



Universidad de Valladolid

ESCUELA TÉCNICA SUPERIOR DE INGENIEROS DE TELECOMUNICACIÓN

DEPARTAMENTO DE TEORÍA DE LA SEÑAL Y COMUNICACIONES E
INGENIERÍA TELEMÁTICA

TESIS DOCTORAL:

TÉCNICAS ADAPTATIVAS Y COGNITIVAS PARA LA GESTIÓN DE RECURSOS EN REDES DE ACCESO ÓPTICAS PASIVAS Y DE TRANSPORTE

Presentada por Tamara Jiménez Garcinuño para optar al grado de
doctora por la Universidad de Valladolid

Dirigida por:
Dra. Noemí Merayo Álvarez



Universidad de Valladolid

DEPARTAMENTO DE TEORÍA DE LA SEÑAL Y COMUNICACIONES E INGENIERÍA
TELEMÁTICA

TÉCNICAS ADAPTATIVAS Y COGNITIVAS PARA LA GESTIÓN DE RECURSOS EN REDES DE ACCESO ÓPTICAS PASIVAS Y DE TRANSPORTE

AUTORA: Tamara Jiménez Garcinuño

DIRECTORA: Dra. Noemí Merayo Álvarez

TRIBUNAL

Presidente: _____

Vocales: _____

Secretario: _____

Suplentes: _____

FECHA DE LECTURA:

CALIFICACIÓN:

*A mis padres,
a mi hermano
y a Javi*

*“Las preguntas que no podemos contestar
son las que más nos enseñan”*

EL TEMOR DE UN HOMBRE SABIO

Patrick Rothfuss

Resumen

El incremento del ancho de banda demandado, junto con la aparición de nuevas aplicaciones y servicios con requisitos muy diversos, hacen que la calidad de servicio (QoS, *Quality of Service*) se convierta en uno de los puntos clave en las redes actuales de telecomunicaciones basadas en fibra óptica. Por ello, en esta Tesis se proponen varios métodos adaptativos y cognitivos para el control eficiente de los recursos en redes de acceso y troncales, con el fin de ofrecer una calidad de servicio y una calidad de transmisión adecuada a los requisitos estipulados por los proveedores de servicio y demandada por los usuarios. En este sentido, el punto fundamental de la Tesis se centra en el desarrollo de nuevas técnicas de control de acceso al medio en redes ópticas pasivas de largo alcance (LR-PON, *Long-Reach Passive Optical Networks*). Estas redes se posicionan como una tecnología muy prometedora para el segmento de acceso, ya que aparte de proporcionar un gran ancho de banda al utilizar la fibra como medio de transmisión, resultan muy económicas. Las redes LR-PON aumentan la distancia extremo a extremo hasta 100 km, lo que implica, por un lado, que son capaces de dar cobertura a un mayor número de abonados, y por otro lado, que simplifican la arquitectura de red, al combinarse la red metropolitana y de acceso en una sola, reduciendo los costes considerablemente. No obstante, a nivel de acceso estas redes tienen que solventar las posibles inefficiencias de los mecanismos de control derivadas del incremento de la distancia. Además, las redes LR-PON tienen que ser capaces de soportar de forma simultánea diferentes tipos de aplicaciones con diversos requisitos de calidad, así como usuarios con perfiles de demanda distintos. Por ello, los mecanismos de control de acceso al medio tienen que tener en cuenta todas estas consideraciones y restricciones. Por este motivo, los algoritmos propuestos en esta Tesis asignan dinámicamente el ancho de banda con el objetivo conjunto de ofrecer una calidad de servicio acorde a las necesidades requeridas y de hacer frente a los problemas surgidos del incremento de la distancia extremo a extremo.

En esta línea, una de las principales aportaciones de esta Tesis es la incorporación de controladores PID (Proporcional, Integral, Derivativo) en los algoritmos de asignación de ancho de banda. Gracias a estos sistemas, se consigue una gestión muy eficiente

del ancho de banda disponible y se controlan parámetros clave para ofrecer calidad de servicio. En este sentido, se han desarrollado algoritmos para asignar unos niveles de ancho de banda garantizado según el perfil del abonado y para mantener el retardo medio por debajo de unas cotas máximas establecidas dependiendo del tipo de tráfico y de la prioridad del abonado. Por tanto, se realiza una doble diferenciación, a nivel de usuario y a nivel de servicio, para conseguir que la calidad de servicio y la calidad de experiencia de los usuarios sea la adecuada. Además, la inclusión de estos sistemas de control PID otorga a los algoritmos independencia frente a las condiciones de red, ya que permiten adaptabilidad a diferentes escenarios de red y patrones de tráfico. También se ha llevado a cabo una optimización de estos algoritmos, mediante el uso de técnicas automáticas, adaptativas y cognitivas para realizar el proceso de sintonización del PID. Este proceso, que consiste en la determinación de los parámetros del controlador PID, tiene una gran importancia en su funcionamiento, ya que tanto la estabilidad como la rapidez del sistema dependen de que los valores se hayan seleccionado correctamente dependiendo del escenario de red. Por tanto, se han desarrollado métodos basados en algoritmos genéticos, redes neuronales y la técnica cognitiva de razonamiento basado en casos, para poder realizar una sintonización automática y adaptativa que ofrezca siempre los parámetros más adecuados del controlador PID, para así obtener una mejora en las prestaciones de ciertos parámetros de red, tales como el retardo y el ancho de banda, en términos de estabilidad, rapidez de convergencia hacia los niveles y precisión.

