar la distancia entre vecinos considerados estables (NB_i^{est}) exigiendo una potencia mínima de señal en recepción ($P_{rx,min}$), de manera que se disponga de un margen adicional de SIR para soportar interferencia[14], reduciendo así las oscilaciones en las reservas, especialmente negativas en el BCH. La pérdida de un vecino estable, como ya se comentó en §4.1.2, supone la desconexión del enlace con la consecuente alerta al protocolo de encaminamiento, que puede identificar así las roturas en los caminos establecidos. Teniendo en cuenta el impacto de éstas sobre las prestaciones (*overhead* de nuevas búsquedas, pérdidas de paquetes) resulta imprescindible realizar una detección fidedigna evitando las "falsas alarmas" (pérdida de un BCH por colisión momentánea). Con este propósito se definió el parámetro $n_{frnolisten}$ como el número máximo de tramas sin recibir con la potencia requerida el BCH de un vecino estable antes de identificar su pérdida. A lo largo de las simulaciones, se ha establecido un valor $n_{frnolisten} = 10$ tramas (0,5 seg.), suficiente para evitar los problemas derivados de las colisiones momentáneas, pero que a la vez permite una rápida reacción del protocolo de encaminamiento ante las roturas.

La definición de un valor de $P_{rx,min}$ superior a la mínima necesaria para alcanzar la SIR de decodificación reduce la cobertura máxima, de manera que la conectividad de la red disminuye (figura 4.14). No obstante, el margen frente a interferencias que supone establecer enlaces con vecinos más cercanos de lo que el límite de decodificación permite, reduce las colisiones en las reservas establecidas, como se observa en el ejemplo ilustrativo de la figura 4.15(a) (estrategia QoSR-CL). Por otra parte, se incrementa el número de saltos de las rutas establecidas (figura 4.15(b)), lo que supone un mayor tráfico ofrecido efectivo al aumentar el número de enlaces ocupados por conexión activa. El valor seleccionado de $P_{rx,min}$ debe resolver el compromiso entre capacidad y eficiencia en la reserva de recursos.

[14]Definición de $P_{rx,min}$ (4.3) como el valor de potencia de señal en recepción necesario para tener una SIR teórica equivalente (SNR) ΔSIR dB superior a SIR_{th}.



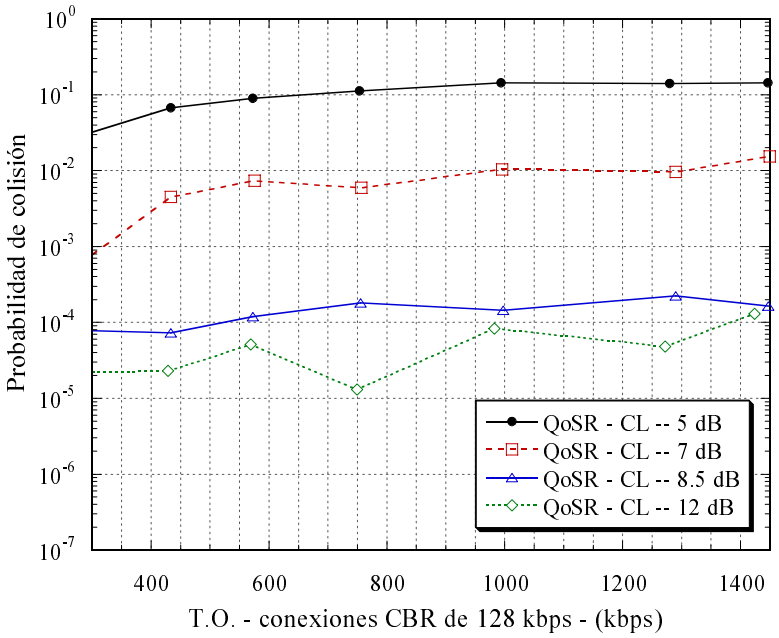
(a) $SNR = 5$ dB

(b) $SNR = 7$ dB

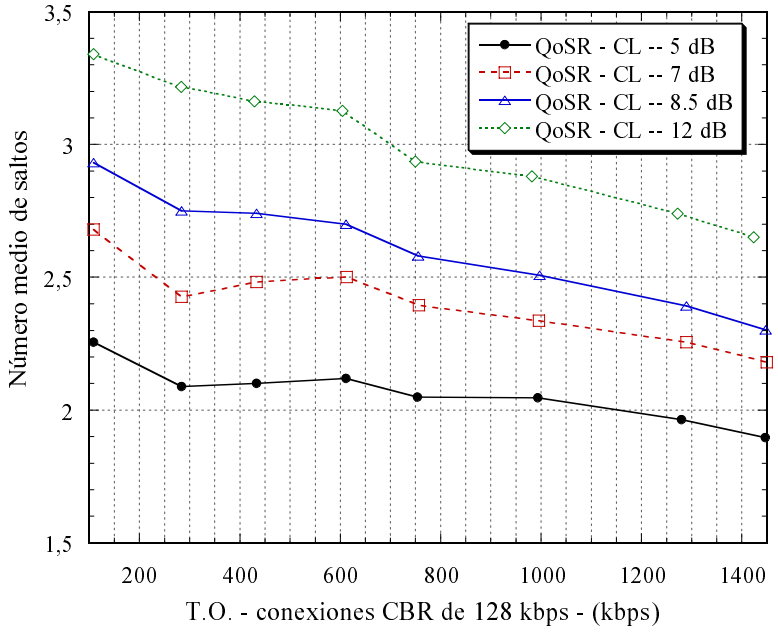
(c) $SNR = 8{,}5$ dB

(d) $SNR = 12$ dB

Figura 4.14. Escenario de muestra de las topologías simuladas. Influencia del valor $P_{rx,min}$ sobre la conectividad de la red – parámetro reflejado: SNR (SIR teórica).



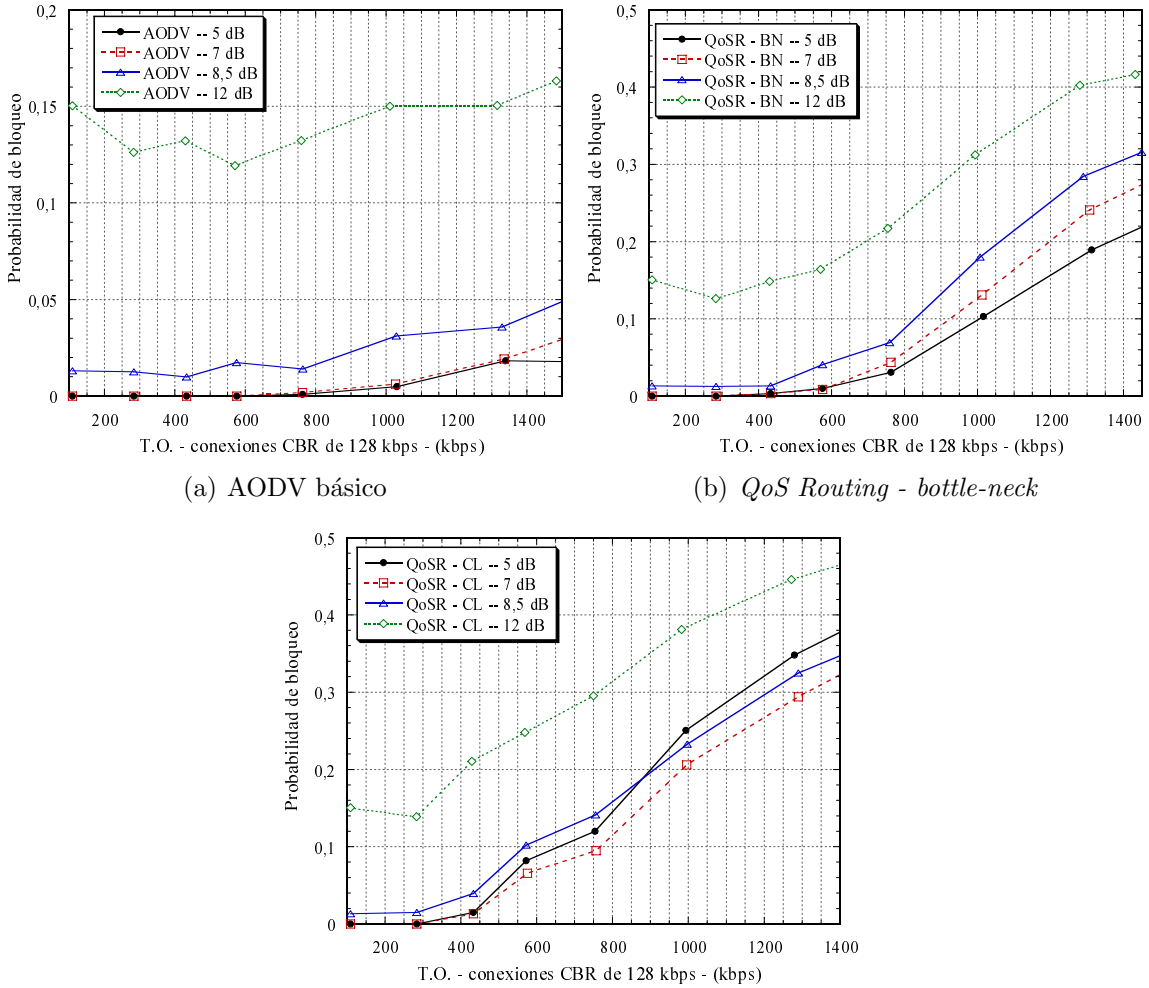
(a) Probabilidad de colisión



(b) Número medio de saltos por ruta

Figura 4.15. Comparativa del valor $P_{rx,min}$ – parámetro reflejado: SNR (SIR teórica). (a) Efecto sobre las colisiones sufridas y (b) número medio de saltos de las rutas establecidas. Estrategia QoSR-CL. Escenario de 50 nodos.

Un valor excesivo de $P_{rx,min}$ puede conducir a particiones de la red que generan una probabilidad de bloqueo mínima para todo tráfico ofrecido a consecuencia del porcentaje de conexiones que no encuentran un camino posible. Como se observa en la figura 4.16, una $P_{rx,min}$ creciente genera una probabilidad de bloqueo base cada vez mayor en la estrategia AODV, incluso a bajas tasas de tráfico ofrecido. Dicho bloqueo puede llegar a situarse entorno a un 15 % cuando se establece un margen de $SNR = 12$ dB, debido a una falta de conectividad excesiva, por lo que dicho valor no resulta apropiado.



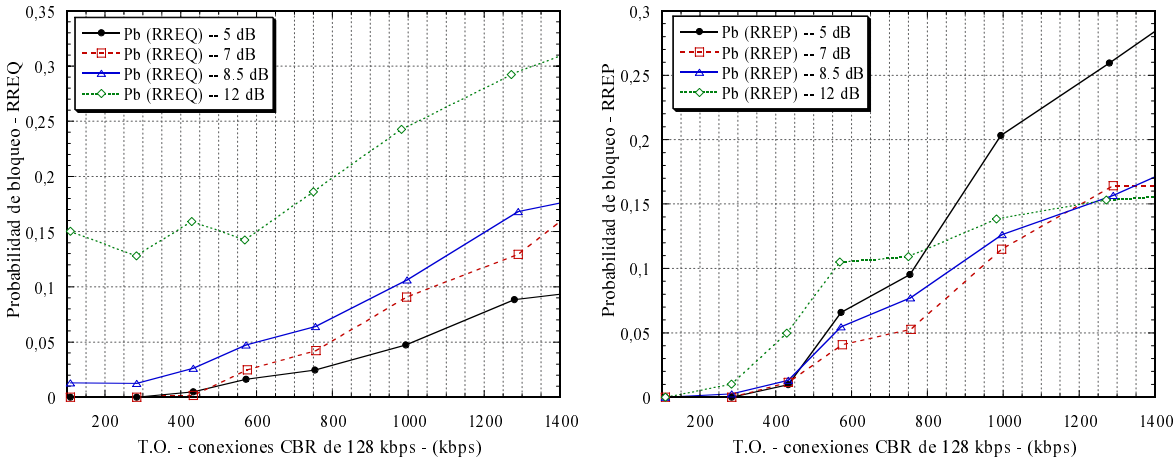
(a) AODV básico

(b) *QoS Routing - bottle-neck*

(c) *QoS Routing - cross-layer*

Figura 4.16. Probabilidad de bloqueo vs. tráfico ofrecido. Estrategias AODV, QoSR-BN y QoSR-CL. Comparativa del valor $P_{rx,min}$ – parámetro reflejado: SNR (SIR teórica). Escenario de 50 nodos.

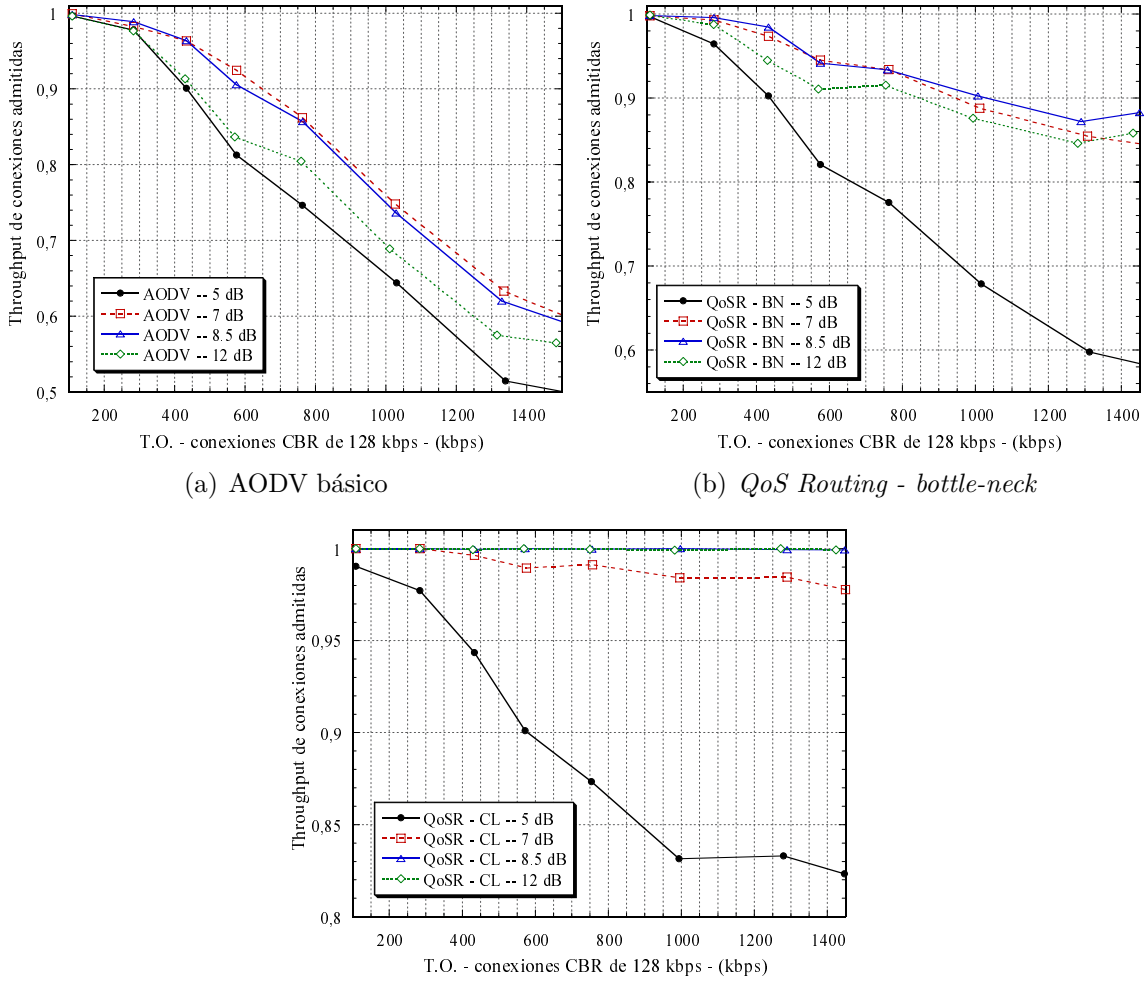
El bloqueo creciente con $P_{rx,min}$ observado en la estrategia AODV se refleja igualmente en las diferencias de bloqueo de la estrategia QoSR-BN. Sin embargo, en este caso, el motivo de bloqueo predominante es el realizado específicamente de acuerdo al criterio de admisión. La falta de recursos medida durante la búsqueda de ruta (fase $RREQ$) limita la activación de conexiones. En la estrategia QoSR-CL se produce el mismo efecto. No obstante, se observa que la probabilidad de bloqueo en ausencia de margen de protección ($SNR = 5$ dB) llega a superar a la obtenida para valores de SNR de 7 y 8,5 dB en condiciones de mayor carga, lo que se debe al aumento de bloqueo en fase $RREP$. Al aumentar el tráfico ofrecido aumenta la probabilidad de realizar una estimación incorrecta de los recursos disponibles como consecuencia de errores en el BCH. Por otra parte, la simultaneidad de conexiones (*race condition*) puede conducir a la pérdida de recursos considerados disponibles en fase $RREQ$. La falta de disponibilidad, en situaciones de alta carga, no puede resolverse con recursos adicionales, lo que se traduce en un número cada vez mayor de conexiones bloqueadas en fase $RREP$. Esta situación se ilustra en la figura 4.17. Este efecto no tiene lugar en la estrategia QoSR-BN, dado que no se realiza una asignación específica de los recursos durante esta fase del procedimiento de búsqueda. Al aumentar el margen de protección frente a interferencias aumenta la precisión de las medidas de ancho de banda y se reduce la probabilidad de colisión. Disminuye así la posibilidad de bloquear la conexión una vez que ya se ha encontrado un camino disponible, es decir, cuando se evalúa de nuevo la disponibilidad al recibir el $RREP$.



(a) Probabilidad de bloqueo ($RREQ$) (b) Probabilidad de bloqueo ($RREP$)

Figura 4.17. Probabilidad de bloqueo según la fase de búsqueda, $RREQ$ – (a) y $RREP$ – (b). Comparativa del valor $P_{rx,min}$ – parámetro reflejado: SNR (SIR teórica). Estrategia QoSR-CL. Escenario de 50 nodos.

De acuerdo a los resultados mostrados, valores de $P_{rx,min}$ que permiten obtener márgenes entre 2 y 3,5 dB (SNR de 7 y 8,5 dB) ofrecen unas prestaciones similares. En vista de la probabilidad de bloqueo obtenida, una SNR requerida de 8,5 dB parece limitar ligeramente la conectividad, dado que se obtiene cierto bloqueo residual incluso a bajas tasas de tráfico ofrecido. No obstante, su efecto sobre el bloqueo introducido por el propio procedimiento de admisión, especialmente cuando el sistema está cargado, no resulta significativo (máximo del 5 %). En cambio, la protección conseguida para las reservas establecidas y la consecuente estabilidad del sistema resultan superiores [Gállego, 2007].



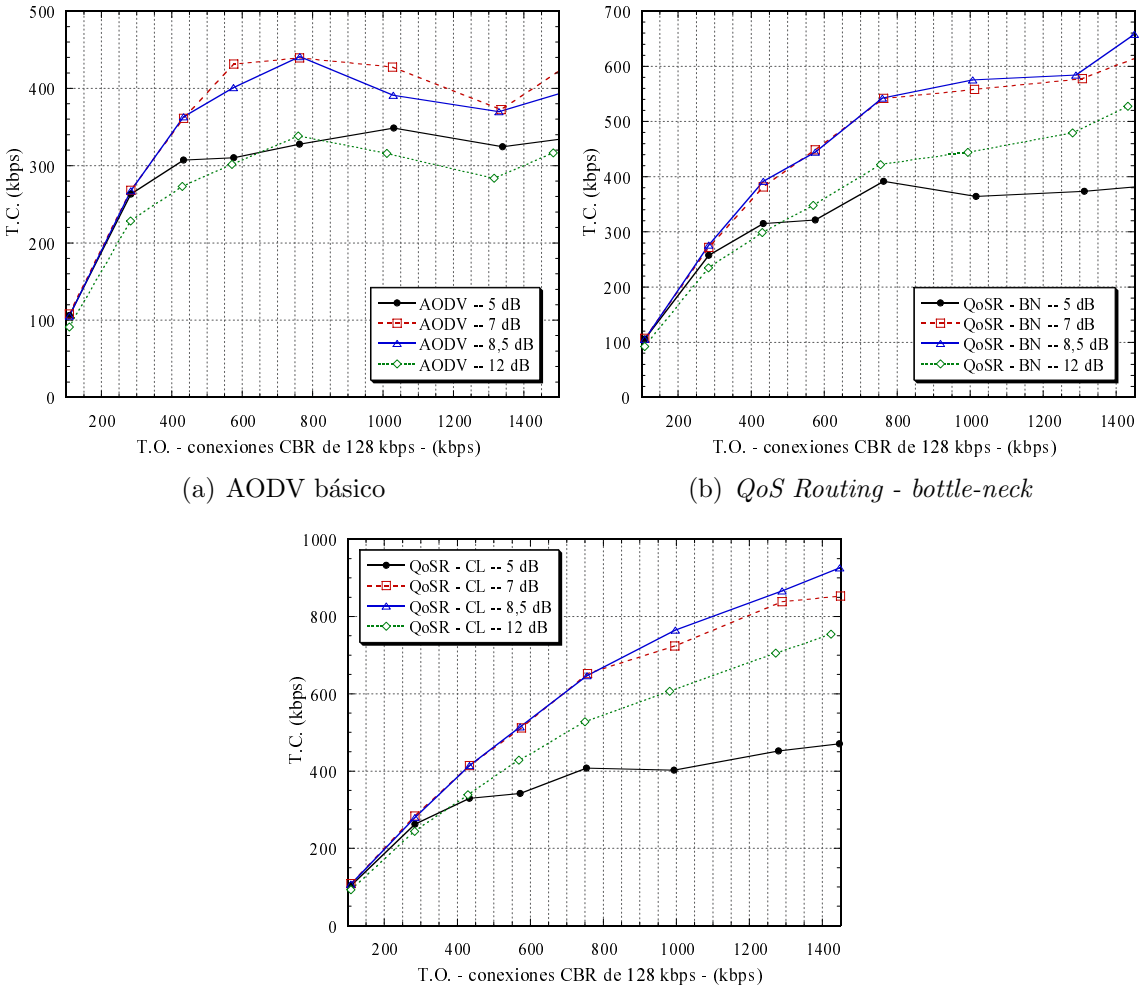
(a) AODV básico

(b) *QoS Routing - bottle-neck*

(c) *QoS Routing - cross-layer*

Figura 4.18. *Throughput* de las conexiones admitidas vs. tráfico ofrecido. Estrategias AODV, QoSR-BN y QoSR-CL. Comparativa del valor $P_{rx,min}$ – parámetro reflejado: SNR (SIR teórica). Escenario de 50 nodos.

Como puede observarse en la figura 4.18, gracias a la utilización de un margen de 3,5 dB ($SNR = 8{,}5$ dB) aumenta el *throughput* obtenido, especialmente en la estrategia diseñada, QoSR-CL, donde las conexiones admitidas mantienen el *throughput* prácticamente en el 100 % deseado. En consecuencia, aumenta el tráfico correctamente cursado, como refleja la figura 4.19.



(a) AODV básico

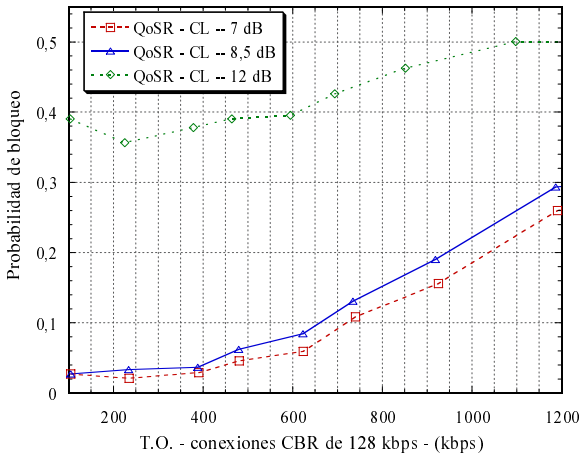
(b) *QoS Routing - bottle-neck*

(c) *QoS Routing - cross-layer*

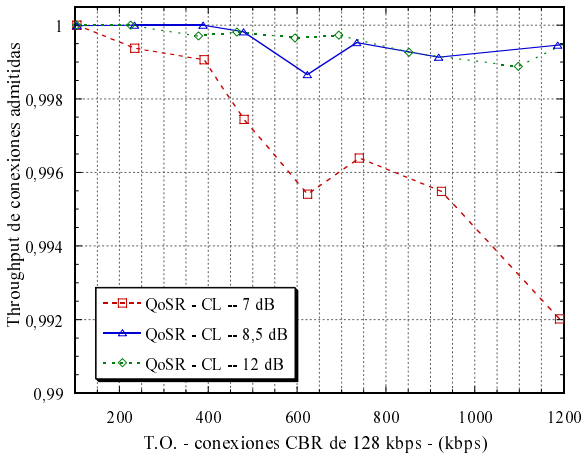
Figura 4.19. Tráfico cursado vs. tráfico ofrecido. Estrategias AODV, QoSR-BN y QoSR-CL. Comparativa del valor $P_{rx,min}$ – parámetro reflejado: SNR (SIR teórica). Escenario de 50 nodos.

Los resultados analizados se han obtenido en condiciones estáticas, puesto que esto permite realizar un estudio centrado en los efectos propios de la interferencia entre usuarios. De este modo, puede parametrizarse el margen mínimo requerido para garantizar la estabilidad del sistema en un entorno de interferencia.

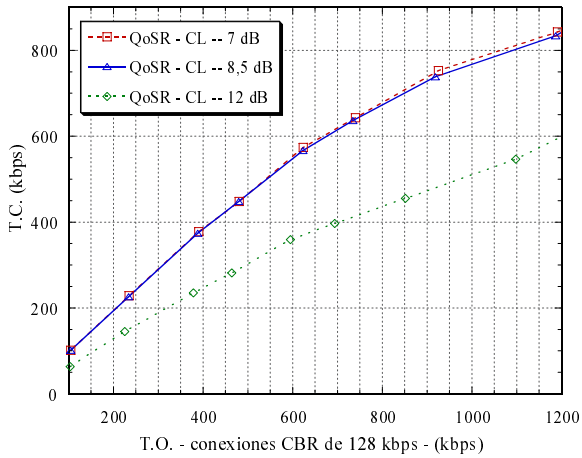
Una vez seleccionado este margen, las variaciones en los niveles de potencia debidas a los desvanecimientos del canal de transmisión deben analizarse con objeto de aumentar en consecuencia dicho margen garantizando así la protección frente a interferencias inicialmente deseada, a pesar de los posibles desvanecimientos de la señal. Dada la variabilidad del entorno, esta garantía únicamente será posible con cierta probabilidad, de modo que será ésta el parámetro de diseño a tener en cuenta [Gállego, 2007].



(a) Probabilidad de bloqueo

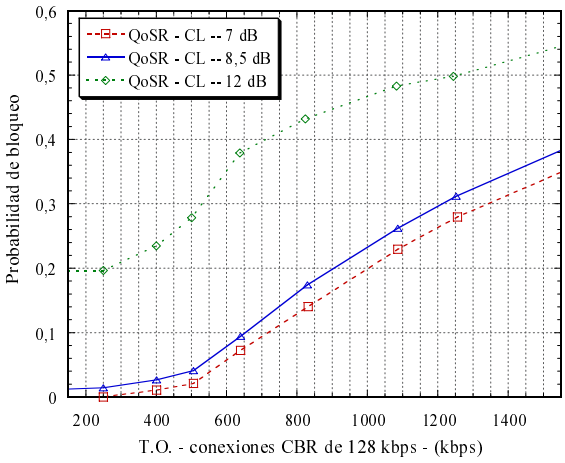


(b) *Throughput*

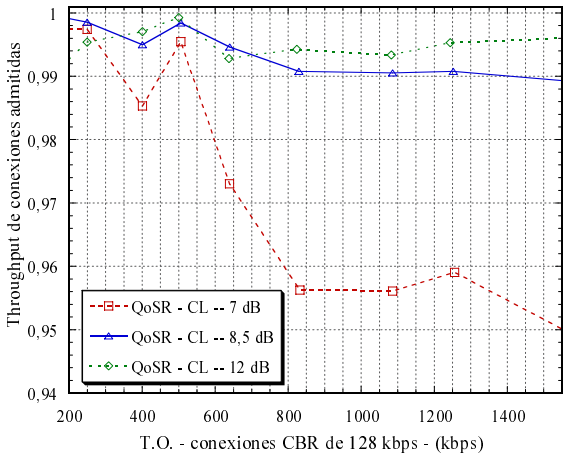


(c) Tráfico cursado

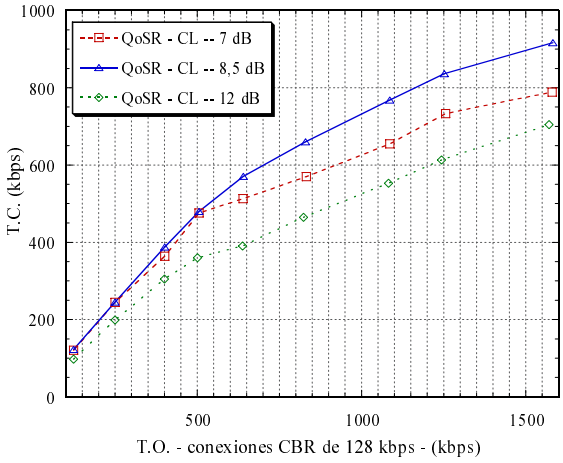
Figura 4.20. Influencia del valor $P_{rx,min}$ sobre las prestaciones de la propuesta QoSR-CL – parámetro reflejado: SNR (SIR teórica). Escenario de 30 nodos.



(a) Probabilidad de bloqueo



(b) *Throughput*



(c) Tráfico cursado

Figura 4.21. Influencia del valor $P_{rx,min}$ sobre las prestaciones de la propuesta QoSR-CL – parámetro reflejado: SNR (SIR teórica). Escenario de 100 nodos.

A la hora de seleccionar el valor adecuado de $P_{rx,min}$ debe tenerse en cuenta el objetivo buscado mediante su utilización. El margen frente a interferencias se establece, principalmente, con la finalidad de evitar que la contribución de la potencia no detectada a la interferencia sufrida en el momento de activar una conexión suponga una caída de la SIR por debajo del umbral de decodificación. Esta influencia depende por tanto del umbral mínimo de decodificación (SIR_{th}) y de la sensibilidad del receptor (umbral de sensado CS_{th}) [Gállego, 2007]. Así pues, el valor $P_{rx,min}$ es generalizable a otros entornos en los que estos parámetros se mantengan, como las topologías estudiadas de 30 y 100 nodos. El análisis realizado en escenarios de menor densidad (30 nodos) y mayor área de despliegue (100 nodos) corroboran dicha conclusión, como se observa

en las figuras 4.20 y 4.21, donde se muestran las prestaciones obtenidas, tanto en bloqueo, como en *throughput* y el consiguiente tráfico cursado, para la estrategia diseñada QoSR-CL. Puesto que el funcionamiento del sistema en ausencia de ningún margen de protección ($SNR = 5$ dB) supone un punto de operación inviable, se ha mostrado únicamente la influencia de los diferentes márgenes de protección propuestos (7, 8,5 y 12 dB). Al igual que lo evaluado con 50 nodos, un valor de $P_{rx,min}$ equivalente a una SNR de 8,5 dB permite mejorar la estabilidad de los enlaces y la efectividad de las reservas establecidas, a costa de incrementar el bloqueo, pero sin provocar desconexiones en la red ($SNR = 12$ dB).

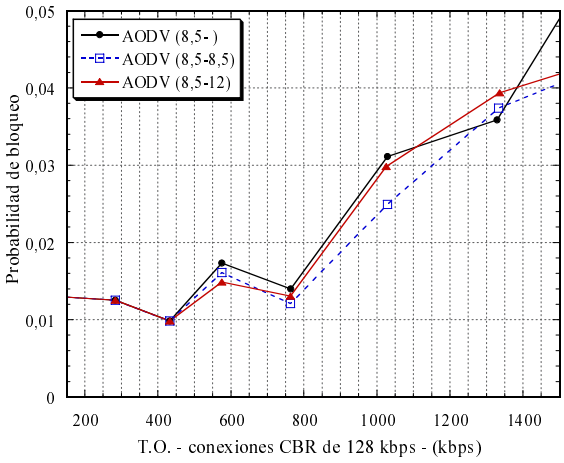
La ganancia observada resulta más significativa en el escenario de 100 nodos. Dado que las rutas establecidas tienen un número mayor de saltos, cada conexión extremo a extremo consta de más enlaces potencialmente vulnerables con un margen de protección insuficiente. El diseño de un margen adecuado, por lo tanto, tiene mayor impacto a la hora de evaluar las prestaciones sobre dichas conexiones.

Evaluación del efecto del control de potencia

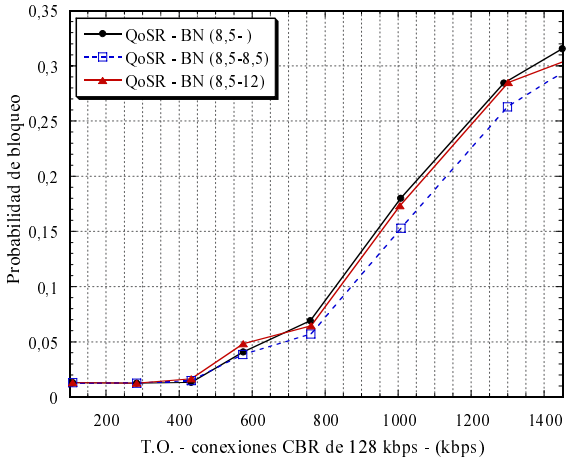
El diseño de mecanismos eficientes de control de potencia y su aplicación en el entorno de las redes móviles ad hoc no se encuentra en el ámbito del trabajo realizado en esta tesis[15]. No obstante, se ha implementado un procedimiento simplificado basado en el ajuste de la potencia de transmisión para obtener una determinada SIR objetivo (SIR_{tar}) según el valor idealmente estimado de las pérdidas de propagación, cuantificadas éstas de acuerdo al modelo de Kammerman descrito en (3.27).

Las estrategias diseñadas: uso del estado DIRTY, transmisión de señales $BEACON$ y diseño de un margen frente a interferencias, limitan la capacidad del sistema debido a la protección de las reservas establecidas que, desde una perspectiva de gestión eficiente de los recursos, puede parecer excesiva. Un mecanismo adecuado de control de potencia puede aumentar la reutilización de los recursos gracias a una disminución de la interferencia de las comunicaciones activas (reducción de slots DIRTY). Su utilización, por otra parte, garantiza un menor consumo energético, lo cual resulta imprescindible en el entorno de las redes móviles ad hoc. Sin embargo, la reducción de las potencias de transmisión conlleva la disminución del margen de protección frente a la interferencia del entorno, a pesar de que ésta sea menor. Resulta por lo tanto necesario mantener la protección mediante las señales $BEACON$, con la finalidad de evitar que el ajuste de potencia realizado desestabilice la garantía de reservas conseguida originalmente.

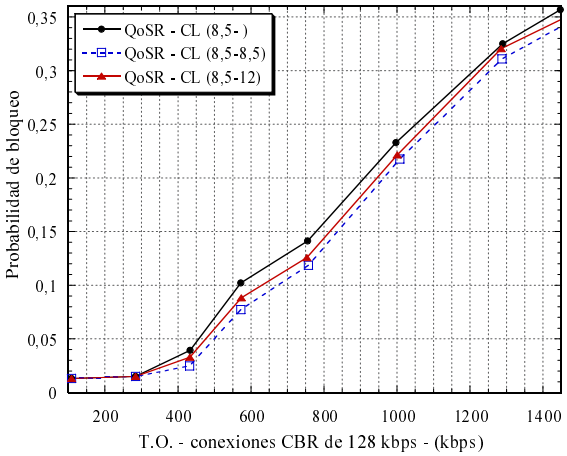
[15]Trabajo complementario realizado en [Gállego, 2007].



(a) AODV básico



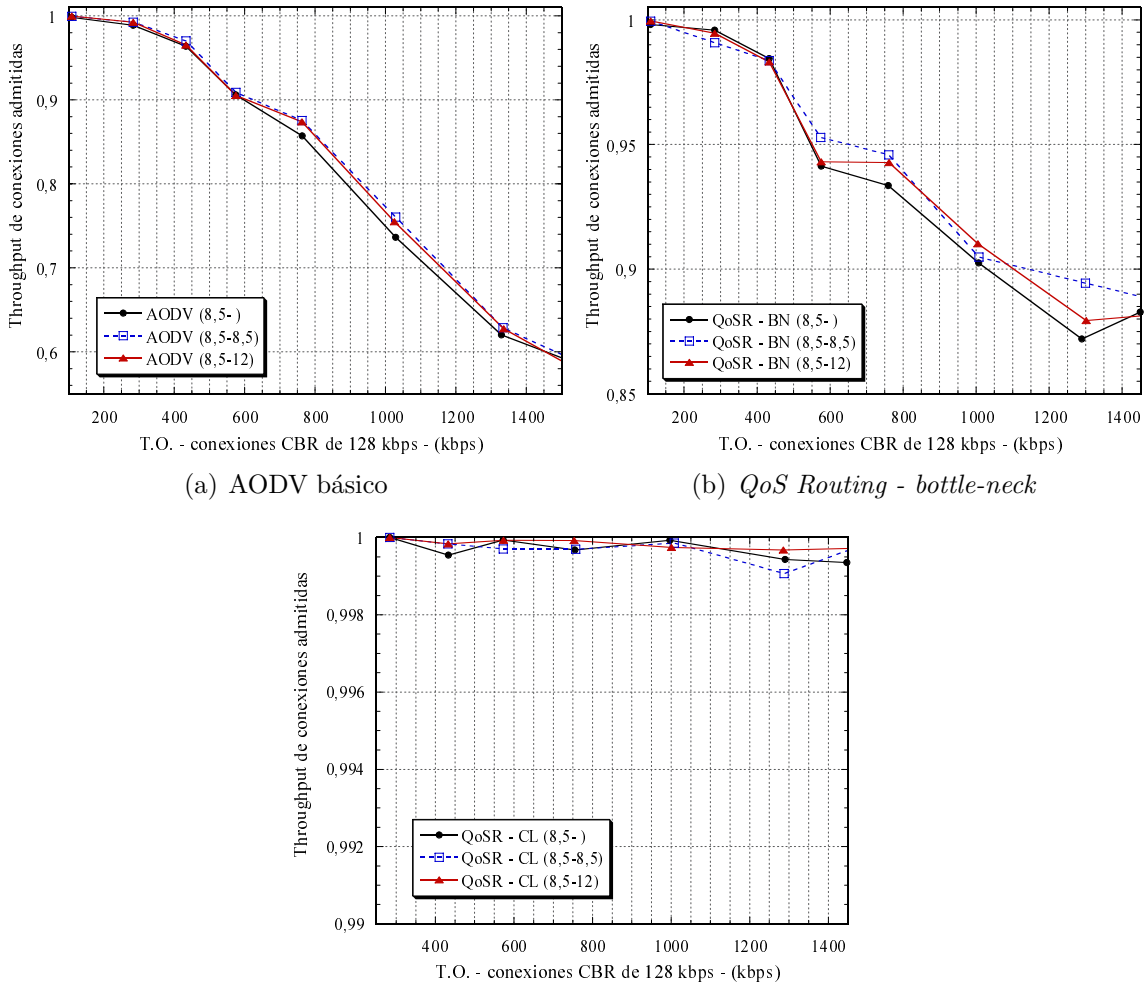
(b) *QoS Routing - bottle-neck*



(c) *QoS Routing - cross-layer*

Figura 4.22. Probabilidad de bloqueo vs. tráfico ofrecido. Estrategias AODV, QoSR-BN y QoSR-CL. Evaluación del control de potencia – notación: (SNR-SIR_{tar}), comparación sin control de potencia (8,5-). Escenario de 50 nodos.

La disminución de la interferencia generada, gracias a las transmisiones de menor potencia, permite aumentar la disponibilidad de recursos. Sin embargo, este aumento se ve limitado por la protección basada en $BEACONS$ que, por otra parte, garantiza la fiabilidad de los nuevos recursos disponibles. Además, la ganancia en disponibilidad está determinada por el incremento de la capacidad de reúso espacial de los recursos, que está condicionada por la geometría del problema. En el escenario de análisis, el área de cobertura de los terminales no es significativamente menor que el área total, lo que también limita la posibilidad real de reutilizar recursos y la consiguiente ganancia del control de potencia.



(a) AODV básico

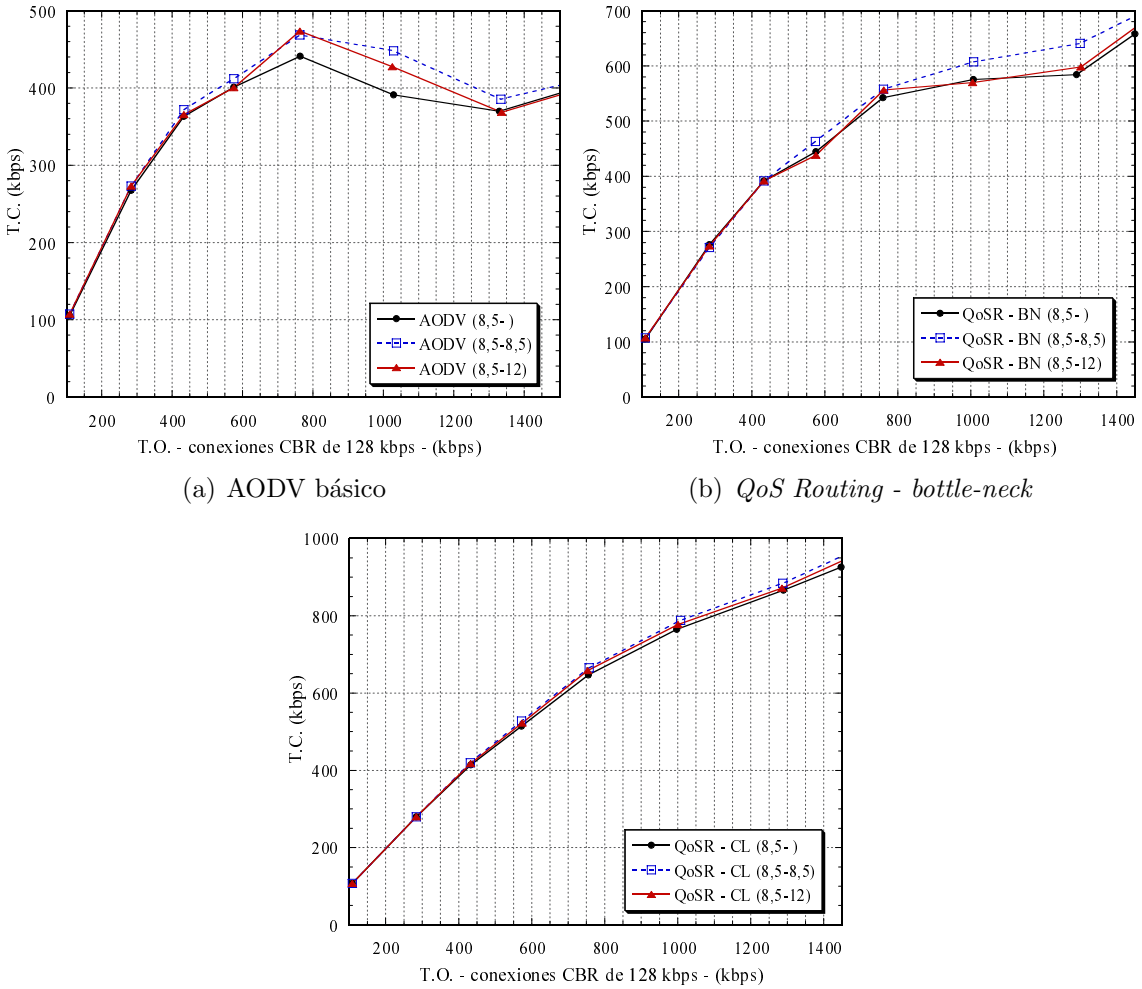
(b) *QoS Routing - bottle-neck*

(c) *QoS Routing - cross-layer*

Figura 4.23. *Throughput* de las conexiones admitidas vs. tráfico ofrecido. Estrategias AODV, QoSR-BN y QoSR-CL. Evaluación del control de potencia – notación: (SNR-SIR_{tar}), comparación sin control de potencia (8,5-). Escenario de 50 nodos.

El aumento en el ancho de banda disponible se manifiesta en una ligera disminución de la probabilidad de bloqueo en las estrategias basadas en admisión – figura 4.22, y un pequeño aumento del *throughput* en situaciones de mayor congestión, provocada por una admisión poco eficiente (AODV, QoSR-BN) – figura 4.23. Como consecuencia, se observa una cierta mejora en el tráfico correctamente cursado en el caso del AODV – figura 4.24(a) – y en el de QoSR-BN – figura 4.24(b). La estrategia QoSR-CL basa su admisión en una medida fiable del ancho de banda extremo a extremo. En este escenario, el aumento poco significativo de recursos disponibles en la red no equivale a un mayor ancho de banda sobre rutas específicas, especialmente teniendo en cuenta el pequeño

grado de reutilización existente, por lo que la ganancia derivada del control de potencia es muy poco apreciable en este caso – figura 4.24(c).



(a) AODV básico

(b) *QoS Routing - bottle-neck*

(c) *QoS Routing - cross-layer*

Figura 4.24. Tráfico cursado vs. tráfico ofrecido. Estrategias AODV, QoSR-BN y QoSR-CL. Evaluación del control de potencia – notación: (SNR-SIR_{tar}), comparación sin control de potencia (8,5-). Escenario de 50 nodos.

De acuerdo a los resultados obtenidos, la definición del valor de SIR objetivo más apropiada se ajusta al margen de estabilidad previamente seleccionado ($SIR_{tar} = 8{,}5$ dB). En definitiva, este valor garantiza una fiabilidad análoga a la conseguida para los slots BCH[16]. La potencia de transmisión se limita al mínimo requerido para garantizar dicha

[16]La selección de $P_{rx,min}$ se ha evaluado en función de las prestaciones globales del sistema. No obstante, el diseño se basó en el objetivo principal de garantizar la estabilidad del sistema, definida ésta de acuerdo a la garantía de las reservas BCH establecidas [Gállego y otros, 2006a]

estabilidad sin generar mayores niveles de interferencia. Así pues, gracias al control de potencia realizado se consigue uno de los objetivos clave en el diseño de redes ad hoc, como es reducir el consumo energético (sobre unos 4 dB de media en la transmisión de los datos), sin provocar una degradación de las prestaciones, manteniéndose así la superioridad de la estrategia *cross-layer* diseñada (QoSR-CL).