Así mismo, se han desarrollado técnicas cognitivas para la estimación de la calidad de transmisión en redes ópticas troncales. En este sentido, la señal óptica, a medida que se propaga por la fibra, se degrada, de forma que puede que no cumpla con los requisitos de calidad establecidos. Por ello, resulta de gran interés determinar antes del establecimiento de una nueva conexión si ésta ofrecerá la calidad requerida o si su establecimiento degradará muy significativamente la calidad de otras conexiones ya establecidas, con el objetivo de hacer un uso eficiente de los recursos existentes en la red troncal. Con este fin, se ha diseñado e implementado un estimador cognitivo de la calidad de transmisión basado en la técnica de razonamiento basado en casos que realiza estas predicciones con un alto grado de exactitud, y que ha sido validado no solo mediante simulación, sino también experimentalmente.

Abstract

The growth of Internet traffic and bandwidth demand, together with the emergence of new applications and services with very different quality requirements make Quality of Service (QoS) a key parameter in current telecommunication networks. Therefore, several adaptive and cognitive methods are proposed in this Thesis to efficiently manage the available resources in order to provide the most suitable quality according to the requirements established by service providers and the needs of end users. As a result, the main line of research in this Thesis is to develop new medium access mechanisms for Long-Reach Passive Optical Networks (LR-PONs). These networks are widely regarded as a promising solution to implement new high-bandwidth access networks due to their high bandwidth capacity and their cost effectiveness. These architectures extend the coverage span of common PONs from the traditional 20 km to 100 km and beyond, by combining the access and metro networks into an integrated system. Thus, the hierarchy network is simplified and significant cost savings are achieved. Moreover, the extended split ratio and the enlargement of the geographic coverage allow a growth in the potential number of users, thereby reducing the cost of the shared equipment. However, at the access level, these networks have to face the inefficiencies that can arise in the control mechanisms due to the enlargement of the coverage. In addition, LR-PON networks have to simultaneously deal with a multi-profile scenario, where a variety of users with different needs exist, and with a plethora of applications with diverse quality requirements. Therefore, medium access control mechanisms have to consider all these issues and restrictions. Consequently, the proposed algorithms in this Thesis dynamically assign bandwidth with the twofold aim of offering the required quality of service and dealing with the issues resulting from the increased distance.

Consequently, one of the main contributions of this Thesis is the integration of PID (Proportional-Integral-Derivative) controllers in the bandwidth assignment algorithms. With these systems, an efficient management of the available bandwidth is achieved and the QoS parameters of the network are controlled. Therefore, dynamic bandwidth allocation algorithms with PID controllers have been developed to guarantee bandwidth levels

and to keep the average delay under a maximum threshold established according to the type of traffic and the priority of the user. Hence, a simultaneous double differentiation is conducted, at user and service level, in order to achieve a suitable QoS and Quality of Experience (QoE) for end users. In addition, the incorporation of these PID controllers provides the algorithms with independence of the network conditions, since they allow adaptability to different network scenarios and traffic patterns. Moreover, optimizations of these algorithms have been accomplished, by means of automatic, adaptive and cognitive techniques to tune the PID. This process, which consists in determining the parameters of the controller, has a great influence in its performance in terms of stability and convergence speed to the stipulated levels. Thus, methods based on genetic algorithms, neural networks and the cognitive technique of case-based reasoning, have been designed and implemented to automatically and adaptively tune the PID in an efficient way, with the general purpose of enhancing the performance of the PID when controlling network parameters under different network conditions.

Furthermore, a cognitive technique to estimate the Quality of Transmission (QoT) in optical core networks has been developed. As optical signals traverse fiber links and nodes, they suffer from a number of physical impairments which degrade the signal quality. Hence, the QoT will be affected and it might not comply with the service requirements. Thus, it is very critical to determine just before the establishment of a new connection if it will offer the required quality and if it will not significantly affect the quality of other connections already established. In this way, an effective management of the resources in core optical networks can be achieved. With this aim, a cognitive estimator of the QoT has been designed and developed, grounded in the case-based reasoning technique, which has demonstrated great accuracy in the predictions. Finally, not only has it been tested by means of simulations, but also experimentally.

Agradecimientos

La investigación desarrollada a lo largo de esta Tesis Doctoral ha sido financiada por la Consejería de Educación de la Junta de Castilla y León, mediante las Ayudas destinadas a la contratación de Personal Investigador de Reciente Titulación Universitaria (EDU/1993/2008 de 11 de Noviembre), cofinanciadas por el Fondo Social Europeo. Así mismo, una parte de la investigación está enmarcada en el proyecto Europeo CHRON (*Cognitive Heterogeneous Reconfigurable Optical Network*) financiado por la Comisión Europea dentro del 7º Programa Marco [FP7/2007-2013] bajo el acuerdo nº 258644. En este sentido, agradecer la colaboración de los socios del proyecto, en especial a AIT y a la Universidad Técnica de Dinamarca (DTU) por la prueba experimental del estimador de calidad de transmisión cognitivo. Así mismo, agradecer a CEDETEL la financiación inicial.

Con la entrega de este Tesis acaba una etapa muy importante en mi vida, y expresar toda la gratitud que siento hacia las personas que me han acompañado en unas pocas líneas se torna muy difícil.

En primer lugar, quiero agradecer al Grupo de Comunicaciones Ópticas la oportunidad de hacer esta Tesis. Sin vosotros hubiese sido imposible afrontar con éxito su elaboración. A Evaristo, a Rubén, a Patricia y a Ramón os quiero agradecer la oportunidad que me habéis brindado de estar hoy aquí, así como la confianza depositada en mí. A Ramón, también quiero darle las gracias por su cercanía y su ayuda.

No obstante, mi mayor gratitud recae sobre mi directora de Tesis, Noemí. Sin su dedicación y colaboración durante todos estos años, la consecución de esta Tesis no hubiera sido posible. Gracias por haber creído en mí, por todos tus consejos, tu tiempo, tu confianza, tu cariño y tu apoyo incondicional. Siempre has estado ahí para guiarme en el camino con una gran ilusión y motivación.

Darle las gracias a Nacho, por ofrecerme su tiempo y ayuda siempre que lo he necesitado. Además, gracias por haber despertado en mí el interés por las Comunicaciones Ópticas durante la carrera. A Juan Carlos, agradecerle toda la ayuda prestada, en especial con el CBR. Pero sobre todo por las tertulias literarias y por poner una nota de

diversión en las comidas.

También quería agradecerlos a todos la confianza que habéis depositado en mí al trabajar en el proyecto CHRON. Aunque en ciertos momentos ha sido duro, ha sido una experiencia muy positiva y he aprendido muchas cosas.

Me acuerdo aquí también de todos mis compañeros de laboratorio durante estos años: Natalia, David, “Nachito”, Javi, Alfonso, Pablo, Santi, Aitor, Bruno, Carolina... A todos ellos tengo que agradecerles los buenos ratos que hemos pasado y su ayuda. En especial a Natalia, que lleva en este viaje conmigo desde el principio, y a la que quiero agradecer su amistad.

Por supuesto, agradecer a mi familia el apoyo prestado. En especial a mi padre y a mi hermano, darles las gracias por todo lo que han hecho por mí. También dedico unas líneas a la peque de la familia, Lydia, porque siempre consigue sacarme una sonrisa. Y de forma muy especial, quiero agradecer a Javi su apoyo incondicional durante todo este tiempo, y como no, su paciencia y comprensión. Gracias por estar siempre ahí.

Y por último, pero no menos importante, quiero agradecer la amistad y el apoyo se tod@s los que me han acompañado a lo largo de estos años. A mis amig@s de Ávila, a mis amigas de la resi, a mis compañeras de piso, a mis compañeras de Alemán... porque sin vosotros estos años no hubieran sido igual. ¡Gracias!

Hay un proverbio Chino que dice así: “El pasado es historia, el futuro un misterio, pero el hoy es un regalo, por eso se llama presente”. Y os doy las gracias a tod@s por formar parte de él.

¡GRACIAS!

Índice general

1. Introducción	1
1.1. Objetivos de la Tesis	4
1.2. Estructura de la Tesis	6
2. Redes de acceso ópticas pasivas	9
2.1. Introducción	9
2.2. Redes de acceso PON	9
2.2.1. Arquitectura básica de una red PON	13
2.2.2. Estándares para el despliegue de redes PON	17
2.2.3. Calidad de servicio en redes PON	20
2.3. Redes de acceso LR-PON	21
2.3.1. Arquitecturas para el despliegue de redes LR-PON	22
2.3.2. Control de acceso al medio en redes LR-PON	33
2.3.3. Algoritmos de asignación dinámica de ancho de banda en redes LR-PON	35
2.3.4. Calidad de servicio en redes LR-PON	40
2.4. Conclusiones	44
3. Algoritmo DBA para garantizar dinámicamente niveles de ancho de banda en LR-EPONs	47
3.1. Introducción	47
3.2. Algoritmo DySGAB	48
3.2.1. Asignación de ancho de banda en DySGAB	48
3.2.2. Adaptación dinámica del ancho de banda en DySGAB	50
3.3. Estudio de simulación del algoritmo DySGAB	60
3.3.1. Escenario de simulación	60
3.3.2. Establecimiento de los parámetros asociados al diseño de DySGAB	62
3.3.3. Análisis de prestaciones de DySGAB	69

3.4.	Análisis de prestaciones de DySGAB bajo patrones de tráfico asimétrico y dinámico	80
3.4.1.	Escenarios con un patrón de tráfico asimétrico	80
3.4.2.	Escenarios con un patrón de tráfico dinámico	82
3.5.	Conclusiones	88
4.	Controladores PID en la gestión de recursos en redes LR-EPON	91
4.1.	Introducción	91
4.2.	Controladores PID (Proporcional-Integral-Derivativo)	92
4.2.1.	Arquitectura general de un PID	94
4.2.2.	Métodos de sintonización de controladores PID	97
4.3.	Controlador PID para garantizar niveles de ancho de banda. Algoritmo SPID	99
4.3.1.	Asignación de ancho de banda en SPID	99
4.3.2.	Adaptación dinámica del ancho de banda en SPID	100
4.3.3.	Diseño de estrategias de asignación de ancho de banda de SPID	103
4.4.	Estudio de simulación del algoritmo SPID	104
4.4.1.	Escenario de simulación	104
4.4.2.	Proceso de sintonización y comparativa de estrategias P-PI-PID	104
4.4.3.	Establecimiento de los parámetros asociados al diseño del controlador en SPID	109
4.4.4.	Análisis de prestaciones de SPID	110
4.5.	Análisis de prestaciones de SPID bajo patrones de tráfico asimétrico y dinámico	124
4.5.1.	Escenarios con un patrón de tráfico asimétrico	124
4.5.2.	Escenarios con un patrón de tráfico dinámico	126
4.6.	Controlador PID en el control del retardo. Algoritmo DaSPID	133
4.6.1.	Diferenciación de servicios y de perfiles de abonado en DaSPID	134
4.6.2.	Asignación y adaptación del ancho de banda en DaSPID para controlar el retardo	136
4.7.	Estudio de simulación del algoritmo DaSPID	138
4.7.1.	Escenario de simulación	138
4.7.2.	Proceso de sintonización y comparativa de estrategias P-PI-PID	141
4.7.3.	Establecimiento de los parámetros asociados al diseño del controlador en DaSPID	145
4.7.4.	Análisis de prestaciones de DaSPID	149