Evaluación del efecto del parámetro Δ (señales $BEACON$)

La restricción en cuanto a disponibilidad que supone la transmisión de señales $BEACON$ puede relajarse aumentando la potencia de señal necesaria para determinar la recepción de las mismas. Sin embargo, la menor protección sobre las reservas prioritarias puede degradar las conexiones admitidas. Los resultados analizados hasta el momento se han obtenido adoptando un criterio conservador, admitiendo cualquier nivel de potencia sensada como señal $BEACON$. Esta suposición se realizó con objeto de garantizar al máximo las reservas más prioritarias como protección frente a las diversas problemáticas descritas, como el hecho de disponer de menores márgenes debido a la aplicación del control de potencia. No obstante, con objeto de realizar una estimación más adecuada del valor de potencia necesario para resolver este compromiso, se han llevado a cabo diversas simulaciones variando el parámetro de diseño Δ, que define el margen de potencia establecido respecto al nivel de sensado utilizado para determinar la recepción de un $BEACON$. De acuerdo a los resultados ya obtenidos, se ha fijado la SIR objetivo del control de potencia en el valor de 8,5 dB.

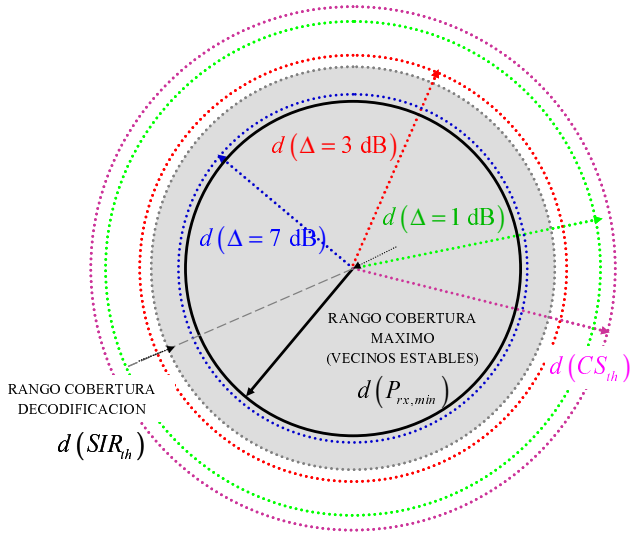
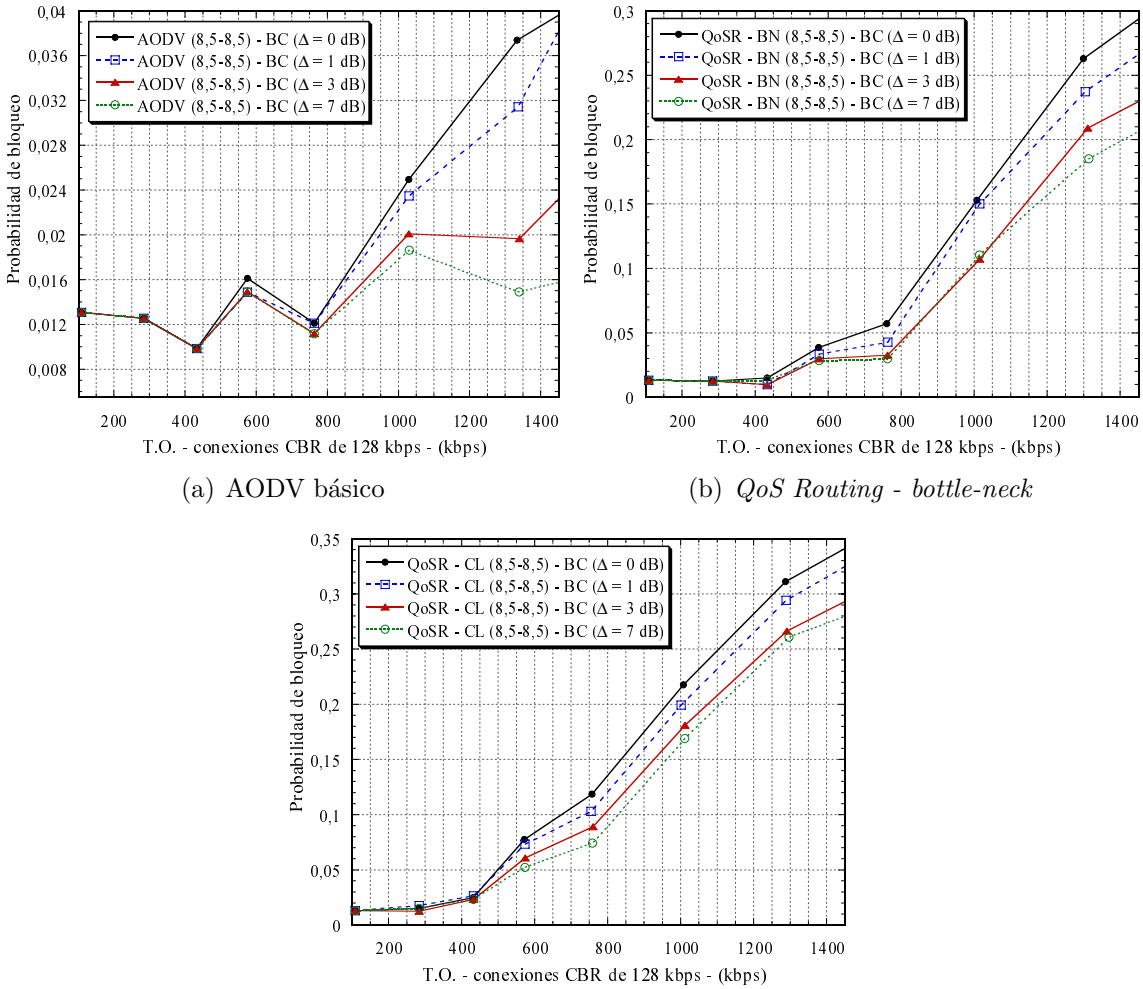


Figura 4.25. Rango de detección de señales $BEACON$ de acuerdo al valor del parámetro Δ: margen en dB sobre el umbral de sensado (CS_{th}).

Una menor sensibilidad definida para las señales *BEACON* (menor Δ) implica un rango de detección mayor para potenciales reservas no decodificadas. Tal y como muestra la figura 4.25, los valores de potencia próximos al umbral de sensado suponen que prácticamente todas las reservas establecidas que el nodo es capaz de detectar, por mínima que sea su potencia, están protegidas frente a la interferencia de éste.



(a) AODV básico

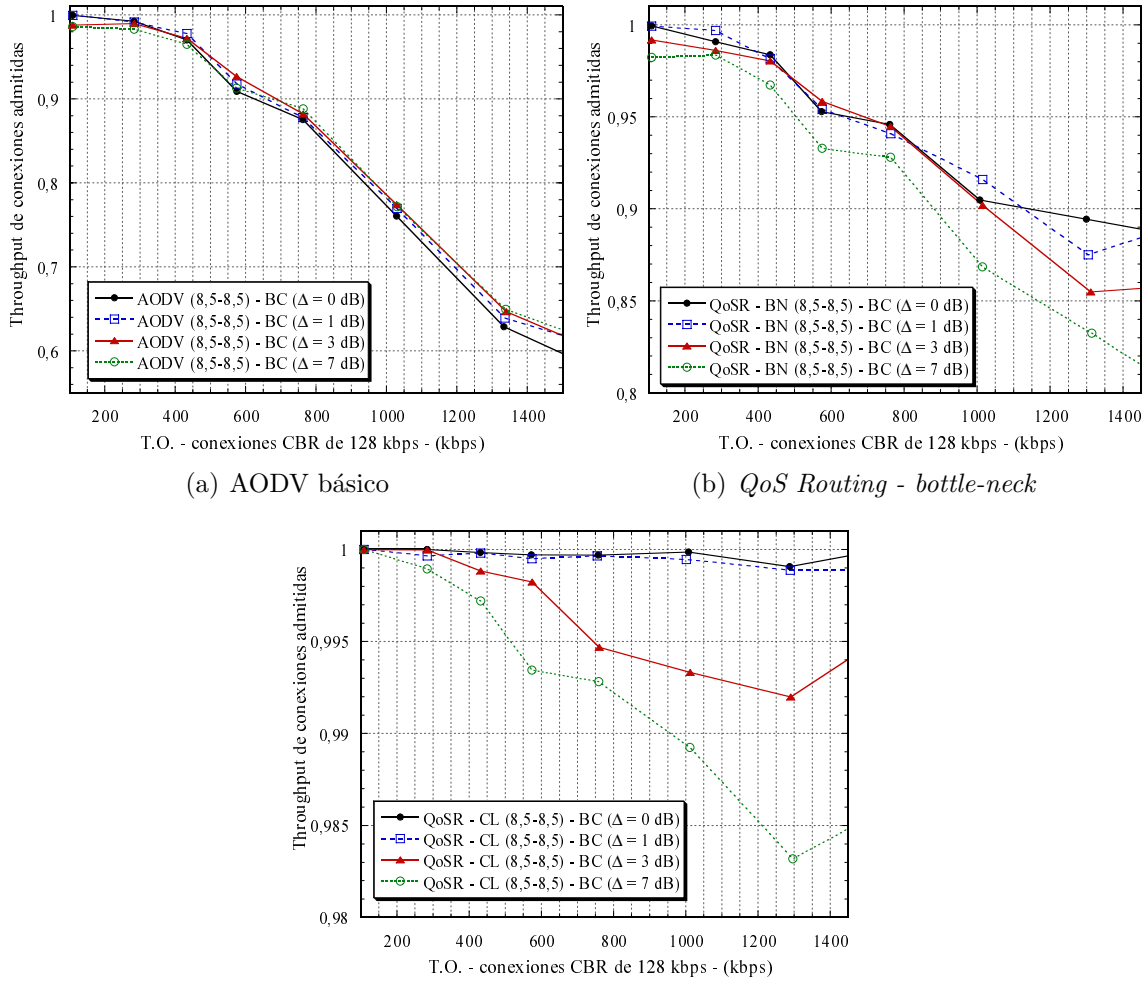
(b) *QoS Routing - bottle-neck*

(c) *QoS Routing - cross-layer*

Figura 4.26. Probabilidad de bloqueo vs. tráfico ofrecido. Estrategias AODV, QoSR-BN y QoSR-CL. Señales *BEACON* – parámetro reflejado: Margen Δ (dB) sobre el umbral de sensado. Escenario de 50 nodos.

La capacidad de reutilizar estos recursos se reduce limitando en consecuencia la disponibilidad y aumentando el bloqueo, como se observa en la figura 4.26. La restricción que supone dar la misma consideración a todos los receptores en el rango de sensado

puede reducirse al aumentar Δ identificando así como receptores efectivos a un número menor de nodos, más cercanos al nodo potencialmente interferente.



(a) AODV básico

(b) *QoS Routing - bottle-neck*

(c) *QoS Routing - cross-layer*

Figura 4.27. *Throughput* de las conexiones admitidas vs. tráfico ofrecido. Estrategias AODV, QoSR-BN y QoSR-CL. Señales $BEACON$ – parámetro reflejado: Margen Δ (dB) sobre el umbral de sensado. Escenario de 50 nodos.

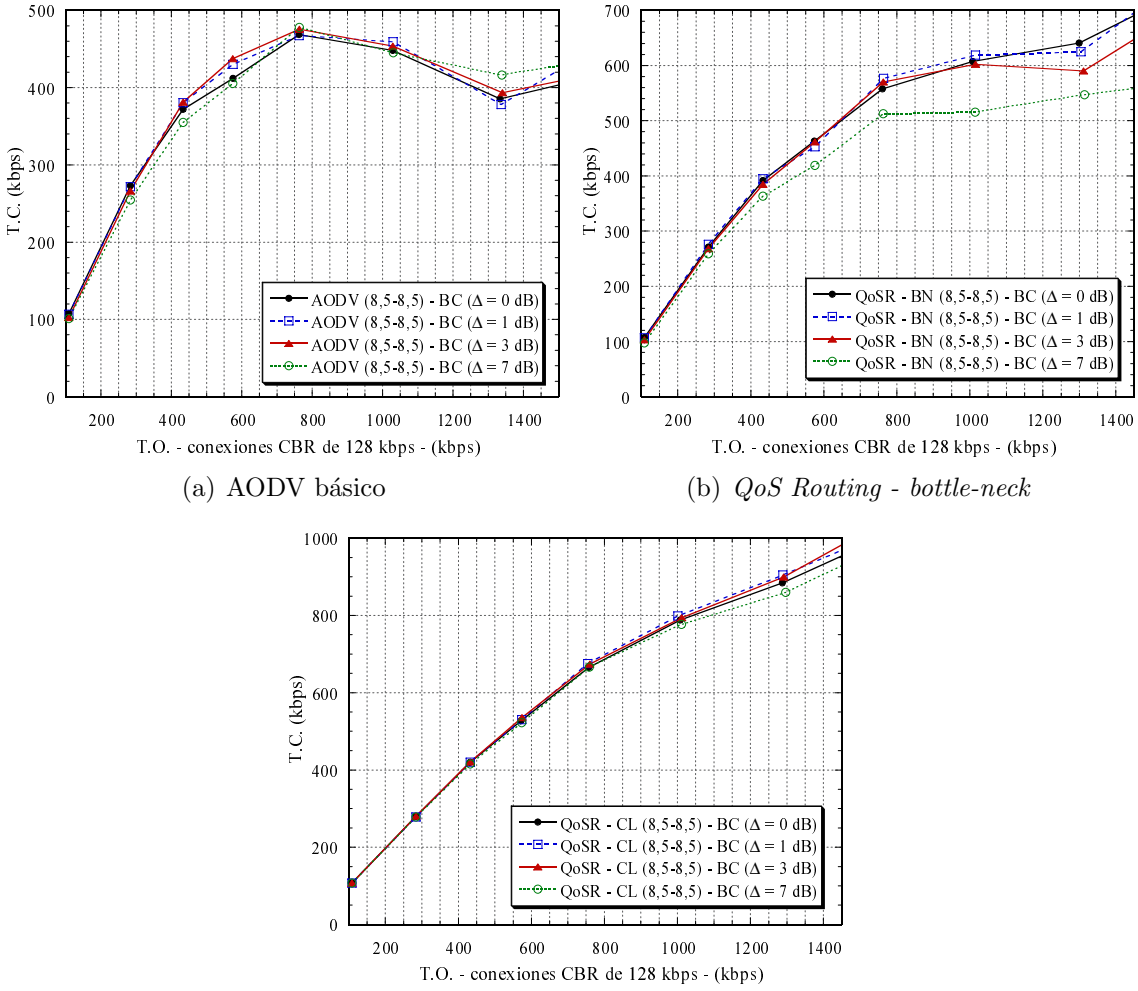
Un margen Δ de 7 dB sobre el sensado supone realmente una potencia equivalente a una SNR de 8 dB, muy próxima pero inferior a la mínima requerida para definir el rango de vecinos estables (8,5 dB). Sin embargo, esta situación no supone una protección real. El umbral mínimo de decodificación es menor (SIR_{th} = 5 dB), de manera que toda la información transmitida en el BCH por los vecinos dentro de dicho rango, aun no siendo éstos considerados estables, es tenida en cuenta, siempre que pueda ser decodificada, a la hora de definir la disponibilidad de recursos. Por esta razón no hay

ninguna aportación real en la recepción de señales $BEACON$. La principal utilidad de dichas señales es precisamente evitar la disponibilidad de los recursos reservados por nodos de los que no puede decodificarse el BCH, bien porque están muy próximos al rango de cobertura (cierta probabilidad de fallo en la recepción del BCH) o porque están a una distancia tal que la transmisión supondría una interferencia no tolerable. Una sensibilidad ligeramente inferior al umbral de detección ($\Delta = 3$ dB) puede considerarse un valor razonable a la hora de conseguir dicha protección manteniendo una restricción menor. Como puede observarse en la figura 4.26, el bloqueo aumenta sólo ligeramente consiguiendo, por otra parte, una mejora en el *throughput* de las conexiones admitidas, como se refleja en la figura 4.27.

Dicha mejora es únicamente apreciable en las estrategias basadas en admisión, dado que en el AODV, la falta de control sobre la activación de conexiones y la consecuente interferencia suponen una situación de congestión inviable. En el caso de la estrategia QoSR-BN, dicha mejora repercute directamente en el tráfico correctamente cursado. El efecto es especialmente significativo debido a la deficiente admisión realizada. Puesto que la medida del enlace "cuello de botella" no identifica correctamente la disponibilidad de la red, el exceso de conexiones admitidas supone niveles de interferencia no tolerables que degradan a dichas conexiones. La utilización de los $BEACONS$ como medida de protección permite aumentar el bloqueo de conexiones mejorando la admisión y de este modo las prestaciones de las conexiones admitidas. En el caso de la estrategia QoSR-CL la mejora es menos significativa, ya que el procedimiento de admisión permite realizar una medida de disponibilidad más acorde con la naturaleza interferente del medio. Los conjuntos de slots disponibles calculados se limitan sucesivamente en cada nodo de acuerdo a su disponibilidad para recibir, de manera que la aportación de las señales $BEACON$ es menor. Así pues, la principal utilidad de esta información no es reducir una admisión incorrecta sino servir de mecanismo de protección frente a nuevas asignaciones de recursos (slots reservados para encaminamiento, conexiones *best-effort*...) limitando la potencial interferencia.

Al aumentar el rango de detección de señales $BEACON$ ($\Delta = 0$ ó 1 dB) aumenta la protección de las reservas establecidas (figura 4.27), aumentando por lo tanto la probabilidad de bloqueo (figura 4.26). A pesar de la reducción de la disponibilidad, finalmente el tráfico correctamente cursado en la estrategia QoSR-BN mejora y se mantiene en la estrategia QoSR-CL, como se observa en la figura 4.28.

La evaluación del grado de servicio (GoS normalizado) obtenido en la estrategia QoSR-CL muestra que a pesar del incremento en la probabilidad de bloqueo, gracias al mantenimiento de las prestaciones deseadas por las conexiones admitidas (*throughput*), se obtiene una mejora global, como se observa en la figura 4.29.



(a) AODV básico

(b) *QoS Routing - bottle-neck*

(c) *QoS Routing - cross-layer*

Figura 4.28. Tráfico cursado vs. tráfico ofrecido. Estrategias AODV, QoSR-BN y QoSR-CL. Señales $BEACON$ – parámetro reflejado: Margen (dB) sobre el umbral de sensado. Escenario de 50 nodos.

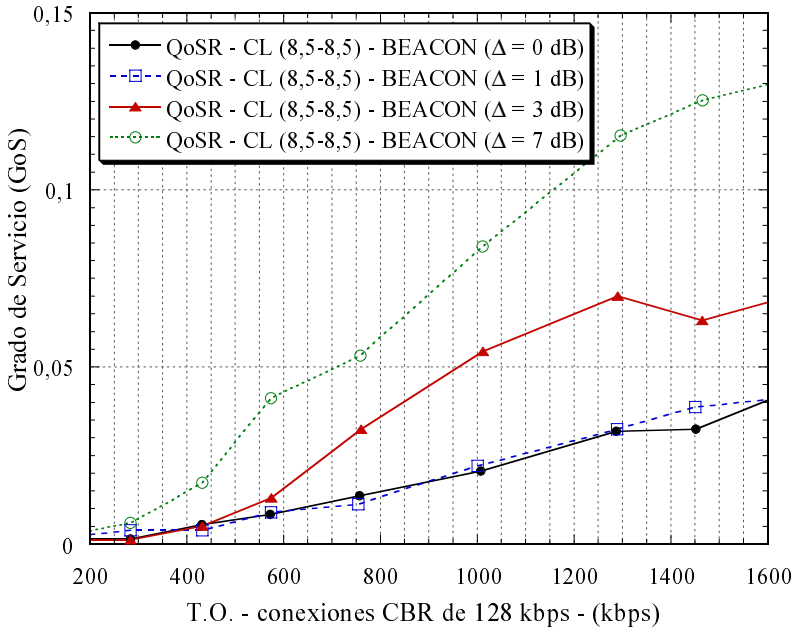


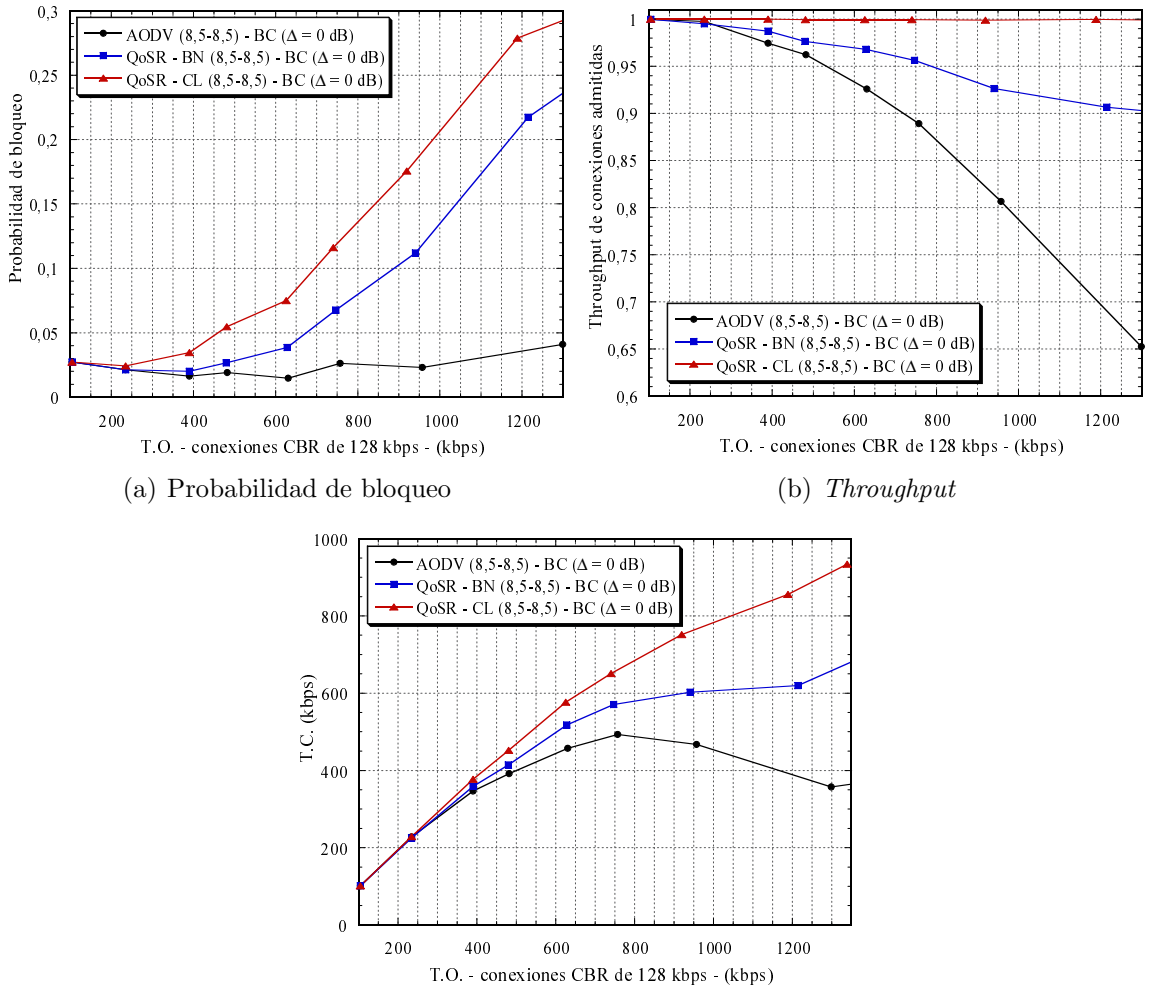
Figura 4.29. Grado de servicio (GoS) vs. tráfico ofrecido. Estrategia QoSR-CL. Señales *BEACON* – parámetro reflejado: Margen (dB) sobre el umbral de sensado. Escenario de 50 nodos.

Evaluación en diferentes condiciones de carga y escenario multiservicio

Con objeto de limitar las inestabilidades en la red, protegiendo al máximo las reservas establecidas para aplicaciones con demandas de QoS, se han establecido inicialmente como parámetros requeridos una $\boldsymbol{SNR} = \mathbf{8{,}5\ dB}$, una $\boldsymbol{SIR_{tar}} = \mathbf{8{,}5\ dB}$ (control de potencia) y un margen $\boldsymbol{\Delta} = \mathbf{0\ dB}$ necesario para detectar la recepción de señales *BEACON*. Dichos parámetros suponen el diseño eficiente del protocolo ADHOC MAC para proporcionar el soporte necesario así como la colaboración adecuada con el encaminamiento que permiten el funcionamiento de la propuesta *cross-layer* diseñada en un escenario de interferencia real [Gállego, 2007]. Así pues, se definen a partir de ahora dichos parámetros como las condiciones básicas de simulación para evaluar las prestaciones del sistema.

A partir del análisis de los resultados obtenidos en §3.5 se determinó la superioridad de la estrategia *cross-layer* para realizar un control de admisión eficiente, dada la correcta estimación de disponibilidad realizada. En un escenario interferente, si bien dicha estimación no es completamente exacta y las prestaciones se ven claramente degradadas por la limitación inherente de capacidad, la estrategia QoSR-CL sigue presentando una clara ventaja sobre un control de admisión simplificado pero ineficiente ("cuello de botella"). Así pues, una vez definidas las condiciones de simulación descritas, se ha

realizado una evaluación más detallada con objeto de verificar el funcionamiento de la arquitectura propuesta ampliando las condiciones de trabajo.



(a) Probabilidad de bloqueo

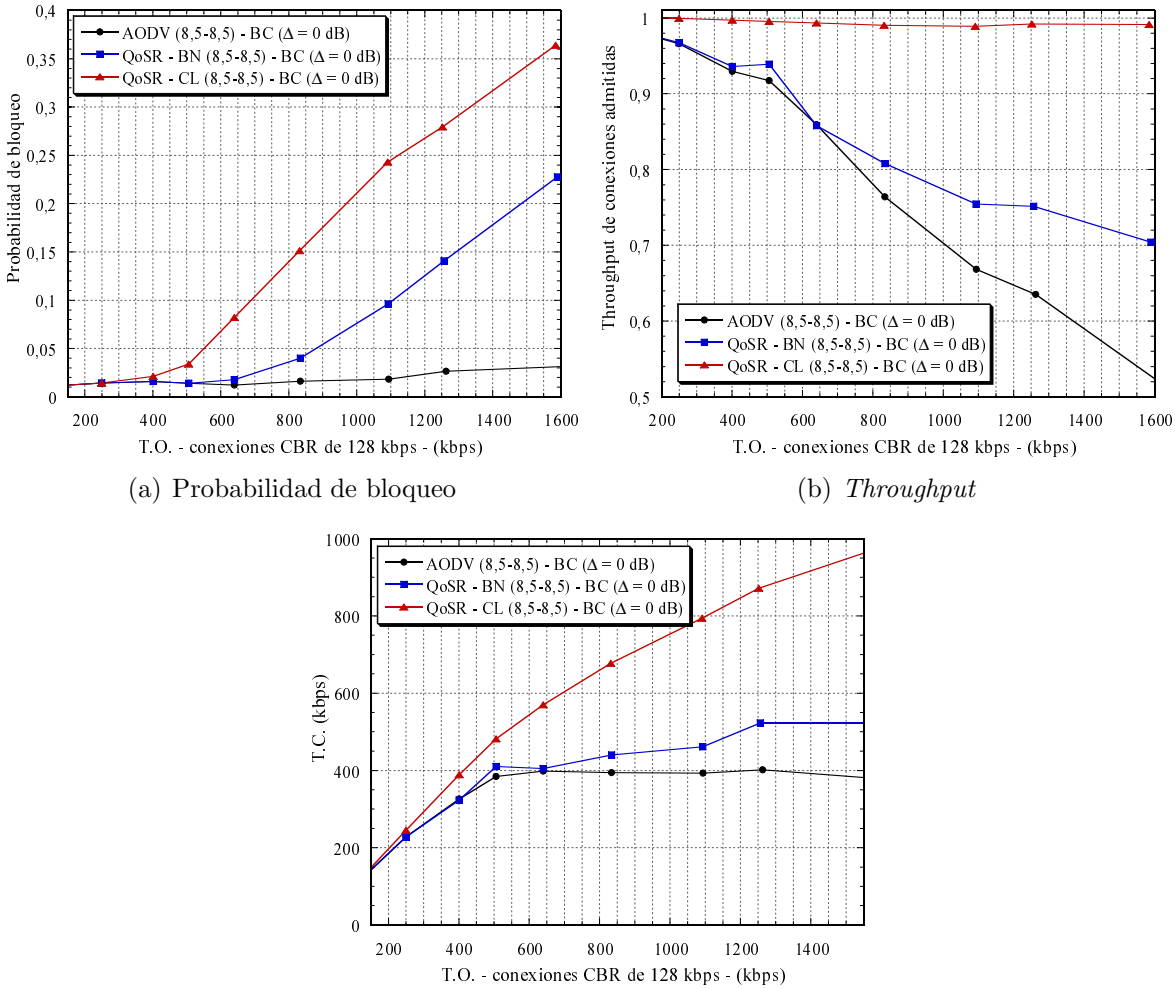
(b) *Throughput*

(c) Tráfico cursado

Figura 4.30. Evaluación de las prestaciones del sistema en las condiciones requeridas de diseño: control de potencia (SIR_{tar} = 8,5 dB) y señales $BEACON$ (Δ = 0 dB). Escenario de 30 nodos.

Los resultados obtenidos en diferentes escenarios verifican la certeza de las conclusiones obtenidas (figura 4.30 – 30 nodos: menor densidad; figura 4.31 – 100 nodos: rutas de mayor número de saltos). La estrategia *cross-layer* supone un bloqueo mayor de conexiones – figuras 4.30(a) y 4.31(a) – que finalmente se traduce en una mejora de las prestaciones obtenidas y un tráfico correctamente cursado superior a la estrategia QoSR-BN, degradada debido a la inexactitud de la estimación realizada mediante la medida de ancho de banda del enlace "cuello de botella". La diferencia observada de-

muestra una mejora más significativa en el escenario de 100 nodos (mayor densidad que el de 30). El mayor bloqueo se debe al incremento de tráfico efectivo ofrecido al considerar un número mayor de enlaces ocupados por conexión. Debido a dicho incremento, la inexactitud de la estrategia QoSR-BN en la medida de ancho de banda supone un mayor exceso de conexiones activadas aumentando la interferencia mutua generada, lo que degrada el *throughput* de las aplicaciones – figura 4.31(b). La asignación efectiva de QoSR-CL, sin embargo, permite el establecimiento de circuitos virtuales garantizados incrementando considerablemente el tráfico correctamente cursado, como se observa en la figura 4.31(c).

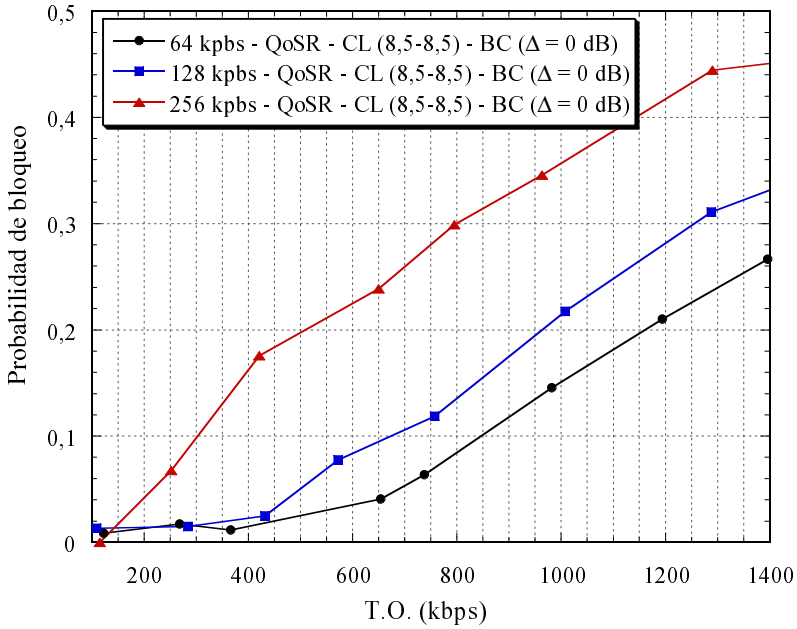


(a) Probabilidad de bloqueo

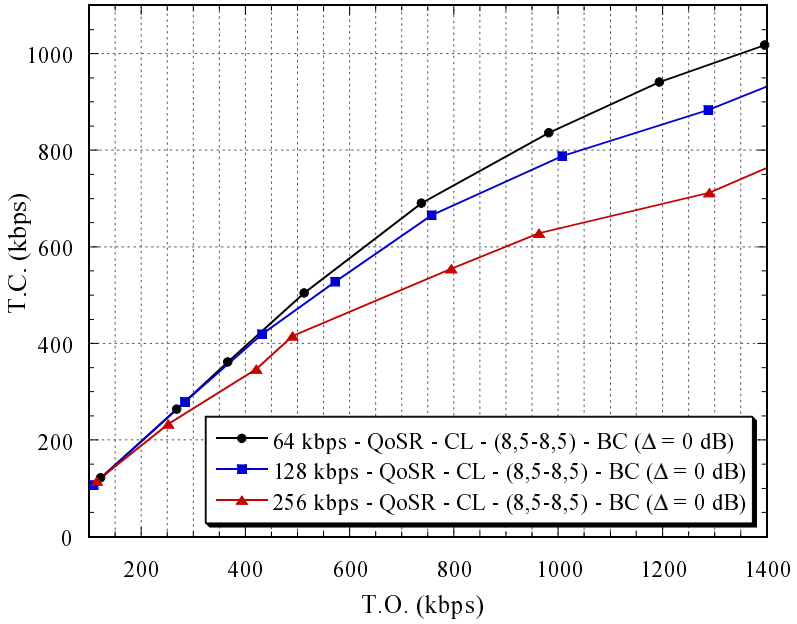
(b) *Throughput*

(c) Tráfico cursado

Figura 4.31. Evaluación de las prestaciones del sistema en las condiciones requeridas de diseño: control de potencia (SIR_{tar} = 8,5 dB) y señales $BEACON$ (Δ = 0 dB). Escenario de 100 nodos.



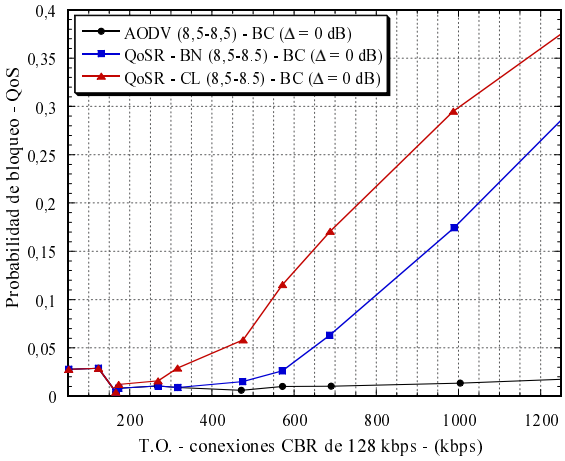
(a) Probabilidad de bloqueo



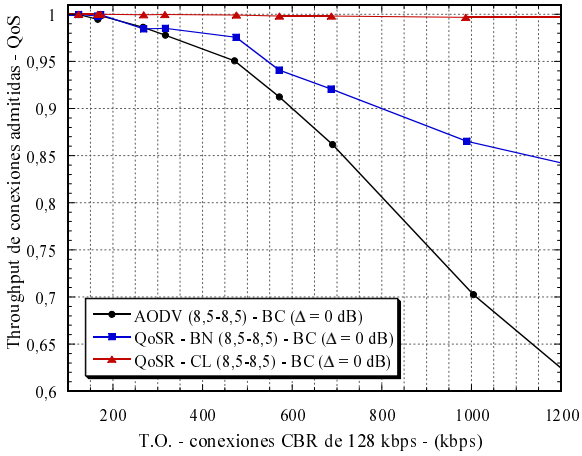
(b) Tráfico cursado

Figura 4.32. Evaluación de la estrategia QoSR-CL en escenarios de tráfico ofrecido total equivalente (kbps) para diversas demandas de ancho de banda por conexión (64, 128 y 256 kbps). BC: señales *BEACON*. Escenario de 50 nodos.

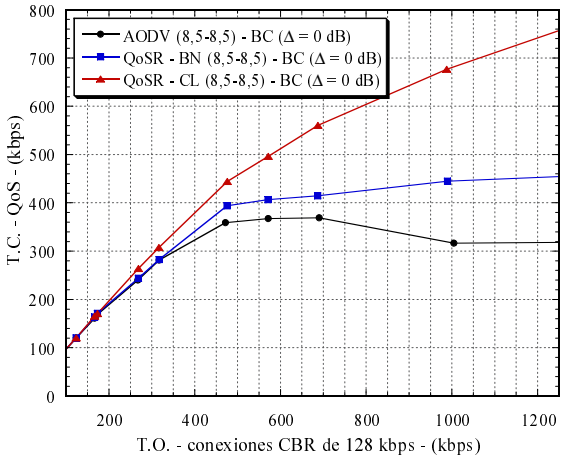
La figura 4.32 muestra el impacto de la demanda individual por conexión sobre las prestaciones de la arquitectura propuesta. Como ya se observó en §3.5, mayores demandas individuales limitan el número de conexiones correctamente admitidas debido a la dificultad intrínseca del protocolo de asignar mayor número de recursos en el mismo número de enlaces. Como se observa en la figura, cuanto menor es la demanda por conexión (64 kbps - 1 slot) menor es la probabilidad de bloqueo – figura 4.32(a) y mayor el tráfico cursado – figura 4.32(b).



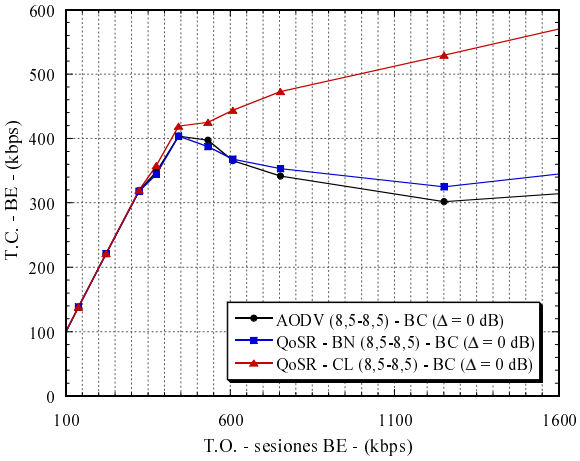
(a) Probabilidad de bloqueo



(b) Throughput



(c) Tráfico QoS cursado



(d) Tráfico *best-effort* cursado

Figura 4.33. Evaluación de las prestaciones del sistema en un escenario multiservicio: Tráfico ofrecido CBR de 128 kbps y sesiones BE (50 %). Control de potencia (SIR_{tar} = 8,5 dB) y señales $BEACON$ (Δ = 0 dB). Escenario de 50 nodos.

Al igual que en el análisis realizado en §3.5, se ha evaluado la capacidad de la propuesta de garantizar un servicio diferenciado acorde con las demandas de las aplicaciones. Para ello, se han evaluado las prestaciones del sistema, asumiendo las condiciones de diseño ya descritas (control de potencia y $BEACONS$), en un entorno multiservicio. El tráfico ofrecido está representado tanto por aplicaciones CBR de 128 kbps (circuitos virtuales de 2 slots TDMA) como tráfico *best-effort*, con una proporción relativa del 50 %. Las sesiones *best-effort* se establecen mediante llamadas, exponencialmente distribuidas (media: 5 seg. entre llamadas) con un número medio de 50 paquetes (distribución geométrica), generados a una tasa de 64 kbps con un tiempo medio entre paquetes exponencialmente distribuido de media 62,5 ms.

La figura 4.33 muestra que la estrategia QoSR-CL mantiene su superioridad sobre QoSR-BN, a pesar del control de admisión también realizado mediante esta última. La interferencia generada por las conexiones *best-effort* admitidas degrada las prestaciones de las conexiones QoS, a pesar de la diferenciación de servicios realizada (prioridades MAC), puesto que la posible reserva de recursos menos prioritarios únicamente está condicionada por la existencia de reservas de mayor prioridad en el entorno a 2 saltos (intercambio de información en el BCH).

El efecto es más llamativo especialmente teniendo en cuenta la sobreestimación de la capacidad disponible realizada mediante el "cuello de botella", que incrementa además la competencia entre aplicaciones prioritarias. El mayor bloqueo realizado en QoSR-CL, como se observa en la figura 4.33(a) limita la interferencia generada por la activación de conexiones. Las reservas prioritarias finalmente establecidas, por otra parte, gozan de una elevada protección frente al tráfico *best-effort* (admitido siempre) gracias a la utilización de las señales $BEACON$. De este modo, los circuitos virtuales mantienen el ancho de banda garantizado, permitiendo obtener unas prestaciones casi ideales en las aplicaciones CBR admitidas (*throughput* – figura 4.33(b)). La gestión de recursos eficiente proporcionada por esta estrategia permite obtener una ganancia considerable tanto en las prestaciones de las conexiones QoS como en el propio tráfico *best-effort*, como se observa en las figuras 4.33(c) y 4.33(d) respectivamente. Gracias a la distribución eficiente de los recursos prioritarios en circuitos virtuales garantizados, las sesiones BE pueden aprovechar mejor el ancho de banda remanente sin perjuicio de las aplicaciones CBR. De hecho, la limitación en el número de conexiones CBR admitidas evita una carga excesiva en el sistema de conexiones prioritarias que realmente no pueden cursarse satisfactoriamente y que, de admitirse (estrategias AODV y QoSR-BN), reducirían innecesariamente la disponibilidad de recursos para el tráfico *best-effort*.

En definitiva, considerando que los resultados mostrados se corresponden con una evaluación preliminar del sistema en condiciones estáticas, las prestaciones obtenidas mediante la arquitectura basada en el diseño *cross-layer* propuesto han demostrado la eficacia de ésta como mecanismo de admisión distribuido incluso en un escenario más realista de interferencia entre usuarios. El diseño eficiente del protocolo ADHOC MAC, adaptado a las problemática de dicho escenario (*interference-awareness*) permite garantizar el soporte adecuado para la arquitectura diseñada. La inestabilidad de la red generada por reasignaciones de BCH [Gállego, 2007] es superada mediante el margen de protección establecido (SNR = 8,5 dB), al reducir las colisiones experimentadas en dichos recursos a consecuencia de niveles de interferencia no sensada. De igual modo, la definición del conjunto de vecinos estables y la limitación de activaciones únicamente en aquellos enlaces fiables garantiza una protección equivalente en las transmisiones de usuario. En estas condiciones, la interferencia acumulada creada por nuevas activaciones es la limitación principal sobre las prestaciones del sistema en comparación con el análisis realizado mediante el *Protocol Model*, una vez asumida la reducción de capacidad inherente a la naturaleza del medio. Las estrategias diseñadas, como el control de potencia ajustado al mismo nivel de protección del BCH y la inclusión de señales $BEACON$, garantizan una elevada protección de las reservas prioritarias establecidas. De este modo, a pesar del incremento en la probabilidad de bloqueo introducido por la estrategia QoSR-CL, se mantienen las prestaciones deseadas en cuanto a *throughput* garantizando así un tráfico correctamente cursado mayor, lo que, desde el punto de vista de la percepción de usuario, supone unas prestaciones globales superiores.

Las conclusiones obtenidas sirven de base para validar el diseño *cross-layer* como mecanismo de admisión. La posterior evaluación en condiciones de mayor variabilidad servirá para analizar la capacidad de los protocolos involucrados de reaccionar adecuadamente en dicho entorno.

4.4. Evaluación de la propuesta en condiciones de movilidad

En §4.3 se ha mostrado el estudio realizado sobre el efecto de la interferencia real en el mecanismo de admisión desarrollado. Como resultado de dicha evaluación, se han establecido las condiciones necesarias, desde la perspectiva del ADHOC MAC, para que este protocolo sirva de soporte fiable del diseño *cross-layer* implementado. Una

vez verificado el funcionamiento en escenarios completamente estáticos, ante diversas condiciones de tráfico ofrecido (CBR de diferentes tasas, *best-effort*), se pretende evaluar la capacidad de adaptación del sistema a un entorno variante condicionado por la movilidad de los terminales. A la hora de caracterizar dicha movilidad, existen multitud de modelos desarrollados en el contexto de redes móviles ad hoc [Camp y otros, 2002]. Uno de los modelos ampliamente utilizado a la hora de evaluar las prestaciones de los protocolos de encaminamiento en condiciones de movilidad, convertido prácticamente en un estándar *de facto*, es el *Random WayPoint mobility model* (RWP), utilizado por primera vez por Johnson y Maltz [1996] en su evaluación del protocolo DSR, y posteriormente refinado por el mismo grupo de trabajo [Broch y otros, 1998]. No obstante, la excesiva aleatoriedad de este modelo así como la problemática asociada al elevado tiempo de convergencia, estudiadas en [Casilari y Triviño, 2005; Resta y Santi, 2002; Yoon y otros, 2003], limitan tanto su funcionalidad a la hora de modelar un entorno variante realista como su aplicación en el entorno de simulación. Si bien existen propuestas que pretenden emular entornos de aplicación más cercanos a la realidad, como se describe en [Bai y otros, 2003; Jardosh y otros, 2005] (desplazamiento en ciudades, entornos de oficinas, etc.), la versatilidad de los mismos resulta mucho menor que la del RWP, al determinar en exceso los escenarios de estudio. Siguiendo un planteamiento más genérico a la hora de desarrollar un entorno de evaluación para la arquitectura propuesta, se ha considerado un modelo de movilidad más simplificado, basado en el RWP. Ciertas modificaciones realizadas, detalladas en el anexo C, pretenden reducir la aleatoriedad en la trayectoria de los móviles, lo que supone una aproximación más cercana al movimiento natural tanto de personas como de vehículos.

La parametrización del modelo responde a la implementación básica del RWP con la incorporación de dos grados de libertad adicionales que permiten introducir cierta dependencia en la trayectoria de los móviles (efecto memoria). Tal y como se ha definido en el anexo C, la caracterización completa depende de los parámetros detallados en la tabla 4.2, donde V_{min} y V_{max} se corresponden con los valores mínimo y máximo que definen el rango de variación uniforme en la velocidad de cada terminal[17]. El T_{pausa} es la duración de los intervalos estáticos entre cada movimiento, definido éste como el desplazamiento, a la velocidad seleccionada, hasta un destino específico, elegido a una distancia aleatoria del punto de partida, uniformemente distribuida y limitada por R_{max}, restringiéndose a un ángulo de variación máximo sobre la dirección actual, α_{max}. En la

[17]El valor V_m, correspondiente a la velocidad media resultante, se desprende del estudio analítico del modelo RWP [Yoon y otros, 2003] – (C.1).

figura 4.34 se observa, a modo de ejemplo, la trayectoria seguida por dos terminales en diferentes escenarios de movilidad. Las velocidades consideradas a la hora de realizar el análisis mostrado en este apartado se corresponden con tres escenarios representativos de movilidad moderada desde una velocidad media de 2 Km/h hasta 20 Km/h, con velocidades comprendidas entre 1 y 40 Km/h.