4.8.	Análisis de prestaciones de DaSPID bajo patrones de tráfico asimétrico y dinámico	159
4.8.1.	Escenarios con un patrón de tráfico asimétrico	160
4.8.2.	Escenarios con un patrón de tráfico dinámico	161
4.9.	Conclusiones	165
5.	Métodos de sintonización automática de PIDs basados en algoritmos genéticos	169
5.1.	Introducción	169
5.2.	Fundamentos básicos de los algoritmos genéticos	170
5.3.	Sintonización de SPID mediante algoritmos genéticos. Algoritmo GA-SPID	175
5.4.	Estudio de simulación del algoritmo GA-SPID	180
5.4.1.	Escenario de simulación	180
5.4.2.	Establecimiento de los parámetros asociados al diseño del algoritmo genético en GA-SPID	181
5.4.3.	Ánalisis de prestaciones de GA-SPID	187
5.5.	Sintonización de DaSPID mediante algoritmos genéticos. Algoritmo GA-DaSPID	201
5.6.	Estudio de simulación del algoritmo GA-DaSPID	202
5.6.1.	Escenario de simulación	202
5.6.2.	Establecimiento de los parámetros asociados al diseño del algoritmo genético en GA-DaSPID	203
5.6.3.	Ánalisis de prestaciones de GA-DaSPID	206
5.7.	Conclusiones	213
6.	Métodos de sintonización adaptativa de PIDs basados en redes neuronales	215
6.1.	Introducción	215
6.2.	Fundamentos básicos de las redes neuronales artificiales	216
6.2.1.	Introducción	216
6.2.2.	Topología de las redes neuronales artificiales	217
6.2.3.	Funciones de activación en las redes neuronales artificiales	218
6.2.4.	Algoritmos de aprendizaje en las redes neuronales artificiales	219
6.2.5.	Modelo del perceptrón multicapa	221
6.2.6.	Estado del arte de redes neuronales para la sintonización de PIDs .	225

6.3. Sintonización adaptativa de SPID mediante redes neuronales. Algoritmo NN-SPID	226
6.4. Estudio de simulación del algoritmo NN-SPID	231
6.4.1. Entorno de simulación	231
6.4.2. Establecimiento de los parámetros asociados al diseño de la red neuronal en NN-SPID	232
6.4.3. Análisis de prestaciones de NN-SPID	235
6.5. Análisis de prestaciones de NN-SPID bajo un patrón de tráfico dinámico .	247
6.6. Conclusiones	250
7. Técnicas cognitivas para optimizar la gestión de recursos en redes ópticas troncales y de acceso	253
7.1. Introducción	253
7.2. Introducción al Razonamiento Basado en Casos (CBR)	254
7.3. Estimador cognitivo de la calidad de transmisión en redes ópticas troncales	256
7.3.1. Introducción a las redes ópticas troncales actuales	256
7.3.2. Descripción del estimador de calidad cognitivo diseñado	259
7.3.3. Diseño del CBR del estimador cognitivo de QoT	261
7.3.4. Optimización de la base del conocimiento	262
7.4. Estudio de simulación del estimador cognitivo de QoT	265
7.4.1. Escenario de simulación	265
7.4.2. Análisis de prestaciones de los estimadores R-CBR y FixE-CBR .	267
7.4.3. Puesta en marcha del estimador cognitivo de QoT en una nueva red	274
7.4.4. Validación experimental del estimador cognitivo de QoT	277
7.5. Desarrollo de técnicas cognitivas para la sintonización de PIDs en redes de acceso LR-EPON	280
7.5.1. Integración de la técnica CBR para la sintonización adaptativa de SPID. Algoritmo CBR-SPID	280
7.5.2. Estudio de simulación del algoritmo CBR-SPID	282
7.5.3. Análisis de prestaciones de CBR-SPID bajo un patrón de tráfico dinámico	288
7.6. Conclusiones	292
8. Conclusiones y Líneas Futuras de Investigación	295
8.1. Conclusiones	295

8.1.1. Red de acceso LR-EPON	295
8.1.2. Redes troncales	301
8.2. Líneas Futuras	301
8.3. Publicaciones	303
Bibliografía	309
Lista de Acrónimos	329