Tabla 4.2. Escenarios de movilidad. Modelo RWPM.

Parámetros del RWP				**Parámetros del RWPM**	
V_{min} (Km/h)	V_{max} (Km/h)	T_{pausa} (seg.)	V_m (Km/h)	α_{max}	R_{max}
1	3	2	~ 2	$\frac{\pi}{12}$	$0{,}1 \times D_{max}$ [a]
3	20	2	~ 9	$\frac{\pi}{12}$	$0{,}1 \times D_{max}$
10	40	2	~ 20	$\frac{\pi}{12}$	$0{,}1 \times D_{max}$

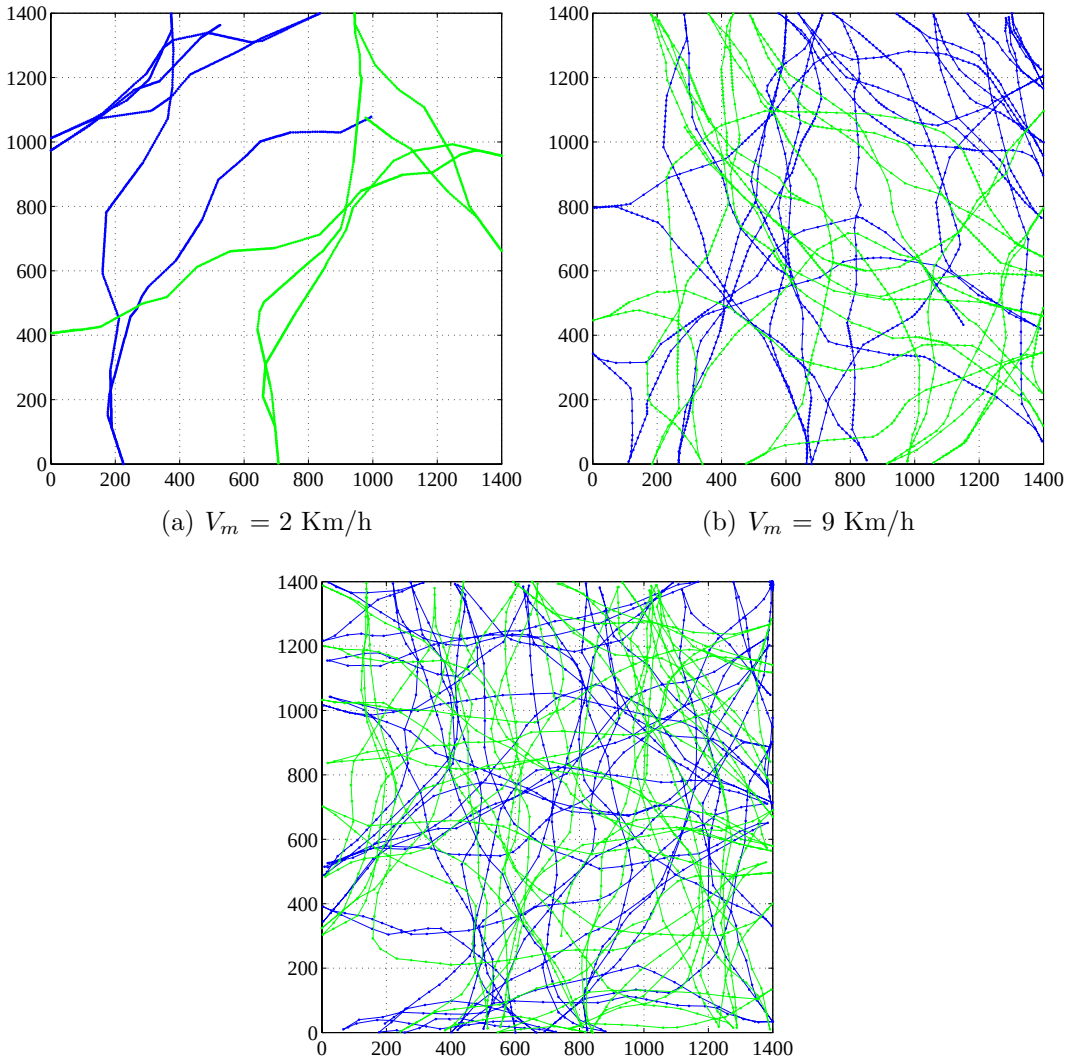
[a] $D_{max} = \sqrt{(X_{max}^2 + Y_{max}^2)}$, donde X_{max} y Y_{max} definen los límites del escenario (1400 m.)

Los parámetros de simulación establecidos se corresponden con la evaluación realizada en condiciones estáticas. Se han considerado diversos escenarios de movilidad, de acuerdo a los modelos mostrados en la tabla 4.2, para topologías de 50 nodos situados en una superficie de 2 Km^2. La tabla 4.3 recoge todos los parámetros de simulación utilizados, tanto a nivel físico como a nivel MAC, de acuerdo a los análisis realizados previamente. La especificación de los parámetros SNR, SIR_{tar} y Δ se corresponde, tal y como se concluyó en §4.3, con la definición de los requisitos necesarios para la estabilidad y eficiencia del sistema en un entorno de interferencia entre usuarios.

Tabla 4.3. Parámetros de simulación.

CS_{th} (sensibilidad)	1 dB
SIR_{th} (decodificación)	5 dB
SNR (de acuerdo a $P_{rx,min}$)	8,5 dB
SIR_{tar} (control de potencia)	8,5 dB
Δ (detección de $BEACON$)	0 dB
Pérdidas de propagación	Kammerman (3.27)
N_{DATA}	25 slots
N_{BCH}	50 slots[a]

[a] Garantizar probabilidad de acceso $p_{access} \sim 1$ en todas las condiciones de simulación.



(a) $V_m = 2$ Km/h

(b) $V_m = 9$ Km/h

(c) $V_m = 20$ Km/h

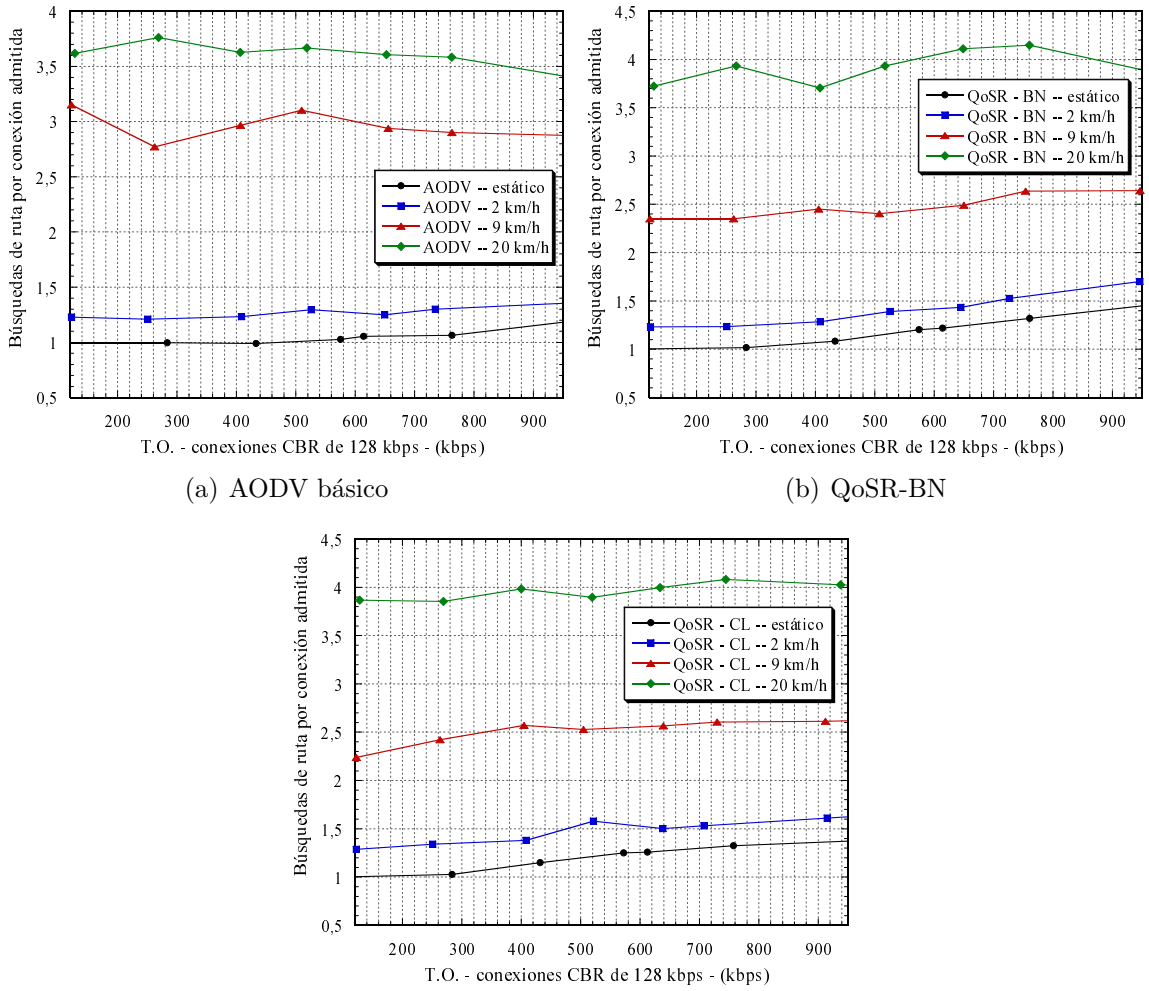
Figura 4.34. Trayectoria de dos terminales móviles (12000 seg.), muestra de las topologías simuladas. Caracterización en base a la velocidad media V_m. $R_{max} = 200$ m. $\alpha_{max} = \frac{\pi}{12}$.

Evaluación del esquema básico en condiciones de movilidad

Como resultados representativos se muestra la evaluación realizada sobre aplicaciones CBR de 128 kbps y longitud constante de 50 segundos. La satisfacción de la QoS se evalúa en función del porcentaje de pérdidas experimentado limitado a un valor P_{loss} máximo del 3 %. A diferencia de los casos estudiados en §3.5 y §4.3, evaluados en condiciones estáticas, donde todas las aplicaciones admitidas son cursadas hasta su finalización, en un escenario variante las desconexiones de enlaces provocadas por la movilidad relativa de los terminales suponen continuas roturas en las rutas que generan en consecuencia intentos de readmisión de las conexiones afectadas. Dicha readmisión, sin embargo, puede resultar imposible dada la variación en la topología de la red (red parcialmente conexa temporalmente). Teniendo en cuenta este fenómeno, se consideran **conexiones cursadas** aquéllas admitidas y finalizadas, independientemente de las roturas sufridas, es decir, que no son bloqueadas en ningún intento de readmisión. En este sentido, los resultados de *throughput* mostrados (ratio entre paquetes correctamente recibidos y paquetes generados) representan la medida relativa únicamente a dichas conexiones, excluyendo por tanto las conexiones fallidas. La representación del tráfico correctamente cursado supone, por otra parte, el tráfico de las aplicaciones cursadas que realmente experimentan la QoS deseada medida de acuerdo al porcentaje P_{loss}. De manera análoga, la no satisfacción se evalúa a partir del GoS normalizado, donde la probabilidad correspondiente P_f tiene una ponderación superior (§4.2.2).

Las figuras 4.35, 4.36 y 4.37 presentan el efecto de los diferentes grados de movilidad sobre las prestaciones de las tres estrategias de encaminamiento analizadas: AODV básico, QoSR-BN o "cuello de botella" y la propuesta *cross-layer* QoSR-CL.

En cualquiera de las estrategias, la creciente movilidad supone un aumento de las readmisiones debidas a roturas, como se refleja en la figura 4.35. La figura 4.36 muestra la consecuente degradación del *throughput* experimentado por las conexiones cursadas debido a las pérdidas de paquetes acumulados en las colas durante los procedimientos de redescubrimiento de ruta, así como los descartados en los enlaces afectados. Como resultado, el tráfico correctamente cursado sufre una degradación muy significativa, como se muestra en la figura 4.37. En definitiva, la capacidad efectiva del sistema se reduce drásticamente.

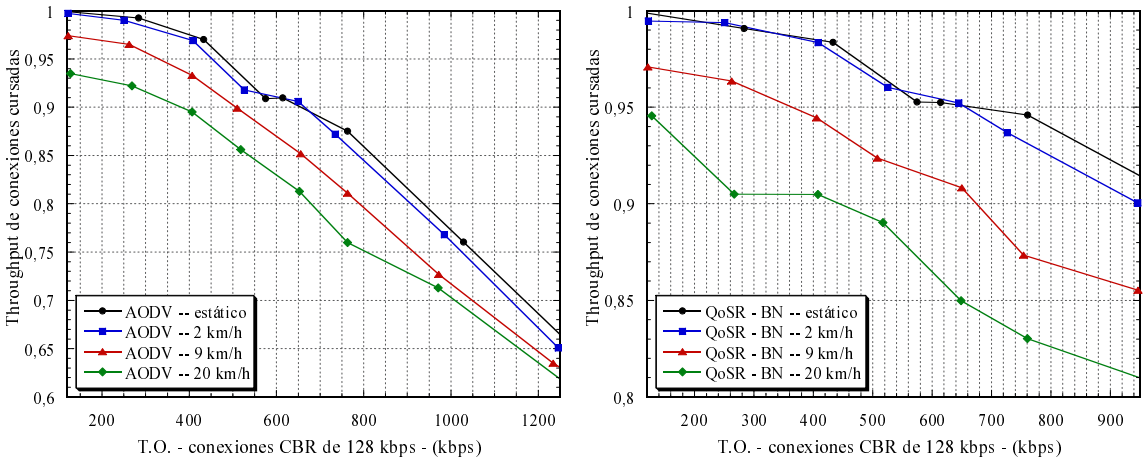


(a) AODV básico

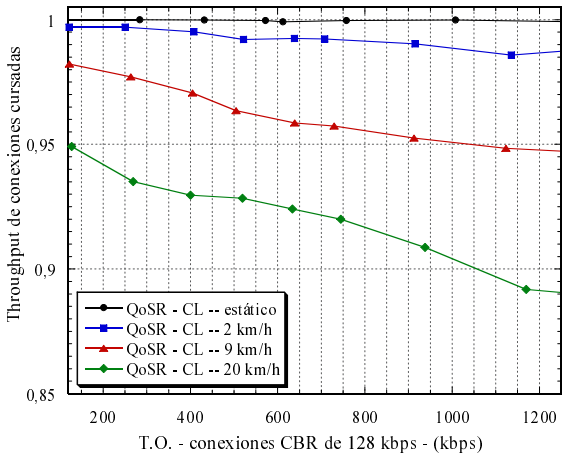
(b) QoSR-BN

(c) QoSR-CL

Figura 4.35. Número de readmisiones (búsquedas de ruta) por conexión vs. tráfico ofrecido. Comparación de estrategias AODV, QoSR-BN y QoSR-CL. Escenarios de 50 nodos, velocidades medias de 2, 9 y 20 Km/h.



(a) AODV básico (b) QoSR-BN

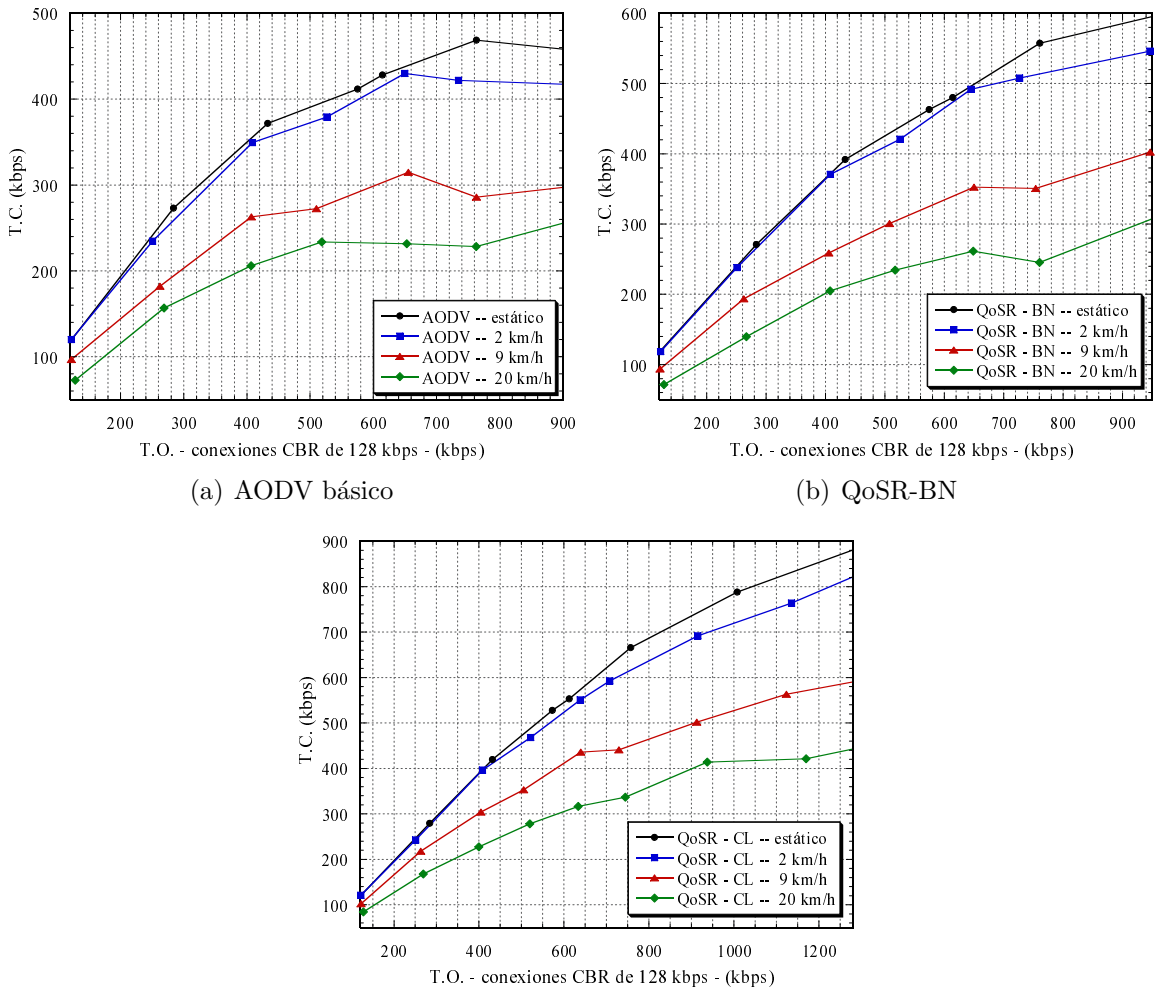


(c) QoSR-CL

Figura 4.36. *Throughput* de conexiones cursadas vs. tráfico ofrecido. Comparación de estrategias AODV, QoSR-BN y QoSR-CL. Escenarios de 50 nodos, velocidades medias de 2, 9 y 20 Km/h.

A pesar de la notable reducción de capacidad efectiva, cualquiera de las estrategias basadas en admisión sigue superando las prestaciones del mecanismo básico AODV. La reducción de colisiones gracias al establecimiento de circuitos virtuales de ancho de banda garantizado permite reducir la pérdida de paquetes siendo el factor limitante las situaciones de rotura inherentes al escenario de movilidad. La estrategia *cross-layer* desarrollada sigue mejorando las prestaciones del encaminamiento basado en una métrica de ancho de banda mínimo ("cuello de botella"), donde las inexactitudes en la definición de recursos disponibles se ven incrementadas por la variabilidad del entorno. No

obstante, los resultados mostrados constatan las evidentes dificultades para garantizar la QoS ante las continuas pérdidas de los enlaces establecidos o la variabilidad de la interferencia sufrida. En consecuencia, es necesario establecer mecanismos de adaptación capaces de flexibilizar la utilización de los recursos evitando en la medida de lo posible la degradación de las conexiones una vez admitidas. En este sentido, de acuerdo a la funcionalidad descrita en §4.2 a continuación se presentan los resultados obtenidos para evaluar y caracterizar los parámetros configurables del sistema diseñados con objeto de permitir la adaptación deseada.



(a) AODV básico

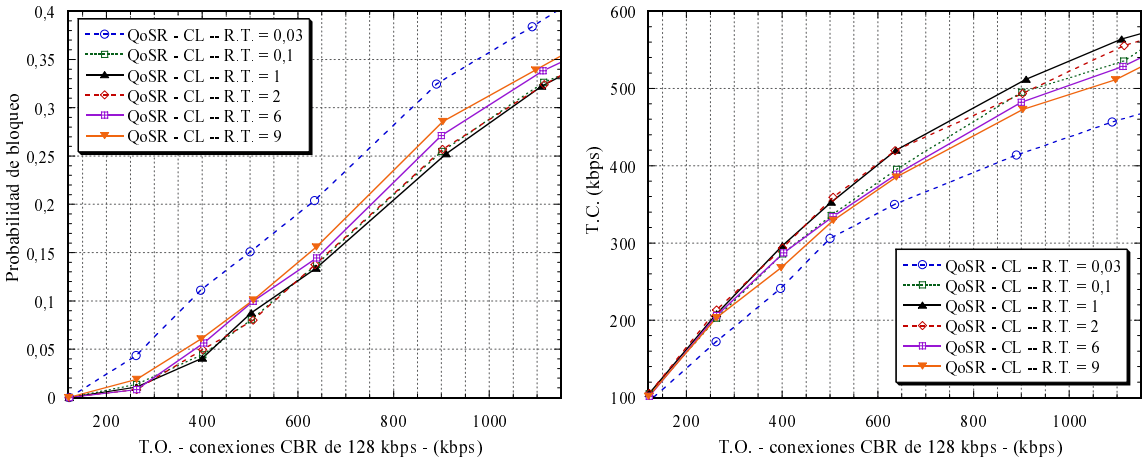
(b) QoSR-BN

(c) QoSR-CL

Figura 4.37. Tráfico correctamente cursado vs. tráfico ofrecido. Comparación de estrategias AODV, QoSR-BN y QoSR-CL. Escenarios de 50 nodos, velocidades medias de 2, 9 y 20 Km/h.

Mecanismo de liberación *soft-rsv*. Efecto del valor de temporización (RT)

Cuando una ruta con QoS se rompe a consecuencia de la movilidad, el protocolo inicia una nueva búsqueda liberando los recursos reservados en el camino fallido para evitar el bloqueo de la nueva readmisión. Sin embargo, la incomunicación entre parte de los nodos que conforman el camino completo (enlaces entre la rotura y el destino) impide la liberación de los recursos en estos enlaces, por lo que se ha implementado un mecanismo de liberación *soft-rsv*, tal y como se describió en §4.2.1, basado en un tiempo de vida – RESERVATION_TIMEOUT (RT) – asociado a las reservas de los slots asignados dentro del BW^n_{CAC}[18]. A la hora de seleccionar un valor RT apropiado se han realizado diversas simulaciones cuyo resultado ha demostrado la necesidad de limitar considerablemente dicho valor para evitar la ocupación innecesaria de recursos, especialmente en aquellos escenarios de elevada movilidad que suponen continuas roturas y reasignaciones. Esta limitación, sin embargo, no debe conducir a un valor excesivamente pequeño que suponga una liberación prematura de los recursos.



(a) Probabilidad de bloqueo (b) Tráfico cursado

Figura 4.38. Efecto del parámetro RESERVATION_TIMEOUT (RT) - seg. Estrategia QoSR-CL. (ART = 10 seg.). Escenario de 50 nodos con velocidad media de 9 Km/h.

En definitiva, una vez evaluados múltiples escenarios de aplicación, de los que se muestra un resultado representativo (figura 4.38), se ha estimado como valor apropiado un RT de 1 seg. utilizado en adelante en la evaluación del resto de parámetros de diseño. Como se muestra en dicha figura, un tiempo de permanencia excesivo de los recursos reservados genera una probabilidad de bloqueo mayor, dada la ocupación innecesaria de éstos, provocando una degradación de las prestaciones respecto al mejor de los casos.

[18]Conjunto de slots reservados para flujos QoS dentro del enlace hacia el destino n – (3.20).

Por otra parte, una liberación temprana genera un escenario de permanente variabilidad en la ocupación de los recursos más próxima a una utilización *best-effort* que a la activación de circuitos virtuales. En consecuencia, la estimación del ancho de banda del mecanismo de admisión (algoritmo *BWC-FA*) puede llegar a resultar completamente ineficaz, generando incluso bloqueos innecesarios – figura 4.38(a) – al no poderse verificar la disponibilidad de los recursos estimados durante la búsqueda (fase *RREQ*) en el momento de confirmar la ruta (fase *RREP*). De este modo, puede llegar a producirse una llamativa pérdida de capacidad efectiva, como se observa con un RT de 0,03 seg.

Mecanismo de reasignación. Efecto de la configuración de parámetros

La degradación de los recursos reservados debido al aumento de la interferencia sufrida como consecuencia del acercamiento de nodos inicialmente capaces de reutilizar los mismos slots sin colisión provoca la necesidad de reasignar recursos dentro de un camino inicialmente seleccionado como circuito virtual. En consecuencia, la admisión realizada durante el encaminamiento únicamente establece un criterio inicial a la hora de estimar como posible la entrada al sistema de una aplicación determinada. Una vez admitida, los recursos establecidos teniendo en cuenta los enlaces específicos dentro del camino pueden quedar completamente invalidados en el transcurso de la comunicación.

A la hora de realizar la necesaria reasignación, de acuerdo al criterio de ancho de banda extremo a extremo sería necesario establecer nuevamente el procedimiento de cálculo y asignación realizado durante la admisión inicial (*BWC-FA*). Mientras la comunicación a nivel de enlace puede mantenerse con el nivel de fiabilidad requerido (vecinos estables) la ruta permanece activada, siendo potencialmente posible seguir encontrando recursos disponibles sin realizar una nueva búsqueda. De acuerdo a este principio, el procedimiento de monitorización implementado (§4.2.1) actúa de manera equivalente al protocolo de encaminamiento incorporando la ejecución del *BWC-FA* gracias a la información de control añadida sobre los propios paquetes de datos, así como mensajes específicos de encaminamiento enviados (*ROUTE_ACK*). Dicho mecanismo, si bien permite recuperar de un modo flexible recursos en un camino establecido, también añade *overhead* de control que puede llegar a ser considerable, reduciendo así la capacidad efectiva del sistema. Así pues, a la hora de establecer los parámetros de diseño adecuados ha de valorarse el compromiso que se presenta.

El mecanismo de reasignación supone el envío de información de control sobre los datos (*piggy-backing*) con una periodicidad establecida t_{update}, lo que directamente condicionará el *overhead* introducido. Cada estimación realizada del ancho de banda dispo-

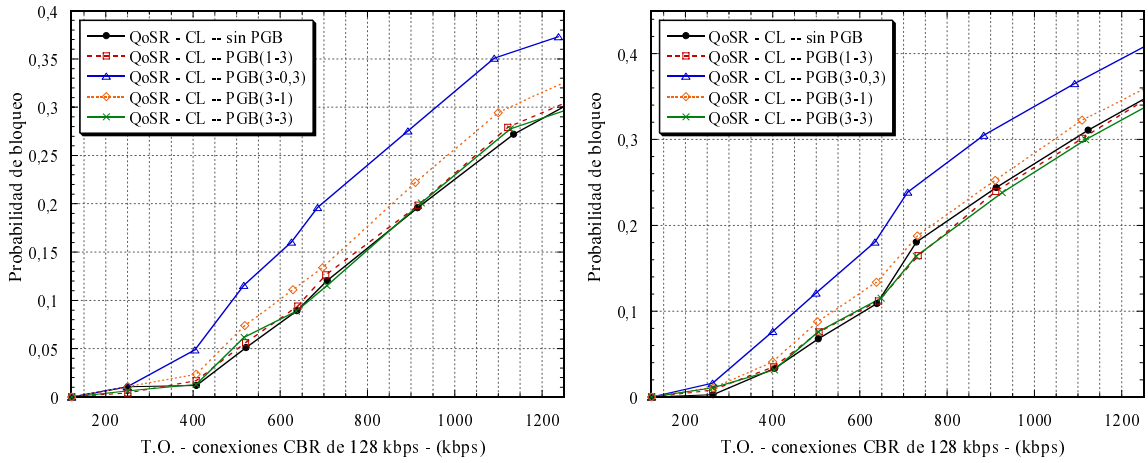
nible adicional supone la asignación de nuevos recursos dentro de dicha disponibilidad en el caso de haber perdido los slots inicialmente reservados. Cuando esta nueva asignación no es posible tras varios intentos $n_{failupd}$, la monitorización realizada permite identificar una pérdida de QoS no recuperable. En definitiva, el tiempo total de reacción disponible para recuperar recursos y en su defecto realizar una nueva búsqueda es de $t_{reacc} = t_{update} \cdot n_{failupd}$. No obstante, hay que destacar igualmente la importancia del parámetro individual $n_{failupd}$, dado que un valor muy reducido puede suponer la detección de "falsas pérdidas de QoS" mientras que un valor excesivo invalida la funcionalidad del mecanismo[19]. A continuación se muestra la importancia de la definición de estos parámetros a través de los resultados obtenidos, en los tres escenarios de movilidad analizados, donde se han considerado las configuraciones detalladas en la tabla 4.4.

Tabla 4.4. Mecanismo de monitorización y reasignación.

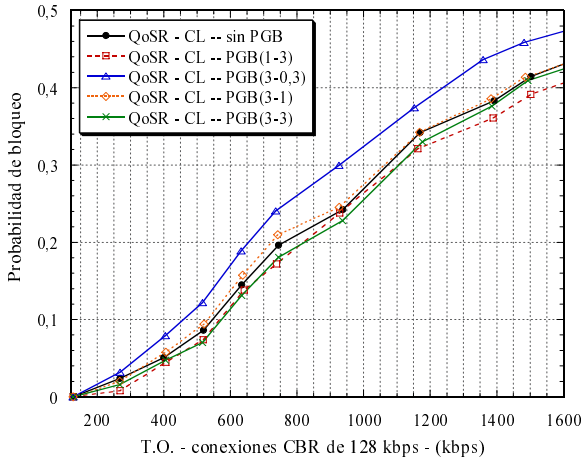
	Parámetros de *PiGgy-Backing* (PGB)		
	$n_{failupd}$	t_{update} (seg.)	t_{reacc} (seg.)
PGB(1-3)	1	3	3
PGB(3-0,3)	3	0,3	0,9
PGB(3-1)	3	1	3
PGB(3-3)	3	3	9

La doble finalidad del mecanismo de actualización-monitorización supone realmente un compromiso que condiciona la aplicabilidad del mismo según el grado de movilidad. Una mayor frecuencia de actualización (t_{update} menor) puede permitir reasignar ancho de banda rápidamente, flexibilizando en gran medida la utilización de los recursos. De igual modo, cuando dicha reasignación no es posible, una pronta reacción ante la pérdida de QoS puede permitir realizar una nueva búsqueda de camino consiguiendo así realojar adecuadamente la conexión en el sistema. Sin embargo, como puede observarse en la figura 4.39, en todos los escenarios de movilidad analizados, el *overhead* que supone introducir periódicamente la información de control asociada al procedimiento de monitorización incrementa la probabilidad de bloquear aplicaciones de QoS, especialmente en el caso del PGB(3-0,3), que supone una actualización cada 0,3 seg. Tanto el propio mecanismo de monitorización (*piggy-backing*) como la posible reacción (nueva búsqueda de ruta) suponen el uso de mayor número de recursos prioritarios, lo que limita la capacidad efectiva disponible para cursar las aplicaciones CBR.

[19]Es especialmente importante tener en cuenta la duración de las aplicaciones (50 seg.). Una actualización demasiado lenta de recursos en comparación con dicha duración no permite reasignar recursos adecuadamente ni reaccionar con la suficiente prontitud ante la pérdida de QoS.



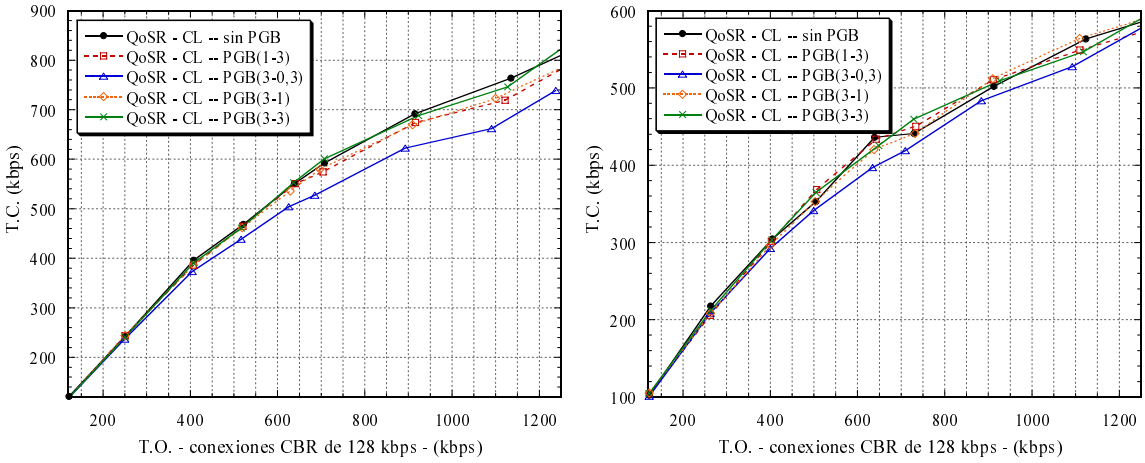
(a) 2 Km/h (b) 9 Km/h



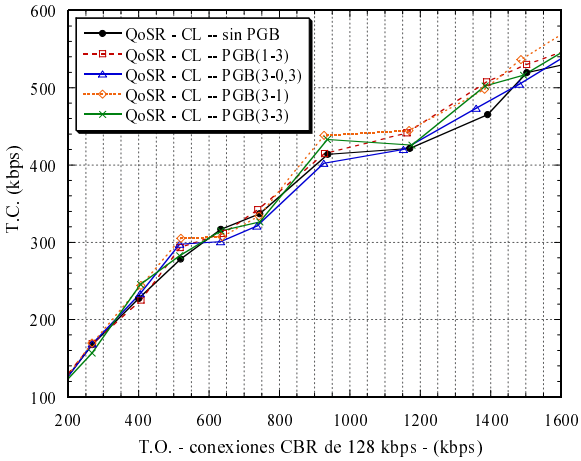
(c) 20 Km/h

Figura 4.39. Probabilidad de bloqueo vs. tráfico ofrecido. Estrategia QoSR-CL. Evaluación del mecanismo de reasignación (*piggy-backing*). Topologías de 50 nodos. Escenarios de 2, 9 y 20 Km/h. – notación: PGB($n_{failupd}$-t_{update}).

Cuando la movilidad del escenario es reducida (2 Km/h), la probabilidad de perder recursos debido a la interferencia variante o a la rotura de los caminos establecidos no es excesivamente elevada, de manera que la reasignación flexible de recursos no supone realmente una ventaja. Por el contrario, la reducción de capacidad debida al tráfico de monitorización y el consiguiente aumento de bloqueo suponen una degradación considerable de las prestaciones, como se observa en la figura 4.40(a). El uso de configuraciones con una frecuencia de actualización menor (t_{update} = 3 seg.), tampoco supone una mejora significativa de las prestaciones, a pesar del limitado *overhead* que permite obtener una probabilidad de bloqueo semejante al caso de no considerar monitorización.



(a) 2 Km/h (b) 9 Km/h

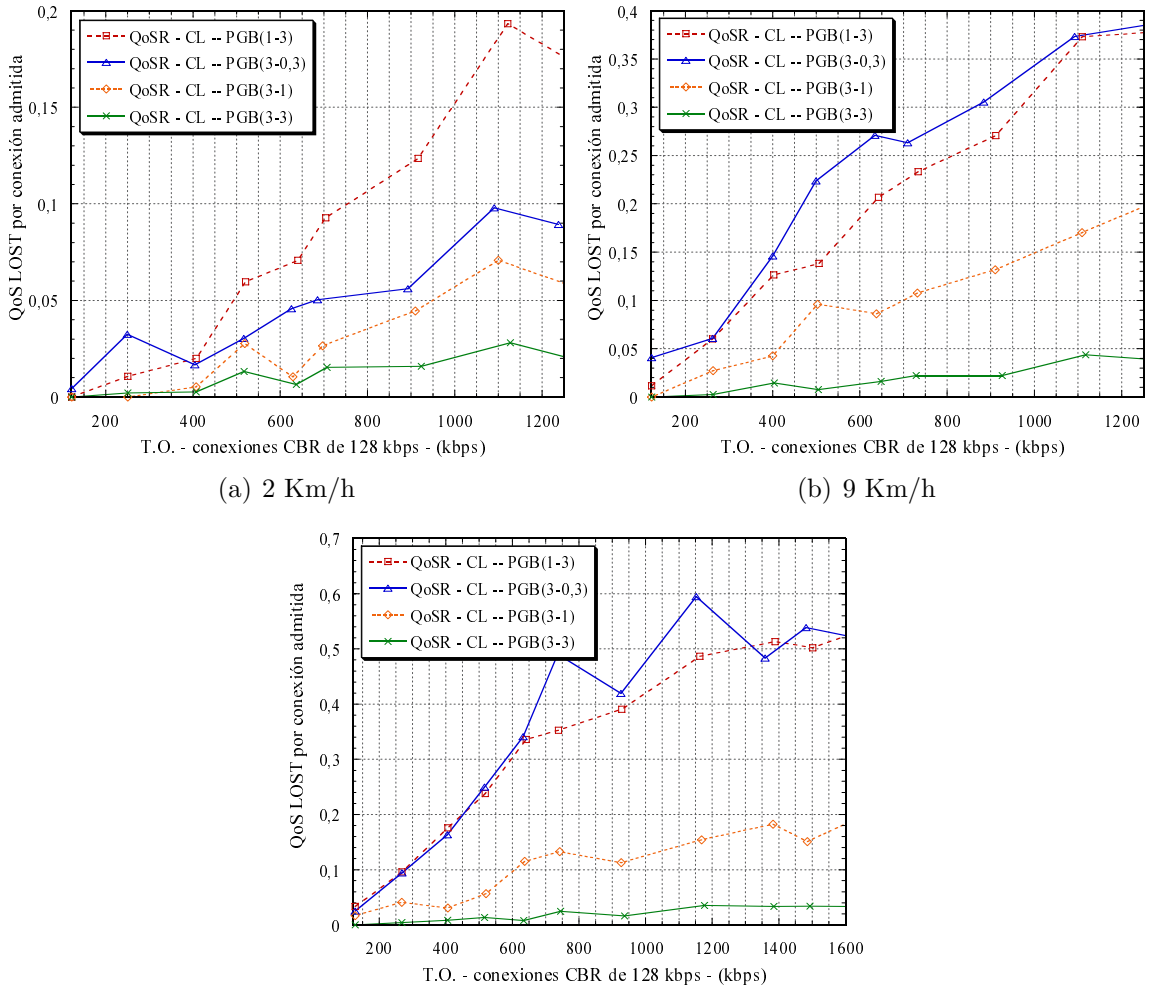


(c) 20 Km/h

Figura 4.40. Tráfico correctamente cursado vs. tráfico ofrecido. Estrategia QoSR-CL. Evaluación del mecanismo de reasignación (*piggy-backing*). Escenarios de 2, 9 y 20 Km/h. – notación: PGB($n_{failupd}$-t_{update}).

En el caso de configurar un intervalo de actualización mayor, resulta igualmente determinante el valor $n_{failupd}$ establecido. Dicho valor se corresponde con el número de actualizaciones fallidas necesarias para activar un nuevo procedimiento de búsqueda ($QLOST$). Si se analiza precisamente el consiguiente mecanismo de readmisión, se observa una diferencia significativa entre las configuraciones PGB(1-3) y PGB(3-3), ambas con un intervalo de actualización de 3 seg. La primera de ellas genera un número de intentos de reasignación mucho mayor – figura 4.41(a) – lo que puede implicar una falsa detección de situaciones de pérdida de la QoS. En consecuencia, el tráfico de encaminamiento correspondiente a las nuevas búsquedas de ruta así como la identificación del

camino virtual asociado como camino *best-effort* (ruta QBE – §4.2.2) suponen tanto un aumento en el bloqueo de nuevas conexiones, como una degradación innecesaria de las conexiones admitidas.



(a) 2 Km/h

(b) 9 Km/h

(c) 20 Km/h

Figura 4.41. Pérdidas de la QoS identificadas mediante monitorización vs. tráfico ofrecido. Estrategia QoSR-CL. Escenarios de 2, 9 y 20 Km/h. – notación: PGB($n_{failupd}$-t_{update}).

Sin embargo, en escenarios de mayor movilidad, si bien el *overhead* de control es igualmente considerable, como se observa en la figura 4.39, la aplicación del mecanismo de monitorización consigue, desde una doble perspectiva, mejorar las prestaciones del sistema. Por una parte, el aumento en el bloqueo de nuevas conexiones permite disponer de mayor flexibilidad a la hora de reasignar recursos para las conexiones admitidas, especialmente en el caso de ser necesario la búsqueda de nuevos caminos, dado que

un menor número de conexiones puede ser más fácilmente redistribuido dentro de la red que, además, no está congestionada. Por otra parte, la actualización de recursos en el propio camino, realizada durante el proceso de monitorización, permite recuperar ancho de banda cuando la variación en la interferencia del entorno degrada los slots inicialmente seleccionados, reduciendo así el número de retransmisiones necesarias, el retardo de los paquetes y consecuentemente su descarte.

Los resultados mostrados en las figuras 4.41(b) y 4.41(c) muestran la similitud en el número de pérdidas identificadas mediante la configuración PGB(1-3) y PGB(3-0,3), a diferencia de lo observado en la figura 4.41(a). Mientras una reacción precipitada en el caso de escenarios de reducida movilidad supone una falsa detección de pérdida de ancho de banda, cuando la movilidad supone una pérdida real de los recursos, la rápida detección es mucho más próxima a la realidad. En dicha comparación, no obstante, la configuración PGB(3-0,3) sigue considerándose más apropiada teniendo en cuenta que una identificación errónea por ser precipitada puede ser mucho más perjudicial. Sin embargo, observando los resultados obtenidos – figura 4.40 – el *overhead* introducido por esta configuración sigue suponiendo una degradación de las prestaciones. En el caso de un escenario de 20 Km/h, el tráfico cursado no disminuye – figura 4.40(c), a pesar de la mayor probabilidad de bloqueo – figura 4.39(c). Sin embargo, la configuración PGB(3-1) supone no sólo una probabilidad de bloqueo comparable sino una mejora en el tráfico cursado, como refleja la figura 4.40(c). Teniendo en cuenta la significativa probabilidad de rotura, como se observó en la figura 4.35(c), y la constante pérdida de recursos que supone un grado de movilidad elevado, reflejada en la degradación del *throughput* experimentado – figura 4.36(c), la mejora obtenida demuestra la respuesta al compromiso *overhead*-actualización proporcionada por esta configuración (t_{update} moderado, de 1 seg.).

Si bien los resultados preliminares parecen indicar la conveniencia de una configuración de monitorización PGB(3-1) únicamente ventajosa en el caso de escenarios de elevada movilidad, no se ha considerado descartable el esquema PGB(3-0,3). Como ha podido observarse en la figura 4.40, la degradación sobre las prestaciones de tráfico correctamente cursado disminuye conforme aumenta el grado de movilidad. En determinadas circunstancias, podría potencialmente llegarse a compensar el efecto de la carga de control asociada. La implementación exclusiva del procedimiento de monitorización supone una herramienta insuficiente para la adaptación de la propuesta a la movilidad del entorno. Como se describió en §4.2.2, el mecanismo completo de readmisión incluye la capacidad de decidir la expulsión del sistema de conexiones cuya QoS se considera

irrecuperable. Esta decisión depende directamente de la configuración de monitorización seleccionada, puesto que ésta determina cuándo se percibe dicha situación de elevada degradación. Por otra parte, la expulsión de conexiones del sistema condiciona el bloqueo de futuras admisiones al modificar las condiciones de tráfico en la red. Al existir un mayor número de parámetros influyentes, la supuesta desventaja del *overhead* asociado a la información de monitorización no resulta evidente. Por lo tanto, a la hora de evaluar el sistema completo se ha analizado la parametrización conjunta tanto de la monitorización como del mecanismo de expulsión.

Efecto de la expulsión del sistema de conexiones degradadas

A continuación se muestra un conjunto de resultados representativos extraídos de los múltiples casos analizados. Se ha considerado la variación de parámetros mostrada en la tabla 4.5, si bien finalmente, de acuerdo a las observaciones realizadas, se han tomado como configuraciones características las destacadas en dicha tabla.

Tabla 4.5. Mecanismo completo de readmisión

	Parámetros de *piggy-backing*		Critero de expulsión
	$n_{failupd}$	t_{update} (seg.)	$t_{qlost,max}$ (seg.)[a]
D3-PGB(3-0,3)	3	0,3	3
D5-PGB(3-0,3)	3	0,3	5
D10-PGB(3-0,3)	3	0,3	10
D5-PGB(3-1)	3	1	5
D10-PGB(3-1)	3	1	10

[a] $t_{noupd} = 2 \cdot n_{failupd} \cdot t_{update}$.

El valor $t_{qlost,max}$, como se describió en §4.2.2, representa el tiempo máximo permitido en estado de QoS insuficiente, antes de considerar que la aplicación asociada no puede mantener la QoS demandada. Debido a los continuos cambios en la topología que supone la movilidad de los terminales, las condiciones de admisión pueden cambiar radicalmente, llegando a resultar imposible la readmisión adecuada de las conexiones afectadas. La expulsión de dichas conexiones pretende garantizar la liberación de recursos, reduciendo la continua competencia con objeto de permitir nuevas admisiones y la satisfacción de sus demandas. La principal dificultad a la hora de seleccionar el valor adecuado de este parámetro estriba en la interpretación de la QoS realizada desde el punto de vista de la red. La medida de la QoS obtenida se basa en la estimación del ancho de banda extremo a extremo en el circuito virtual establecido. La exactitud de

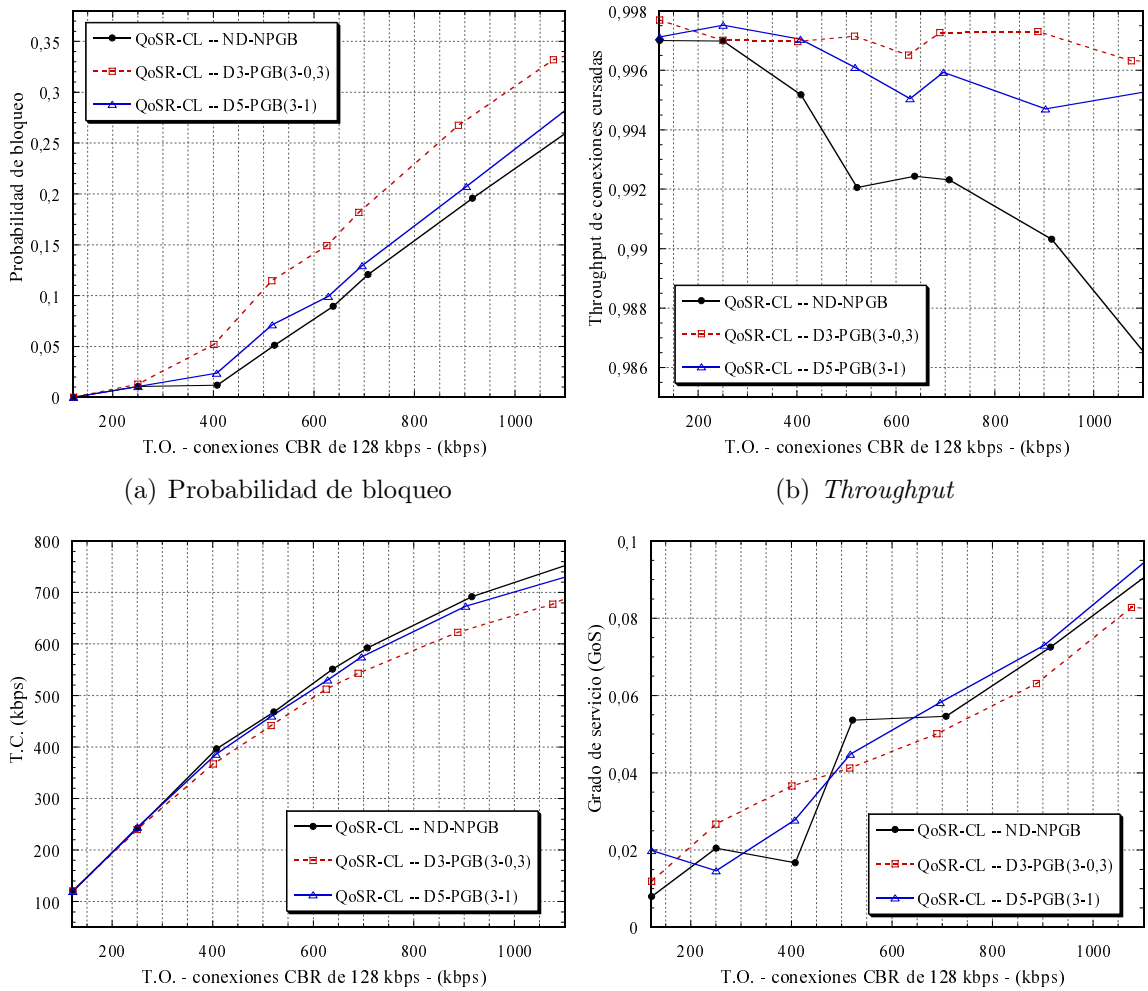
dicha medida depende de diversos factores. Por una parte, la propia red es incapaz de medir de una manera completamente exacta el ancho de banda durante toda la duración de la conexión. Si bien la pérdida total de recursos durante las roturas supone un ancho de banda nulo, los posibles fallos en el procedimiento de monitorización o el ancho de banda desconocido durante la readmisión por $QLOST$ (ruta QBE) implican la indeterminación en dicha medida. Por otra parte, un mecanismo eficiente de expulsión acorde con la satisfacción de la QoS demandada por los usuarios supone la necesidad de una correcta interrelación entre las aplicaciones reales y la red (*Service Level Agreement* (SLA)).

Como ha podido observarse en las simulaciones realizadas, una tolerancia excesiva a pérdidas de ancho de banda, reflejada en un $t_{qlost,max}$ elevado, puede llegar a suponer grandes pérdidas de paquetes, como es el caso de 10 seg. en comparación con la longitud de las aplicaciones consideradas (CBR de 50 seg. de duración). Un tiempo reducido, sin embargo, supone la necesidad de realizar una monitorización frecuente, con objeto de disponer de una medición suficientemente actualizada del ancho de banda proporcionado, así como de facilitar una rápida identificación de la posible degradación sufrida. Por otra parte, con la intención de evitar expulsiones innecesarias, el mecanismo de reacción debe evitar las falsas identificaciones. Un $t_{qlost,max}$ reducido, acompañado de una monitorización comparativamente poco frecuente sería contraproducente. Este sería el caso, por ejemplo, de un $t_{qlost,max} = 3$ seg. con una monitorización PGB(3-1), dado que el tiempo total de reacción de éste es igualmente 3 seg., de manera que no existirían oportunidades reales de readmitir la conexión. Teniendo en cuenta estas consideraciones, con las dos configuraciones de monitorización utilizadas se han obtenido unos resultados sensiblemente mejores con la tolerancia $t_{qlost,max}$ más restrictiva: PGB(3-1) – 5 seg. (D5-PGB(3-1)), y PGB(3-0,3) – 3 seg (D3-PGB(3-0,3))[20].

En escenarios con una movilidad media poco representativa (2 Km/h) los resultados han seguido demostrando la ineficacia de un procedimiento de monitorización, y en este caso también de expulsión, en comparación con el funcionamiento más básico de la propuesta. Debido al mayor bloqueo experimentado – figura 4.42(a) – y a pesar del ligero incremento en *throughput* medio de las conexiones cursadas – figura 4.42(b) – teniendo en cuenta las buenas prestaciones obtenidas en ausencia de adaptación, el tráfico finalmente cursado realmente disminuye, como se observa en la figura 4.42(c). De hecho, la ligera ganancia observada en el grado de servicio para altas tasas de

[20]El funcionamiento básico, sin configuración de monitorización ni expulsión se ha denominado en los resultados ND-NPGB.

tráfico ofrecido – figura 4.42(d) – no supone una clara ventaja teniendo en cuenta la disminución de la capacidad efectiva del sistema.



(a) Probabilidad de bloqueo

(b) *Throughput*

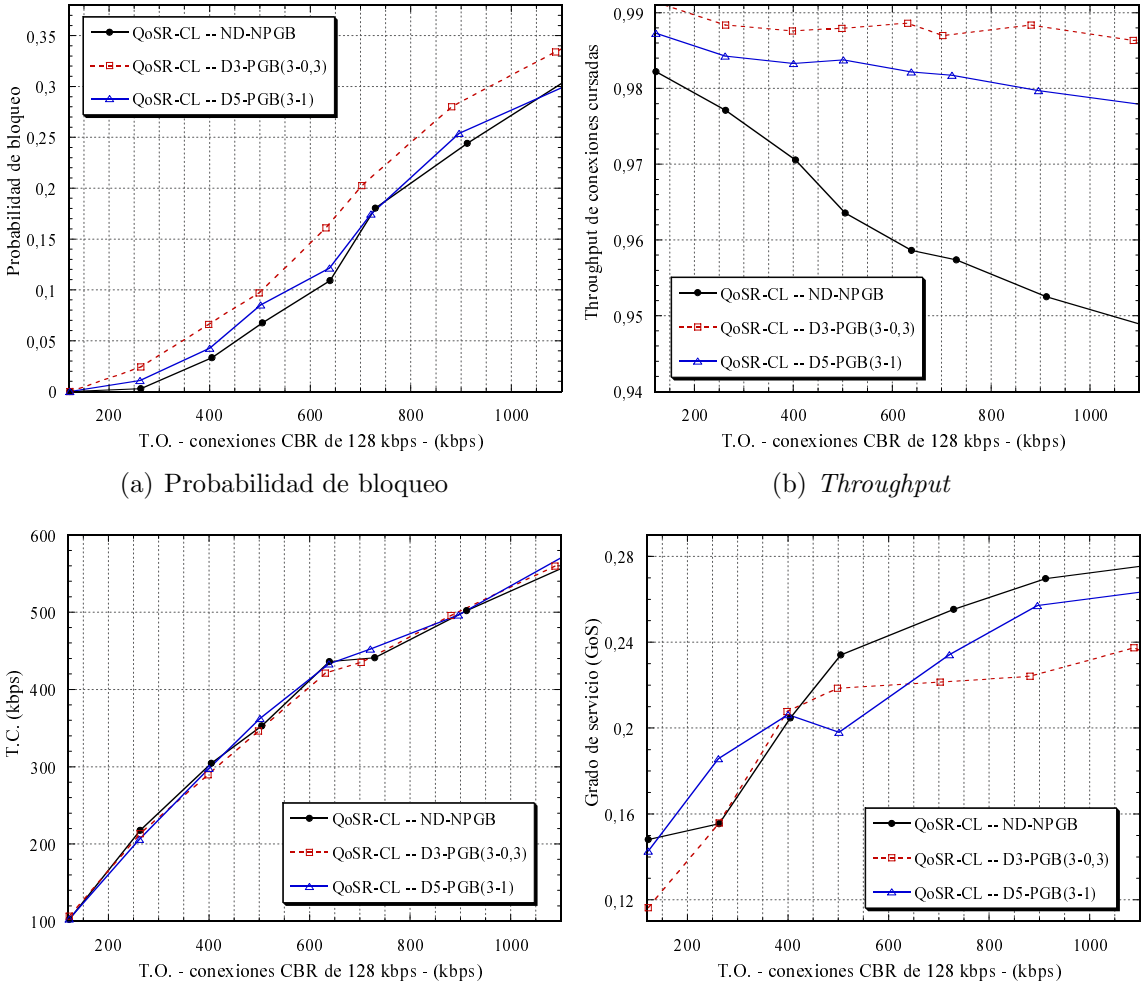
(c) Tráfico cursado

(d) Grado de servicio

Figura 4.42. Prestaciones de la estrategia QoSR-CL. Mecanismo completo de readmisión: monitorización y expulsión. Topología de 50 nodos con velocidad media de 2 Km/h.

Cuando la movilidad en el entorno comienza a ser considerable (9 Km/h), la ganancia obtenida gracias a la expulsión de conexiones degradadas supera la desventaja del *overhead* de control generado. Las dos configuraciones mostradas en la figura 4.43 permiten mejorar las prestaciones obtenidas sin el mecanismo de readmisión. Una configuración de monitorización menos frecuente permite mantener una probabilidad del bloqueo análoga, como se observa en la figura 4.43(a). Sin embargo, una rápida actualización de los recursos, así como una mayor expulsión de conexiones, suponen al final

unas prestaciones medias superiores, como se observa en el *throughput* experimentado por las conexiones cursadas – figura 4.43(b). Ambas configuraciones proporcionan un tráfico correctamente cursado equivalente, sin una pérdida de capacidad efectiva respecto a la situación inicial, como se refleja en la figura 4.43(c). Teniendo en cuenta el menor bloqueo conseguido con la configuración D5-PGB(3-1), para el mismo tráfico correctamente cursado, esta configuración podría considerarse más ventajosa. Sin embargo, si ponderamos realmente con una importancia mucho mayor la degradación experimentada por las conexiones cursadas, gracias a una configuración D3-PGB(3-0,3) se consigue mejorar el grado de servicio, como refleja la figura 4.43(d), especialmente en situaciones de alta carga.



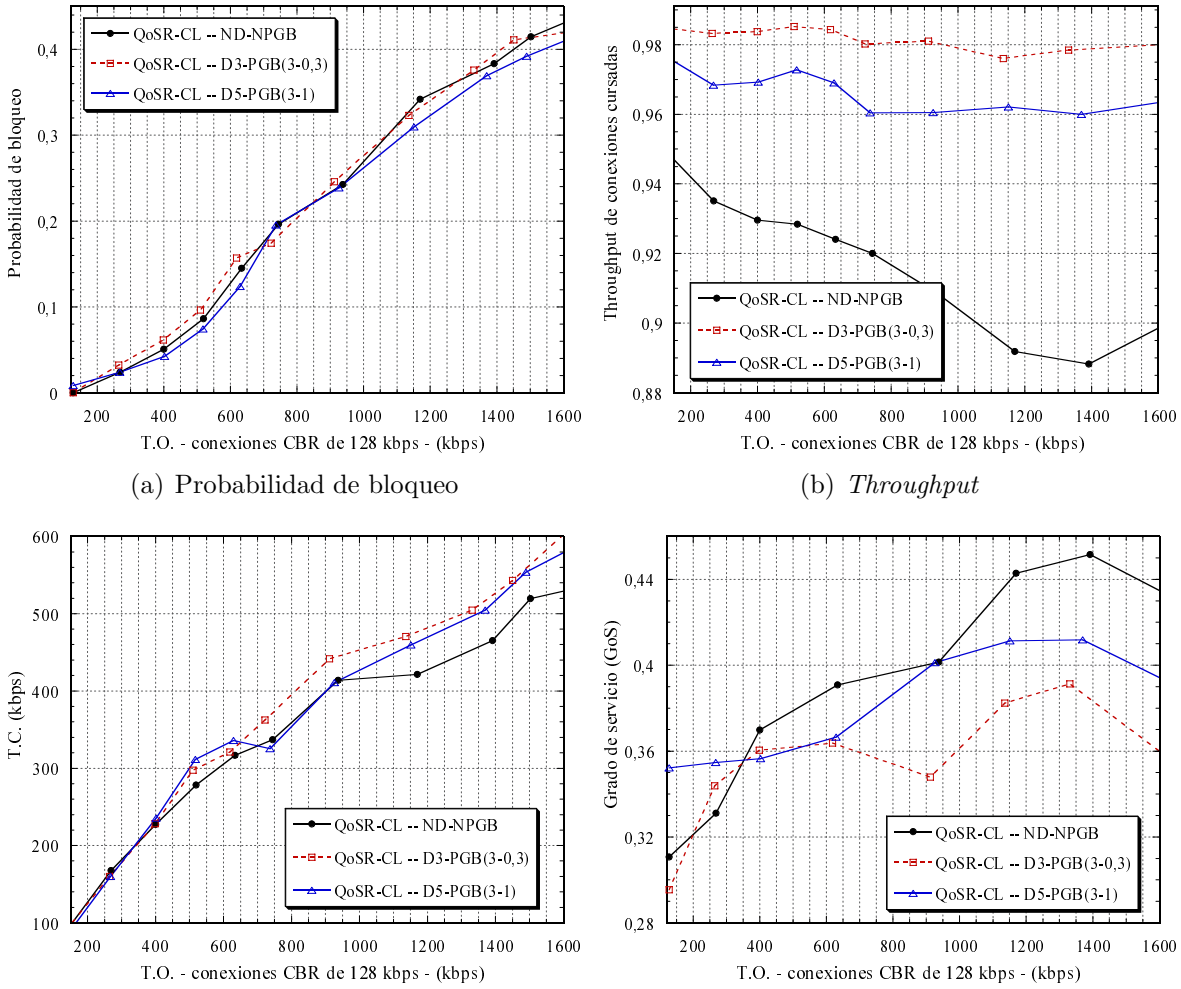
(a) Probabilidad de bloqueo

(b) *Throughput*

(c) Tráfico cursado

(d) Grado de servicio

Figura 4.43. Prestaciones de la estrategia QoSR-CL. Mecanismo completo de readmisión: monitorización y expulsión. Topología de 50 nodos con velocidad media de 9 Km/h.



(a) Probabilidad de bloqueo

(b) *Throughput*

(c) Tráfico cursado

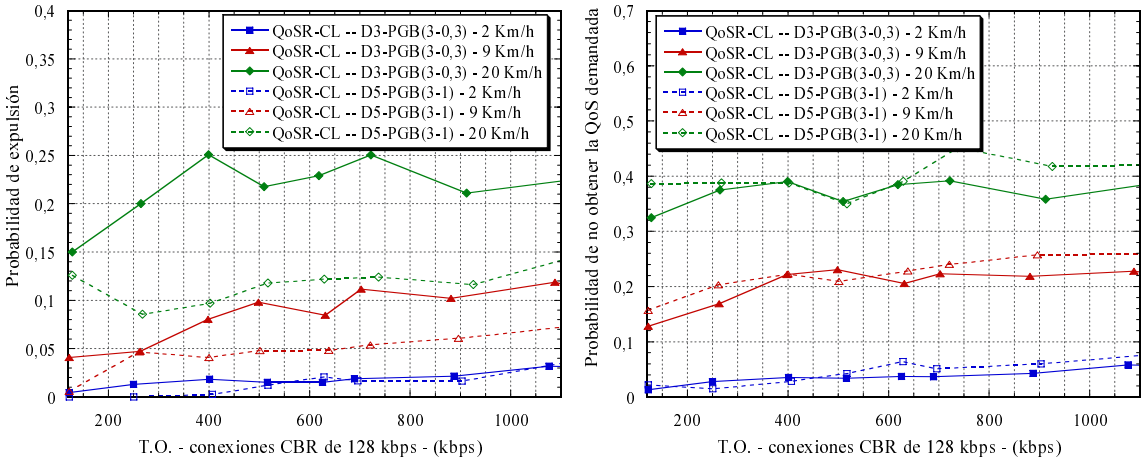
(d) Grado de servicio

Figura 4.44. Prestaciones de la estrategia QoSR-CL. Mecanismo completo de readmisión: monitorización y expulsión. Topología de 50 nodos con velocidad media de 20 Km/h.

Los resultados observados en un escenario de 20 Km/h ponen de manifiesto la clara ventaja del mecanismo completo de readmisión/expulsión. El *overhead* introducido (en las dos configuraciones analizadas) se compensa gracias a la mejor redistribución de las conexiones admitidas y, especialmente, a la expulsión de conexiones degradadas que permite la liberación de recursos ocupados ineficientemente. Una mayor posibilidad de admitir nuevas conexiones genera finalmente un bloqueo similiar – figura 4.44(a). En el caso de la configuración D3-PGB(3-0,3), el *throughput* medio, como se aprecia en la figura 4.44(b), es sensiblemente más elevado, lo que constata la efectiva expulsión de conexiones. En este caso, gracias a la mejor utilización de los recursos y a la reducción de continuas colisiones entre conexiones que se mantendrían en continua competencia

debido a las múltiples búsquedas de ruta, el tráfico correctamente cursado supera incluso al obtenido originalmente, como demuestra la figura 4.44(c). En consecuencia, considerando además la mejora obtenida en el grado de servicio – figura 4.44(d) – la configuración D3-PGB(3-0,3), especialmente apropiada gracias a la rápida reacción del protocolo, mejora globalmente las prestaciones del sistema.

El beneficio observado en escenarios de movilidad gracias al empleo de una estrategia de expulsión apropiada y la mala adecuación de ésta a un escenario quasi-estático pueden entenderse mejor a partir de los resultados reflejados en la figura 4.45.



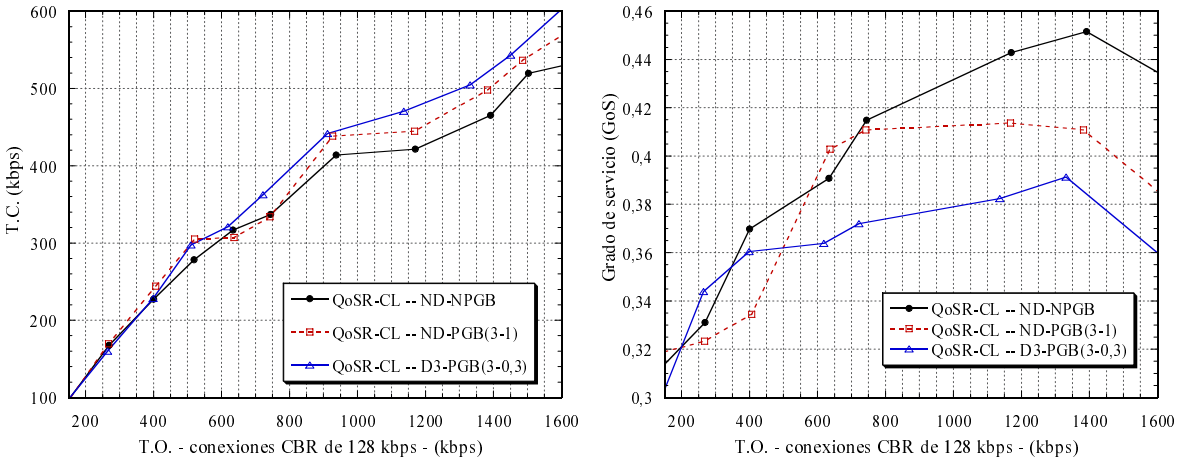
(a) Probabilidad de expulsión, P_d (b) Probabilidad de no obtener la QoS, P_f

Figura 4.45. Análisis del mecanismo de expulsión sobre la estrategia QoSR-CL. Escenarios de 2, 9 y 20 Km/h.

La probabilidad de expulsión P_d, como se muestra en la figura 4.45(a), sólo tiene una representatividad importante en el caso de escenarios de 9 y 20 Km/h. Sin embargo, cuando la movilidad media se encuentra entorno a los 2 Km/h, el ratio de conexiones expulsadas es sensiblemente menor, de manera que la influencia de dicha eliminación en las prestaciones globales del sistema no es realmente significativa. En el caso de escenarios móviles, una reacción más rápida por parte del sistema (estrategia D3-PGB(3-0,3)) incrementa la probabilidad de expulsión, lo que, a priori, podría considerarse contraproducente en los clásicos términos de GoS[21]. Sin embargo, como ya se ha explicado anteriormente, de acuerdo a la valoración realizada en el contexto ad hoc (§4.2.2) la probabilidad P_f, considerada con una ponderación mayor en el GoS, se ha definido realmente teniendo en cuenta no sólo la expulsión real sino la probabilidad de

[21] P_d equivalente a la P_h: probabilidad de bloquear un *handoff* en redes celulares, (4.9).

no experimentar la QoS demandada (4.10). La liberación de recursos realizada gracias a una adecuada expulsión de las conexiones que no pueden readmitirse con garantías en cuanto a satisfacción de la QoS permite reducir dicha probabilidad, como se muestra en la figura 4.45(b). Como consecuencia, como se observó en los resultados previos, se mantiene la capacidad efectiva – figura 4.43(c) – o incluso se obtiene cierta ganancia en las prestaciones – figura 4.44(c). A consecuencia de la reducción demostrada sobre la P_f, el GoS mejora en cualquiera de los escenarios, como pudo comprobarse en las figuras 4.42(d), 4.43(d) y 4.44(d). Dicha ganancia, sin embargo, no es tal en el escenario de 2 Km/h, debido al incremento sobre la P_b resultado del *overhead* introducido que, a diferencia de los escenarios de alta movilidad, no se ve compensado en términos de garantías de la QoS, resultando en una disminución sobre la capacidad efectiva del sistema, como se observó en la figura 4.42(c).



(a) Tráfico cursado (b) Grado de servicio

Figura 4.46. Comparativa entre monitorización aislada y readmisión completa. Estrategia QoSR-CL Topología de 50 nodos con velocidad media de 20 Km/h.

Los resultados obtenidos han mostrado la capacidad conjunta del mecanismo de monitorización y expulsión en condiciones de movilidad. Sin embargo, la expulsión de conexiones activas puede parecer, desde una perspectiva conservadora, una garantía limitada de provisión de QoS. La monitorización en sí misma, en condiciones de elevada movilidad (20 Km/h) ya pareció proporcionar una mejora considerable en las prestaciones del sistema, gracias a la flexibilización en la utilización de los recursos, como se mostró en la figura 4.40(c). Sin embargo, como se observa en los resultados comparados entre un mecanismo de monitorización aislado y el procedimiento conjunto de readmisión/expulsión, mostrado en la figura 4.46, la capacidad efectiva conseguida

aumenta gracias a la estrategia completa de readmisión, así como se consigue mejorar el grado de servicio. Es particularmente destacable la ganancia conseguida con la configuración PGB(3-0,3) gracias a la inclusión del procedimiento de expulsión, ya que permite superar las prestaciones, inicialmente mejores, de la configuración PGB(3-1) sin expulsión.

Teniendo en cuenta la naturaleza variante inherente a las redes móviles ad hoc, pretender garantizar de una manera continuada la QoS sin pérdidas asumibles resulta extremadamente complicado, si no imposible. Desde esta perspectiva, mantener en activo conexiones cuya satisfacción no puede garantizarse no se ha considerado un planteamiento de gestión eficiente a nivel de sistema. Una negociación adecuada entre las aplicaciones y la red permitiría, finalmente, adecuar éstas a la naturaleza variante del entorno, incluso modulando las prestaciones ofrecidas sobre las conexiones activas (priorización de servicios). Permitir una implementación más flexible de este proceso de negociación implica, como primera aproximación, facilitar la monitorización y expulsión de conexiones desde el punto de vista de la propia red, lo que se consigue gracias a las estrategias propuestas.

Dependencia de la configuración respecto a las aplicaciones específicas

Un análisis más exhaustivo, realizado con objeto de establecer una configuración de parámetros adecuada, ha mostrado la gran dependencia de ésta con la naturaleza de las aplicaciones específicas. Concretamente, como ya se comentó anteriormente, la medida de la tolerancia máxima ante la pérdida de la QoS depende directamente de la caracterización de dicha QoS. La definición del $t_{qlost,max}$ se relaciona directamente con la P_{loss} considerada (3 %) así como la duración de las aplicaciones. La idoneidad de la configuración D3-PGB(3-0,3) para conexiones CBR de 50 seg. puede no ser tal en otras circunstancias. La evaluación de un mayor número de resultados ha permitido concluir la estrecha relación entre el $t_{qlost,max}$ y la longitud de la conexión. Una QoS basada en el porcentaje de paquetes perdidos implica necesariamente cierta indefinición dada la relatividad de dicho parámetro. La figura 4.47 muestra un ejemplo del tráfico correctamente cursado, para conexiones CBR de 200 seg., donde la configuración observada como más apropiada es en este caso D5-PGB(3-1). Una tolerancia mayor (5 seg.) permite garantizar el porcentaje de paquetes perdidos, generando además una carga de control asociada menor, lo cual permite mejorar la capacidad efectiva global. Por otra parte, la prematura reacción del sistema ante pérdidas de la QoS que supone la estrategia D3-PGB(3-0,3) puede generar un número elevado de falsas detecciones. En condiciones de elevada carga de tráfico, el beneficio obtenido gracias a la expulsión de

conexiones, dada la reducción en la ocupación del sistema que se consigue, permite una mayor flexibilidad en la gestión de recursos que redunda en el aumento de capacidad efectiva observado. Sin embargo, una expulsión excesiva de conexiones en una situación de baja congestión, cuando las conexiones afectadas podrían ser readmitidas adecuadamente, se traduce en una degradación significativa de las prestaciones obtenidas.

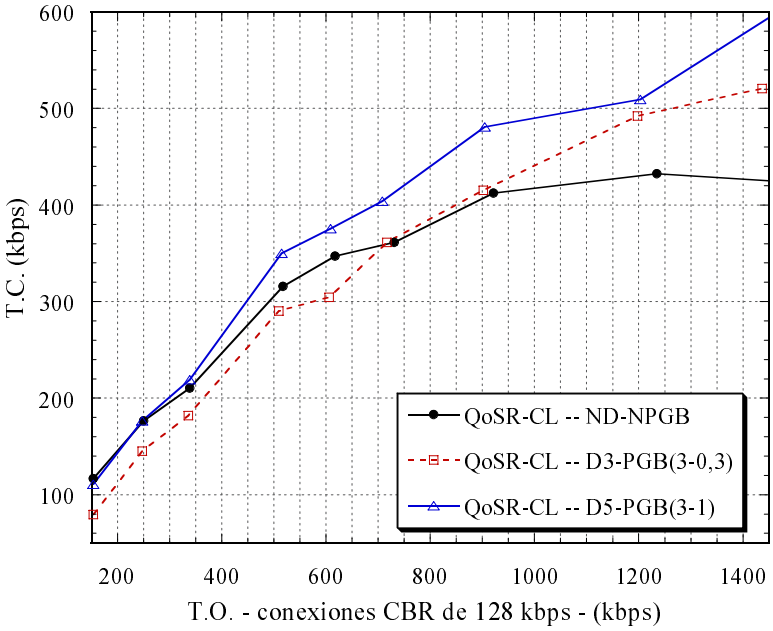


Figura 4.47. Tráfico correctamente cursado vs. tráfico ofrecido (CBR de 128 kbps). Estrategia QoSR-CL. Mecanismo completo de readmisión: monitorización y expulsión. Topología de 50 nodos con velocidad media de 9 Km/h. Conexiones de 200 seg de duración.

La principal conclusión de tan estrecha dependencia entre los parámetros a nivel de red y la QoS observada desde la aplicación es la necesidad de incluir una interrelación más avanzada en la arquitectura *cross-layer* propuesta. Puesto que el objetivo perseguido con el trabajo desarrollado se ha centrado en la cooperación entre encaminamiento y nivel de enlace, no se ha profundizado en la evaluación a nivel de aplicación, si bien se ha considerado necesario remarcar su importancia. Asumiendo una correcta comunicación entre las aplicaciones y la arquitectura propuesta, el valor finalmente apropiado del $t_{qlost,max}$ se ha considerado un parámetro resultante de la asociación aplicación - clase de servicio conocido a nivel de red como parte de la demanda de QoS realizada (ancho de banda). En cualquier caso, los resultados constatan la necesidad de una correcta identificación entre las necesidades de la aplicación y los requerimientos a nivel de red.

Como ha podido observarse en los resultados presentados, la respuesta del sistema debe adecuarse al dinamismo del entorno, siendo especialmente importante establecer

una configuración adaptada tanto a la naturaleza de las aplicaciones como al grado de movilidad. El análisis sobre los diversos escenarios propuestos permite concluir la necesidad de diferenciar dos modos de operación básicos:

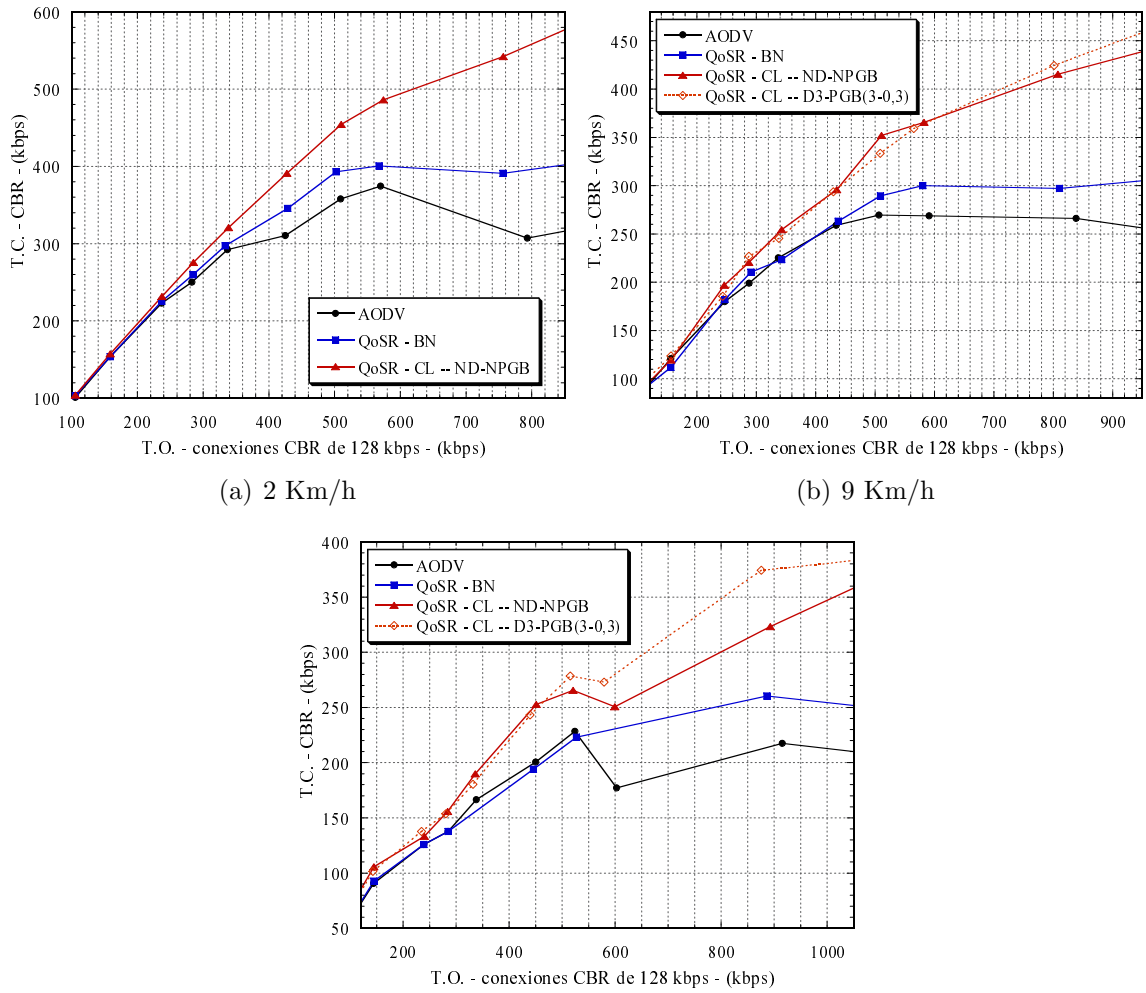
- **Funcionamiento en entorno quasi-estático (movilidad reducida)** – La estabilidad de los circuitos virtuales reservados en el procedimiento de admisión inicial es suficiente para obtener unas prestaciones cercanas a las obtenidas en un escenario completamente estático. Si bien puede existir cierto grado de degradación, resultado tanto de una interferencia variante como de la rotura esporádica de conexiones, el *overhead* introducido por el mecanismo de monitorización-expulsión puede limitar en exceso la capacidad efectiva del sistema. Así pues, la configuración apropiada para este entorno se basa en el funcionamiento más básico de la propuesta: **ND-NPGB**.

- **Funcionamiento en movilidad** – La continua variabilidad en las condiciones iniciales de admisión, resultante en una elevada probabilidad tanto de rotura como de pérdida de los recursos, requiere una redistribución permanente de las conexiones admitidas. El mecanismo de actualización-monitorización permite flexibilizar la utilización de los recursos así como reaccionar con prontitud a la hora de activar nuevas búsquedas. Por otra parte, la liberación de recursos ineficientemente ocupados, facilitada por la expulsión de conexiones degradadas, permite incrementar la admisión de nuevas conexiones compensando la ocupación adicional que supone el *overhead* de control. En vista de los resultados obtenidos, la configuración **D3-PGB(3-0,3)** se ha considerado la más adecuada en las condiciones de estudio consideradas (aplicaciones CBR de 128 kbps, con una duración constante de 50 seg.).

Evaluación en escenario multiservicio

Una vez extraídos los parámetros adecuados para los dos modos de funcionamiento finalmente propuestos, a continuación se muestra la evaluación del sistema realizada en un escenario multiservicio. Al igual que en las evaluaciones previas, se ha considerado un tráfico ofrecido total compuesto en igual proporción de aplicaciones CBR de 128 kbps (flujos QoS) y sesiones *best-effort* con estos parámetros: distribución geométrica con media 50 paquetes, tasa de 64 kbps, tiempo medio entre paquetes e intervalo entre llamadas exponenciales de medias 62.5 ms y 5 seg. respectivamente.

La mayor capacidad obtenida gracias a la correcta admisión y la diferenciación de servicios se mantiene a pesar de la variabilidad del entorno, como muestra la figura 4.48.



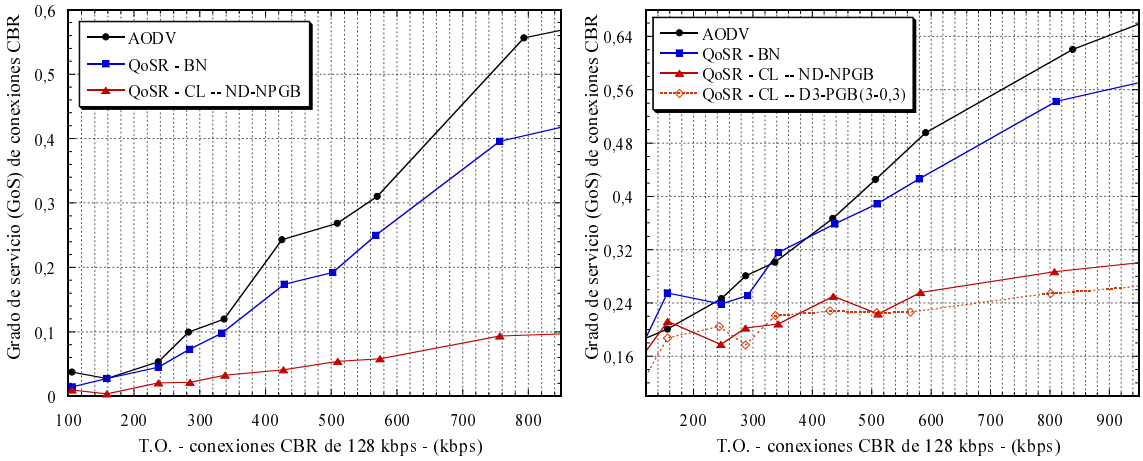
(a) 2 Km/h

(b) 9 Km/h

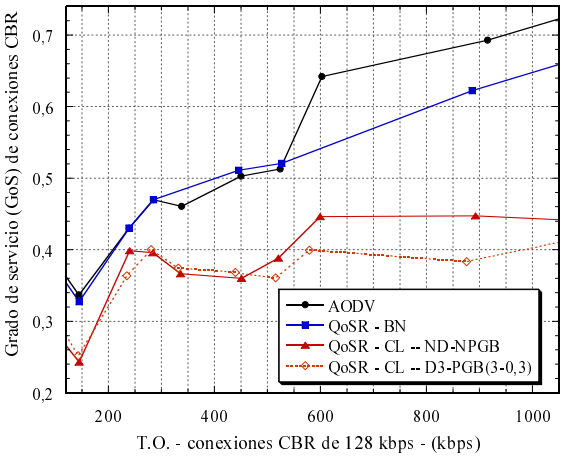
(c) 20 Km/h

Figura 4.48. Tráfico CBR correctamente cursado vs. tráfico ofrecido en escenario multiservicio (50 % CBR de 128 kbps y 50 % de tráfico *best-effort*). Configuraciones de readmisión básica ND-NPGB y completa D3-PGB(3-0,3). Topologías de 50 nodos. Escenarios de 2, 9 y 20 Km/h.

El modo de operación D3-PGB(3-0,3) sigue resultando beneficioso en entornos de mayor movilidad, no sólo igualando la capacidad efectiva del sistema sino incluso aumentando ésta en escenarios con un grado de movilidad medio de 20 Km/h – figura 4.48(c). El resultado obtenido en términos de grado de servicio constata la mejora global obtenida, como se observa en la figura 4.49. A pesar de la coexistencia de tráfico *best-effort*, la competencia por los recursos sigue favoreciendo a los servicios CBR, incluso en situaciones de readmisión. Gracias a la priorización flexible realizada en el nivel MAC (§4.2.3) las conexiones afectadas tanto por roturas como por pérdidas de ancho de banda son preferentes respecto a nuevas admisiones y los recursos ocupados menos prioritarios (sesiones BE) pueden serles reasignados.



(a) 2 Km/h (b) 9 Km/h

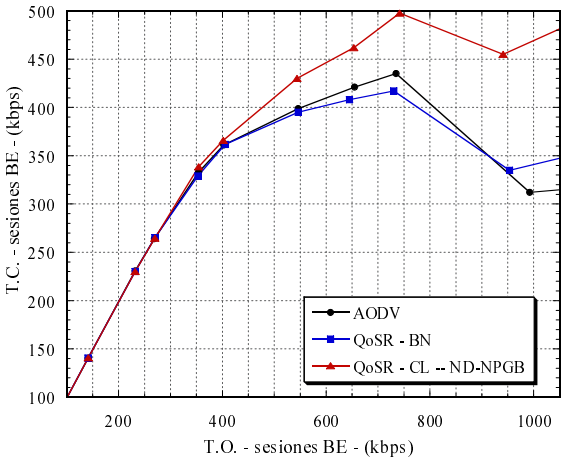


(c) 20 Km/h

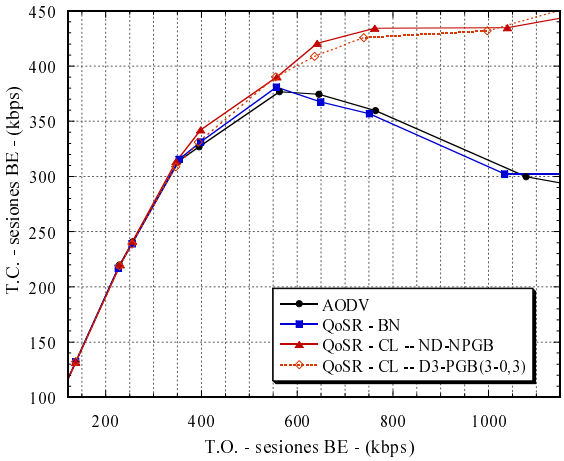
Figura 4.49. Grado de servicio de aplicaciones CBR vs. tráfico ofrecido en escenario multiservicio (50 % de ráfico *best-effort*). Configuraciones de readmisión básica ND-NPGB y completa D3-PGB(3-0,3). Topologías de 50 nodos. Escenarios de 2, 9 y 20 Km/h.

La reducción de la capacidad efectiva resultado del aumento de la movilidad afecta igualmente al tráfico cursado *best-effort*. Sin embargo, a pesar de la menor priorización en la asignación de recursos, la gestión eficiente realizada por la estrategia QoSR-CL permite seguir proporcionando mejores prestaciones que el encaminamiento QoSR-BN o el AODV sin admisión, como se demuestra en la figura 4.50. Las sesiones BE nunca son bloqueadas ni expulsadas del sistema salvo por desconexiones inevitables o excesiva congestión en la red. No obstante, su competencia con las aplicaciones CBR se limita a la utilización de los recursos sobrantes. Así pues, dado que el modo de operación D3-PGB(3-0,3) permite flexibilizar la utilización de los recursos, su funcionalidad en

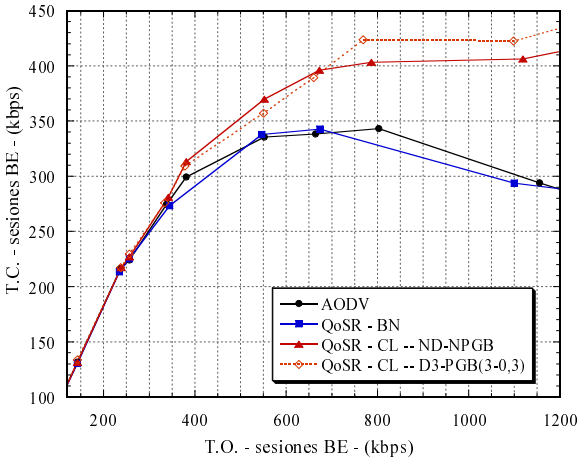
un escenario de alta movilidad permite igualmente un mejor aprovechamiento de dicho ancho de banda sobrante, gracias a lo cual incluso se incrementa la capacidad efectiva para cursar tráfico menos prioritario, como se observa en la figura 4.50(c).



(a) 2 Km/h



(b) 9 Km/h



(c) 20 Km/h

Figura 4.50. Tráfico *best-effort* correctamente cursado vs. tráfico ofrecido en escenario multiservicio (50 % CBR de 128 kbps y 50 % de tráfico *best-effort*). Configuraciones de readmisión básica ND-NPGB y completa D3-PGB(3-0,3). Topologías de 50 nodos. Escenarios de 2, 9 y 20 Km/h.

Efecto del ancho de banda de guarda ΔBW

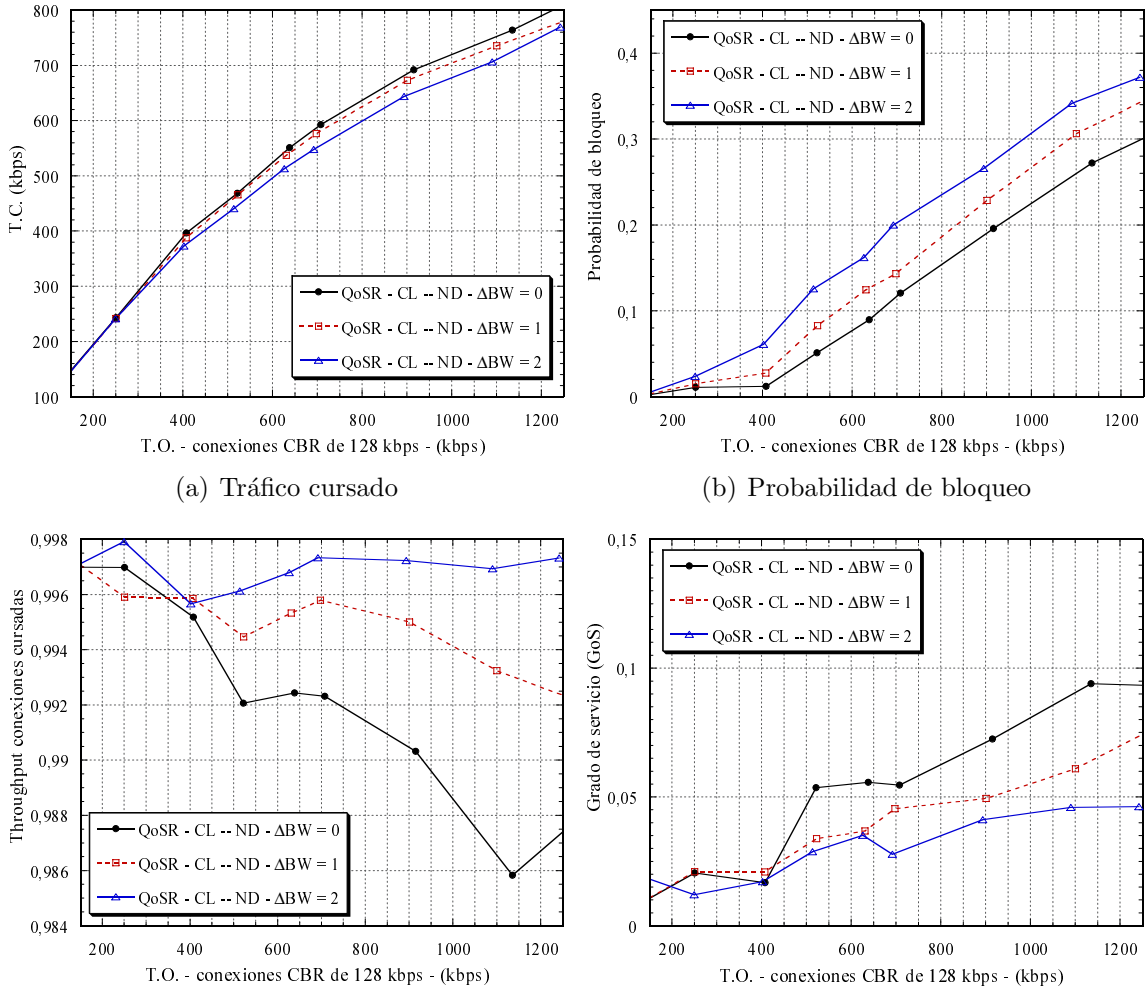
La expulsión de las conexiones cuya tolerancia a la pérdida de la QoS es sobrepasada supone un considerable perjuicio desde el punto de vista del usuario. Por lo tanto, desde una perspectiva de diseño del sistema, mejorar las prestaciones en relación a la QoS

proporcionada, pensando no sólo en la capacidad efectiva, implica establecer dentro de los objetivos clave la reducción de la probabilidad de expulsión. Como ya se explicó en §4.2.2, la probabilidad de fallo (P_f), significativamente ponderada en el GoS, incluye además las conexiones erróneamente mantenidas dentro del sistema, teniendo en cuenta el equivalente perjuicio que supone para el usuario. Por lo tanto, una reducción eficiente de esta probabilidad requiere no sólo minimizar la expulsión de conexiones sino asegurar además una clara identificación de éstas, manteniendo así las conexiones activadas en el sistema con garantías de QoS.