Índice de figuras

2.1.	Esquema de una red de comunicaciones	10
2.2.	Escenarios de distintos tipos de despliegues FTTx	12
2.3.	Topologías propuestas en el despliegue de tecnologías FTTx	13
2.4.	Arquitectura básica de una red PON con topología en árbol	14
2.5.	Ejemplo de arquitectura de una red PON con protocolo TDMA	15
2.6.	Simplificación de la red de acceso mediante redes LR-PON	22
2.7.	Arquitectura de la red SuperPON de 100 km de distancia	23
2.8.	Arquitectura de una red LR-PON TDMA de 10 Gbit/s	24
2.9.	Arquitectura general de una red híbrida DWDM-TDM	25
2.10.	Arquitectura general de la red híbrida WDM/TDMA del proyecto PIEMAN .	27
2.11.	Arquitectura general de la red PON del proyecto DISCUS	28
2.12.	Arquitectura básica de una red LR-PON en anillo	29
2.13.	Arquitectura general de SUCCESS	30
2.14.	Arquitectura general del prototipo de red XL-PON	31
2.15.	Arquitectura del prototipo de red WE-PON para FTTH	31
2.16.	Arquitectura de red SARDANA	32
2.17.	Arquitectura del prototipo de red Wx-PON	33
2.18.	Esquema de funcionamiento de un algoritmo centralizado en una red LR-PON	36
2.19.	Esquema de funcionamiento del algoritmo centralizado TSD	36
2.20.	Esquema de funcionamiento del algoritmo <i>Multi-Thread polling</i>	37
2.21.	Esquema de funcionamiento de los algoritmos de testeo continuo o algo- ritmos de <i>polling</i> en una red LR-PON	39
3.1.	Funcionamiento de la ventana deslizante de la ONU i para el cálculo del ancho de banda medio asignado a dicha ONU	50
3.2.	Esquema explicativo de la adaptación de máximos realizada por el algo- ritmo DySGAB	51

3.3. Esquema con los pasos de la adaptación que realiza DySGAB.	59
3.4. Variación en tiempo real del ancho de banda medio para tiempos de ventana relativamente altos. (a) SLA ₀ (b) SLA ₁ (c) SLA ₂	63
3.5. Variación en tiempo real del ancho de banda medio para tiempos de ventana relativamente bajos. (a) SLA ₀ (b) SLA ₁ (c) SLA ₂	64
3.6. Evolución en tiempo real del ancho de banda medio asignado a una ONU de cada SLA para valores elevados del término $B_{reduction}$ (en bytes). (a) SLA ₀ (b) SLA ₁ (c) SLA ₂	65
3.7. Evolución en tiempo real del ancho de banda medio asignado a una ONU de cada SLA para valores pequeños del término $B_{reduction}$ (en bytes). (a) SLA ₀ (b) SLA ₁ (c) SLA ₂	66
3.8. Tiempo de convergencia hacia los niveles de ancho de banda garantizados considerando valores pequeños del término $B_{reduction}$. (a) SLA ₀ (b) SLA ₁ (c) SLA ₂	67
3.9. Variación en tiempo real del ancho de banda medio para una ONU de cada SLA con distintos valores de T_{alg} . (a) SLA ₀ (b) SLA ₁ (c) SLA ₂	69
3.10. Ancho de banda demandado y ofrecido para una ONU de cada SLA bajo los tres escenarios considerados en la Tabla 3.4. (a) SLA ₀ (b) SLA ₁ (c) SLA ₂	71
3.11. Comparación de la evolución en tiempo real del ancho de banda medio en LIPSA y DySGAB considerando diferentes condiciones iniciales de ancho de banda para el SLA ₀	72
3.12. Comparación de la evolución en tiempo real del ancho de banda medio en LIPSA y DySGAB considerando diferentes condiciones iniciales de ancho de banda para el SLA ₁	73
3.13. Comparación de la evolución en tiempo real del ancho de banda medio en LIPSA y DySGAB considerando diferentes condiciones iniciales de ancho de banda para el SLA ₂	74
3.14. Evolución en tiempo real del ancho de banda medio asignado por DySGAB para una ONU de cada SLA y para carga de red de 1.6 considerando los dos escenarios de la Tabla 3.5. (a) Escenario 1 (b) Escenario 2.	75
3.15. Evolución a lo largo del tiempo del ancho de banda medio asignado a una ONU de cada SLA en DySGAB cuando varían los anchos de banda garantizados en tiempo real. (a) SLA ₀ (b) SLA ₁ (c) SLA ₂	77

3.16. Ancho de banda demandado y ofrecido a una ONU de cada SLA compa- rando DySGAB y LIPSA frente la carga de red para los escenarios de la Tabla 3.7. (a) Escenario 1 (b) Escenario 2.	78
3.17. Ancho de banda demandado y ofrecido a una ONU de cada SLA para redes EPON y LR-EPON de distinta distancia.	79
3.18. Ancho de banda demandado y ofrecido a cada ONU de la red LR-EPON bajo las condiciones del Escenario 1 de la Tabla 3.8.	81
3.19. Ancho de banda demandado y ofrecido a cada ONU de la red LR-EPON bajo las condiciones del Escenario 2 de la Tabla 3.8.	82
3.20. Ancho de banda demandado para una ONU de cada SLA (azul) y varia- ción en tiempo real del ancho de banda medio asignado (negro) conside- rando una frecuencia de cambio de 1800 s. (a) SLA ₀ (b) SLA ₁ (c) SLA ₂	84
3.21. Ancho de banda demandado para una ONU de cada SLA (azul) y varia- ción en tiempo real del ancho de banda máximo (negro) considerando una frecuencia de cambio de 1800 s. (a) SLA ₀ (b) SLA ₁ (c) SLA ₂	85
3.22. Ancho de banda demandado para una ONU de cada SLA (azul) y varia- ción en tiempo real del ancho de banda medio asignado (negro) conside- rando una frecuencia de cambio de 900 s. (a) SLA ₀ (b) SLA ₁ (c) SLA ₂	86
3.23. Ancho de banda demandado para una ONU de cada SLA (azul) y varia- ción en tiempo real del ancho de banda máximo (negro) considerando una frecuencia de cambio de 900 s. (a) SLA ₀ (b) SLA ₁ (c) SLA ₂	87
4.1. Diagrama de bloques de un proceso controlado por un PID.	94
4.2. Diagrama de bloques del proceso controlado por un PID propuesto para el control del ancho de banda en una red LR-EPON.	100
4.3. Variación del ancho de banda medio asignado para distintos valores de K_u . (a) Valores bajos de K_u (b) Valores altos de K_u	106
4.4. Evolución en tiempo real del (a) ancho de banda medio hacia nivel garan- tizado del SLA ₀ (b) ancho de banda máximo permitido para el SLA ₀	107
4.5. Evolución en tiempo real del (a) ancho de banda medio hacia nivel garan- tizado del SLA ₁ (b) ancho de banda máximo permitido para el SLA ₁	108
4.6. Evolución en tiempo real del (a) ancho de banda medio hacia nivel garan- tizado del SLA ₂ (b) ancho de banda máximo permitido para el SLA ₂	109
4.7. Variación en tiempo real del ancho de banda medio asignado del SLA ₂ considerando (a) valores bajos de T (b) valores altos de T	110
4.8. Evolución en tiempo real del (a) ancho de banda medio hacia nivel garan- tizado del SLA ₀ (b) ancho de banda máximo permitido para el SLA ₀	112