Como se describió en §4.2.2, la principal dificultad a la hora de mantener las prestaciones de una conexión admitida en el sistema estriba en la capacidad de readmitir adecuadamente la conexión cuando la elevada ocupación de recursos existente, unida a la variación en la topología de la red, dificulta al algoritmo de encaminamiento encontrar un nuevo camino virtual disponible. Asumiendo que una mayor disponibilidad global de recursos facilitaría dicha readmisión se propone la inclusión de un margen de recursos ΔBW en la admisión inicial que limita la aceptación de nuevas conexiones. Una implementación desarrollada desde una perspectiva global del sistema supondría la definición de un porcentaje genérico del ancho de banda total disponible reservado para su utilización en el momento de readmitir conexiones ya activas. Sin embargo, la naturaleza distribuida de las redes ad hoc y la complejidad que supone disponer de dicha "visión global" conducen a la búsqueda de soluciones aproximadas más apropiadas para la indeterminación del entorno. De este modo, se ha propuesto una consideración de ancho de banda adicional medido sobre la demanda individual de cada conexión. Obviamente, los recursos desocupados no tienen una asignación real con dichas conexiones. La dificultad en la definición del margen ΔBW reside, precisamente, en el compromiso entre un posible exceso de recursos inutilizados debido a su dependencia con la activación de conexiones y la flexibilidad de readmisión perseguida.

Los resultados obtenidos en los diversos escenarios demuestran nuevamente que las degradaciones sufridas debido a la interferencia variable y las roturas de rutas no suponen una pérdida de eficiencia muy significativa en un escenario de reducida movilidad, siendo en consecuencia innecesaria la aplicación de un ancho de banda de guarda. La capacidad efectiva se ve reducida debido a la infrautilización de los recursos, como se observa en la figura 4.51(a). Teniendo en cuenta que el modo de funcionamiento apropiado para un escenario de estas características no implementa la funcionalidad de expulsión, el aumento en la probabilidad de bloqueo – figura 4.51(b), acompañado de una mejora poco representativa del *throughput* – figura 4.51(c), sin efecto considerable sobre la

probabilidad P_f, supone una mejora poco significativa en el GoS – figura 4.51(d) que no compensa la reducción de capacidad efectiva.

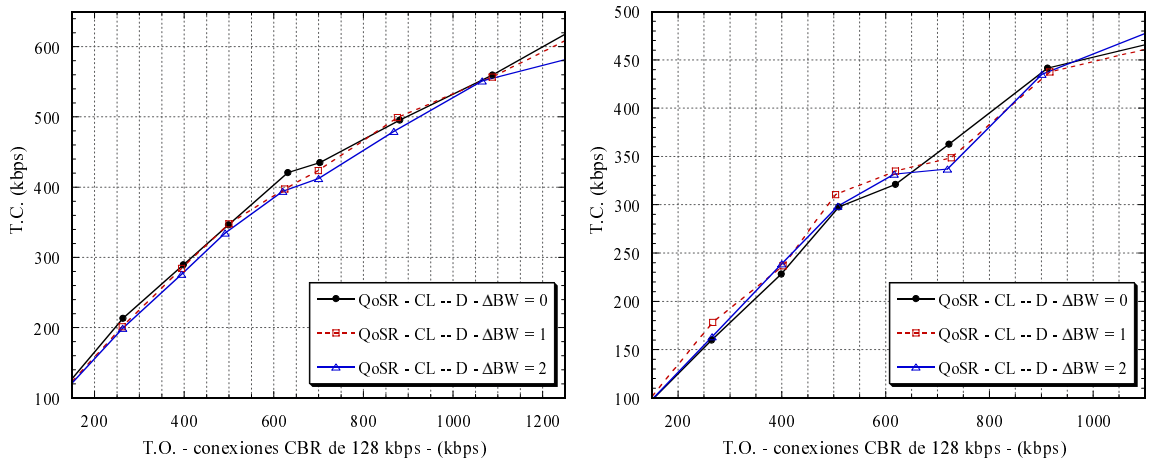


(a) Tráfico cursado (b) Probabilidad de bloqueo

(c) *Throughput* (d) Grado de servicio

Figura 4.51. Margen adicional de recursos en admisión, ΔBW expresado en número de slots. Prestaciones globales del sistema. Mecanismo básico sin monitorización-expulsión ND: ND-NPGB. Topología de 50 nodos con velocidad media de 2 Km/h.

Sin embargo, la utilización de un ancho de banda de guarda en un escenario realmente dinámico no degrada las prestaciones en términos de capacidad efectiva, como se observa en la figura 4.52. El tráfico correctamente cursado, tanto a 9 Km/h como a 20 Km/h es prácticamente equivalente al obtenido sin ancho de banda de guarda. El análisis del grado de servicio, por otra parte, demuestra la mejora real en la prestaciones del sistema.

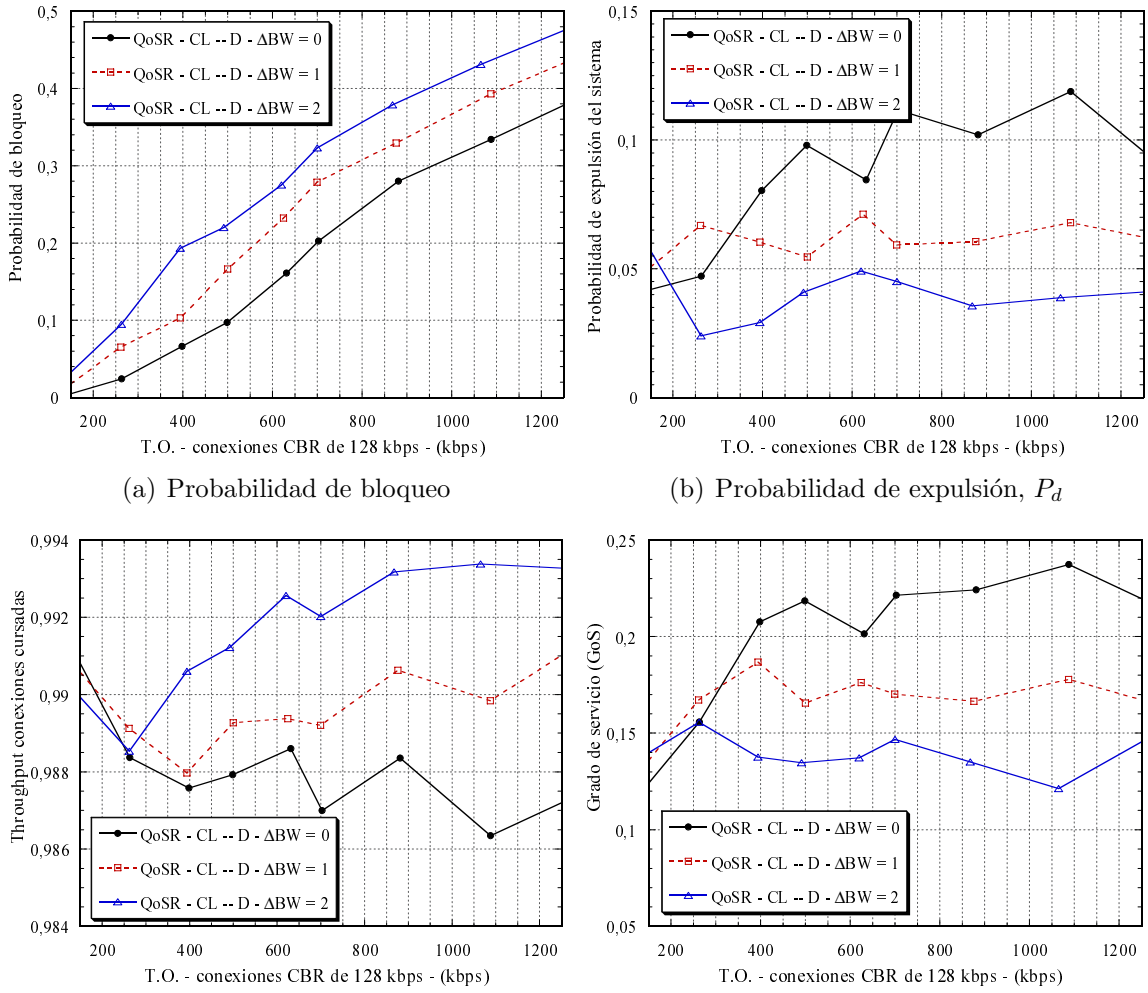


(a) 9 Km/h (b) 20 Km/h

Figura 4.52. Margen adicional de recursos en admisión, ΔBW expresado en número de slots. Tráfico correctamente cursado vs. tráfico ofrecido. Configuración de readmisión completa D: D3-PGB(3-0,3). Topología de 50 nodos. Escenarios de 9 Km/h y 20 Km/h.

Como se observa en la figura 4.53, un margen de guarda de 2 slots proporciona mejores prestaciones que un margen de tan sólo un slot. El mayor número de conexiones bloqueadas debido a la mayor infrautilización de recursos supone, sin embargo, un aumento en la disponibilidad de recursos para readmisiones ocasionadas tanto por rotura como por pérdidas de la QoS. Se reduce así la probabilidad de expulsión de conexiones (P_d) como demuestra la figura 4.53(b). Por otra parte, gracias al menor número de conexiones potencialmente coincidentes (colisiones debidas a movilidad) así como a la mayor disponibilidad de recursos en el procedimiento de permanente actualización (*piggy-backing*) las conexiones admitidas experimentan un aumento significativo en el *throughput* medio – figura 4.53(c) – que finalmente se traduce en una probabilidad total de fallo menor (P_f) mejorando así el grado de servicio obtenido, como se refleja en la figura 4.53(d).

El comportamiento del sistema observado en la figura 4.54, para un escenario de 20 Km/h, es análogo al descrito para 9 Km/h. En cualquiera de los dos casos se ha comprobado la mayor ganancia obtenida con un margen ΔBW superior (2 slots).



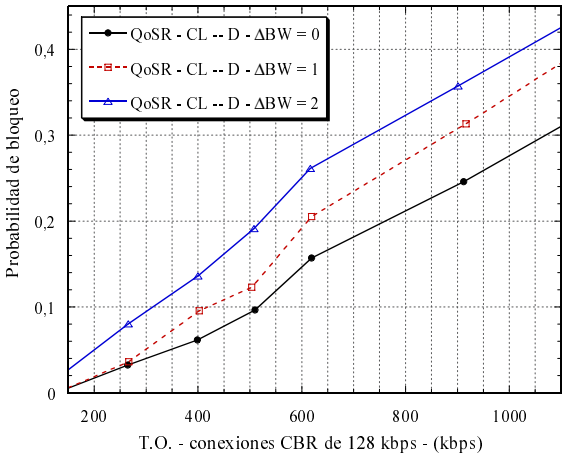
(a) Probabilidad de bloqueo

(b) Probabilidad de expulsión, P_d

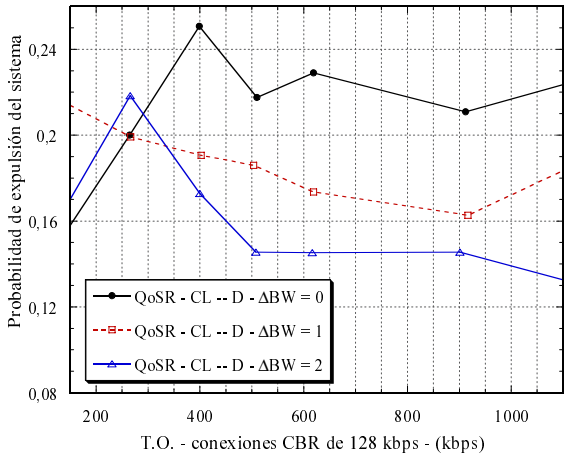
(c) *Throughput*

(d) Grado de servicio

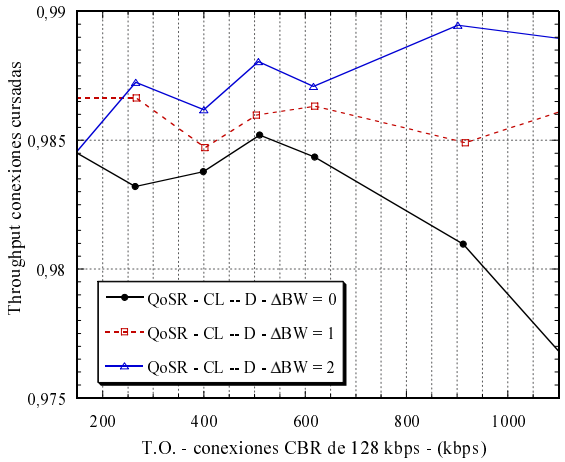
Figura 4.53. Margen adicional de recursos en admisión, ΔBW expresado en número de slots. Prestaciones globales del sistema. Configuración completa del mecanismo de readmisión D: D3-PGB(3-0,3). Topología de 50 nodos con velocidad media de 9 Km/h.



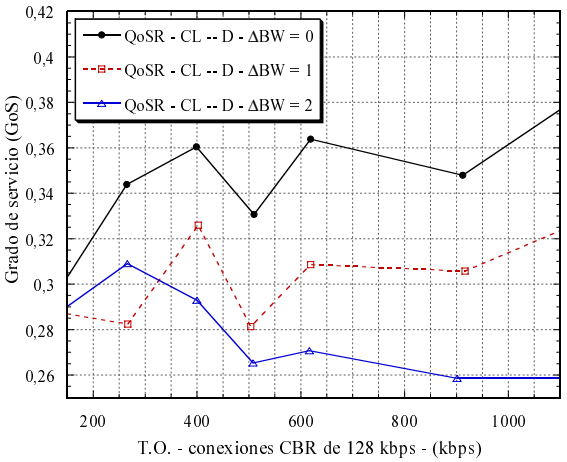
(a) Probabilidad de bloqueo



(b) Probabilidad de expulsión, P_d



(c) *Throughput*

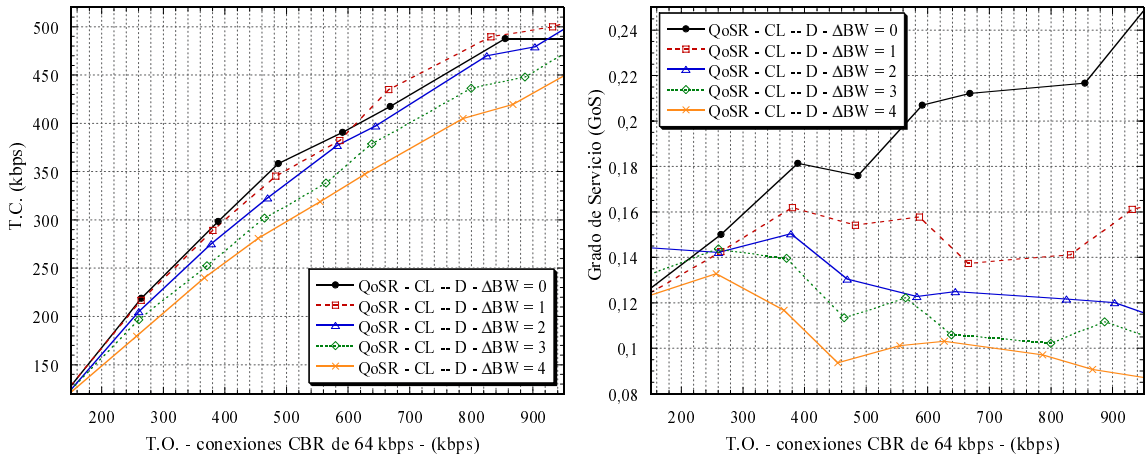


(d) Grado de servicio

Figura 4.54. Margen adicional de recursos en admisión, ΔBW expresado en número de slots. Prestaciones globales del sistema. Configuración completa del mecanismo de readmisión D: D3-PGB(3-0,3). Topología de 50 nodos con velocidad media de 20 Km/h.

Los resultados analizados muestran el compromiso existente entre capacidad y satisfacción de usuario. Resolver dicho compromiso exige un estudio más detallado del comportamiento del sistema en diversos escenarios de tráfico. A continuación se muestra el análisis realizado sobre conexiones CBR ofrecidas de 64 kbps, 128 kbps y 192 kbps, lo que supone una demanda individual de 1, 2 y 3 slots respectivamente. En la evaluación realizada se han considerado únicamente los escenarios de mayor movilidad: 9 Km/h y 20 Km/h.

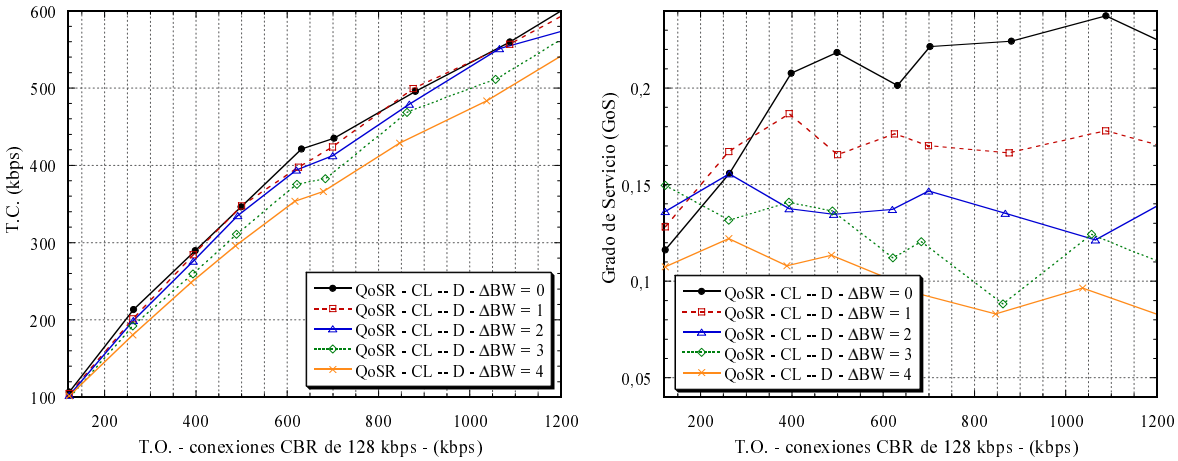
El margen de guarda aplicable está limitado por la capacidad efectiva obtenida finalmente, es decir, por el tráfico que el sistema es capaz de cursar satisfactoriamente. Obviamente, una limitación natural es la capacidad máxima del sistema, de manera que el ancho de banda de guarda máximo tolerable debe ser aquél que mantenga un porcentaje de los recursos desocupado en admisión sin reducir demasiado los recursos utilizables. Por otra parte, la implementación distribuida realizada relaciona directamente las aplicaciones ofrecidas y el margen obtenido, al imponer una demanda adicional de ΔBW slots sobre los requerimientos individuales de cada conexión. Como ya se mostró en §3.5, conexiones de mayor número de slots demandados tienen mayores dificultades para ser admitidas, especialmente si requieren de un número de saltos elevado. Esta dificultad impone, incluso en ausencia de márgenes adicionales, un porcentaje de recursos desocupados creciente ante mayores demandas individuales. Sin embargo, dicha desocupación surge de la dificultad de asignar los recursos a las conexiones dada la gran demanda por enlace que éstas suponen, lo que implica igualmente la dificultad de readmitir a aquéllas para las que, una vez dentro del sistema, cambian las condiciones originales de admisión. Teóricamente, un escenario de aplicaciones de mayor demanda individual requeriría, en valor absoluto, mayores márgenes para posibilitar, en caso de fallo, la readmisión de las conexiones. En el caso de aplicaciones de bajo ancho de banda, el requerimiento podría ser menor dada la menor dificultad de reasignación. Además, la relación existente entre el número de conexiones ofrecidas y el porcentaje de recursos libres adicionales que se demandan condiciona directamente el margen global conseguido.



(a) Tráfico Cursado (b) GoS

Figura 4.55. Evaluación de la elección del margen ΔBW. Tráfico ofrecido: conexiones CBR de 64 kbps – 1 slot. Topología de 50 nodos con velocidad media 9 Km/h.

Cuando el tráfico ofrecido total equivale a un mayor número de conexiones que demandan un ancho de banda menor, el sistema tiene más facilidad para asignarles recursos, aumentando el número de conexiones admitidas y por tanto el tráfico cursado global. Sin embargo, al asignar a cada conexión un margen adicional, el porcentaje total infrautilizado es mayor que su equivalente para mayor tasa. En consecuencia, cuando el tráfico ofrecido se corresponde con aplicaciones de menor demanda individual el margen aplicable se encuentra muy limitado por la consecuente pérdida de capacidad efectiva, como se muestra en la figura 4.55(a). Si bien la reducción en la probabilidad de fallo permite obtener un mejor GoS conforme aumenta el margen establecido, como se observa en la figura 4.55(b), las prestaciones globales del sistema, en términos de tráfico correctamente cursado, disminuyen. Así pues, demandas individuales mínimas (1 slot) requieren márgenes de ancho de banda reducidos (1 slot) que mejoran sensiblemente el GoS manteniendo la capacidad efectiva.

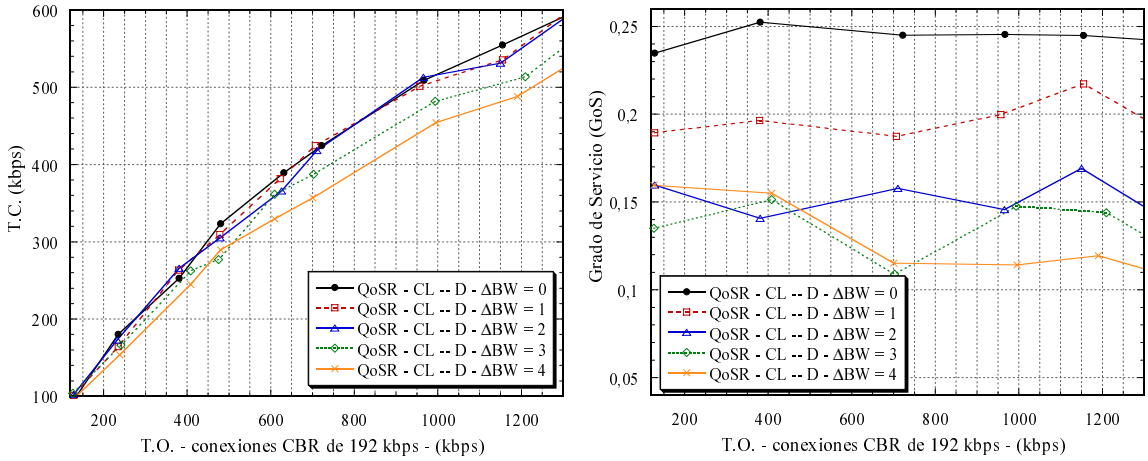


(a) Tráfico Cursado (b) GoS

Figura 4.56. Evaluación de la elección del margen ΔBW. Tráfico ofrecido: conexiones CBR de 128 kbps – 2 slots. Topología de 50 nodos con velocidad media 9 Km/h.

Al aumentar la demanda individual por conexión, aumenta la dificultad de encontrar caminos con suficiente ancho de banda. Un mayor ΔBW permite aumentar la probabilidad de readmitir las conexiones al disponer de más recursos desocupados. Para el mismo tráfico ofrecido global, el número de conexiones ofrecidas disminuye, reduciéndose en consecuencia el porcentaje de ancho de banda demandado en exceso. Sin embargo, se logra mantener la proporción real de ancho de banda respecto a cada conexión lo que facilita su futura readmisión. La reducción en la probabilidad de fallo y la consecuente ganancia en GoS permiten compensar el bloqueo creciente manteniendo las prestaciones

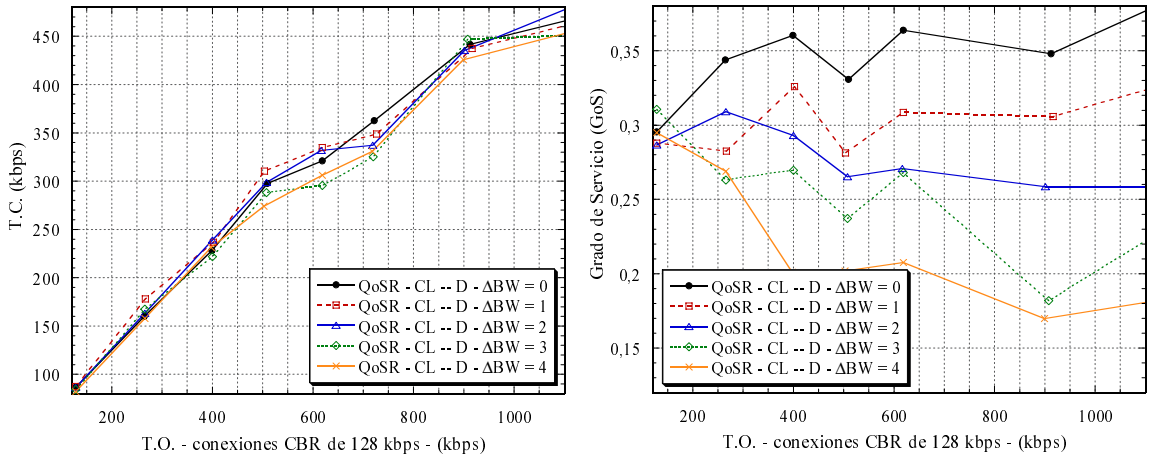
en cuanto al tráfico correctamente cursado. Como se observa en la figura 4.56(b), la mejora en GoS resulta mucho más apreciable a partir de un margen de 2 slots, con el que se garantiza un tráfico correctamente cursado comparable – figura 4.56(a), degradándose éste para márgenes mayores. El análisis realizado sobre tasas de 192 kbps, con una elevada exigencia de ancho de banda (3 slots) muestra igualmente la conveniencia en la utilización de un margen de 2 slots, que permite obtener una mejora apreciable en GoS – figura 4.57(b) – sin degradación en el tráfico correctamente cursado – figura 4.57(a).



(a) Tráfico Cursado (b) GoS

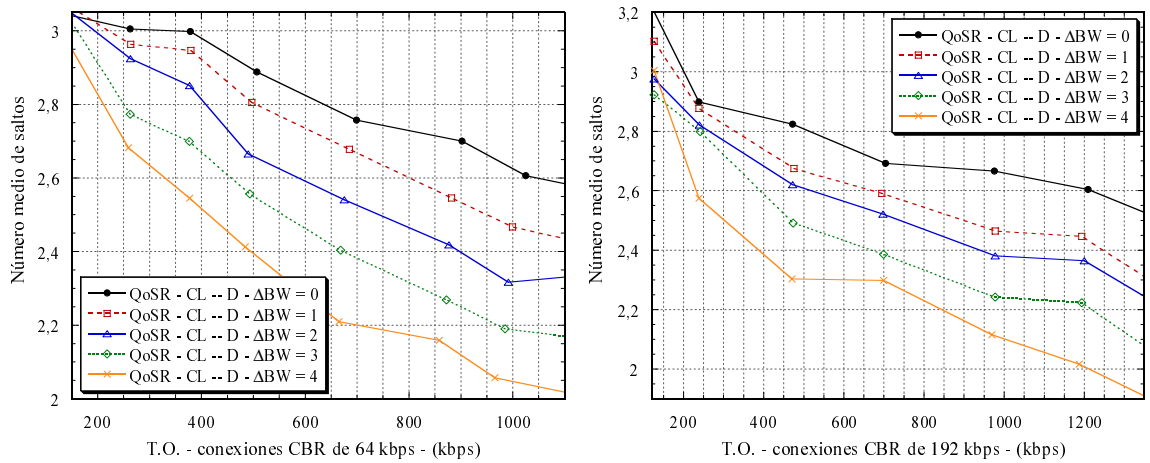
Figura 4.57. Evaluación de la elección del margen ΔBW. Tráfico ofrecido: conexiones CBR de 192 kbps – 3 slots. Topología de 50 nodos con velocidad media 9 Km/h.

Por otra parte, a pesar de que el tráfico correctamente cursado en un escenario de 20 Km/h es equivalente para los diferentes márgenes de ancho de banda analizados – figura 4.58(a), la ganancia en el GoS – figura 4.58(b) – no debe ser el único criterio a la hora de establecer como apropiada la utilización de uno u otro margen. El número medio de saltos de las rutas establecidas se reduce significativamente con márgenes mayores, como se observa en la figura 4.59. Esta reducción, debida al menor número de recursos disponibles, supone realmente un mantenimiento no igualitario de conexiones dentro del sistema. Aquellas aplicaciones entre nodos cercanos, es decir, que requieran rutas más cortas, tienen mayores probabilidades de encontrar ancho de banda tanto en su admisión inicial como en posibles readmisiones. Así pues, a pesar de una capacidad global equivalente, el comportamiento del sistema ante márgenes de ancho de banda diferentes condiciona las prestaciones que realmente éste es capaz de ofrecer a las aplicaciones. Así pues, a igual capacidad efectiva, se establece un nuevo compromiso entre el GoS y la longitud media de las conexiones admitidas, en términos de saltos.



(a) Tráfico Cursado (b) GoS

Figura 4.58. Evaluación de la elección del margen ΔBW. Tráfico ofrecido: conexiones CBR de 128 kbps – 2 slots. Topología de 50 nodos con velocidad media 20 Km/h.



(a) 64 kbps (b) 192 kbps

Figura 4.59. Efecto del margen ΔBW sobre la capacidad analizado en términos de longitud de las rutas (número medio de saltos). T.O. CBR de 64 y 192 kbps. Topología de 50 nodos con velocidad media 20 Km/h.

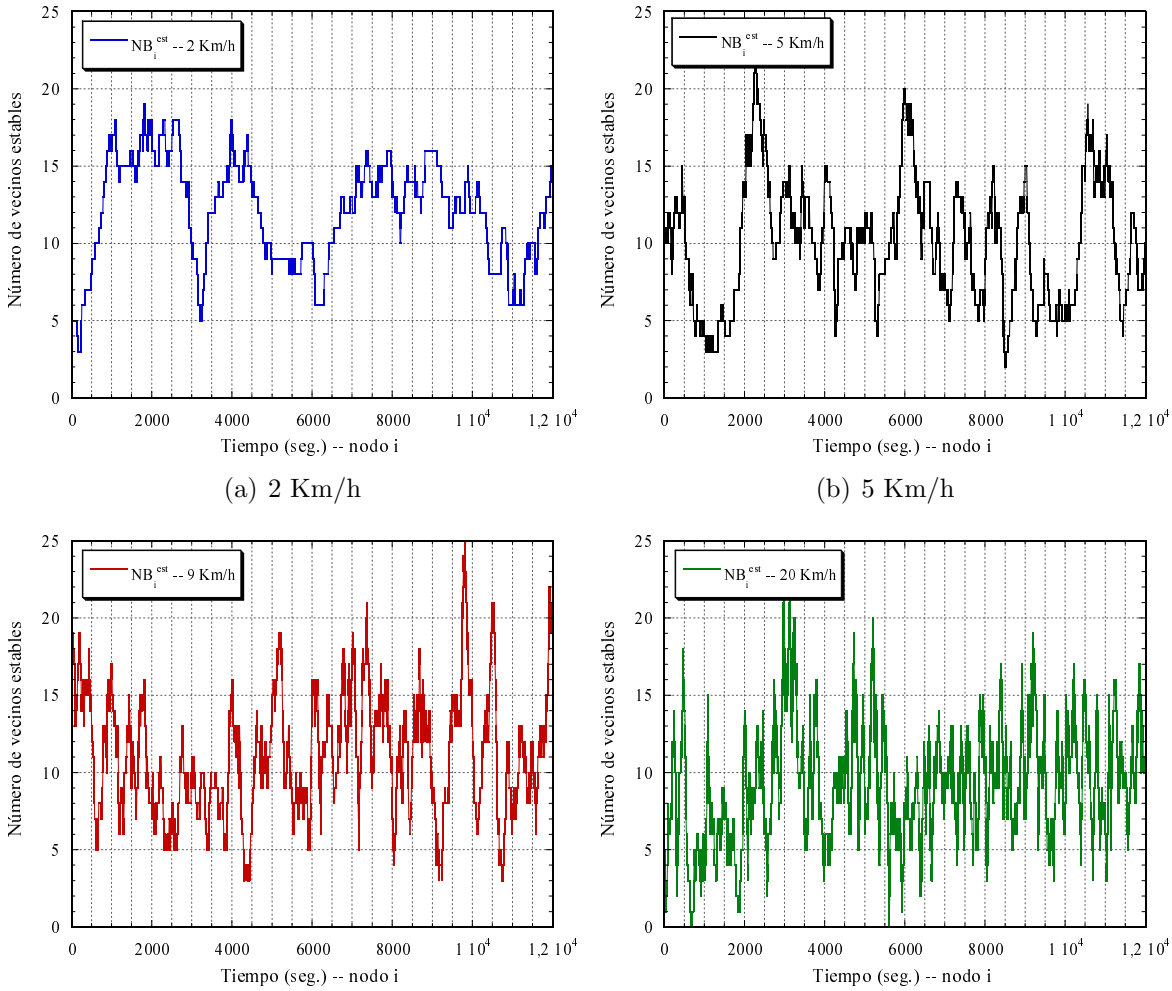
Teniendo en cuenta la generalización buscada a la hora de obtener una parametrización apropiada para escenarios con condiciones de tráfico y grados de movilidad variable, se han considerado adecuados márgenes de 1 ó 2 slots. Observando la mejora en GoS obtenida con márgenes mayores y los valores comparables de capacidad efectiva se ha establecido un margen $\Delta BW = 2$ slots, que no supone una reducción excesiva en términos de saltos[22], siendo así utilizado a partir de ahora como parte intrínseca del modo de operación "**Funcionamiento en movilidad**".

Configuración adaptativa en función del grado de movilidad

La configuración apropiada de los parámetros de admisión depende en gran medida, como se ha extraído del análisis de los resultados, del grado de movilidad del escenario. Esta dependencia sugiere la necesidad de diseñar un mecanismo eficiente de adaptación al entorno capaz de estimar adecuadamente la movilidad modificando en consecuencia la respuesta del sistema.

Los escenarios considerados generan trayectorias comparables entre todos los terminales móviles, lo que supone la formación de topologías variantes uniformes en las que no existe diferenciación de nodos estacionarios (como podrían ser puntos de acceso) o distribuciones geográficas de diferentes grados de movilidad. Esta generalización, sin embargo, permite un mayor nivel de abstracción en el análisis desarrollado que facilita la evaluación de la propuesta realizada. En contraposición, la falta de un grado de determinismo suficiente en los escenarios, a pesar del efecto memoria introducido en el RWPM, limita la aplicabilidad de los clásicos modelos de predicción. La uniformidad en la movilidad del escenario, sin embargo, permite realizar una estimación localizada y por tanto simplificada de la misma, especialmente apropiada en un entorno distribuido. Esta estimación puede resultar una herramienta particularmente útil a la hora de permitir una configuración del sistema convenientemente adaptada a la movilidad del entorno. La principal dificultad que se presenta consiste en la adecuada determinación de los parámetros del entorno capaces de identificar correctamente el grado de movilidad existente.

[22]Esta reducción puede permitir, además, una disminución natural en el retardo medio experimentado por las aplicaciones, lo que a su vez mejora las prestaciones en cuanto a la QoS ofrecida, especialmente en aplicaciones en tiempo real.



(a) 2 Km/h

(b) 5 Km/h

(c) 9 Km/h

(d) 20 Km/h

Figura 4.60. Evolución temporal del número de vecinos estables. Muestra representativa de un terminal concreto. Topologías de 50 nodos. Escenarios de 2, 5, 9 y 20 Km/h.

Gracias al intercambio de información de conectividad proporcionado por el AD-HOC MAC, una de las medidas del entorno más inmediatas es el número de vecinos estables. Dado que la movilidad relativa entre los diferentes nodos genera continuas desconexiones de los enlaces establecidos, en mayor grado cuanto mayores son las respectivas velocidades, la variación en el número de vecinos puede constituir una muestra significativa de la variabilidad local, como se observa en la figura 4.60, donde puede apreciarse el incremento de oscilaciones de acuerdo a la velocidad media de los diversos escenarios. Con objeto de evaluar adecuadamente la respuesta de la arquitectura propuesta ante la variabilidad del entorno, se han considerado los tres escenarios ya analizados de 2, 9 y 20 Km/h así como un escenario de 5 Km/h con los mismo parámetros (R_{max}, α_{max} y T_{pausa}) y unas velocidades mínima y máxima de 3 y 8 Km/h.

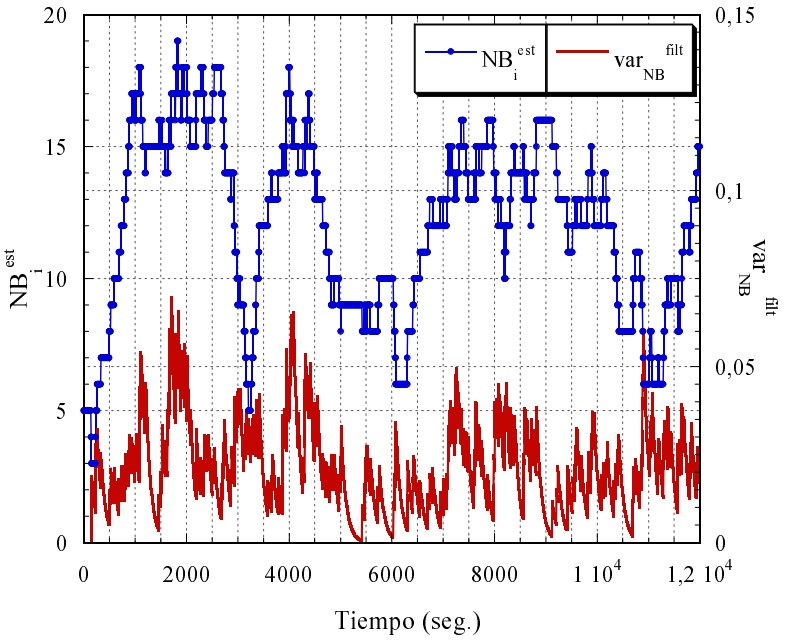


Figura 4.61. Evolución del número de vecinos y métrica $var_{NB,i}^{filt}$ para un nodo i. $\alpha = 0{,}99$. Simulación de 12000 segundos. Topología de 50 nodos con velocidad media de 2 Km/h.

Con el fin de medir objetivamente el grado de movilidad en relación a la variación de vecinos se ha definido la métrica $var_{NB,i}^{filt}$ (4.11). El valor medio en la rapidez de oscilación del número vecinos – $var_{NB,i}$ (4.12) medio – se relaciona directamente con el grado de movilidad de cada escenario. La figura 4.61 muestra un ejemplo, para un escenario de 2 Km/h, de la evolución del número de vecinos estables y la consecuente métrica $var_{NB,i}^{filt}$ obtenida.

$$var_{NB,i}^{filt}(t) = \alpha \cdot var_{NB,i}^{filt}(t - \Delta t) + (1 - \alpha) \cdot var_{NB,i}(t) \tag{4.11}$$

$$var_{NB,i}(t) = \left| \frac{NB_i^{est}(t) - NB_i^{est}(t - \Delta t)}{\Delta t} \right| \tag{4.12}$$

Como se observa en la figura 4.62, que representa el promedio del valor medio observado por cada uno de los 50 nodos de la topología , existe una clara diferenciación de acuerdo a la velocidad media del escenario. Cada terminal, a la hora de estimar este valor medio, realiza un promediado de $var_{NB,i}$ mediante un filtro IIR paso bajo de primer orden con parámetro α, que representa el efecto memoria del filtro. Su diseño debe adecuarse al propósito perseguido, que consiste por un lado, en realizar una rápida adaptación al entorno, y por otro, filtrar adecuadamente las variaciones transitorias propias de las oscilaciones que se pretende promediar.

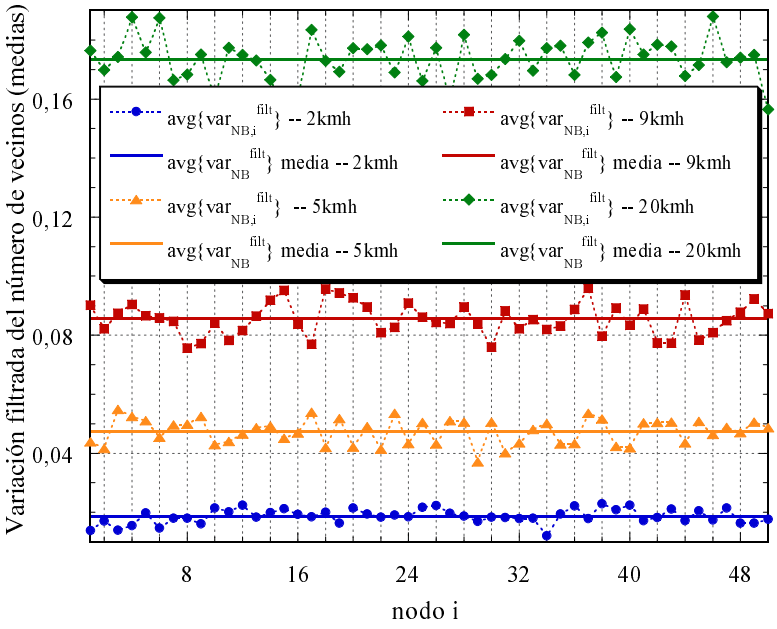
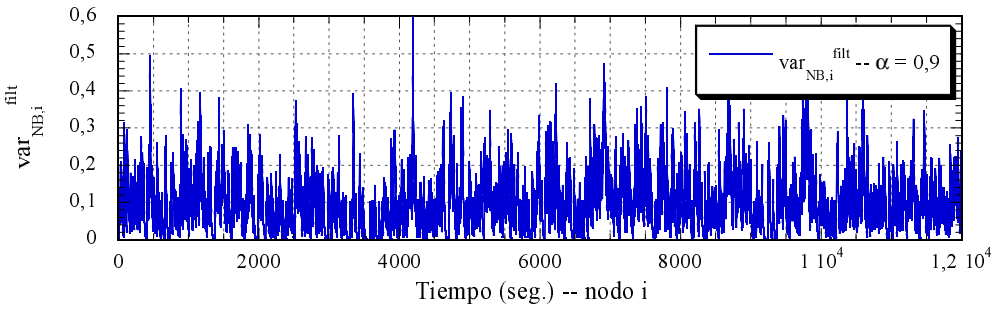
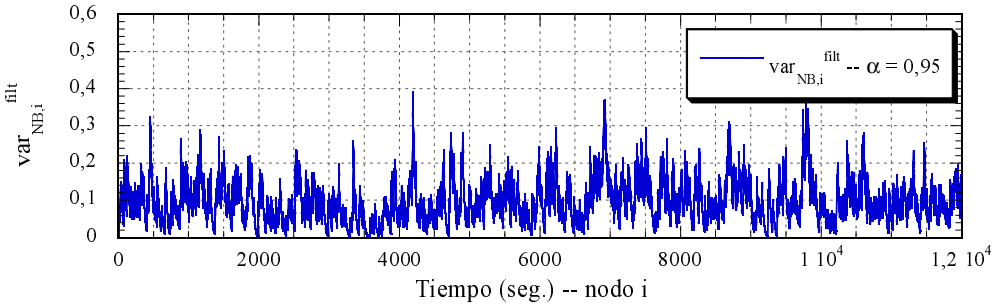


Figura 4.62. Variación filtrada del número de vecinos, $\alpha = 0{,}99$. Valor medio para cada terminal en topologías de 50 nodos. Simulación de 12000 segundos. Escenarios de 2, 5, 9 y 20 Km/h.

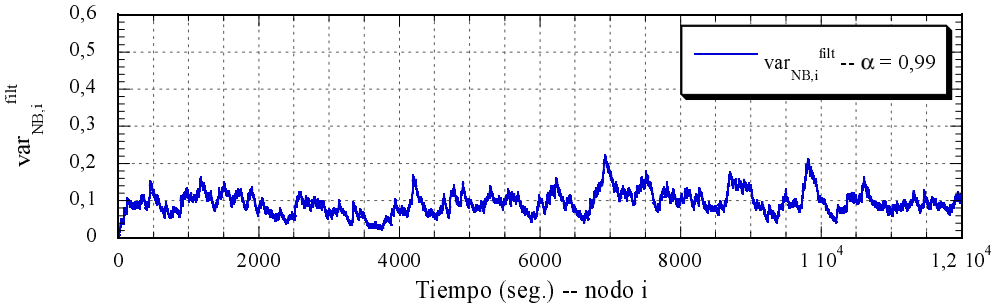
La figura 4.63 muestra el seguimiento de $var_{NB,i}$ realizado por un nodo particular i, a partir de diversos valores α, donde si bien la actualización de vecinos estables (NB_i^{est}) se realiza trama a trama, la periodicidad seleccionada Δt se corresponde con 1 seg. (20 tramas).