- 4.9. Evolución en tiempo real del (a) ancho de banda medio hacia nivel garantizado del SLA₁ (b) ancho de banda máximo permitido para el SLA₁. 113
- 4.10. Evolución en tiempo real del (a) ancho de banda medio hacia nivel garantizado del SLA₂ (b) ancho de banda máximo permitido a lo largo del tiempo para el SLA₂. 113
- 4.11. Desviación estándar del ancho de banda medio asignado sobre el ancho de banda garantizado. (a) SLA₀ (b) SLA₁ (c) SLA₂. 115
- 4.12. Evolución en tiempo real del ancho de banda medio asignado para los tres perfiles de abonado y los tres escenarios considerados en la Tabla 4.5.
(a) SLA₀ (b) SLA₁ (c) SLA₂. 117
- 4.13. Variación en tiempo real del ancho de banda medio asignado a una ONU cuando los niveles garantizados varían a lo largo del tiempo. (a) SLA₀
(b) SLA₁ (c) SLA₂. 119
- 4.14. Variación en tiempo real del ancho de banda medio asignado a una ONU para diferentes tiempos de ejecución en los algoritmos (T_{alg}, T). (a) DySGAB
(b) SPID. 120
- 4.15. Evolución en tiempo real del ancho de banda medio asignado al SLA₂ para DySGAB y SPID considerando distintos tamaños de ventana (T_{window}).
(a) 7.5 s (b) 30 s. 121
- 4.16. Ancho de banda ofrecido a cada ONU para las políticas conservativa y agresiva bajo las condiciones de los tres escenarios de la Tabla 4.7. 122
- 4.17. Ancho de banda demandado y ofrecido a una ONU de cada SLA para redes EPON y LR-EPON de distinta distancia. 123
- 4.18. Ancho de banda demandado y ofrecido a cada ONU por SPID y DySGAB para las condiciones del Escenario 1 de la Tabla 4.8. 125
- 4.19. Ancho de banda demandado y ofrecido a cada ONU por SPID y DySGAB para las condiciones del Escenario 2 de la Tabla 4.8. 125
- 4.20. Ancho de banda demandado por una ONU (azul) y variación en tiempo real del ancho de banda medio asignado en DySGAB (negro) y SPID (verde) para una frecuencia de cambio de 1800 s. (a) SLA₀ (b) SLA₁
(c) SLA₂. 127
- 4.21. Ancho de banda demandado por una ONU (azul) y variación en tiempo real del ancho de banda máximo en DySGAB (negro) y en SPID (verde) para una frecuencia de cambio de 1800 s. (a) SLA₀ (b) SLA₁ (c) SLA₂. 128

4.22. Desviación del ancho de banda medio asignado con respecto a nivel garantizado o nivel demandado en DySGAB (negro) y SPID (verde) para una frecuencia de cambio de 1800 s. (a) SLA ₀ (b) SLA ₁ (c) SLA ₂	129
4.23. Ancho de banda demandado por una ONU (azul) y variación en tiempo real del ancho de banda medio asignado en DySGAB (negro) y SPID (verde) para una frecuencia de cambio de 900 s. (a) SLA ₀ (b) SLA ₁ (c) SLA ₂	130
4.24. Ancho de banda demandado por una ONU (azul) y variación en tiempo real del ancho de banda máximo en DySGAB (negro) y en SPID (verde) para una frecuencia de cambio de 900 s. (a) SLA ₀ (b) SLA ₁ (c) SLA ₂	131
4.25. Desviación del ancho de banda medio asignado con respecto a nivel garantizado o nivel demandado en DySGAB (negro) y SPID (verde) para una frecuencia de cambio de 900 s. (a) SLA ₀ (b) SLA ₁ (c) SLA ₂	132
4.26. Método de extracción de paquetes basado en el esquema de prioridad de colas estrictas (<i>Strict Priority Queue</i>).	135
4.27. Diagrama de bloques del proceso controlado por un PID propuesto para el control del retardo máximo del tráfico prioritario en redes LR-EPON.	137
4.28. Variación en tiempo real del retardo medio asignado para distintos valores de K_u . (a) SLA ₁ (b) SLA ₂	141
4.29. Variación en tiempo real del retardo medio de P ₁ para los tres SLAs considerados cuando se comparan los controladores (a) PID con PI (b) PID con P.	142
4.30. Evolución en tiempo real del término integral. (a) Controlador PID (b) Controlador PI.	143
4.31. Evolución en el tiempo del error instantáneo $e[n]$ comparando un controlador P con un (a) controlador PID (b) controlador PI.	144
4.32. Variación del retardo medio de la clase de servicio P ₁ del SLA ₁ cuando se considera un número diferente de errores en el término integral.	145
4.33. Evolución en tiempo real del retardo medio de P ₁ para distintos valores de T_{window} (a) SLA ₁ valores bajos de T_{window} (b) SLA ₁ valores altos de T_{window} (c) SLA ₂ valores bajos de T_{window} (d) SLA ₂ valores altos de T_{window}	146
4.34. Evolución en tiempo real del retardo medio de P ₁ para distintos valores de T (a) SLA ₁ valores bajos de T (b) SLA ₁ valores altos de T (c) SLA ₂ valores bajos de T (d) SLA ₂ valores altos de T.	148