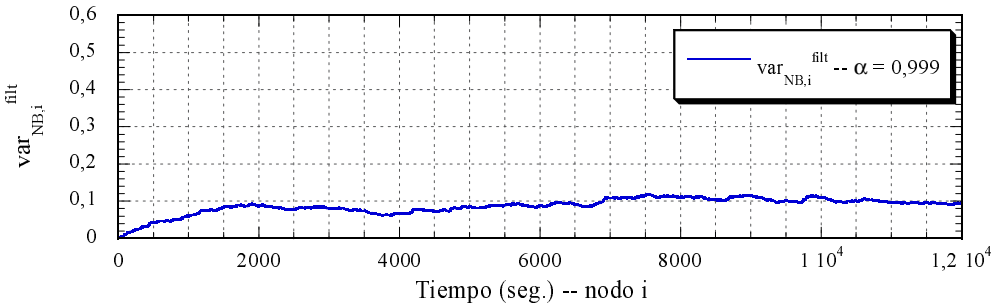
(a) $\alpha = 0{,}9$



(b) $\alpha = 0{,}95$



(c) $\alpha = 0{,}99$

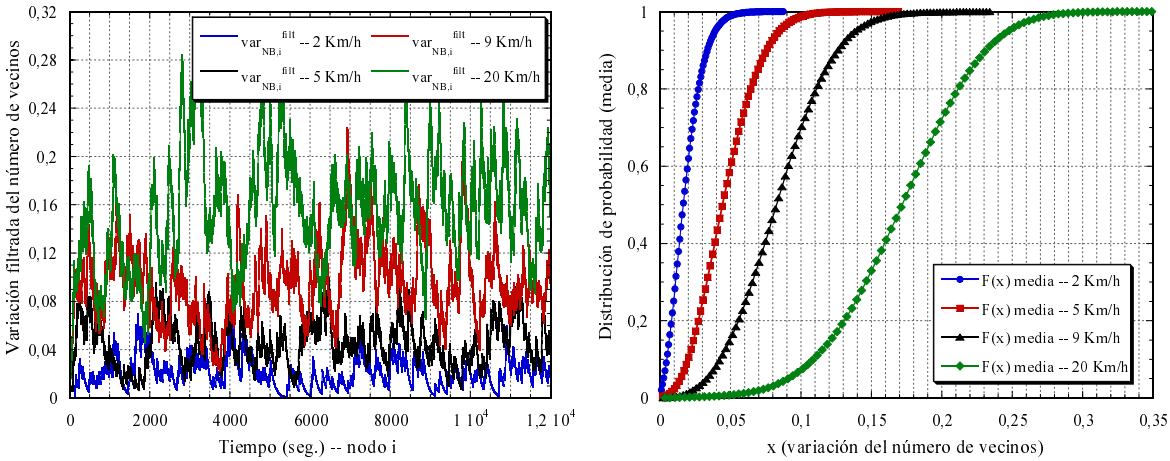


(d) $\alpha = 0{,}999$

Figura 4.63. Variación filtrada del número de vecinos. Comparativa respecto a la constante del filtro, α. Valores entre 0,9 y 0,999. Muestra de la medida de un nodo. Topología de 50 nodos con velocidad media de 9 Km/h.

Como muestran los resultados, valores de α más pequeños suponen un menor filtrado que genera un métrica oscilante poco informativa – figura 4.63(a). No obstante, una memoria excesiva – figura 4.63(d) – aunque facilita una mejor estimación del valor medio, supone una convergencia demasiado lenta para adaptarse a los posibles cambios de movilidad en el entorno. A la vista de los resultados obtenidos, se ha considerado apropiado un valor $\alpha = 0{,}99$, que proporciona una buena aproximación al valor medio. Este valor genera una oscilación que puede parecer excesiva, dificultando la interpretación del grado de movilidad asociado. Sin embargo, a su vez permite una pronta adaptación a cambios en la movilidad del entorno, como se desprende del corto transitorio observado en la figura 4.63(c).

Una vez definida una métrica capaz de identificar el grado de movilidad del entorno, es necesario diseñar adecuadamente el mecanismo de decisión que permite activar la reacción esperada del sistema. Dado que resulta imposible establecer una relación totalmente exacta entre la movilidad y dicha métrica, teniendo en cuenta las evidentes oscilaciones, como muestra la figura 4.64(a), se he considerado una política basada en umbrales para definir el estado del sistema: **quasi-estático** o en **movilidad**. El objetivo del procedimiento es establecer una separación apropiada entre ambos estados de manera que se configuren los modos de operación adecuados a cada entorno, ND-NPGB y D3-PGB(3-0,3) – $\Delta BW = 2$ respectivamente, sin desestabilizar el sistema.



(a) $var_{NB,i}^{filt}$ (b) Distribución de probabilidad (F_x de $var_{NB,i}^{filt}$)

Figura 4.64. Variación filtrada del número de vecinos. Filtro de promediado: $\alpha = 0{,}99$. Simulación de 12000 segundos. Topologías de 50 nodos con velocidades medias de 2, 5, 9 y 20 Km/h.

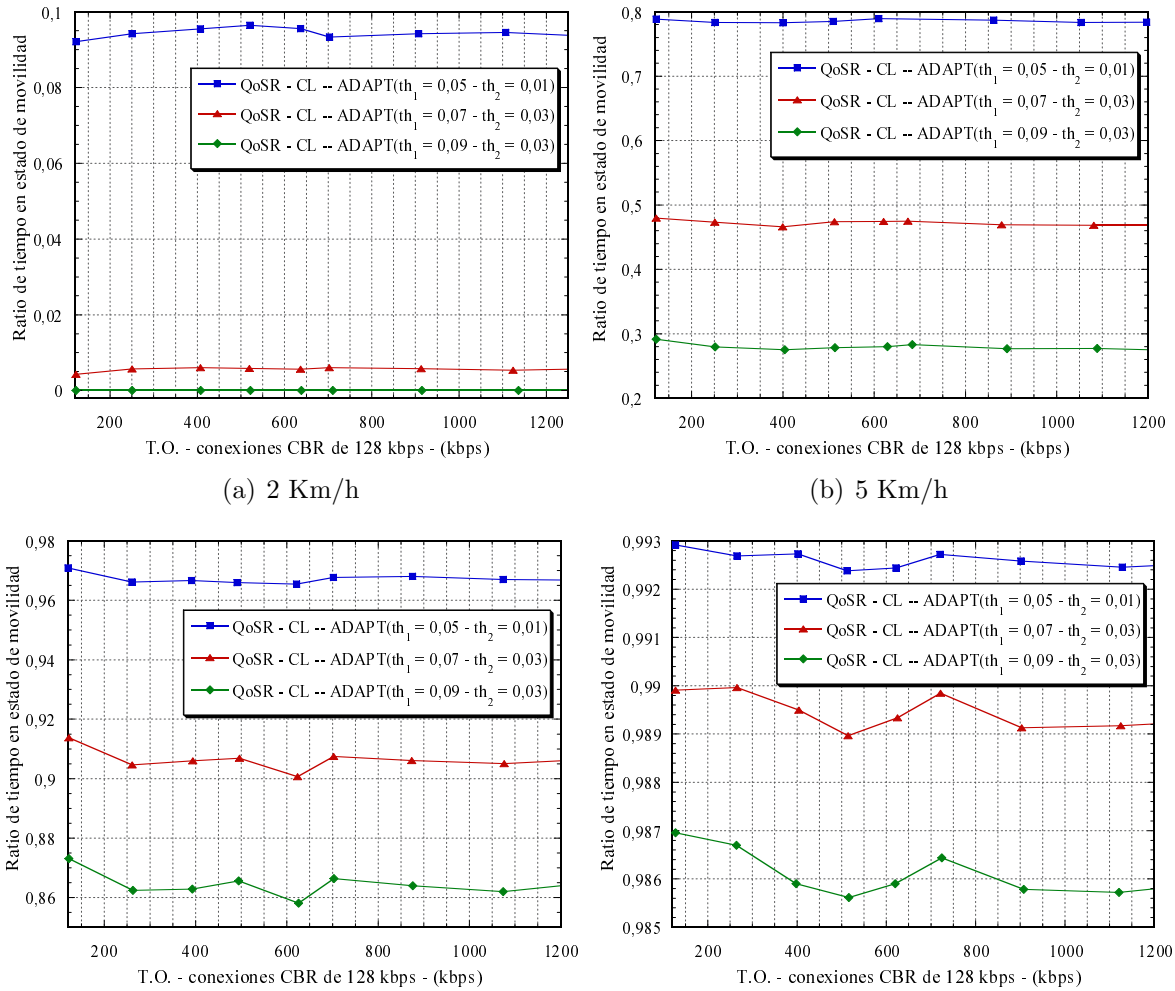
Como puede observarse en la figura 4.64(a) la variabilidad de la métrica, como consecuencia del valor α considerado, dificulta una clara diferenciación de los escenarios. No obstante, si se considera la dispersión observada, pueden establecerse unos umbrales que proporcionen unas probabilidades de acierto en la definición del estado suficientemente elevadas, atendiendo a las distribuciones de la métrica especificada para cada velocidad – figura 4.64(b).

La respuesta del sistema se ha basado en una configuración adaptada de los parámetros de acuerdo a dos umbrales. La utilización de estos dos valores responde a una doble necesidad: por una parte se pretende establecer el paso adecuado entre un escenario quasi-estático y uno de movilidad (umbral th_1), y por otra, se implementa un procedimiento de histéresis con el objetivo de garantizar una permanencia más o menos estable en este último estado evitando la oscilación continua de los parámetros de configuración del mecanismo de readmisión (umbral th_2). Considerando los resultados obtenidos previamente, se ha tomado como referencia la diferenciación entre un escenario de 9 Km/h (movilidad) y un escenario de 2 Km/h (quasi-estático). De acuerdo a esta consideración, se han analizado los valores medios y desviaciones estándar de ambos escenarios, correspondientes a lo siguiente: 2 Km/h (media: 0,0183 – desviación: 0,0102) y 9 Km/h (media: 0,0865 – desviación: 0,035). Las tres configuraciones finalmente mostradas en la evaluación de la propuesta se corresponden con la siguiente parametrización: $th_1 = 0{,}05$ – $th_2 = 0{,}01$; $th_1 = 0{,}07$ – $th_2 = 0{,}03$ y $th_1 = 0{,}09$ – $th_2 = 0{,}03$. La evaluación sobre un escenario de 5 Km/h pretende estudiar el impacto de la selección realizada sobre un escenario de velocidad intermedia (media: 0,0474 – desviación: 0,0209).

La figura 4.65 muestra el ratio de tiempo en el que el entorno se estima en estado de movilidad y, consecuentemente, el modo de operación del sistema equivale a la estrategia D3-PGB(3-0,3) – $\Delta BW = 2$. La figura representa el promedio realizado sobre los 50 nodos de la topología para el valor medio calculado a lo largo de las simulaciones realizadas[23]. Un umbral th_1 bajo (0,05) supone una estimación del movimiento excesiva, que genera ratios significativos incluso en los escenarios de 2 Km/h y 5 Km/h. Por el contrario, un umbral mayor (0,09) supone una identificación más tardía del movimiento que favorece la clara diferenciación entre los escenarios de 2 Km/h y 20 Km/h. Sin embargo, en un escenario de 9 Km/h que, en vista de los resultados previos requiere un modo de operación en movilidad, un ratio del orden del 0,8 puede suponer intervalos de

[23]Tiempos de simulación (T_{sim}) según el escenario, de acuerdo a la estimación realizada en el análisis del modelo de movilidad RWPM (ver anexo C). 2 y 5 Km/h: $T_{sim} = 12.000$ seg., 9 Km/h: $T_{sim} = 9.000$ seg. y 20 Km/h: $T_{sim} = 6.000$ seg.

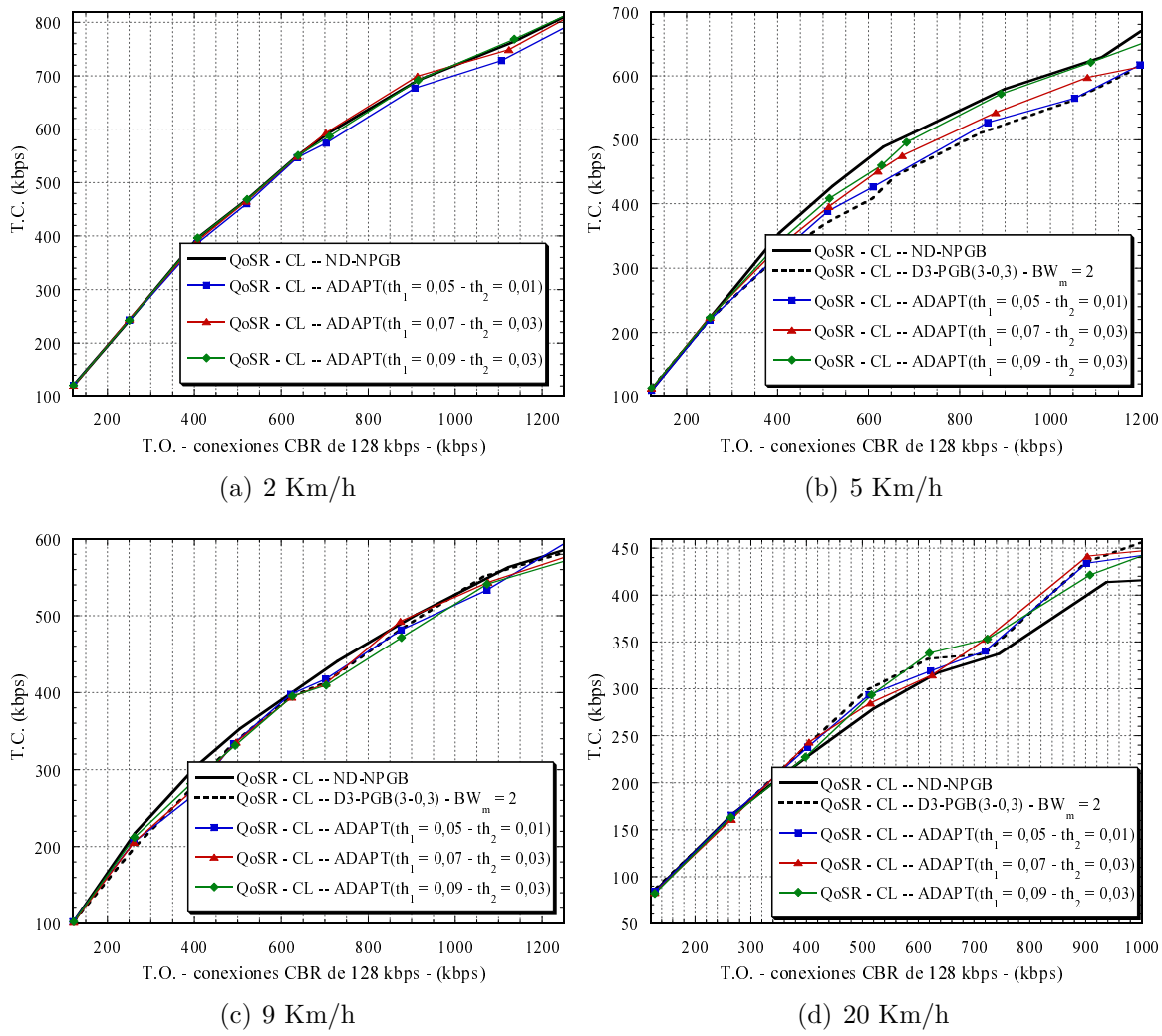
tiempo considerables con una configuración quasi-estática. La oscilación en los parámetros de configuración y la indeterminación en las condiciones de admisión/readmisión que se presentan (como por ejemplo, expulsión invocada en un escenario sin recursos disponibles dada una excesiva admisión previa) generan un escenario poco estable de resultados impredecibles. Una reducción del umbral, hasta un valor del orden del 0,07, si bien reduce esta inestabilidad, genera por el contrario una consideración de movilidad excesiva en escenarios de menor velocidad (5 Km/h). Las implicaciones reales de esta estimación del movimiento se observan a continuación.



(a) 2 Km/h (b) 5 Km/h

(c) 9 Km/h (d) 20 Km/h

Figura 4.65. Ratio de tiempo en modo de funcionamiento "movilidad" vs. tráfico ofrecido. Mecanismo adaptativo de readmisión, análisis de los umbrales que definen la movilidad (th_1 y th_2). Topologías de 50 nodos. Escenarios de 2, 5, 9 y 20 Km/h.



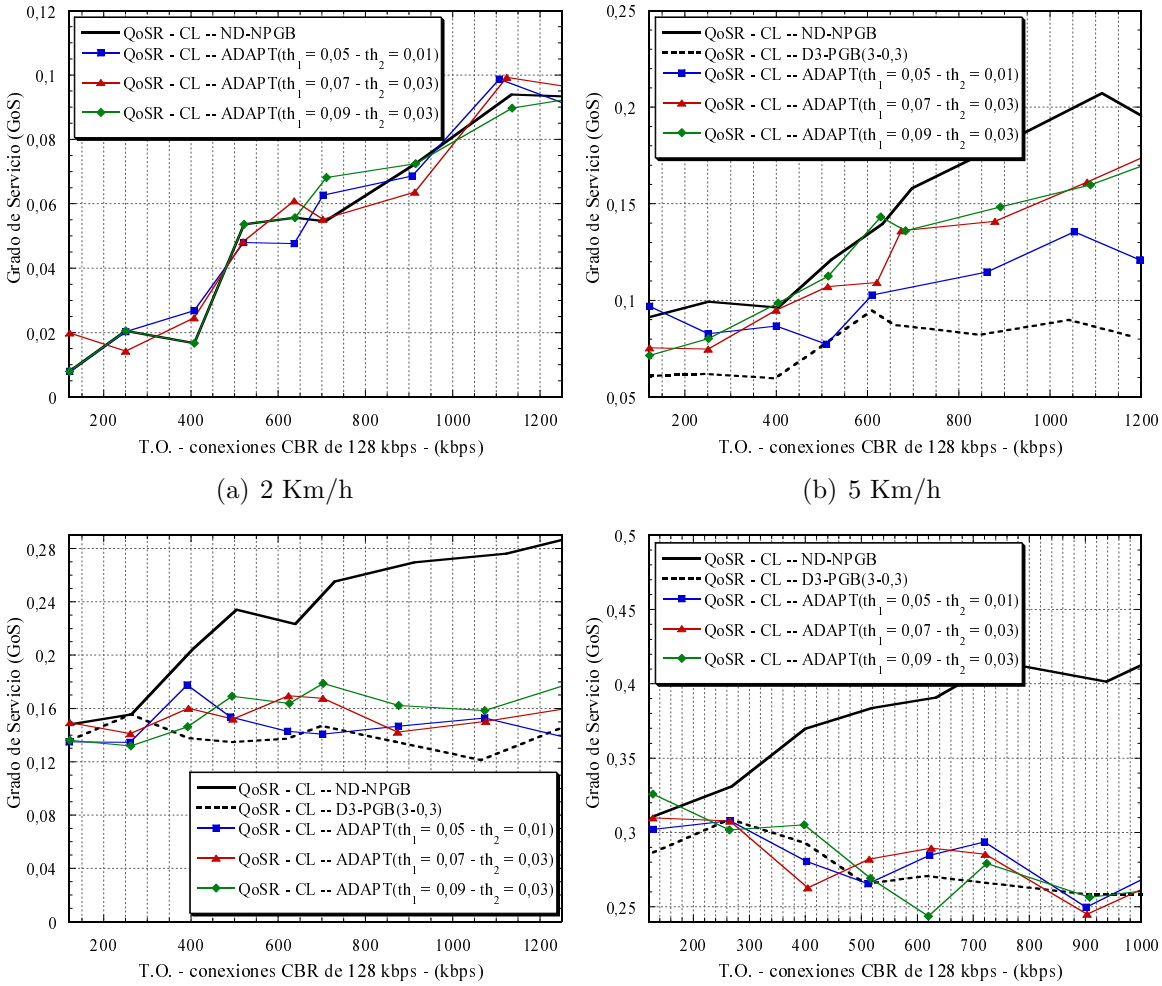
(a) 2 Km/h (b) 5 Km/h

(c) 9 Km/h (d) 20 Km/h

Figura 4.66. Tráfico cursado vs. tráfico ofrecido. Mecanismo adaptativo de readmisión, análisis de los umbrales que definen la movilidad (th_1 y th_2). Topologías de 50 nodos. Escenarios de 2, 5, 9 y 20 Km/h.

Las figuras 4.66 y 4.67 muestran el efecto de los diversos umbrales seleccionados sobre el tráfico correctamente cursado y el GoS respectivamente. Las estrategias que suponen un mayor tiempo de movimiento generan unas prestaciones más cercanas a las obtenidas mediante la aplicación continuada del modo de funcionamiento en movilidad. Dichas configuraciones generan una degradación apreciable en escenarios de 5 Km/h – figura 4.66(b), mientras que la aplicación de un margen th_1 de 0,09 permite mantener la capacidad efectiva implicando una pérdida poco significativa a 9 Km/h – figura 4.66(c). Se obtiene además en este caso un GoS comparable a la situación de completa movilidad – figura 4.67(c). Un umbral de 0,09 es ligeramente superior al nivel medio de $var_{NB,i}^{filt}$

obtenido con una velocidad media de 9 Km/h (0,0865). En estas condiciones, si bien puede haber cierto transitorio hasta la identificación real de la movilidad, es altamente probable el mantenimiento del modo de operación, teniendo en cuenta además el umbral th_2 seleccionado (0,03) que permite absorber en gran medida las oscilaciones (desviación estándar 0,035).



(a) 2 Km/h (b) 5 Km/h

(c) 9 Km/h (d) 20 Km/h

Figura 4.67. Grado de servicio vs. tráfico ofrecido (CBR de 128 kbps). Mecanismo adaptativo de readmisión, análisis de los umbrales que definen la movilidad (th_1 y th_2). Topologías de 50 nodos. Escenarios de 2, 5, 9 y 20 Km/h.

En definitiva, de acuerdo a los resultados obtenidos, un posible compromiso en el establecimiento de los umbrales es la selección de los valores $th_1 = 0{,}09$ y $th_2 = 0{,}03$. De este modo, se asegura una rápida identificación y respuesta ante escenarios de mayor movilidad, con unas prestaciones tan sólo ligeramente inferiores a las deseadas en las

situaciones límite entre escenario quasi-estático y con movilidad (9 Km/h), garantizando igualmente la configuración más básica (ND-NPGB) en condiciones quasi-estáticas (velocidades medias menores de 9 Km/h), en las que la configuración de movilidad produce pérdidas en la capacidad efectiva.

A pesar de la dificultad de optimizar completamente la respuesta dada por el sistema teniendo en cuenta cada escenario particular, un diseño adaptativo permite por el contrario generalizar el mecanismo propuesto, garantizando una adecuación automática al entorno sin necesitar una configuración prediseñada. La evaluación del mecanismo realizada en escenarios de movilidad variable ha constatado la superioridad de un esquema adaptado al entorno frente a la configuración a priori de los distintos modos de operación, como se observa a continuación.

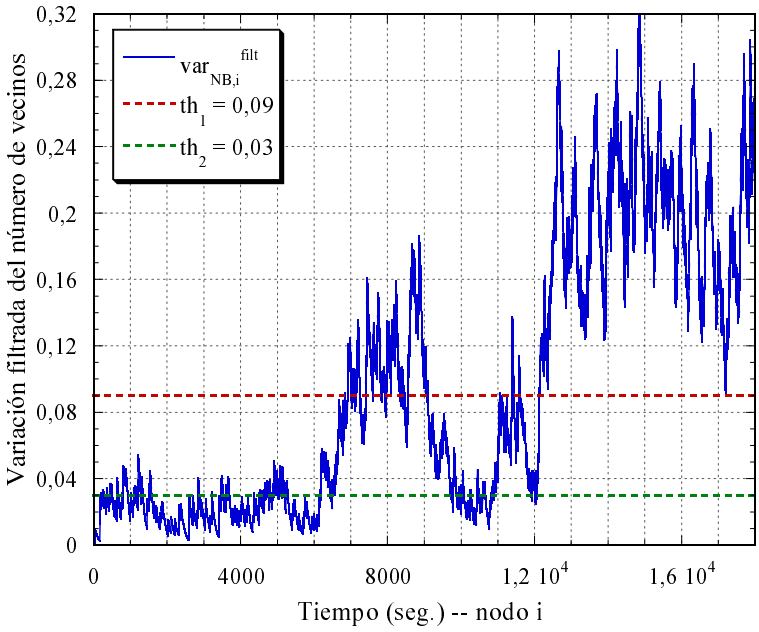
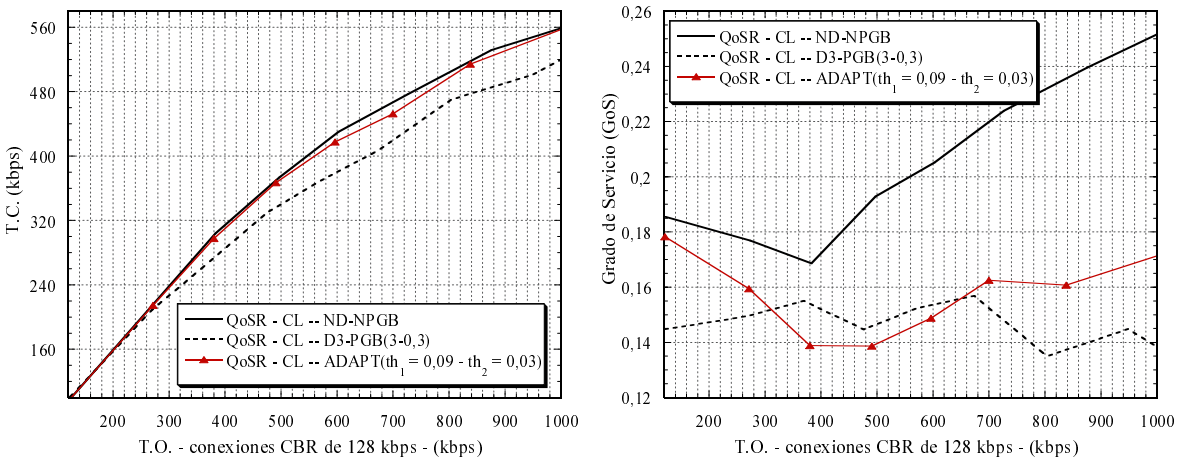


Figura 4.68. Variación filtrada del número de vecinos. Filtro de promediado: $\alpha = 0{,}99$. Simulación de 18000 segundos. Escenario de 50 nodos con movilidad variable (tramos de 6000 seg. y velocidades medias de 2, 9 y 20 Km/h).

La figura 4.4 muestra, a modo representativo, la estimación de la métrica considerada (variación filtrada del número de vecinos) para un nodo concreto en un escenario de movilidad variable. Dicho escenario se corresponde con una topología de 50 nodos en movimiento, de acuerdo al modelo RWPM, durante una simulación de 18000 seg. en la que la velocidad media observada varía en tramos sucesivos de 6000 seg. de acuerdo a los valores de 2, 9 y 20 Km/h respectivamente.

Como se observa en la figura, a pesar de las oscilaciones en el valor estimado, en relación a los umbrales establecidos ($th_1 = 0{,}09$ y $th_2 = 0{,}03$) la estimación del modo de operación apropiado se corresponde mayoritariamente con lo establecido previamente (modo quasi-estático para 2 Km/h y operación en movilidad en caso contrario).

La figura 4.69 muestra las prestaciones globales obtenidas en términos de tráfico correctamente cursado y grado de servicio. Como se observa en la figura 4.69(a), la configuración del modo de operación en movilidad (D3-PGB(3-0,3)) reduce la capacidad efectiva obtenida. El *overhead* introducido, si bien puede verse compensado cuando la movilidad del entorno es mayor, no supone ninguna mejora observable en el caso contrario. A pesar de que este efecto negativo es teóricamente más representativo en un porcentaje menor del tiempo (los 6000 segundos de velocidad media de 2 Km/h) redunda en una degradación considerable al no verse compensado por un aumento de capacidad efectiva para velocidades mayores (como se observó anteriormente – figura 4.66, en términos de tráfico cursado, una mejora real es únicamente observada para escenarios de movilidad elevada, de 20 Km/h, mientras que para 9 Km/h los resultados son comparables). El grado de servicio obtenido, como se refleja en la figura 4.69(b), es sensiblemente mejor gracias a la expulsión del sistema de las conexiones cuya QoS no puede ser garantizada. Sin embargo, dada la menor capacidad efectiva, no puede considerarse una mejora real.



(a) Tráfico Cursado (b) Grado de Servicio

Figura 4.69. Evaluación del mecanismo de adaptación al entorno: $th_1 = 0{,}09$ y $th_2 = 0{,}03$. Comparativa con los modos básicos de operación. Topología de 50 nodos. Escenario de movilidad variable (velocidad mínima de 1 Km/h y máxima de 40 Km/h).

La aplicación del mecanismo de adaptación, por el contrario, garantiza el mantenimiento de la capacidad efectiva (tráfico correctamente cursado comparable) llegando incluso a observarse cierta mejora para altas tasas de tráfico, como muestra la figura 4.69(a). El grado de servicio, por otra parte, se encuentra próximo al obtenido en el mejor caso – figura 4.69(b), de manera que, globalmente, las aplicaciones experimentan una mejora significativa en las prestaciones obtenidas.

La definición de los umbrales realizada responde al análisis llevado a cabo en las condiciones de estudio propuestas. No obstante, la generalización del escenario ha permitido establecer los mecanismos de estudio y parametrización necesarios de acuerdo a los cuales se puede realizar una configuración adecuada. Los resultados presentados han demostrado tanto la necesidad de adecuarse a la movilidad del entorno como la capacidad de adaptación de la propuesta para conseguir precisamente este objetivo.

Capítulo 5

Conclusiones y líneas futuras de investigación

5.1. Conclusiones

El trabajo realizado en esta tesis, enmarcada en el contexto de la provisión de la calidad de servicio en redes móviles ad hoc (MANETs), se ha centrado en la propuesta y evaluación de una arquitectura *cross-layer* basada en la cooperación entre el encaminamiento con QoS y el nivel MAC. Con esta motivación se ha diseñado un mecanismo de control de admisión distribuido capaz de proporcionar QoS extremo a extremo adaptándose al entorno variante característico de las MANETs. De acuerdo al propósito establecido se han abarcado aspectos relacionados con el encaminamiento en redes ad hoc, los protocolos de acceso al medio, la provisión de QoS en términos de ancho de banda garantizado extremo a extremo, la gestión distribuida de los recursos y la diferenciación de servicios. De igual modo, el dinamismo de los escenarios de aplicación ha motivado el desarrollo de mecanismos de adaptación de los procedimientos de gestión de acuerdo a la movilidad del entorno.

La provisión de calidad de servicio depende directamente de la capacidad de realizar una asignación diferenciada de los recursos entre las aplicaciones de acuerdo a un mecanismo apropiado de control de admisión. El compromiso que se presenta con la eficiencia en la utilización de los recursos en un entorno inalámbrico tiene una solución especialmente compleja en el contexto ad hoc, condicionada por la necesidad de establecer mecanismos de control distribuidos sobre topologías de red variantes.

El papel del encaminamiento en un entorno cooperativo caracterizado por el establecimiento de rutas multisalto es especialmente relevante a la hora de definir los criterios de admisión. El sistema debe garantizar no sólo el establecimiento de un camino de conexión, sino también la satisfacción de los requerimientos demandados. Con esta premisa, el trabajo desarrollado se ha concretado en coordinar el establecimiento de la conectividad necesaria (ruta) con la provisión de la QoS demandada gracias al funcionamiento conjunto del procedimiento de búsqueda con la evaluación de la QoS proporcionada. De este modo, se ha diseñado un procedimiento de admisión fundamentado en las decisiones de encaminamiento basadas en una métrica representativa de la medida de ancho de banda disponible extremo a extremo.

Uno de los condicionantes más relevantes a la hora de definir la disponibilidad real de los recursos, finalmente representada en la nueva métrica de encaminamiento, es la dificultad de realizar una medida fiable de acuerdo a la utilización real de los recursos controlada por el nivel de acceso al medio. La propuesta realizada en este trabajo, lejos de independizar la funcionalidad del encaminamiento y la gestión de los recursos, interrelaciona ambos aspectos mediante una correcta comunicación entre niveles capaz de proporcionar una medida fiable que garantiza una admisión realista.

La falta de una infraestructura fija y la necesidad de implementar procedimientos distribuidos presentan una clara dificultad en el diseño de protocolos específicos para redes ad hoc. La arquitectura propuesta se ha validado primeramente en un escenario de aplicación simplificado, capaz de identificar los problemas asociados a la gestión de recursos, diferenciación de servicios, señalización e intercomunicación entre niveles aportando soluciones específicas. Sin embargo, la adaptación de los esquemas a la variabilidad inherente a la red supone el elemento de diseño más crítico, que ha centrado gran parte del esfuerzo investigador de esta tesis. Un análisis más profundo, en condiciones más realistas de interferencia entre usuarios y movilidad de los terminales, ha permitido diseñar las estrategias de adaptación adecuadas para hacer frente a la variabilidad del entorno.

Si bien existen en la literatura diversas propuestas de encaminamiento con QoS en el entorno de las redes móviles ad hoc, incluyendo alternativas de cooperación entre niveles, el estudio desarrollado en esta tesis ha pretendido identificar de una manera mucho más detallada la problemática asociada a la implementación real de un diseño *cross-layer* específico con objeto de analizar la viabilidad de este tipo de soluciones. El modelado de la arquitectura propuesta a través de un simulador, desarrollado en C++, sobre el que se han implementado y evaluado las diversas estrategias diseñadas a lo largo

de la realización de la tesis, ha permitido identificar los problemas derivados tanto de la propia cooperación, como de su aplicación en escenarios más realistas. La identificación mediante simulación de las necesidades concretas ha facilitado el planteamiento de las soluciones presentadas en este trabajo. Por otra parte, gracias al diseño modular realizado, las fuentes de tráfico, modelos de propagación o patrones de movilidad incluidos pueden ser fácilmente sustituibles en futuras evaluaciones.

Las conclusiones concretas del trabajo realizado pueden desglosarse en tres aspectos diferenciados:

Estudio, caracterización y selección de los algoritmos de encaminamiento y nivel MAC que constituyen una arquitectura de cooperación adecuada.

El estudio llevado a cabo en relación a los protocolos MAC desarrollados en el contexto ad hoc se ha centrado en las estrategias de acceso TDMA, ampliamente analizadas por diversos autores y consideradas especialmente atractivas gracias a la asignación dinámica y la reserva de recursos que proporcionan. Se ha tenido en cuenta como criterio diferenciador el soporte que la capa de acceso debe constituir para la arquitectura *cross-layer* y la necesaria capacidad para proporcionar QoS en el nivel de enlace. La selección realizada no pretende constituir un diseño óptimo en sí mismo sino una cimentación adecuada para la implementación de la arquitectura finalmente propuesta. Las características del protocolo ADHOC MAC en su versión modificada para proporcionar QoS [Gállego, 2007], responde a estas pretensiones, especialmente considerando el diseño del mecanismo de reserva diferenciada de recursos (BIAS) sustentado en la distribución eficiente de información tanto de conectividad como de ocupación de los recursos. Esta funcionalidad resulta particularmente atractiva a la hora de proporcionar una estimación fiable de la disponibilidad de recursos al nivel superior.

Tras el análisis crítico de las múltiples alternativas de encaminamiento estudiadas, el algoritmo AODV se ha considerado un punto de partida apropiado para desarrollar un mecanismo distribuido de control de admisión. Su naturaleza bajo demanda, con una activación inmediata ante el establecimiento de una nueva conexión, y la consecuente difusión de información de encaminamiento a lo largo de la red, se han estimado adecuadas para determinar tanto la disponibilidad de recursos como el impacto sobre las conexiones admitidas, reduciendo el *overhead* asociado a un mantenimiento periódico de la rutas. Su estructura plana, sin jerarquización de las funcionalidades, puede facilitar la generalización del diseño, en la medida que particularizar en protocolos es-

pecíficos lo hace posible, no asociándolo así a un contexto de aplicación determinado y proporcionando mayor flexibilidad al diseño. A pesar de que la propuesta final no ha considerado una estrategia de encaminamiento basada en multicamino, la versión correspondiente, AOMDV, se ha estudiado con objeto de incorporar las facilidades asociadas a la búsqueda de diversas rutas de manera que la definición de una métrica de QoS permita realizar una mejor selección en función del criterio establecido.

Definición de una métrica de encaminamiento basada en el ancho de banda extremo a extremo. Propuesta y evaluación del procedimiento de admisión distribuido.

La asignación de recursos necesaria de acuerdo a los requerimientos establecidos por las aplicaciones se ha concretado en la definición de una métrica de encaminamiento basada en la medida del ancho de banda extremo a extremo. La particularización de dicha medida en el acceso TDMA considerado se ha desarrollado mediante el diseño de un procedimiento de cálculo de los slots disponibles en cada uno de los enlaces a lo largo del camino buscado (algoritmo *BWC-FA*). La consecuente reserva de los recursos demandados durante la confirmación de la ruta seleccionada permite el establecimiento de circuitos virtuales garantizados, constituyendo el mecanismo de admisión distribuido.

La evaluación de la propuesta en un escenario simplificado, donde la conectividad se basa en distancias Euclídeas, desestimando la interferencia de usuarios más allá del rango de cobertura (*Protocol Model*), ha permitido identificar los problemas derivados del carácter compartido del medio de transmisión y la naturaleza distribuida del procedimiento. A pesar de la capacidad el protocolo ADHOC MAC de establecer con exactitud la potencial utilización de los recursos, resolviendo los problemas del terminal oculto y el terminal expuesto, la búsqueda de la máxima disponibilidad se ve dificultada en un escenario de búsquedas de ruta paralelas. En este sentido, se han planteado estrategias aproximadas que solucionan parcialmente la problemática, flexibilizando la posterior reserva de los recursos y garantizando una mayor admisión, especialmente en situaciones de baja carga en el sistema, como ha evidenciado el análisis comparativo realizado al respecto.

La identificación específica de los recursos utilizables en cada enlace posibilita la estimación del ancho de banda evaluado extremo a extremo, mucho más próxima a la realidad que la clásica medida local del enlace "cuello de botella". Una decisión de admisión acorde con esta medida, así como la consecuente reserva real de los recur-

sos especificados, aumentan la probabilidad de admitir las conexiones con garantías de satisfacción en la QoS demandada, como han constatado los resultados obtenidos en condiciones estáticas, donde se ha observado la ganancia obtenida, en términos de capacidad efectiva (tráfico correctamente cursado), gracias al establecimiento de circuitos virtuales garantizados. La ampliación de las condiciones de estudio a un escenario multiservicio ha mostrado la capacidad de diferenciación de la propuesta gracias a la priorización de los circuitos virtuales sobre el tráfico *best-effort*, a pesar de la ligera degradación experimentada como consecuencia de la naturaleza distribuida del procedimiento y los problemas inherentes al medido compartido.

Análisis del impacto de la variabilidad del entorno sobre el control de admisión. Propuesta y evaluación de un mecanismo de readmisión y estrategias de adaptación.

La inclusión de un modelo de conectividad más realista basado en la interferencia real entre usuarios (*Physical Model*) ha permitido evaluar el efecto real de la SIR recibida sobre la fiabilidad de las comunicaciones establecidas así como sobre la propia medida de la disponibilidad. Las inestabilidades en el sistema a consecuencia de una información de estado imprecisa en el nivel MAC han tenido que resolverse desde la perspectiva de dicho nivel, con objeto de seguir constituyendo un soporte fiable para la arquitectura propuesta. Los problemas observados gracias a la valoración de la arquitectura completa han focalizado los puntos débiles del protocolo ADHOC MAC. Las posibles soluciones [Gállego, 2007], concretadas en la modificación adecuada de los procedimientos y la señalización involucrada, así como en la definición de un conjunto de parámetros configurables, se han evaluado desde un punto de vista global, analizando el impacto sobre las prestaciones obtenidas en términos de admisión y tráfico correctamente cursado. Los diversos escenarios de test, analizados en condiciones de interferencia variante (variabilidad en el tráfico) han permitido extraer la parametrización necesaria del protocolo MAC para garantizar la fiabilidad y estabilidad de la propuesta.

La evaluación en condiciones de movilidad ha evidenciado las dificultades para garantizar la QoS ante un escenario de conectividad e interferencia variantes. La modificación de las condiciones iniciales de admisión, resultado de la variabilidad observada, puede producir escenarios de tráfico admitido y utilización de recursos poco eficientes. La nueva topología puede implicar la desconexión y el redescubrimiento de rutas o el

incremento de la interferencia en los circuitos virtuales ya establecidos de manera que no pueda mantenerse el ancho de banda garantizado.

Con objeto de flexibilizar la utilización de los recursos se ha diseñado y evaluado un mecanismo de readmisión capaz de priorizar el mantenimiento de las comunicaciones activas mediante una adecuada redistribución de los recursos expulsando las aplicaciones que no puedan permanecer con la QoS demandada en el nuevo escenario de admisión. Los procedimientos existentes en el propio AODV para identificar roturas en las rutas se han complementado con un procedimiento de monitorización de la QoS basado en el mecanismo *BWC-FA* propuesto para la admisión inicial. La identificación del ancho de banda residual disponible permite realizar reasignaciones inmediatas sobre los caminos establecidos activando la readmisión únicamente en caso de estimar una QoS no recuperable. Con objeto de mejorar las prestaciones en términos de grado de servicio se ha diseñado una estrategia de sobredimensionado de las demandas iniciales (búsqueda de ancho de banda adicional ΔBW) que gracias al aumento en el bloqueo de nuevas admisiones aumenta la capacidad disponible en readmisión reduciendo la expulsión de conexiones activadas. Por otra parte, la priorización flexible realizada en el nivel MAC, permite el trato diferenciado entre aplicaciones activas y en proceso de nueva admisión para garantizar una mayor satisfacción a nivel de usuario.

Una detallada evaluación de las estrategias propuestas ha tratado de identificar la correcta configuración de parámetros del mecanismo completo. Sin embargo, los resultados observados han constatado la clara dependencia de la respuesta del sistema con la naturaleza de las aplicaciones y el grado de movilidad. Si bien no resulta viable una optimización generalizable, han podido establecerse dos modos de operación básicos: funcionamiento en entorno quasi-estático (movilidad reducida) y funcionamiento en movilidad. El *overhead* introducido por los mecanismos de monitorización y readmisión supone una pérdida de capacidad efectiva que únicamente se ve compensada en escenarios de elevada movilidad donde una admisión más limitada y una probabilidad de expulsión representativa de conexiones interferentes permite mejorar sensiblemente las prestaciones de las aplicaciones admitidas, de manera que la activación del procedimiento únicamente se configura en dicho escenario. De igual modo, una infrautilización de los recursos debida a una menor admisión (margen ΔBW) es únicamente ventajosa en estas mismas condiciones.

La influencia determinante del grado de movilidad del entorno sobre las prestaciones obtenidas, y en especial sobre la configuración adecuada de la respuesta del sistema, ha impulsado el diseño de un mecanismo eficiente de adaptación al entorno capaz de

estimar adecuadamente la movilidad modificando en consecuencia los parámetros del sistema. Mediante un análisis del impacto de la movilidad sobre la conectividad de la red (duración de los enlaces, número de vecinos) se ha identificado una métrica representativa basada en la variación del número de vecinos. Gracias al funcionamiento intrínseco del protocolo ADHOC MAC, basado en la identificación de los vecinos estables, se han diseñado un mecanismo de seguimiento y una configuración de umbrales que permiten activar la reacción esperada del sistema.

El planteamiento propuesto no pretende ser el óptimo, especialmente considerando la falta de generalización que supone la implementación particular de la arquitectura, sino que pretende dar una respuesta realista a la necesidad de adaptación a la variabilidad del entorno. Se ha realizado una parametrización basada en la observación de diversos escenarios de estudio que resuelve el compromiso entre una rápida identificación y respuesta ante escenarios de mayor movilidad, con unas prestaciones tan sólo ligeramente inferiores a las deseadas en las situaciones límite entre alta y baja movilidad. La gran ventaja en la implementación de un diseño adaptativo, constatada en los resultados analizados, reside en una flexibilización en la configuración del sistema que permite obtener unas prestaciones razonables en diversos escenarios de aplicación.

5.2. Líneas futuras de investigación

En relación a los tres ámbitos de estudio descritos anteriormente, el trabajo desarrollado en esta tesis sugiere las siguientes líneas futuras:

Arquitectura cross-layer basada en la selección específica de algoritmos

A pesar de la decisión tomada en la selección del encaminamiento, en la que se ha considerado ventajosa la utilización de un protocolo reactivo basado en una métrica vector-distancia, podría realizarse una análisis de la viabilidad de la aplicación de las estrategias propuestas sobre un protocolo de encaminamiento proactivo basado en una métrica de estado (como puede ser el OLSR).

La identificación de las adaptaciones requeridas ante el nuevo planteamiento y el análisis comparativo con la propuesta de partida pueden permitir obtener una caracterización aún más completa de los requerimientos necesarios para garantizar la viabilidad

de una propuesta *cross-layer* genérica. En esta misma línea de trabajo, sería conveniente realizar una comparación realista en cuanto a las prestaciones obtenidas con nuevas propuestas desarrolladas en el mismo ámbito de estudio.

Métrica QoS. Control de admisión distribuido

Las dificultades identificadas durante el desarrollo de la tesis en la utilización eficiente de los recursos, como consecuencia de la naturaleza distribuida del procedimiento de admisión, plantean la posibilidad de realizar estudios en profundidad de nuevas alternativas. Resultaría interesante optimizar la ocupación de la red gracias a una distribución del ancho de banda en caminos complementarios (*BW splitting*). La dificultad de identificar el ancho de banda disponible a través de rutas multicamino plantea como futura línea la adaptación de la propuesta a este nuevo escenario.

Por otra parte, si bien la selección de una métrica basada en ancho de banda responde a la necesidad de establecer una admisión eficiente de acuerdo a la utilización de recursos en el sistema, la inclusión de nuevas métricas dotaría de mayor generalidad a la propuesta, permitiendo realizar un ajuste mucho más adecuado a las necesidades de los diversos escenarios. En este sentido, la realización de medidas energéticas y su aplicación al diseño de encaminamientos eficientes en potencia se presenta especialmente interesante dadas las limitaciones reales de los terminales móviles, lo que supone uno de los puntos críticos de diseño en el contexto ad hoc.

Adaptación a la variabilidad del entorno

En la evaluación realizada en esta tesis se han considerado escenarios realistas en cuanto a la definición de conectividad o la medida de la interferencia entre usuarios. Sin embargo, sería conveniente ampliar los escenarios de estudio mediante la inclusión de nuevos modelos de propagación, añadiendo el efecto de los desvanecimientos lentos y rápidos y modelando patrones de movilidad más realistas.

La inclusión de las modificaciones realizadas en [Gállego, 2007] para garantizar el correcto funcionamiento del protocolo de acceso al medio en presencia de desvanecimientos supone la continuación natural del trabajo realizado. Además, podría evaluarse la mejor reutilización de recursos obtenida mediante la incorporación de nuevas técnicas de control de potencia.

Por otra parte, la utilización de modelos de movilidad más deterministas, más representativos del movimiento natural tanto de personas como de vehículos en escenarios de aplicación específicos (ciudades, oficinas, etc.), pueden servir de base para la aplicación de mecanismos de predicción de movilidad que faciliten el establecimiento de rutas estables que minimicen la pérdida de paquetes gracias a la reducción del número de roturas. La potencial anticipación en la búsqueda de nuevos caminos basada en la detección de futuras roturas podría asimismo incorporarse sobre los procedimientos de readmisión propuestos. De este modo, el diseño realizado de los mecanismos de adaptación a la movilidad del entorno podría complementarse con una estimación mucho más particular que permitiría una mejora todavía más significativa de las prestaciones del sistema.

Apéndice A

ADHOC MAC

ADHOC MAC es un protocolo de acceso al medio basado en acceso síncrono TDMA con un formato de trama estructurada en slots de igual tamaño. Dicha igualdad permite aliviar parcialmente la complejidad de la sincronización. El protocolo implementa una técnica distribuida capaz de establecer dinámicamente un slot *broadcast* fiable para cada terminal activo: *Broadcast CHannel* (BCH). Cada BCH transporta información de señalización hacia sus vecinos a un salto, proporcionando una distribución rápida y fiable de información de conectividad, a nivel de enlace, para todos los terminales. Esta información es el soporte para proporcionar servicios de usuario punto a punto – *Point-To-Point* (PTP) – constituyendo la base para una reserva diferenciada de recursos mediante la gestión eficiente de prioridades. Estas características hacen del ADHOC MAC una estructura apropiada para proporcionar QoS a nivel de enlace, a su vez que atractiva para cooperar con el algoritmo de encaminamiento para garantizar conexiones con QoS extremo a extremo.