4.35. Evolución en tiempo real del retardo medio de P_1 de los dos SLAs de menor prioridad considerando diferentes valores de α en DySLa. (a) SLA ₁ (b) SLA ₂	150
4.36. Evolución en tiempo real del ancho de banda máximo permitido a cada ONU del SLA ₁ y SLA ₂ considerando diferentes valores de α en DySLa. (a) SLA ₁ (b) SLA ₂	150
4.37. Retardo medio de la clase de servicio P_0 frente a la carga de red al comparar DySLa y DaSPID.	151
4.38. Retardo medio de la clase de servicio P_1 frente a la carga de red al comparar DySLa y DaSPID.	152
4.39. Retardo medio de la clase de servicio P_2 frente a la carga de red al comparar DySLa y DaSPID.	154
4.40. Evolución en tiempo real del retardo medio de la clase de servicio P_1 para las condiciones iniciales de ancho de banda de la Tabla 4.11 en DySLa y DaSPID. (a) SLA ₀ (b) SLA ₁ (c) SLA ₂	156
4.41. Evolución en tiempo real del ancho de banda máximo permitido para los escenarios de la Tabla 4.11 en DySLa y DaSPID. (a) SLA ₀ (b) SLA ₁ (c) SLA ₂	157
4.42. Evolución en tiempo real del retardo medio de la clase de servicio P_1 para DySLa y DaSPID cuando las ONUs transmiten a una tasa de 87.5 Mbit/s considerando las cotas de retardo máximo de la Tabla 4.12	158
4.43. Retardo medio obtenido en DaSPID para las clases de servicio prioritarias considerando redes LR-EAPON y EPON de diferente longitud. (a) P_0 (b) P_1	159
4.44. Retardo medio de P_0 y P_1 para cada ONU para el Escenario 1 de la Tabla 4.13.	160
4.45. Retardo medio de P_0 y P_1 para cada ONU para el Escenario 2 de la Tabla 4.13.	161
4.46. Retardo medio obtenido en DaSPID para la clase de servicio P_0 cuando varía el ancho de banda demandado a lo largo del tiempo. (a) SLA ₀ (b) SLA ₁ (c) SLA ₂	162
4.47. Retardo medio obtenido en DaSPID para la clase de servicio P_1 y el SLA ₀ cuando varía el ancho de banda demandado a lo largo del tiempo.	163
4.48. Retardo medio obtenido en DaSPID para la clase de servicio P_1 y el SLA ₁ cuando el ancho de banda demandado varía a lo largo del tiempo.	163

4.49. Retardo medio obtenido en DaSPID para la clase de servicio P ₁ y el SLA ₂ cuando el ancho de banda demandado varía a lo largo del tiempo.	164
4.50. Variación del ancho de banda máximo en DaSPID para los tres SLAs considerados cuando el ancho de banda demandado varía a lo largo del tiempo.	165
5.1. Ejemplo del operador de cruce de dos individuos en un algoritmo genético. (a) Con un único punto de cruce (b) Con dos puntos de cruce.	172
5.2. Ejemplo del operador mutación en un algoritmo genético cuando solo un gen del cromosoma es mutado.	173
5.3. Diagrama de bloques del algoritmo GA-SPID incorporando un algoritmo genético para la sintonización del controlador PID.	175
5.4. Diseño del cromosoma y codificación en el algoritmo genético desarrollado.	176
5.5. Diagrama de flujo del proceso del algoritmo genético para la sintonización de controladores PID en la gestión de recursos en redes LR-EAPON.	179
5.6. Media del error del mejor individuo para los dos intervalos analizados para las cotas de los parámetros y considerando distinto tamaño de población.	182
5.7. Evolución del ancho de banda medio asignado para distintos tiempos de muestreo T del PID en el algoritmo SPID.	183
5.8. Media del error del mejor individuo al variar el número de iteraciones del PID. (a) Para una población de 10 individuos (b) Para una población de 30 individuos.	184
5.9. Evolución de la media del error del mejor individuo para distintos tamaños de población considerados (10, 15, 20, 25 y 30) a lo largo de sucesivas generaciones.	185
5.10. Evolución de la media del error cometido por el mejor individuo de la población de cada generación para distintos tamaños de población (15, 20) e iteraciones del PID (2, 3).	186
5.11. Esquema para la sintonización automática siguiendo el método de Ziegler-Nichols en SPID.	189
5.12. Evolución en tiempo real del ancho de banda medio asignado para el algoritmo de sintonización automática basado en el método de Ziegler-Nichols considerando un intervalo de fluctuaciones de ancho de banda del 20 % por encima y por debajo del nivel garantizado del SLA ₂ (50 Mbit/s). (a) K_p inicial igual a 0.9 (b) K_p inicial igual a 2.7 (c) K_p inicial igual a 4.65.	190