A.1. Funcionamiento básico del protocolo

ADHOC MAC funciona sobre una estructura de trama TDMA donde los slots se agrupan en tramas virtuales – *Virtual Frame* (VF) – de longitud N slots, sin necesidad de una alineación de trama entre los terminales (figura A.1). Para un funcionamiento correcto del protocolo se necesita que cada terminal activo tenga asignado un slot BCH fiable, correspondiente con un slot de la VF, en el que transmitir en modo *broadcast* a

todos sus vecinos a un salto. El slot BCH contiene un campo de control[1], denominado *Frame Information* (FI), donde se transmite la información de estado de los N slots que preceden a la transmisión del BCH (el FI puede interpretarse como un vector de N elementos, asociados a la información de cada slot). Dicho estado puede ser OCUPADO o LIBRE: un slot es marcado como OCUPADO por un terminal si éste ha recibido correctamente en él un paquete[2], en cualquier otro caso, el slot está LIBRE. En el caso de marcar el slot como OCUPADO, ha de indicarse también la identidad del terminal transmisor.

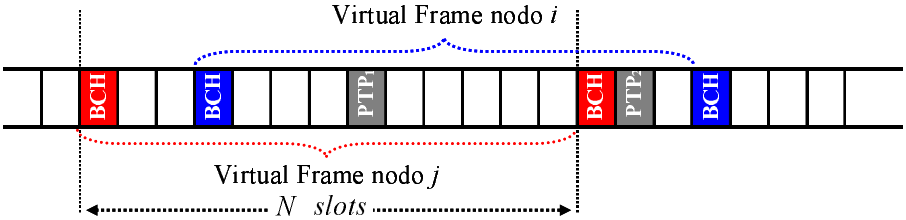


Figura A.1. ADHOC MAC, estructura básica de trama: slots de propósito general (BCH o datos de usuario). Tramas virtuales de N slots.

Basándose en la información recibida en los BCHs de los diversos vecinos (su FI), cada terminal puede marcar un slot, por ejemplo, el slot k, como RESERVADO o DISPONIBLE. Dicho slot estará RESERVADO si el slot $k - N$ está marcado como OCUPADO en al menos uno de los FIs de los BCHs recibidos desde $k - N$ hasta $k - 1$. En caso contrario, el slot estará DISPONIBLE. Si un slot está DISPONIBLE, puede ser usado para intentar nuevos accesos. A la hora de acceder a un slot DISPONIBLE, un terminal j reconocerá, tras N slots (una VF), si la transmisión ha resultado exitosa, si el slot es marcado como "OCUPADO por el terminal j" en todos los FIs recibidos, o fallida en caso contrario.

Una vez que el terminal ha adquirido su BCH, pueden establecerse comunicaciones punto a punto fiables entre los terminales aprovechando la señalización distribuida en los FIs. Cada elemento del vector FI incluye, además de la información de estado y la identificación del transmisor, un flag denominado *Point-To-Point Flag* (PTP_F). Un terminal activa el PTP_F de un slot dado si ha recibido correctamente un paquete *broadcast*, o *unicast* donde el destino era él mismo. A la hora de establecer una comunicación

[1]El BCH puede contener igualmente datos de usuario *broadcast*.

[2]La recepción correcta de un paquete en un slot implica que éste está libre de colisiones. Recibido correctamente, sin embargo, no significa necesariamente ser el destinatario de la información.

punto a punto, además de poder usar todos los slots DISPONIBLES, pueden usarse algunos de los slots RESERVADOS, siempre y cuando se cumpla:

- El PTP_F está señalizado a 0 en todos los FIs recibidos.
- En el FI recibido del terminal destinatario de la comunicación que quiere establecerse, el slot está señalizado como LIBRE.

De este modo, el conjunto de slots $\text{DISPONIBLE}^i_{\text{PTP}}$ identifica a todos los slots que un terminal puede seleccionar para establecer comunicaciones PTP hacia un destino i. Gracias a la utilización del PTP_F se pueden establecer conexiones paralelas compartiendo el mismo slot de transmisión cuando no existe colisión en los receptores, evitando así el problema del terminal expuesto y maximizando en la medida de lo posible (dada la naturaleza distribuida del protocolo) el factor de reutilización de los slots.

A.2. Reparto de recursos para datos de control y de usuario

A la hora de realizar un reparto eficiente de recursos, hay que tener en cuenta que éstos son compartidos por las transmisiones *broadcast* del BCH (obligatorio, desde el punto de vista de un terminal, para considerarse dentro del sistema) y las transmisiones de datos de usuario. En el modo de operación básico, cada slot de la trama puede usarse para cualquier propósito. En esta situación, cuando crece el número de comunicaciones de datos, el número de slots DISPONIBLES para nuevos terminales que intenten acceder al sistema decrece, reduciendo el número de terminales que el sistema puede mantener activos para un número dado de slots. Puesto que la adquisición de un canal básico de *broadcast* es obligatoria para acceder al sistema, un dimensionado apropiado de la red debe garantizar ciertos recursos para las transmisiones BCH si se quiere asegurar un número determinado de usuarios. En este sentido, se considera como métrica para medir las prestaciones del sistema la **probabilidad de bloqueo de nuevos terminales**: *un terminal se bloquea si no adquiere un canal BCH en cierto número de tramas tras el inicio de su actividad*. En esta situación, existe un compromiso entre la minimización de dicha probabilidad de bloqueo (un número suficiente de slots únicamente utilizables para transmisiones BCH) y la maximización de la capacidad disponible para datos (slots reservados para conexiones de usuarios). En [Gállego y otros, 2005a] se ha propuesto la

subdivisión de la estructura de trama (N slots) en dos regiones de slots diferenciadas de acuerdo al propósito de utilización de los mismos (A.1), como se muestra en la figura A.2. Para esta suposición, a diferencia de la propuesta original del ADHOC MAC, se requiere la sincronización temporal a nivel de trama y slot de cada terminal de la red, que puede obtenerse mediante GPS u otras soluciones [Rentel y Kunz, 2005]. Dicha subdivisión, no obstante, permite incrementar la probabilidad de acceso de nuevos terminales al sistema, como se demuestra en [Gállego y otros, 2005a].

$$N = N_{BCH} + N_{DATA} \tag{A.1}$$

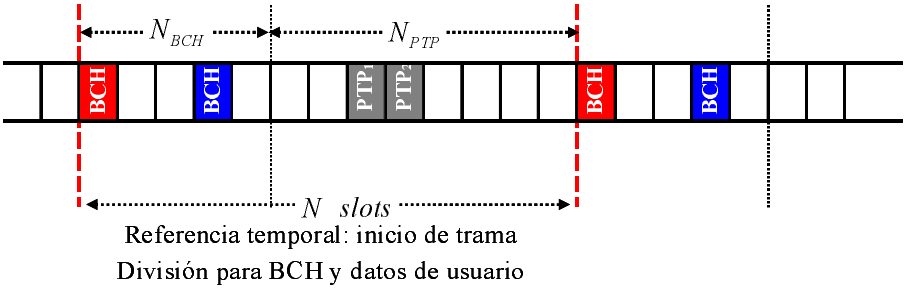


Figura A.2. ADHOC MAC, división estática de la trama: N_{BCH} slots para acceder al sistema, y N_{DATA} slots para datos adicionales (*broadcast* o punto a punto).

El análisis desarrollado por Gállego y otros [2005a] establece la necesidad de flexibilizar la separación de recursos para su utilización de un modo más eficiente. El empleo de una subdivisión estática, con un N_{BCH} fijo, limita la densidad máxima de terminales que soporta el sistema. Esto conduce, por un lado, a una infrautilización de recursos ante densidades de usuarios menores a las estimadas (slots reservados para BCH podrían aumentar la capacidad reducida de PTP). Por otra parte, una subestimación respecto a la demanda de nuevos usuarios incrementa la probabilidad de bloqueo que podría solventarse reduciendo la capacidad PTP a costa de admitir más usuarios. Para superar estas limitaciones, se propone una división adaptativa de la trama, en la que los terminales son capaces de adaptar los límites de su trama a la variabilidad de la densidad de terminales en sus alrededores.

A pesar de la ventaja que supone una solución adaptativa, especialmente teniendo en cuenta la naturaleza dinámica de las redes móviles ad hoc [Gállego y otros, 2005a], se ha optado por una subdivisión estática de la trama, con un valor N_{BCH} capaz de soportar las densidades de usuarios con las que se ha sido evaluada la propuesta final. El propósito de esta simplificación reside en centrar los esfuerzos de desarrollo en solucionar

los problemas reales derivados de la cooperación entre niveles, adaptación al entorno o gestión de recursos en un escenario de terminales activos, sin la limitación impuesta por la posibilidad de bloqueo de ninguno de ellos. El modelo de red representa, de este modo, a un conjunto de N_t terminales activos (con su BCH correspondiente), siendo N_{BCH} un número de slots suficiente para garantizar dicha actividad.

En el funcionamiento básico del protocolo, a la hora de establecer una comunicación PTP con uno de sus vecinos, un terminal decide en qué slot desea transmitir, de entre los $\text{DISPONIBLES}_{\text{PTP}}$ para dicho destino, e intenta acceder directamente en él, en la siguiente trama virtual (RR-ALOHA). Dado que diferentes terminales pueden intentar concurrentemente acceder al mismo slot (elección de slots aleatoria), pueden darse colisiones que reducen las prestaciones del mecanismo de acceso.

Con objeto de limitar las colisiones de acceso y garantizar una reserva asegurada de los recursos, en [Gállego y otros, 2005c] se propone un método de acceso basado en la petición del recurso en lugar del acceso directo, lo que se ha denominado *Book In Advance Scheme* (BIAS). Dicho esquema aprovecha la capacidad del BCH para distribuir eficientemente información de control entre los vecinos a un salto. Se propone la utilización de dicho slot BCH para señalizar la petición de la reserva antes del acceso.

A.3. Reserva de recursos PTP. Método de acceso BIAS

El esquema de acceso BIAS se apoya en la inclusión de un nuevo estado, ACCEDIENDO, además de los anteriormente mencionados de OCUPADO y LIBRE. Dicho estado, de una manera análoga al estado OCUPADO, señaliza el terminal transmisor, y a su vez añade el destino de la comunicación (la petición). En estas condiciones, cuando un terminal distinto al destinatario recibe diversas peticiones para un mismo recurso PTP en su trama virtual debe decidir, de un modo coherente con el resto de los vecinos[3], para qué terminal transmisor señalizar el slot como ACCEDIENDO en su siguiente FI (es decir, a quién "conceder" la reserva, resolviendo así la colisión). Únicamente el destinatario señalizará el PTP_F a 1 en su FI. Finalmente, el terminal transmisor, digamos j, podrá considerar su reserva exitosa si se cumple:

[3]Realizar una decisión coherente entre los diferentes vecinos en el entorno de un salto únicamente supone haber establecido el mismo criterio de selección (orden de llegada de las peticiones, dirección de los transmisores, número aleatorio asociado a la petición del transmisor) [Gállego, 2007].

- El slot está marcado como "ACCEDIENDO por el nodo j" en el FI del destinatario, con el PTP_F a 1.
- El PTP_F de dicho slot está a 0 en los FIs del resto de terminales.

En cualquier otro caso, la reserva se considerará fallida y deberá reintentarse. Si la reserva ha concluido satisfactoriamente, el transmisor puede señalizar el slot como OCUPADO en el FI de su siguiente BCH, estando en condiciones de transmitir, libre de colisiones, en la siguiente VF. La figura A.3 muestra un ejemplo simplificado del procedimiento descrito.

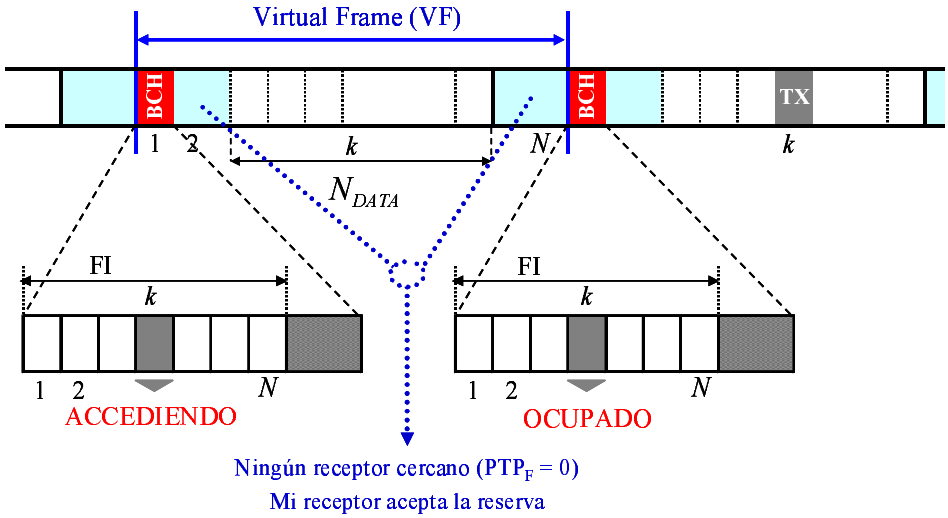


Figura A.3. Reserva previa de recursos PTP. Acceso BIAS.

Con el soporte de este método de acceso, puede realizarse una reserva diferenciada de recursos, de acuerdo a ciertos requerimientos de QoS, basándose en el uso de prioridades. Cada elemento del vector de slots FI incorpora un campo adicional relativo a la prioridad del servicio. Dicho campo permite, por una parte, que a la hora de resolver conflictos en el acceso, la reserva sea favorable para aquellos servicios más prioritarios. Por otro lado, servicios prioritarios pueden realizar peticiones de acceso sobre recursos PTP ya reservados de menor prioridad, es decir, existe la posibilidad de reasignar recursos reservados como menos prioritarios para servicios de mayor prioridad. Esto sólo será posible, siempre y cuando se cumpla:

- El PTP_F está señalizado a 0, ó a 1, pero con prioridad menor, en todos los FIs recibidos.

- El FI recibido del destinatario indica que el slot está LIBRE, u OCUPADO o ACCEDIENDO de prioridad menor.

La definición de estas reglas de acceso es generalizable para cualquier número de prioridades (todo aquel servicio de prioridad mayor puede utilizar recursos ya reservados de cualquiera de las prioridades inferiores). Considerando un conjunto PR de prioridades, de dimensión $N_p = |PR|$:

$$PR = \{p_i, 0 \leq i \leq N_p\} \tag{A.2}$$

se ha adoptado el criterio de numeración creciente en orden inverso a la prioridad del servicio, de manera que se cumple:

$$p_{N_p} \leq ... \leq p_1 \leq p_0 \tag{A.3}$$

En cualquier caso, a la hora de establecer correctamente los criterios de admisión será importante tener en cuenta la prioridad que se desea conceder a las nuevas conexiones a la hora de definir qué recursos ya utilizados pueden ser reasignados para darles servicio. Puesto que la decisión que finalmente va a determinar dicha admisión se basa en la posibilidad de encontrar un camino con ancho de banda suficiente, la medida de dicha disponibilidad debe tener en cuenta todos los potenciales recursos, incluyendo aquéllos asignados a servicios menos prioritarios y susceptibles de reasignación. En definitiva, es necesario redefinir los dos conjuntos base (SRT_i y SRR_i)[4] utilizados en la medida de slots disponibles en el enlace (i, j) (§3.1.3). El conjunto TS_i de slots reservados para transmitir[5] en el nodo n_i puede descomponerse de acuerdo a las prioridades utilizadas, definiendo así TS_i^{pr} como el conjunto de slots que el nodo i tiene reservados con prioridad pr. De igual modo puede hacerse la distinción de los slots ocupados en recepción[6], diferenciando cada posible RS_i^{pr}. De acuerdo a esta nueva clasificación, la disponibilidad de recursos vendrá condicionada por la prioridad del nuevo servicio a admitir (pr) (A.4) – (A.5).

$$SRT_i^{pr} = \left\{ s_k \in S : s_k \notin TS_i^p, s_k \notin RS_i^p, s_k \notin \bigcup_{n_j \in NB_i} RS_j^p, \forall p \geq pr \right\} \tag{A.4}$$

[4] *Slots Ready to Transmit* (SRT), *Slots Ready to Receive* (SRR).
[5] *Transmission Schedule* (TS) para el nodo i.
[6] *Reception Schedule* (RS) para el nodo i.

$$SRR_i^{pr} = \left\{ s_k \in S : s_k \notin TS_i^p, s_k \notin RS_i^p, s_k \notin \bigcup_{n_j \in NB_i} TS_j^p, \forall p \geq pr \right\} \tag{A.5}$$

Gracias al método de acceso BIAS, una reasignación de un recurso ocupado realmente consiste en realizar la petición del recurso pese al conocimiento de su ocupación de manera que en aplicación de las reglas de competencia todos los vecinos determinen la reserva como efectiva para la nueva comunicación. Dado que una reserva previamente establecida no proporciona una información completa del potencial estado de dicho slot si en el futuro la comunicación se detuviera, el resultado del intento de reasignación no tiene por qué concluir correctamente. No obstante, considerar dichos recursos como potencialmente libres puede incrementar la probabilidad de admisión de nuevas conexiones.

A.4. Disponibilidad de recursos. Información básica para el control de admisión.

Cada vez que un terminal i envía su slot BCH con la información actualizada del FI, señaliza en éste cómo ha resuelto los posibles conflictos de reservas, cómo ha visto ocupados los slots durante la última trama virtual (VF), y sus propias reservas confirmadas. Para un slot determinado k, se define su estado como la quíntupla $FI_i^k = \langle st, id_{TX}, pr, PTP_F, id_{RX} \rangle$ (Tabla A.1). Para hacer referencia a un subconjunto de elementos concreto, por ejemplo el par estado y prioridad, se usará la notación $FI_i^k\,[st, pr] = \langle st, pr \rangle$.

De acuerdo al procedimiento de acceso BIAS, una reserva solicitada en la trama virtual previa $(m-1)$:

$$FI_i^k\big|_{VF=m-1}\,[st, id_{TX}, pr, id_{RX}] = [\text{ACCEDIENDO}, i, p, j] \tag{A.6}$$

se verá confirmada en la trama m siempre y cuando el destino de la comunicación acepte la reserva:

$$FI_j^k\big|_{VF=m-1}\,[st, id_{TX}, pr, PTP_F] = [\text{ACCEDIENDO}, i, p, 1] \tag{A.7a}$$

no exista ningún receptor entre los vecinos a un salto:

$$FI_n^k\big|_{VF=m-1}[PTP_F] = 0, \forall n \neq j, n \in NB_i \tag{A.7b}$$

y el propio nodo fuente no deba aceptar, en aplicación de las reglas del protocolo, una conexión más prioritaria como destino:

$$\nexists n \in NB_i : FI_n^k\big|_{VF=m-1}[st, pr, id_{RX}] = [ACC, p_n, i], p_n > p \tag{A.7c}$$

Como resultado, el BCH transmitido en la trama m incluirá la información relativa a la reserva efectiva (A.8), pudiendo finalmente transmitir los datos de usuario en el slot k de dicha trama (subtrama de datos).

$$FI_i^k\big|_{VF=m}[st, id_{TX}, pr, id_{RX}] = [\text{OCUPADO}, i, p, j] \tag{A.8}$$

Tabla A.1. Estado del slot k visto por el nodo i, FI_i^k

st^a	id_{TX}	pr	PTP_F	id_{RX}	
LIBRE	-	-	0	-	...
OCUPADO	i	pr	0	j	i transmite en k
OCUPADO	$n \neq i$	pr	0	j	n transmite en k
OCUPADO	$n \neq i$	pr	1	i	i recibe en k
ACCEDIENDO	i	pr	0	j	i quiere transmitir a j en k
ACCEDIENDO	$n \neq i$	pr	0	j	n podrá transmitir a j en k^b
ACCEDIENDO	$n \neq i$	pr	1	i	i acepta la reserva de n en k

[a] st = estado del slot, id_{TX} = nodo transmisor, pr = prioridad de la reserva, PTP_F = flag PTP, id_{RX} = nodo receptor

[b] El nodo i puede haber resuelto un conflicto en la reserva de k (varios transmisores) a favor de n

Tras actualizar su "visión" del entorno (ocupación del canal de transmisión), un terminal debe a su vez tener en cuenta la información recibida en todos los FIs de dicha VF para añadir a dicha "visión" la existencia de receptores cercanos (puesto que van a limitar su disponibilidad para transmitir). Así pues, antes de analizar sus necesidades de nuevos recursos, con objeto de poder determinar la disponibilidad para esas futuras reservas, debe tener una idea clara de su entorno, lo que se ha denominado $Status-Channel$ (SC). Así pues, la definición del SC depende del estado percibido por

el propio terminal, $FI_i^k, \forall k \in N$, y de la información recibida de sus vecinos, $FI_n^k, \forall k \in N, \forall n \in NB_i$. Como resultado, el SC_i^k representa el estado del slot k, percibido por el terminal i, definido como $SC_i^k = \langle st, pr \rangle$, donde hay que destacar el significado de pr como la máxima prioridad para la que tiene significado el estado st, el cual representa la situación más restrictiva a la hora de definir la disponibilidad de un slot. La tabla A.2 muestra las reglas con las que éste se define.

Como resultado de la definición del SC puede determinarse la disponibilidad de recursos para nuevas reservas, lo que se resume en la tabla A.3. En dicha tabla, se definen los slots disponibles para transmitir a un destino desconocido, como aquellos slots en los que un determinado nodo i podría potencialmente transmitir, teniendo en cuenta las prioridades (a falta de conocer el destino). En el caso de los servicios con QoS, se incluyen tanto los DISPONIBLES como los REASIGNABLES, lo que constituye el SRT_i definido en (A.4). En el caso del tráfico *best-effort* sólo se consideran los DISPONIBLES, es decir el SRT_i definido en (3.2). De igual modo, en recepción, el SRR_i para tráfico prioritario está constituido por los DISPONIBLES y los REASIGNABLES – definición (A.5), mientras que únicamente los DISPONIBLES son considerados para *best-effort* – definición (3.3). La tabla refleja, por otra parte, la disponibilidad específica en transmisión para un destino concreto, ya que una vez que es necesario hacer efectiva la reserva de recursos, la definición de DISPONIBLE o REASIGNABLE depende realmente del nodo concreto al que se quiere transmitir. En cualquier caso, la disponibilidad mostrada en dicha tabla se corresponde con los recursos que un terminal no está empleando directamente como transmisor o receptor. En cualquiera de estos casos (estados RSVT y RSVR respectivamente, descritos en la tabla A.2), es necesario definir un SC secundario que recoja el estado del entorno de dicho terminal, para poder determinar la disponibilidad de estos recursos, según los criterios de la tabla A.3, en caso de que deban ser reasignados.

Tabla A.2. $Status - Channel$ del slot k en el terminal i

$\boldsymbol{SC_i^k}$	$\boldsymbol{pr}$	**Regla de decisión**
Información extraída del FI_i^k		
LIBRE	-	$FI_i^k\,[st] = \text{FREE}$
RSVT[a]	pr_i	$FI_i^k\,[st, id_{TX}, pr] = [\text{OCUPADO}, i, pr_i]$
RSVR	pr_i	$FI_i^k\,[st, id_{TX}, pr, PTP_F] = [\text{OCUPADO}, n, pr_i, 1]$ $n \neq i$
TX	pr_n	$FI_i^k\,[st, id_{TX}, pr, PTP_F] = [\text{OCUPADO}, n, pr_n, 0]$ $n \neq i$
PRE_TX	pr_n	$FI_i^k\,[st, id_{TX}, pr, PTP_F] = [\text{ACCEDIENDO}, n, pr_n, 0]$ $n \neq i$
Información actualizada de acuerdo a los $FI_n^k,\ n \neq i$ (Actualización realizada sobre slots no RSVT/RSVR)		
RX(0)	$\text{máx}\,(pr_n)$	$FI_j^k\,[PTP_F] = 0\ \forall j \in NB_i$ $\exists n : FI_n^k\,[st, id_{TX}] = [\text{OCUPADO}, m]\,,\ n \neq m$ $n \neq i$
PRE_RX(0)	$\text{máx}\,(pr_n)$	$FI_j^k\,[PTP_F] = 0\ \forall j \in NB_i$ $\exists n : FI_n^k\,[st, id_{TX}] = [\text{ACCEDIENDO}, m]\,,\ n \neq m$ $n \neq i$ $\forall x : FI_x^k\,[st] = \text{OCUPADO},\ FI_x^k\,[id_{TX}] = x$
RX(1)	$\text{máx}\,(pr_n)$	$\exists n : FI_n^k\,[st, PTP_F] = [\text{OCUPADO}, 1]\,,$ $FI_n^k\,[id_{TX}] \neq i,\ n \neq i$
PRE_RX(1)	$\text{máx}\,(pr_n)$	$\exists n : FI_n^k\,[st, PTP_F] = [\text{ACCEDIENDO}, 1],$ $FI_n^k\,[id_{TX}] \neq i,\ n \neq i$ $\forall x : FI_x^k\,[st] = \text{OCUPADO},\ FI_x^k\,[id_{TX}] = x$

[a]RSVT: reservado para transmitir; RSVR: reservado para recibir; TX: existen transmisores cerca; PRE_TX: existen peticiones de transmisión cerca; RX(0): existen nodos cercanos que oyen transmisiones para otros; PRE_RX(0): existen nodos cercanos que oyen peticiones para otros; RX(1): existen receptores cerca; PRE_RX(1): existen futuros receptores cerca.

Tabla A.3. Disponibilidad de recursos a partir del $Status - Channel$ (SC)

<table>
<tr><td colspan="2"></td><td colspan="7">DISPONIBILIDAD sobre el slot $\boldsymbol{k}$ con prioridad $\boldsymbol{pr}$</td></tr>
<tr><td colspan="2"></td><td colspan="6">TRANSMITIR</td><td rowspan="3">RECIBIR</td></tr>
<tr><td colspan="2"></td><td>broadcast</td><td colspan="2">unicast</td><td colspan="3">unicast</td></tr>
<tr><td>$\boldsymbol{SC_i^k}$</td><td>prior</td><td>todos destinos</td><td colspan="2">destino desconocido[a]</td><td colspan="3">destino j</td></tr>
<tr><td>LIBRE</td><td rowspan="7">pr_{SC}</td><td>DISP[b]</td><td colspan="2" rowspan="4">$\mathrm{DISP}_{\mathrm{PTP}}$</td><td colspan="3">$\mathrm{DISP}^j_{\mathrm{PTP}}$</td><td>DISP</td></tr>
<tr><td>TX[c]</td><td rowspan="4">RSV
$pr < FI_n^k\,[pr]$
$\forall n \in NB_i$</td><td rowspan="2">$FI_j^k\,[status] = \mathrm{LIBRE}$
$\mathrm{DISP}^j_{\mathrm{PTP}}$</td><td colspan="2">$FI_j^k\,[status] = \mathrm{OCU/ACC}$</td><td rowspan="2">REAS
$pr > FI_i^k\,[pr]$</td></tr>
<tr><td>PRE_TX</td><td>$pr > FI_j^k\,[pr]$
REAS</td><td>$pr \leq FI_j^k\,[pr]$
RSV</td></tr>
<tr><td>RX(0)</td><td rowspan="2">$FI_j^k\,[status] = \mathrm{LIBRE}$
$\mathrm{DISP}^j_{\mathrm{PTP}}$</td><td colspan="2">$FI_j^k\,[status] = \mathrm{OCU/ACC}$</td><td rowspan="4">DISP</td></tr>
<tr><td>PRE_RX(0)</td><td>$pr > pr_{SC}$
REAS</td><td>$pr \leq pr_{SC}$
RSV</td></tr>
<tr><td>RX(1)</td><td>$pr > pr_{SC}$
REAS</td><td>$pr \leq pr_{SC}$
RSV</td><td>$pr > pr_{SC}$
REAS</td><td colspan="2">$pr \leq pr_{SC}$
RSV</td></tr>
<tr><td>PRE_RX(1)</td><td colspan="2"></td><td colspan="2"></td></tr>
</table>

[a]Evaluación genérica de la disponibilidad para transmisiones punto a punto, sin especificar destino. Su uso se especifica en el apartado §3.3.2.3.

[b]DISP = DISPONIBLE, REAS = REASIGNABLE, RSV = RESERVADO, OCU = OCUPADO, ACC = ACCEDIENDO

[c]Estados definidos en la tabla A.2.

El control de admisión implementado se basa en la medida de la disponibilidad trasladada del nivel MAC al encaminamiento, para evaluar el ancho de banda existente en un posible camino extremo a extremo para la comunicación a admitir. Como se describe en §3.3.2, dicha medida se basa en la correcta definición de los conjuntos SRT y SRR que, según las reglas específicas del protocolo ADHOC MAC y de acuerdo a su definición en el $Status - Channel$, pueden resumirse en las siguientes ecuaciones:

$$\begin{aligned} SRT_i^p = \{k \in N :& \forall n \in NB_i, FI_n^k\left[PTP_F\right] = 0, \\ & \forall n : FI_n^k\left[st\right] = (\text{OCUP.} \vee \text{ACCED.}), \\ & FI_n^k\left[id_{TX}\right] = n, \\ & \forall n : FI_n^k\left[id_{TX}\right] = m \neq n, FI_n^k\left[pr\right] < p\} \end{aligned} \tag{A.9}$$

$$\begin{aligned} SRR_i^p = \{k \in N :& \forall n \in NB_i : FI_n^k\left[st\right] = (\text{OCUP.} \vee \text{ACCED.}), \\ & FI_n^k\left[id_{TX}\right] = m \neq n, \\ & \forall n : FI_n^k\left[id_{TX}\right] = n, FI_n^k\left[pr\right] < p\} \end{aligned} \tag{A.10}$$

Apéndice B

Funciones para el cálculo del ancho de banda

Algoritmo B.1 Tomar n slots del conjunto IN

function $OUT = BW_1(IN, n)$
  $assert\,(n \leq |IN|)$
  elegir en OUT n elementos de IN aleatoriamente
  return
end function

Algoritmo B.2 Dados IN_1 y IN_2, determinar los subconjuntos correspondientes OUT_1 y OUT_2 disjuntos.

function $(OUT_2, OUT_1) = BW_2(IN_2, IN_1)$
  $C = IN_1 \cap IN_2$
  $E_1 = IN_1 \cap \overline{IN_2}$
  $E_2 = IN_2 \cap \overline{IN_1}$
  if $|E_2| \geq |IN_1|$ **then**
    $OUT_2 = BW_1\,(E_2, |IN_1|)$
    $OUT_1 = IN_1$
    return
  else if $|E_1| \geq |IN_2|$ **then**
    $OUT_1 = BW_1\,(E_1, |IN_2|)$
    $OUT_2 = IN_2$
    return
  else
    $T = floor\,(|IN_1 \cup IN_2|/2)$
    $C_2 = BW_1\,(C, T - |E_2|)$
    $C_1 = C \cap \overline{C_2}$
    $OUT_1 = BW_1\,(C_1 \cup E_1, T)$
    $OUT_2 = BW_1\,(C_2 \cup E_2, T)$
    return
  end if
  return
end function

Algoritmo B.3 Dados IN_1, IN_2 y IN_3, determinar los subconjuntos correspondientes OUT_1, OUT_2 y OUT_3 disjuntos.

function $(OUT_3, OUT_2, OUT_1) = BW_3\ (IN_3, IN_2, IN_1)$
  $assert\,(|IN_3| = |IN_2| \,\&\&\, IN_2 \cap IN_3 = \emptyset)$
  $C_{21} = IN_2 \cap IN_1$
  $C_{31} = IN_3 \cap IN_1$
  $E_1 = IN_1 \cap \overline{C_{21}} \cap \overline{C_{31}}$
  $E_2 = IN_2 \cap \overline{C_{21}}$
  $E_3 = IN_3 \cap \overline{C_{31}}$
  if $|E_1| \geq |IN_2|$ **then**
    $OUT_1 = BW_1\,(E_1, |IN_2|)$
    $OUT_2 = IN_2$
    $OUT_3 = IN_3$
    return
  else if $|E_3| \geq |BW_2\,(IN_2, IN_1)|$ **then**
    $(OUT_2, OUT_1) = BW_2\,(IN_2, IN_1)$
    $OUT_3 = BW_1\,(E_3, |OUT_1|)$
    return
  else if $|E_2| \geq |BW_2\,(IN_3, IN_1)|$ **then**
    $(OUT_3, OUT_1) = BW_2\,(IN_3, IN_1)$
    $OUT_2 = BW_1\,(E_2, |OUT_1|)$
    return
  else
    $T = floor\,(|IN_1 \cup IN_2 \cup IN_3|/3)$
    $C_{31}^3 = BW_1\,(C_{31}, T - |E_3|)$
    $C_{31}^1 = C_{31} \cap \overline{C_{31}^3}$
    $C_{21}^2 = BW_1\,(C_{21}, T - |E_2|)$
    $C_{21}^1 = C_{21} \cap \overline{C_{21}^2}$
    $OUT_1 = BW_1\,(C_{21}^1 \cup C_{31}^1 \cup E_1, T)$
    $OUT_2 = E_2 \cup C_{21}^2$
    $OUT_3 = E_3 \cup C_{31}^3$
    return
  end if
  return
end function

Apéndice C

Modelo de Movilidad

Existen multitud de modelos de movilidad desarrollados en el contexto de la evaluación de redes móviles ad hoc [Camp y otros, 2002]. Uno de los modelos más ampliamente utilizado en este entorno, convertido prácticamente en un estándar *de facto* es el *Random WayPoint mobility model* (RWP) [Broch y otros, 1998; Johnson y Maltz, 1996], que se comporta del modo siguiente:

Inicialmente, todos los terminales se posicionan de manera aleatoria en el área de simulación. Cada terminal comienza permaneciendo estático en dicha posición inicial durante un cierto periodo de tiempo denominado "Tiempo de pausa" (T_{pausa}), habitualmente constante para todos los terminales. Una vez que expira dicho tiempo, el terminal elige un destino aleatorio dentro del área de simulación así como una velocidad que se encuentra uniformemente distribuida entre V_{min} y V_{max}, velocidades mínima y máxima de desplazamiento, y se dirige en línea recta hacia dicho destino a la velocidad seleccionada. Tras alcanzar el destino, el terminal permanece en dicha posición un tiempo T_{pausa} antes de iniciar de nuevo el mismo proceso.

Puede demostrarse [Lin y Huang, 2004] que la velocidad media de los terminales en la red no resulta del valor medio V_{min} y V_{max} sino que viene dada por :

$$V_{media} = \frac{V_{max} - V_{min}}{\ln\left({V_{max}}/{V_{min}}\right)} \tag{C.1}$$

Este comportamiento se justifica por el hecho de que los terminales que se desplazan con velocidades más lentas tardan más tiempo en alcanzar su destino y detenerse y, por tanto, tienen más peso en la determinación de la velocidad media.

Este modelo presenta dos problemas fundamentales: por un lado, su elevado tiempo de convergencia, que impone tiempos de simulación muy grandes para limitar el sesgo de los resultados obtenidos [Casilari y Triviño, 2005; Resta y Santi, 2002; Yoon y otros, 2003] y por otro, la falta de realismo de los patrones de movimiento que describen los terminales, debido fundamentalmente a la gran simplicidad del modelo.

Si bien existen propuestas que pretenden emular entornos de aplicación más cercanos a la realidad, como se describe en [Bai y otros, 2003; Jardosh y otros, 2005] (desplazamiento en ciudades, entornos de oficinas, etc.), el modelo RWP presenta también dos importantes ventajas respecto a modelos más sofisticados: la propia simplicidad del mismo facilita la parametrización, ya que básicamente el movimiento queda totalmente caracterizado mediante tres valores, T_{pausa}, V_{min} y V_{max}. Además, al no estar enfocado a ningún escenario de aplicación concreto, su alto grado de abstracción lo convierte en un modelo generalista, útil para evaluar en primera instancia nuevas propuestas de protocolos o mecanismos de gestión de recursos en redes ad hoc.

A continuación, se propone un modelo de movilidad basado en el RWP (RWP modificado (RWPM)) que trata de mantener las ventajas del mismo en cuanto a simplicidad y generalidad, pero con ciertas modificaciones que pretenden reducir la aleatoriedad en la trayectoria de los terminales, proporcionando una aproximación más cercana al movimiento natural tanto de personas como de vehículos.

Se mantienen los tres parámetros fundamentales del modelo RWP (T_{pausa}, V_{min} y V_{max}), pero se incorporan dos grados de libertad adicionales que permiten introducir cierta dependencia en la trayectoria de los móviles (efecto memoria). El destino de cada nuevo movimiento no se elige de manera totalmente aleatoria dentro del área de simulación, sino que debe encontrarse a una distancia del punto de partida uniformemente distribuida entre 0 y R_{max}. Además, el ángulo de variación sobre la dirección del último movimiento realizado por el terminal debe estar uniformemente distribuido entre $-\alpha_{max}$ y α_{max}, es decir, se limita el ángulo de variación entre movimientos consecutivos a un ángulo α_{max}.

La figura C.1 ilustra este comportamiento. En el caso de que recorrer la distancia seleccionada en el ángulo correspondiente implique salir del área de simulación, el terminal rebota en sentido contrario hasta que recorre la distancia total asignada.

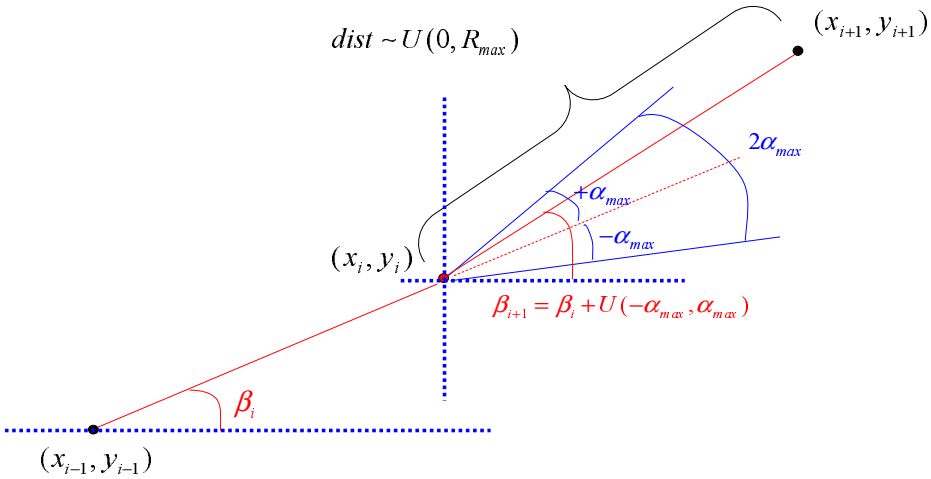


Figura C.1. Modelo de movilidad RWP modificado con parámetros R_{max} y α_{max}.

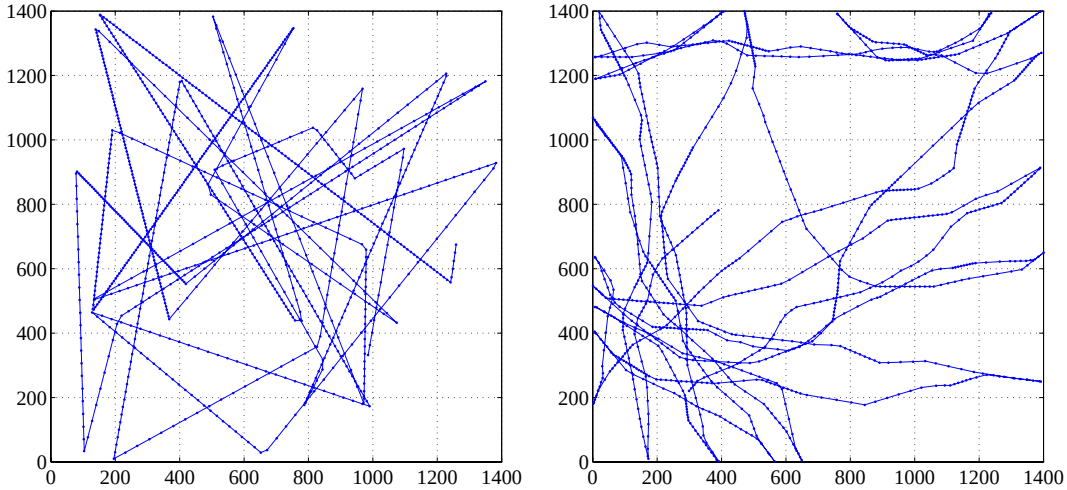
Debido al alto tiempo de convergencia del modelo RWP, que condiciona la longitud de las simulaciones para poder obtener resultados fiables, se ha estudiado también la velocidad de convergencia del modelo RWPM propuesto para determinar tiempos de simulación adecuados para los distintos escenarios que se evalúan en este trabajo.

Las velocidades medias analizadas en la memoria de 2, 9 y 20 Km/h se corresponden con unos escenarios en los que las velocidades mínima (V_{min}) y máxima (V_{max}) toman valores de 1 y 3, 3 y 20, y 10 y 40 Km/h respectivamente. El tiempo de pausa, T_{pausa}, es de 2 segundos. Los terminales se desplazan dentro de una región cuadrada con una superficie de 2 Km^2. El ángulo máximo de variación entre movimientos, α_{max}, se ha fijado en $\pi/12$, mientras que el parámetro R_{max}, correspondiente al desplazamiento máximo en una misma dirección, se ha definido como una fracción de la distancia máxima D_{max} que puede recorrerse dentro de la cuadrícula:

$$R_{max} = 0{,}1 \cdot D_{max} = 0{,}1 \cdot \sqrt{X_{max}^2 + Y_{max}^2} \tag{C.2}$$

donde X_{max} e Y_{max} definen los límites del escenario.

La figura C.2 muestra un ejemplo del movimiento seguido por un terminal mediante los dos modelos de movilidad, RWP y RWPM, con una velocidad media de 9 Km/h. Puede comprobarse que la trayectoria descrita en el caso del RWP presenta unas variaciones muy bruscas en la dirección de los movimientos, que son claramente suavizadas en el modelo RWPM, lo que permite una aproximación más cercana a un movimiento natural.



(a) RWP (b) RWPM

Figura C.2. Ejemplo de la trayectoria descrita por un terminal siguiendo los dos modelos de movilidad, RWP y RWPM, con una velocidad media de 9 Km/h. $R_{max} = 200$ m. $\alpha_{max} = \pi/12$. 50 terminales.

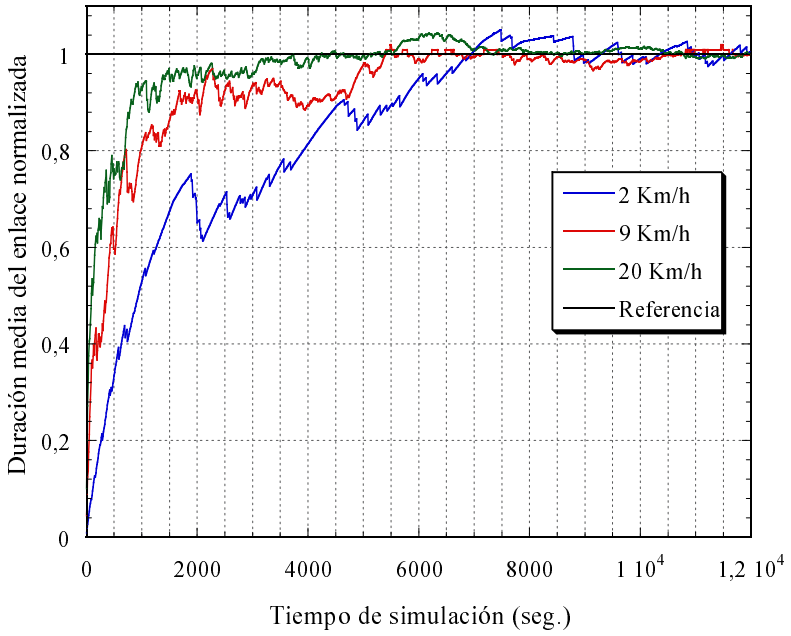
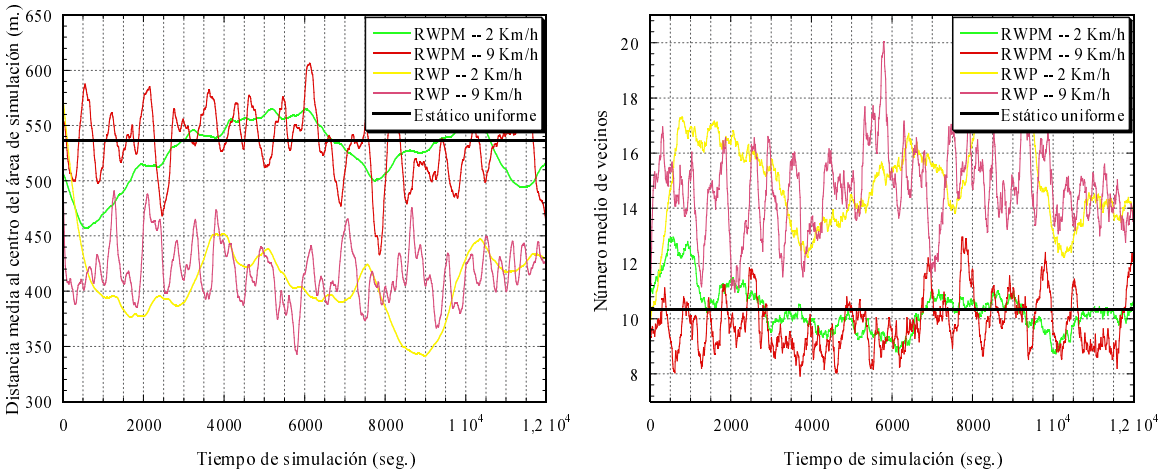


Figura C.3. Evolución de la duración media de los enlaces para distintas velocidades de desplazamiento con el modelo RWPM. $R_{max} = 200$ m. $\alpha_{max} = \pi/12$. 50 terminales.

En [Boleng y otros, 2002] se justifica que la duración de los enlaces es una de las métricas más eficientes para representar la movilidad de una red ad hoc. Definimos la duración de un enlace como el tiempo que transcurre desde que un nodo entra en el radio de cobertura de otro hasta que vuelve a escapar del mismo. La figura C.3 muestra la evolución de la duración media de los enlaces para las tres velocidades de análisis con el modelo RWPM. Puede comprobarse que a medida que aumenta la velocidad de los terminales, más rápida es la convergencia y, por tanto, menor el tiempo de simulación requerido para alcanzar un régimen estacionario. De acuerdo con estos resultados se han escogido unos tiempos de simulación de 6000, 9000 y 12000 segundos para las velocidades de 2, 9 y 20 Km/h. En la figura C.3 se observa que con estos valores, la duración media de los enlaces ha logrado alcanzar el estacionario para las tres velocidades.

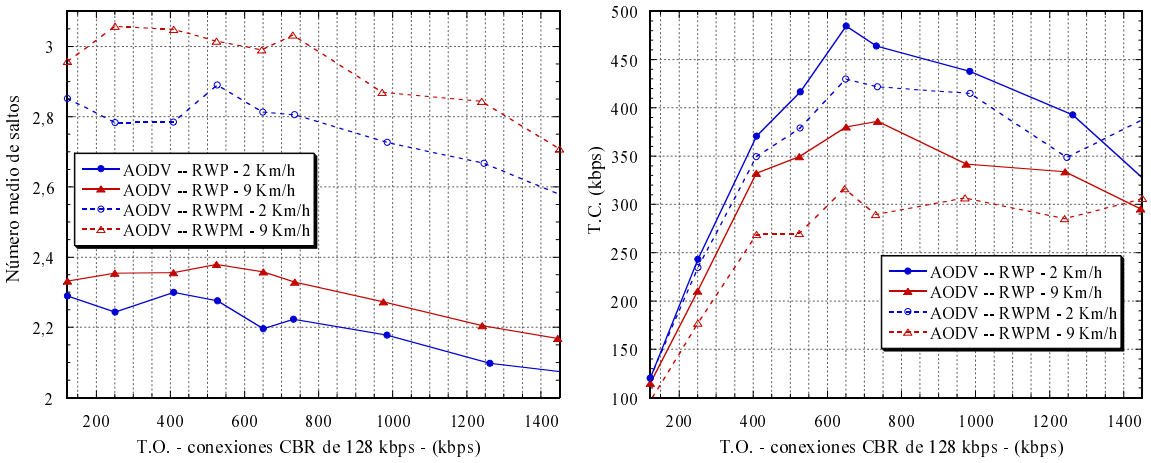
Otra de las características peculiares del modelo de movilidad RWP es que las posiciones de los nodos en el régimen estacionario no se corresponden con una distribución uniforme, que es la que se considera habitualmente en la evaluación en escenarios estáticos, sino que los terminales tienden a concentrarse en la región central del área de simulación. Este fenómeno tiene implicaciones importantes en las prestaciones de los protocolos de encaminamiento, puesto que se reduce el número medio de saltos para establecer una comunicación extremo a extremo, lo que incide directamente en el tráfico total que la red es capaz de cursar.



(a) Distancia media al centro. (b) Número medio de vecinos.

Figura C.4. Distancia media al centro geométrico del área de simulación y número medio de vecinos. Modelos RWP y RWPM (R_{max} = 200 m. $\alpha_{max} = \pi/12$) – 2 Km/h y 9 Km/h – y escenario estático uniforme. 50 terminales.

Sin embargo, el modelo RWPM no presenta este comportamiento, sino que la distribución en el régimen estacionario sí sigue un modelo uniforme. Estas diferencias en el comportamiento de ambos modelos pueden observarse en la figura C.4, donde se presenta la distancia media de cada terminal al centro geométrico del área de simulación así como su número medio de vecinos, para los dos modelos de movilidad y en escenario estático. Se comprueba que con el modelo RWP la distancia media al centro es menor (los nodos se concentran) y como consecuencia, el número medio de vecinos de cada terminal es mayor. En cambio, los valores obtenidos con el RWPM en ambos casos coinciden con los conseguidos mediante el promediado de múltiples (50000) escenarios estáticos con distribución uniforme.



(a) 2 Km/h (b) 5 Km/h

Figura C.5. Comparativa entre RWP y RWPM (R_{max} = 200 m. $\alpha_{max} = \pi/12$). Funcionamiento del protocolo AODV. Topologías de 50 nodos, 2 Km/h y 9 Km/h.

La figura C.5 muestra un ejemplo del impacto de la utilización de uno u otro modelo de movilidad sobre las prestaciones del algoritmo de encaminamiento AODV. Dada la tendencia de los terminales a concentrarse en el centro en el modelo RWP, los nodos tienen mayor probabilidad de encontrarse cerca reduciéndose así el número medio de saltos de las rutas activadas, como se observa en la figura C.5(a). Como resultado, es posible cursar más tráfico total – figura C.5(b) – ya que un menor número de enlaces ocupados por conexión supone, aún teniendo en cuenta la posible reutilización de los recursos en enlaces lejanos dentro de una misma ruta[1], una utilización global de los recursos menor. La pérdida de capacidad efectiva mostrada en la figura se corresponde

[1]Slots TDMA utilizados en un enlace sólo pueden reutilizarse pasados tres saltos.

con la falta de un mecanismo de admisión, puesto que los resultados presentados se corresponden con el funcionamiento básico del protocolo.

En resumen, el modelo propuesto mantiene la condición de modelo sencillo y generalista del RWP, pero que gracias a la introducción de dos nuevos grados de libertad, R_{max} y α_{max}, permite obtener unos patrones de movimiento más cercanos a la realidad, manteniendo además una distribución uniforme de terminales en el área de simulación durante el régimen estacionario.

Conclusions

The work presented in this PhD thesis, developed in the context of quality of service provision in mobile ad hoc networks, has been focused in the proposal and evaluation of a cross-layer architecture based on the cooperation between a Quality of Service (QoS) routing protocol and the Medium Access Control (MAC) level. With this motivation, it has been designed a distributed admission control mechanism capable of providing the required end to end QoS adapting the operation to the characteristic variant environment of Mobile Ad hoc NETworks (MANETs). According to this purpose, different aspects have been tackled, related to routing in ad hoc networks, medium access control protocols, QoS provision in terms of end to end guaranteed bandwidth, distributed resource management and service differentiation. In addition, the dynamism of the application scenarios has motivated the development of adaptation mechanisms for the management procedures according to the observed mobility in the environment.

The quality of service provision directly depends on the capacity of performing a differentiated resource allocation among the demanding applications according to an appropriate admission control policy. The consequent tradeoff between this provision and an efficient resource utilization in the wireless environment has an especially complex solution in the ad hoc context, conditioned by the necessity of establishing distributed control mechanisms over variant network topologies.

The role of routing in a cooperative environment, characterized by the establishment of multihop routes is particularly relevant for defining the admission criteria. The system has to guarantee not only a reliable establishment of a connection path, but also the satisfaction of the demanded requirements. With this premise, the developed work has been concentrated in coordinating the establishment of the required connectivity (route) with the provision of the demanded QoS thanks to the joint operation of the discovery process with the evaluation of the provided QoS. Therefore, an admission procedure has been designed supported by routing decisions based on a metric that represents the available end to end bandwidth.

One of the most relevant determining factors to define the real resource availability, finally represented by the new routing metric, is the difficulty to carry out a reliable measurement according to the real resource utilization, which is controlled by the medium access level. The proposal of this presented work, far from making the routing operation and the resource management independent, links both functionalities by means of a correct inter-level communication capable of providing a reliable measurement that guarantees a realistic admission.

The lack of a fixed infrastructure and the necessity of implementing distributed procedures pose a clear difficulty to design specific protocols for ad hoc networks. The proposed architecture has been firstly validated in a simplified application scenario, capable of identifying the problems stemmed from the resource management, service differentiation, signaling and inter-level communication, contributing specific solutions. However, the adaptation of these schemes to the inherent variability in the network implies the most critical element of the design, which has focused a great extent of the researcher effort in this thesis. A deeper analysis, in more realistic conditions of interference among users and terminals mobility, has allowed to design the adequate adaptation strategies to face the variability of the environment.

Although there are several QoS routing proposals for mobile ad hoc networks in the literature, including some alternatives about inter-layer cooperation, the study developed in this thesis has tried to identify with more detail the problems stemmed from a real implementation for a specific cross-layer design, in order to analyze the feasibility of this kind of solutions. The modeling of the proposed architecture has been carried out by means of a C++ simulator where all the designed strategies have been implemented and evaluated during the thesis. The specific approach has allowed to identify the problems derived from the cooperation itself and its application in more realistic scenarios. The identification of the particular requirements, performed by simulation, has facilitated the proposal of the solutions presented in this work. In addition, thanks to the modular design, the different traffic generators, propagation models or mobility patterns included in the simulator can be easily replaced during future evaluations. Next, the main conclusions of this work can be detailed differentiating three development stages:

Study, characterization and selection of the routing and MAC level protocols that constitute an appropriate cooperation architecture.

With regard to the study of MAC protocols in the context of ad hoc networks, the analysis has been focused on TDMA access strategies, widely analyzed by several authors

and considered especially attractive due to the dynamic resource selection and reservation provided. In order to select a suitable MAC protocol, the ability to become a reliable support for the cross-layer architecture and the capacity of QoS provision in the MAC level have been considered the main criteria. The specific selection does not intend to represent the optimum design, but a strong base for the implementation of the finally proposed architecture.

The characteristics of the ADHOC MAC protocol, modified in order to provide QoS [Gállego, 2007], respond to this aims, especially considering the developed differentiated resource reservation mechanism (Book In Advance Scheme (BIAS)), which is based on the efficient distribution of both connectivity and resource occupation information. This functionality is particularly attractive to provide a reliable estimation of the resources availability to the higher level.

After the critical analysis of the wide range of routing approaches that have been studied, the Ad hoc On-demand Distance Vector (AODV) protocol has been considered an appropriate start point to develop a distributed admission control mechanism. Its on-demand nature, which involves an immediate activation according to the establishment of new communications, and the subsequent routing information diffusion throughout the network, have been considered adequate in order to estimate both the resource availability and the effect of the new admission over the activated connections, reducing the overhead generated by a periodic route maintenance. Its plain structure, without a hierarchical functionality, can facilitate the generalization of the design, as far as the particularization with specific protocols makes it possible, without associating the proposal with a specific context providing higher flexibility. Although the final proposal does not consider a multipath approach, the corresponding version for the AODV (Ad hoc On-demand Multipath Distance Vector (AOMDV)) has been studied in order to incorporate some facilities related to the discovery of several routes with the aim of selecting the best path according to the QoS metric, based on the end-to-end bandwidth.

Proposal and evaluation of the distributed admission control

The appropriate resource allocation according to the requirements demanded by the applications has been fixed in the definition of a routing metric based on the measurement of the end-to-end bandwidth. The application of this estimation in the considered TDMA access has been developed thanks to the design of a procedure for calculating the available slots in every link along the discovered path (BandWidth Calculation – Forward Algorithm (BWC – FA)). The subsequent reservation for the demanded resources

is performed during the confirmation for the selected route, with allows the establishment of guaranteed virtual circuits, thus performing the distributed admission.

The evaluation carried out in a simplified scenario where connectivity is only based on Euclidean distances, rejecting the interference generated by users out of the coverage range (Protocol Model), has allowed to identify the problems stemmed from the shared nature of the transmission medium and the distributed implementation of the procedure. Despite the ability of the ADHOC MAC protocol to correctly estimate the potential resources utilization, the estimation of the maximum resources availability in a scenario with several parallel discoveries becomes more difficult. The proposal of some approximated strategies partially alleviates the problem due to a more flexible reservation that guarantees an increase in the admission probability, especially with lower traffic loads, as it has been proved with the comparative analysis carried out on the matter.

The specific identification of the available resources in every link makes it possible an estimation of the end-to-end bandwidth closer to the reality than the classical local link measurement of the bottle-neck. An admission decision according to this measurement and the reservation of the specified resources increase the probability of admitting connections with the guarantee of a fulfillment of the QoS requirements, as the results obtained in static conditions have shown. The observed improvement in the effective capacity (correctly dispatched connections) is the result of the guaranteed virtual circuits. The extension of the analysis in a multiservice scenario has shown the differentiation capacity of the proposal thanks to the prioritization of the virtual circuits over the best-effort traffic, despite a slightly degradation due to the distributed nature of the procedure and the problems inherent to the shared medium.

Analysis of the effect of the variability on the environment over the admission control. Proposal and evaluation of a readmission mechanism and adaptation strategies

The inclusion of a more realistic connectivity model, based o the real interference among users (Physical Model), has allowed to evaluate the real effect of the received Signal to Interference Ratio (SIR) over the reliability of the established communications and the availability measurement. The instability in the system as a consequence of the imprecise status information in the MAC level has been solved from the perspective of this level, in order to keep the reliable support for the proposed architecture. The problems detected thanks to the evaluation of the whole architecture have identified the weakness of the ADHOC MAC protocol. The possible solutions [Gállego, 2007] resulting

in the appropriate modification of the procedures, the involved signaling and the definition of a set of configurable parameters, have been studied from a global point of view, analyzing the effect over the system performance, evaluated in terms of admission and correctly dispatched traffic. The diverse test scenarios, analyzed in variable interference conditions (traffic variability), have allowed to extract the required parameters definition for the MAC protocol in order to guarantee the reliability and stability of the proposal.

The evaluation including mobility has shown the difficulties to guarantee QoS in a scenario with variable connectivity and interference conditions. The modification in the initial admission conditions, as a result of the variability, can generate new scenarios of admitted traffic and resource utilization that can become little efficient. The new topology can involve the disconnection and the rediscovery of new routes for the previously admitted applications. On the other hand, the increase in the experienced interference in the already established virtual circuits can degrade the end-to-end bandwidth, which cannot be guaranteed any more. In order to allow a more flexible resource utilization, a readmission mechanism capable of prioritizing the maintenance of the admitted and correctly dispatched connections, even dropping the ones that cannot be reallocated with the desired QoS, has been proposed and evaluated. The existent procedures in the AODV itself, which allow to identify the broken paths, have been complemented with a monitoring and updating procedure based on the BWC – FA proposed for the initial admission. The identification of the additional available bandwidth allows to perform an immediate reallocation over the already connected paths or it can trigger the readmission procedure when a new QoS path is required. In order to improve the performance in terms of grade of service, an over-dimensioning strategy based on the inclusion of a guard bandwidth in the initial admission (ΔBW) has been developed, which increases the blocking probability leaving an extra unused capacity, available during readmission, thus reducing the dropping of admitted connections. Moreover, the flexible prioritization in the MAC level allows a differentiate treatment between the active applications and the ones to be admitted, which guarantees a higher user satisfaction.

A detailed evaluation of the proposed strategies has been carried out in order to identify the correct configuration of the set of parameters for the whole mechanism. However, the observed results have shown a clear dependency for the system response on the nature of the applications or the grade of mobility. Although it is unviable a generalized optimization, two operation modes have been defined: operation in a quasi-static environment (low mobility) and operation in high mobility. The overhead introduced by the monitoring and readmission procedures involves the lost of effective capacity that is only

compensated in high mobility scenarios where the limited admission and a representative dropping for interfering applications allow to obtain a significant improvement in the performance of the admitted connections. The unused resources due to the establishment of a bandwidth guard (ΔBW) is also an advantage only in the same conditions. Therefore, the configuration of the whole process is only carried out in that mobile scenario.

The determining influence of the grade of mobility in the environment over the obtained performance, especially the effect on the correct configuration for the response of the system, has motivated the design of an efficient adaptation mechanism capable of correctly estimating the mobility, modifying accordingly the system parameters. By analyzing the mobility effect over the network connectivity (link duration, number of neighbors), a representative metric based on the variation in the number of neighbors has been defined. Thanks to the intrinsic operation of the AHDOC MAC, based on the identification of the stable neighbors, a monitoring mechanism and a configuration based on thresholds have been designed in order to appropriately activate the desired reaction of the system.

The proposed approach is not intended to be the optimum, taking into account the lack of generality of a particular architecture implementation, but it tries to point to a realistic answer to the need of adaptation to the variability of the environment. According to the observation of several scenarios of study , the configurable parameters have been set in order to solve the tradeoff between the rapid identification and response to a high mobility environment with a slightly degradation in the performance in the border situation between both operation modes. The great advantage of an adaptive design, proved with the analyzed results, lies in the flexible system configuration that allows to obtain a reasonable performance in diverse application scenarios.

Bibliografía

Ahn, C.W.; Kang, C.G. y Cho, Y.Z.: «Soft Reservation Multiple Access with Priority Assignment (SRMA/PA): A Distributed MAC Protocol for QoS-Guaranteed Integrated Services in Mobile Ad-Hoc Networks». *IEICE Transactions on Communications*, 2003, **E86-B(1)**, pp. 50–59.

Ahn, G.S.; Campbell, A.T.; Veres, A. y Sun, L.-H.: «Supporting Service Differentiation for Real-time and Best-effort Traffic in Stateless Wireless Ad Hoc Networks (SWAN)». *IEEE Transactions on Mobile Computing*, 2002, **1(3)**, pp. 192–207.

Badis, H. y Agha, K.A: «QOLSR, QoS Routing for Ad Hoc Wireless Networks using OLSR». *European Transactions on Telecommunications*, 2005, **15(4)**, pp. 427–442.

Bai, F.; Sadagopan, N. y Helmy, A.: «IMPORTANT: A Framework to Systematically Analyze the Impact of Mobility on Performance of Routing Protocols for Ad hoc Networks». En: *Proc. of the 22nd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM'03)*, volumen 22, pp. 825–835. San Francisco, CA. EE.UU., 2003.

Basagni, S.; Chlamtac, I.; Syrotiuk, V.R. y Woodward, B.A.: «A Distance Routing Effect Algorithm for Mobility (DREAM)». En: *Proc. of ACM/IEEE International Conference on Mobile Computing and Networking, (Mobicom'98)*, pp. 76–84. Dallas, TX., 1998.

Belding-Royer, E. y Perkins, C.: «Multicast Ad hoc On Demand Distance Vector (MAODV) Routing». *IETF internet draft*, The IETF Network Working Group, 2000. draft-ietf-manet-maodv-00.txt.

Belding-Royer, E.M. y Toh, C.: «A Review of Current Routing Protocols for Ad Hoc Mobile Wireless Networks». *IEEE Personal Communications*, 1999, **6(2)**, pp. 46–55.

Bharghavan, V.; Lee, K.; Lu, S.; Ha, S.; Li, J. y Dwyer, D.: «The TIMELY Adaptive Resource Management Architecture». *IEEE Personal Communications Magazine*, 1998, **5(4)**.

Bianchi, G.; Campbell, A. y Liao, R.: «On Utility-Fair Adaptive Services in Wireless Networks». En: *Proc. of the IEEE/IFIP IWQoS 1998*, Napa, CA, 1998.

Blake, S.; Black, D.; Carlson, M.; Davies, E.; Wang, Z. y Weiss, W.: «An Architecture for Differentiated Services». *RFC 2475*, The Internet Engineering Task Force, 1998. `http://www.ietf.org/rfc/rfc2475.txt`.

Boleng, J.; Navidi, W. y Camp, T.: «Metrics to Enable Adaptive Protocols for Mobile Ad Hoc Networks». En: *Proc. of ICWN'02*, pp. 293–298, 2002.

Borgonovo, F.; Capone, A.; Cesana, M. y Fratta, L.: «ADHOC MAC: a New MAC Architecture for Ad Hoc Networks Providing Efficient and Reliable Point-to-Point and Broadcast Services». *ACM/Springer Wireless Networks (WINET)*, 2004, **10(4)**, pp. 359–366.

Braden, R.; Clark, D. y Shenker, S.: «Integrated Services in the Internet Architecture - an Overview». *RFC 1633*, The Internet Engineering Task Force, 1994. `http://www.ietf.org/rfc/rfc1633.txt`.

Braden, R.; Zhang, L.; Berson, S.; Herzog, S. y Jamin, S.: «Resource ReSerVation Protocol (RSVP)». *RFC 2205*, The Internet Engineering Task Force, 1997. `http://www.ietf.org/rfc/rfc2205.txt`.

Broch, J.; Maltz, D.A.; Johnson, D.B.; Hu, Y. y Jetcheva, J.: «A Performance Comparison of MultiHop Wireless Ad Hoc Network Routing Protocols». En: *Proc. of the IV ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom'98)*, Dallas, EE.UU., 1998.

Camp, T.; Boleng, J. y Davies, V.: «A Survey of Mobility Models for Ad Hoc Network Research». *Wiley Wireless Communications and Mobile Computing*, 2002, **2(5)**, pp. 483–502.

Canales, M.; Gállego, J.R.; Hernández-Solana, A. y Valdovinos, A.: «Cross-Layer Proposal for QoS Routing in Mobile Ad-Hoc Networks». En: *Proc. of 8th International Symposium on Wireless Personal Multimedia Communications (WPMC'05)*, pp. 1325–1329. Aalborg, Dinamarca, 2005a.

——: «Interference-aware Routing with Bandwidth Requirements in Mobile Ad Hoc Networks». En: *Proc. of 62nd Semiannual IEEE Vehicular Technology Conference (VTC'05-fall)*, pp. 2556–2560. Dallas, EE.UU., 2005b.

——: «Cross-layer Routing for QoS Provision in Multiservice Mobile Ad Hoc Networks». En: *Proc. of 17th Annual IEEE International Personal, Indoor and Mobile Radio Communications (PIMRC'06)*, Helsinki, Finlandia, 2006a.

——: «Performance Analysis of Cross-layer QoS Routing for Mobile Ad Hoc Networks». En: *Proc. of 9th International Symposium on Wireless Personal Multimedia Communications (WPMC'06)*, pp. 946–950. San Diego, EE.UU., 2006b.

——: «Performance Evaluation of Cross-layer Routing for QoS Support in Mobile Ad Hoc Networks». *Springer Lecture Notes in Computer Science (11th IFIP International Conference on Personal Wireless Communications)*, 2006c, **4217**, pp. 322–333.

Casilari, E. y Triviño, A.: «Análisis de la Estabilidad de Modelos de Movilidad en Simulaciones de Redes Ad Hoc». En: *Proc. V Jornadas Telemáticas, Jitel'05*, pp. 33–40. Vigo, España, 2005.

Chakrabarti, S. y Mishr, A.: «QoS Issues in Ad Hoc Wireless Networks». *IEEE Communications Magazine*, 2001, **39(2)**, pp. 142–148.

Chen, S. y Nahrstedt, K.: «An Overview of Quality-of-Service Routing for the next Generation High-speed Networks: Problems and Solutions». *IEEE Network Magazine, special issue on Transmission and Distribution of Digital Video*, 1998, **12(6)**, pp. 64–79.

——: «Distributed Quality-of-Service Routing in Ad Hoc Networks». *IEEE Journal on Selected Areas in Communications*, 1999, **17(8)**, pp. 1488–1505.

Chen, T.-W. y Gerla, M.: «Global State Routing: A New Routing Scheme for Ad-hoc Wireless Networks». En: *Proc. of the IEEE International Conference on Communications ICC'98*, pp. 171–175. Atlanta, GA, USA, 1998.

Chen, T.-W.; Tsai, J.T. y Gerla, M.: «QoS Routing Performance in Multihop, Multimedia, Wireless Networks». En: *6th IEEE International Conference on Universal Personal Communications Record*, volumen 2, pp. 557–561. San Diego, CA. USA, 1997.

Chiang, C.-C.; Wu, H.-K.; Liu, W. y Gerla, M.: «Routing in Clustered Multihop, Mobile Wireless Networks with Fading Channel». En: *Proc. of the IEEE Singapore International Conference on Networks, SICON'97*, pp. 197–211. Singapore, 1997.

Clark, D.D.: «The Structuring of Systems using Upcalls». En: *Proc. of 10th ACM Symposium on Operating Systems Principles (SOSP'85)*, pp. 171–180. Orcas Island, Wash, USA, 1985.

Clark, D.D. y Tennenhouse, D.L.: «Architectural Considerations for a new Generation Protocols». En: *Proc. of ACM Symposium on Communications, Architectrues and Protocols (SGCOMM'90)*, pp. 200–208. Philadelphia, Pa, USA, 1990.

Clausen, T. y Jacquet, P.: «Optimized Link State Routing Protocol (OLSR)». *Experimental RFC 3626*, The IETF Network Working Group, 2003. `http://www.ietf.org/rfc/rfc3626.txt`.

Conti, M.; Maselli, G.; Turi, G. y Giordano, S.: «Cross-Layering in Mobile Ad Hoc Network Design». *IEEE Computer*, 2004, **37(2)**, pp. 48–51.

Corson, S. y Macker, J.: «Mobile Ad Hoc Networking (MANET): Routing Protocol Performance Issues and Evaluation Considerations». *RFC 2501*, The IETF Network Working Group, 1999. `http://www.ietf.org/rfc/rfc2501.txt`.

Crawley, E.; Nair, R.; Rajagopalan, B. y Sandick, H.: «A Framework for QoS-based Routing in the Internet». *RFC 2386*, The IETF Network Working Group, 1998. `http://www.ietf.org/rfc/rfc2386.txt`.

Dimitriou, N.; Tafazolli, R. y Sfikas, G.: «Quality of Service for Multimedia CDMA». *IEEE Communications Magazine*, 2000, **38(7)**, pp. 88–94.

Dong, Y.; Yang, T. y Makrakis, D.: «SRL-Enabled QoS Model for Mobile Ad Hoc Networks». En: *Proc. of the International Conference of Communications, Circuits and Systems and West Sino-Exposition 2002 (ICCCAS-WESINO EXPO'02)*, Chengdu, China, 2002.

Dube, R.; Rais, C.D.; Wang, K. y Tripathi, S.K.: «Signal Stability based Adaptive routing (SSA) for ad hoc mobile networks». *IEEE Personal Communication*, 1997, pp. 36–45.

Fasolo, E.; Furiato, R. y Zanella, A.: «Smart Broadcast Algorithm for Inter-vehicular Communications». En: *Proc. of the IWS-WPMC'05*, Aalborg, Dinamarca, 2005.

Feeney, L.: «A Taxonomy for Routing Protocols in Mobile Ad Hoc Networks». *SICS Technical Report T99/07*, HP/SICS Internet Research Insitute, 1999.

Gállego, J.R: *Contribución al Desarrollo de Protocolos de Acceso al Medio (MAC) en Redes Inalámbricas Ad Hoc Multiservicio.* Tesis doctoral, Centro Politécnico Superior, Universidad de Zaragoza, Zaragoza, España, 2007.

Gállego, J.R.; Campelli, L.; Cesana, M.; Capone, A.; Borgonovo, F.; Hernández-Solana, A.; Canales, M. y Valdovinos, A.: «Efficient Bandwidth Allocation for Basic Broadcast and Point-to-Point Services in the AD-HOC MAC Protocol». En: *Proc. of 10th IFIP International Conference on Personal Wireless Communications (PWC'05)*, pp. 87–98. Colmar, Francia, 2005a.

Gállego, J.R.; Canales, M.; Hernández-Solana, A.; Campelli, L.; Cesana, M. y Valdovinos, A.: «Performance Evaluation of an Interference-Aware Access Scheme for Point-to-Point Services in the ADHOC MAC Protocol». En: *Proc. of 7th IFIP/IEEE International Conference on Mobile and Wireless Communications Networks (MWCN'05)*, Marrakech, Marruecos, 2005b.

——: «Performance Evaluation of Point-to-Point Scheduling Strategies for the AD-HOC MAC Protocol». En: *Proc. of 8th International Symposium on Wireless Personal Multimedia Communications (WPMC'05)*, pp. 1380–1384. Aalborg, Dinamarca, 2005c.

Gállego, J.R.; Canales, M.; Hernández-Solana, A. y Valdovinos, A.: «Performance Analysis of an Interference-Aware MAC Protocol with Power Control for Wireless Ad Hoc Networks». En: *Proc. of 17th Annual IEEE International Personal, Indoor and Mobile Radio Communications (PIMRC'06)*, Helsinki, Finlandia, 2006a.

——: «A TDMA Power Controlled MAC Protocol for Wireless Ad Hoc Networks». En: *Proc. of 3rd IEEE International Symposium on Wireless Communication Systems (ISWCS'06)*, Valencia, 2006b.

García-Luna-Aceves, J.J. y Spohn, M.: «Source Tree Adaptive Routing». *IETF Mobile Ad Hoc Networking Working Group. INTERNET DRAFT*, 1999a. draft-ietf-manet-star-00.txt, work in progress.

——: «Source-Tree Routing in Wireless Networks». En: *Proc. of the 7th International Conference on Network Protocols, IEEE ICNP'99*, pp. 273–282. Toronto, Candada, 1999b.

Garcia-Luna-Aceves, J.J. y Rangarajan, H.: «A New Framework for Loop-Free On-Demand Routing Using Destination Sequence Numbers». En: *Proc. 1st IEEE International Conference on Mobile Ad-Hoc and Sensor Systems (MASS'04)*, Fort Lauderdale, Florida, EE.UU., 2004.

Garey, M.R. y Johnson, D.S.: *Computers and Intractability: A Guide to the Theory of Np-Completeness*. Freeman, W.H., 1979.

Gerasimov, I. y Simon, R.: «A Bandwidth-Reservation Mechanism for On-Demand Ad hoc Path Finding». En: *Proc. of the IEEE-SCS 35th Annual Simulation Symposium*, pp. 27–34. San Diego,CA, 2002.

Gerharz, M.C. y de Waal, M.F.: «A Practical View on Quality-of-Service Support in Wireless Ad Hoc Networks». En: *Proc. of the 3rd IEEE Workshop on Applications and Services in Wireless Networks (ASWN)*, pp. 185–196. Berne, Switzerland, 2003.

Gerla, M.; Hong, X.; Ma, L. y Pei, G.: «Landmark Routing Protocol (LANMAR)». *IETF Mobile Ad Hoc Networking Working Group. INTERNET DRAFT*, 2001a. draft-ietf-manet- lanmar-01.txt, work in progress.

Gerla, M.; Pei, G.; Hong, X. y Chen, T.-W.: «Fisheye State Routing Protocol (FSR) for Ad Hoc Networks». *IETF Mobile Ad Hoc Networking Working Group. INTERNET DRAFT*, 2001b. draft-ietf-manet-fsr-00.txt, work in progress.

Gupta, P. y Kumar, P.R.: «The Capacity of Wireless Networks». *IEEE Transactions on Information Theory*, 2000, **46(2)**, pp. 388–404.

Haas, Z.J. y Deng, J.: «Dual Busy Tone Multiple Access (DBTMA): A Multiple Access Control Scheme for Ad Hoc Networks». *IEEE Transactions on Communications*, 2002, **50(6)**, pp. 975–985.

Haas, Z.J.; Pearlman, M.R. y Samar, P.: «The Zone Routing Protocol (ZRP) for Ad Hoc Networks, Internet Draft». *Mobile Ad Hoc Networking Working Group. INTERNET DRAFT*, 2002. draft-ietf-manet-zone-zrp-04.txt, work in progress.

IEEE 802.11 Task Group e: «Draft Supplement to Standard for Telecommunications and Information Exchange between Systems-LAN/MAN Specific Requirements - Part 11: Wireless Medium Access Control (MAC) and PHYsical layer (PHY) Specification. Medium Access Control (MAC) Quality of Service (QoS) Enhancements». *IEEE 802.11e/draft 13.0*, 2005.

IEEE 802.11 Working Group: «Part 11: Wireless LAN Medium Access Control (MAC) and PHYsical Layer (PHY) Specifications». *ANSI/IEEE Std. 802.11*, 1999, pp. 46–55.

ITU-T: «Terms and Definitions related to Quality of Service and Network Performance including Dependability». *Recomendation E.800*, ITU-T, 0894. `http://www.itu.int/rec/T-REC-E.800-199408-I/E`.

Iwata, A.; Chiang, C.-C.; Peiand, G.; Gerla, M. y Chen, T.-W.: «Scalable Routing Strategies for Ad Hoc Wireless Networks». *IEEE Journal on Selected Areas in Communications, Special Issue on Ad-Hoc Networks*, 1999, **17(8)**, pp. 1369–1379.

Jacquet, P.; Muhlethaler, P.; Qayyum, A.; Laouiti, A.; Viennot, L. y Clausen, T: «Proyecto HIPERCOM (High Performance Communications)– OLSR related work. Institut National de Recherche en Informatique et Automatique (INRIA)», 2005. `http://hipercom.inria.fr/olsr/`.

Jardosh, A.P.; Belding-Royer, E.M.; Almeroth, K.C. y Suri, S.: «Real-world Environment Models for Mobile Network Evaluation». *IEEE Journal on Selected Areas in Communications*, 2005, **23(3)**, pp. 622–632.

Jawhar, I. y Wu, J.: «Race-Free Resource Allocation for QoS Support in Wireless Networks». *Ad Hoc and Sensor Wireless Networks*, 2005a, **1(1)**, pp. 179–206.

——: *Resource Management in Wireless Networking*. capítulo Quality of Service Routing in Mobile Ad Hoc Networks. Springer, 2005b.

Jiang, M.; Li, J. y Tay, C.: «Cluster Based Routing Protocol (CBRP) Functional Specification». *Mobile Ad Hoc Networking Working Group. INTERNET DRAFT*, 1999. draft-ietf-manet-cbrp.txt, work in progress.

Joa-Ng, M. y Lu, I.-T.: «A Peer-to-Peer Zone-Based Two-Level Link State Routing for Mobile Ad Hoc Networks». *IEEE Journal on Selected Areas In Communication*, 1999, **17(8)**, pp. 1415–1425.

Johnson, D.; Maltz, D. y Hu, Y.-C.: «The Dynamic Source Routing Protocol for Mobile Ad Hoc Networks». *Mobile Ad Hoc Networking Working Group. INTERNET DRAFT*, 2004, pp. 153–181. draft-ietf-manet-dsr-10.txt, work in progress.

Johnson, D.B. y Maltz, D.A.: *Mobile Computing.* volumen 353, capítulo 5: Dynamic Source Routing in Ad Hoc Wireless Networks, pp. 153–181. Kluwer Academic, 1996. Imielinski T. and Korth H. (Eds.).

Jurdak, R.; Videira Lopes, C. y Baldi, P.: «A Survey, Classification and Comparative Analysis of Medium Access Control Protocols for Ad Hoc Networks». *IEEE Communications Surveys and Tutorials*, 2006, **6(1)**, pp. 2–16.

Karp, B.N. y Kung, H.T.: «Greedy Perimeter Stateless Routing for Wireless Networks». En: *Proc. of the 6th annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom'00)*, pp. 243–254. Boston, MA., 2000.

Kim, D.; Min, C.-H. y Kim, S.: «On-demand SIR and Bandwidth-guaranteed Routing with Transmit Power Assignment in Ad Hoc Mobile Networks». *IEEE Transactions on Vehicular Technology*, 2004, **53(4)**, pp. 1215–1223.

Ko, Y.-B. y Vaidya, N.H.: «Location-Aided Routing in Mobile Ad hoc Networks». En: *Proc. of the ACM/IEEE Conference on Mobile Computing and Networking Mobicom'98*, pp. 66–75. Dallas, TX., 1998.

Kumar, S.; Raghavan, V.S. y Deng, J.: «Medium Access Control Protocols for Ad Hoc Wireless Netwroks: A survey». *Elsevier Ad Hoc Networks*, 2006, **4(3)**, pp. 326–358.

Lang, D.: «A Comprehensive Overview about Selected Ad Hoc Networking Routing Protocols». *Technical Report TUM-I0311*, Technische Universitat Munchen, 2003.

Lee, S-B. y Campbell, A.T.: «INSIGNIA: In-Band Signaling Support for QoS in Mobile Ad Hoc Networks». En: *Proc. of 5th Intl. Workshop on Mobile Multimedia Communication MoMuc'98*, Berlín, 1998.

Liao, W.-H.; Tseng, Y.-C. y Shih, K.-P.: «A TDMA-based Bandwidth Reservation Protocol for QoS Routing in a Wireless Mobile Ad Hoc Network». En: *Proc. of the IEEE International Conference on Communications 2002, (ICC'02)*, volumen 25–1, pp. 3186–3190, 2002a.

Liao, W.-H.; Wang, S.-L.; Sheu, J-P. y Tseng, Y.-C.: «A Multi-Path QoS Routing Protocol in a Wireless Mobile Ad Hoc Network». *Springer Telecommunication Systems*, 2002b, **19(3–4)**, pp. 329–347.

Lin, C.-H.R. y Huang, Y.S.: «Distributed Bandwidth-Constrained Routing for Data Streaming Across Multiple Routes for Improving Call Blocking Rate and Channel Utilization in Wireless Ad Hoc Networks». *Journal of Information Science and Engineering*, 2004, **20(1)**, pp. 85–103.

Lin, C.R. y Liu, J.-S.: «QoS Routing in Ad Hoc Wireless Networks». *IEEE Journal on Selected Areas in Communications*, 1999, **17(8)**, pp. 1426–1438.

Marina, M.K. y Das, S.R.: «On-Demand Multi Path Distance Vector Routing in Ad Hoc Networks». En: *Proc. of the International Conference on Network Protocols ICNP'01*, pp. 14–23. Riverside, California, EE.UU., 2001.

——: «Ad hoc On-demand Multipath Distance Vector Routing». *Communications Review, Special Feature on the First AODV Next Generation Workshop ACM SIGMOBILE*, 2002, pp. 92–93.

——: «Ad hoc On-demand Multipath Distance Vector Routing», 2006. To appear in the Wiley Wireless Communications and Mobile Computing (WCMC) Journal Special Issue on Wireless Ad hoc Networks: Technologies and Challenges.

Mirhakkak, M.; Shult, N. y Thomson, D.: «Dynamic Bandwidth Management and Adaptive Applications for a Variable Bandwidth Wireless Environment». *IEEE Journal of Selected Areas in Communications*, 2001, **19(10)**.

Mohapatra, P.; Li, J. y Gui, C.: «QoS in Mobile Ad Hoc Networks». *IEEE Wireless Communications*, 2003, pp. 44–52.

Mosko, M. y Garcia-Luna-Aceves, J.J.: «Ad hoc Routing with Distributed Ordered Sequences». En: *Proc. IEEE INFOCOM'06*, Barcelona, España, 2006.

Ogier, R.; Templin, F. y Lewis, M.: «Topology Dissemination Based on Reverse-Path Forwarding (TBRPF)». *Experimental RFC 3684*, The IETF Network Working Group, 2004. `http://www.ietf.org/rfc/rfc3684.txt`.

O'Neil, A. y Li, H.: «Hierarchical State Routing Protocol, Internet Draft». *Mobile Ad Hoc Networking Working Group. INTERNET DRAFT*, 2000. draft-oneill-li-hsr-00.txt.

Park, V. y Corson, M.S.: «Temporally-Ordered Routing Algorithm (TORA), a Link Reversal Routing (LRR) algorithm». *Mobile Ad Hoc Networking Working Group. INTERNET DRAFT*, 2001. draft-ietf-manet-tora-spec- 03.txt, work in progress.

Park, V.D. y Corson, M.S.: «A Highly Adaptive Distributed Routing Algorithm for Mobile Wireless Networks». En: *Proc. of the IEEE Conference on Computer Communications (INFOCOM'97)*, Kobe, Japan, 1997.

Pei, G.; Gerla, M. y Hong, X.: «LANMAR: Landmark Routing for Large Scale Wireless Ad Hoc Networks with Group Mobility». En: *Proc. of the ACM/IEEE Workshop on Mobile Ad Hoc Networking and Computing (MOBIHOC'00)*, pp. 11–18. Boston, MA., 2000.

Perkins, C.; Belding-Royer, E. y Das, S.: «Ad hoc On-Demand Distance Vector (AODV) Routing». *Experimental RFC 3561*, The IETF Network Working Group, 2003. `http://www.ietf.org/rfc/rfc3561.txt`.

Perkins, C.E. y Belding-Royer, E.M.: «Quality of Service for Ad hoc On-Demand Distance Vector Routing». *Mobile Ad Hoc Networking Working Group. INTERNET DRAFT*, 2000. draft-ietf-manet-aodvqos-00.txt.

——: «Performance Comparison of Two On-Demand Routing Protocols for Ad Hoc Networks». *IEEE Personal Communications Magazine, special issue on Ad hoc Networking*, 2001, **8(1)**, pp. 16–28.

Perkins, C.E. y Bhagwat, P.: «Highly Dynamic Destination-Sequenced Distance Vector (DSDV) for Mobile Computers». En: *Proc. of the ACM/SIGCOMM'94 Conference on Communications Architectures, Protocols and Applications*, pp. 234–244. Londres, Inglaterra, 1994.

Raisinghani, V.T. y Iyer, S.: «Cross-Layer Design Optimizations in Wireless Protocol Stacks». *Computer Communications*, 2004, **27(8)**, pp. 720–724.

Rangarajan, H. y Garcia-Luna-Aceves, J.J.: «Making On-demand Routing Protocols Based on Destination Sequence Numbers Robust». En: *Proc. IEEE International Conference on Communications (ICC'05)*, Seúl, Korea, 2005a. BEST PAPER AWARD.

——: «On-demand Loop-Free Routing in Ad hoc Networks Using Source Sequence Numbers». En: *Proc. 2nd IEEE International Conference on Mobile Ad-Hoc and Sensor Systems (MASS'05)*, Washington, D.C., EE.UU., 2005b. BEST PAPER AWARD.

Reddy, T.B.; Karthigeyan, I.; Manoj, B.S. y Siva Ram Murthy, C.: «Quality of Service Provisioning in Ad Hoc Wireless Networks: a Survey of Issues and Solutions». *Ad Hoc Networks*, 2006, **4**, pp. 83–124.

Rentel, C.H. y Kunz, T.: «A Clock-sampling Mutual Network Time-synchronization Algorithm for Wireless Ad Hoc Networks». En: *Proc. of IEEE WCNC'05*, Nueva Orleáns, EE.UU., 2005.

Resta, G. y Santi, P.: «An Analysis of the Node Spatial Distribution of the Random Waypoint Mobility Model for Ad Hoc Networks». En: *Proc. of the 2nd ACM Workshop on Principles of Mobile Computing*, pp. 44–50. Toulouse, Francia, 2002.

Shah, S.H.; Chen, K. y Nahrstedt, K.: «Dynamic Bandwidth Management in Single-hop Ad Hoc Wireless Networks». *Mobile Networks and Applications (MONET) Journal, Special Issue on Algorithmic Solutions for Wireless, Mobile, Ad Hoc and Sensor Networks*, 2005, **10**, pp. 199–217.

Shu, Y.-T.; Wang, G.-H.; Wang, L.; Yang, O.W.W. y Fan, Y.-J.: «Provisioning QoS Guarantee by Multipath Routing and Reservation in Ad Hoc Networks». *Journal of Computer Science and Technology*, 2004, **19(2)**, pp. 128–137.

Sivakumar, R.; Sinha, P. y Bharghavan, V.: «CEDAR: a Core-Extraction Distributed Ad Hoc Routing Algorithm». *IEEE Journal on Selected Areas in Communications, Special Issue on Wireless Ad Hoc Networks*, 1999, **17(8)**, pp. 1454–1465.

Srivastava, V. y Motani, M.: «Cross-layer Design: A Survey and the Road Ahead». *IEEE Communications Magazine*, 2005, **43(12)**, pp. 112–119.

Stojmenovic, I. y Lin, X.: «GEDIR: Loop-free Location based Routing in Wireless Networks». En: *Proc. of the International Conference on Parallel and Distributed Computing Systems*, Cambridge, MA, USA, 1999.

Su, W. y Gerla, M.: «IPv6 Flow Handoff in Ad-Hoc Wireless Networks Using Mobility Prediction». En: *Proc. of IEEE Global Telecommunications Conference GLOBECOM'99*, pp. 271–275. Rio de Janeiro, Brasil, 1999.

Tang, Z. y García-Luna-Aceves, J.J.: «A Protocol for Topology-Dependent Transmission Scheduling in Wireless Networks». En: *Proc. of EEE WCNC'99*, volumen 3, pp. 1333–1337. Nueva Orleans, EE.UU., 1999.

Tobagi, F. A. y Kleinrock, L.: «Packet Switching in Radio Channels. Part 2. The Hidden Terminal Problem in Carrier Sense Multiple-Access and the Busy-Tone Solution». *IEEE Transactions on Communications*, 1975, **23(12)**, pp. 1417–1433.

Toh, C.-K.: «Associativity Based Routing For Ad Hoc Mobile Networks». *Springer Wireless Personal Communications*, 1997, **4(2)**, pp. 103–139. US Patent.

Tseng, Y.-C.; Ni, S.-Y.; Chen, Y.-S. y Sheu, J.-P.: «The Broadcast Storm Problem in a Mobile Ad Hoc Network». *ACM/Springer Wireless Networks, Selected Papers from Mobicom'99*, 2002, **8(2–3)**, pp. 153–167.

Tsuchiya, P.F.: «The Landmark Hierarchy: a new Hierarchy for Routing in very large Networks». *Computer Communications Review*, 1988, **18(4)**, pp. 35–42.

Wang, Y.-H.; Lin, H.-Z. y Chang, S.-M.: «Interfering-aware QoS Multipath Routing for Ad Hoc Wireless Network». En: *18th International Conference on Advanced Information Networking and Applications (AINA'04)*, volumen 1. Fukuoka, Japan, 2004.

Wang, Z. y Crowcroft, J.: «Quality of Service Routing for Supporting Multimedia Applications». *IEEE Journal on Selected Areas in Communications*, 1996, **14(7)**, pp. 1228–1234.

Wu, K. y Harms, J.: «QoS Support in Mobile Ad Hoc Networks». *Crossing Boundaries*, 2001, **1(1)**, pp. 92–106.

Xiao, H.; Deah, A.L. y Chua, K.C.: «A Flexible Quality of Service Model for Mobile Ad-Hoc Networks». En: *Proc. of the IEEE Vehicular Technology Conference Spring (VTCS'00)*, Tokyo, Japón, 2000.

Xue, Q. y Ganz, A.: «Ad Hoc QoS On-Demand Routing (AQOR) in Mobile Ad Hoc Networks». *Journal on Parallel and Distributed Computing*, 2003, **63(2)**, pp. 154–165.

Xylomenos, G. y Polyzos, G.C.: «Internet Protocol Performance over Networks with Wireless Links». *IEEE Network*, 1999, **13(4)**, pp. 55–63.

Yoon, J.; Liu, M. y Noble, B.: «Random Waypoint Considered Harmful». En: *Proc. of the 22nd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM'03)*, pp. 1312–1321. San Francisco, CA, EE.UU., 2003.

Zhou, B.; Marshall, A.; Wu, J.; Lee, T.-H. y Liu, J.: «A Cross-Layer Route Discovery Framework for Mobile Ad Hoc Network». *EURASIP Journal on Wireless Communications and Networking*, 2005, **5**, pp. 645–660.

Zhu, C. y Corson, M.S.: «A Five-Phase Reservation Protocol (FPRP) for Mobile Ad Hoc Networks». *Wireless Networks (WINET)*, 2001, **7(4)**, pp. 371–384.

——: «QoS routing for Mobile Ad Hoc Networks». En: *Proc. of the 21st Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM'02)*, volumen 2, pp. 958–967. Nueva York, EE.UU., 2002.

Zou, X.; Ramamurthy, B. y Magliveras, S.: «Routing Techniques for Wireless Ad Hoc Networks - Classification and Comparison». En: *Proc. of the Sixth World Multiconference on Systemics, Cybernetics, and Informatics–SCI*, Orlando, USA, 2002.