5.13. Evolución en el tiempo del ancho de banda medio asignado para el algoritmo de sintonización automático basado en el método de Ziegler-Nichols con un intervalo de fluctuaciones de ancho de banda del 10 % por encima y por debajo del nivel garantizado del SLA ₂ (50 Mbit/s). (a) K_p inicial igual a 0.9 (b) K_p inicial igual a 2.7 (c) K_p inicial igual a 4.65.	192
5.14. Media del error del mejor individuo en GA-SPID para distinto tamaño de población y número de generaciones en función del tiempo de sintonización.	193
5.15. Evolución en tiempo real del ancho de banda medio asignado a una ONU de cada SLA para las sintonizaciones de la Tabla 5.3. (a) SLA ₀ (b) SLA ₁ (c) SLA ₂	195
5.16. Evolución en tiempo real del ancho de banda máximo permitido a una ONU de cada SLA para las sintonizaciones de la Tabla 5.3. (a) SLA ₀ (b) SLA ₁ (c) SLA ₂	196
5.17. Evolución en tiempo real de la desviación estándar del ancho de banda medio asignado sobre el ancho de banda garantizado para las sintonizaciones de la Tabla 5.3. (a) SLA ₀ (b) SLA ₁ (c) SLA ₂	197
5.18. Variación en tiempo real del ancho de banda máximo para los distintos escenarios de la Tabla 5.4. (a) Escenario 1 (b) Escenario 2 (c) Escenario 3 (d) Escenario 4.	199
5.19. Variación en tiempo real de la desviación del ancho de banda medio asignado sobre el ancho de banda garantizado para los distintos escenarios de la Tabla 5.4. (a) Escenario 1 (b) Escenario 2 (c) Escenario 3 (d) Escenario 4.	200
5.20. Diagrama de bloques del algoritmo GA-DaSPID con sintonización del controlador P mediante un algoritmo genético.	201
5.21. Media del error del mejor individuo considerando distinto número de iteraciones del controlador P.	204
5.22. Media del error del mejor individuo considerando distinto tamaño de población.	205
5.23. Media del error del mejor individuo a lo largo de 20 generaciones.	205
5.24. Evolución en tiempo real del retardo medio de P ₁ del SLA ₂ para el algoritmo de sintonización automático basado en el método de Ziegler-Nichols considerando un intervalo de oscilaciones del retardo del 30 % por encima y por debajo de la cota máxima de retardo para este perfil y clase de servicio. (a) K_p inicial igual a 0.2 (b) K_p inicial igual a 2.1 (c) K_p inicial igual a 4.5.	207

5.25. Media del error del mejor individuo considerando distinto tamaño de población y número de generaciones en función del tiempo de sintonización en GA-DaSPID.	208
5.26. Evolución en tiempo real del retardo medio de la clase de servicio P_0 de cada SLA para las sintonizaciones de la Tabla 5.7. (a) SLA ₀ (b) SLA ₁ (c) SLA ₂	210
5.27. Evolución en tiempo real del retardo medio de la clase de servicio P_1 de cada SLA para las sintonizaciones de la Tabla 5.7. (a) SLA ₀ (b) SLA ₁ (c) SLA ₂	211
5.28. Evolución en tiempo real del retardo medio para la clase de servicio P_1 de los tres SLAs considerados comparando la sintonización de DaSPID y GA-DaSPID para las cotas máximas de retardo recogidas en la Tabla 5.8.	212
6.1. Modelo de neurona artificial.	217
6.2. Esquema de una red neuronal multicapa totalmente interconectada.	218
6.3. Diagrama de bloques del controlador PID para la gestión de ancho de banda en SPID con sintonización adaptativa mediante redes neuronales.	227
6.4. Diseño final de la red neuronal para la sintonización adaptativa del controlador PID.	228
6.5. Diagrama de flujo del proceso de funcionamiento del algoritmo NN-SPID.	230
6.6. Evolución en tiempo real del ancho de banda medio asignado a una ONU de cada SLA para las sintonizaciones de la Tabla 6.6. (a) SLA ₀ (b) SLA ₁ (c) SLA ₂	237
6.7. Evolución en tiempo real del ancho de banda máximo permitido a una ONU de cada SLA para las sintonizaciones de la Tabla 6.6. (a) SLA ₀ (b) SLA ₁ (c) SLA ₂	238
6.8. Evolución en tiempo real de la desviación estándar del ancho de banda medio asignado sobre el ancho de banda garantizado para las sintonizaciones de la Tabla 6.6. (a) SLA ₀ (b) SLA ₁ (c) SLA ₂	239
6.9. Variación en tiempo real del ancho de banda máximo para los escenarios de la Tabla 6.7 comparando SPID, GA-SPID y NN-SPID. (a) Escenario 1 (b) Escenario 2 (c) Escenario 3 (d) Escenario 4.	241
6.10. Evolución en tiempo real de los parámetros de sintonización K_p , T_i y T_d para los cuatro escenarios considerados en la Tabla 6.7 comparando NN-SPID y GA-SPID. (a) Escenario 1 (b) Escenario 2 (c) Escenario 3 (d) Escenario 4.	242

6.11. Evolución de la desviación estándar en tiempo real del ancho de banda medio asignado sobre el ancho de banda garantizado para los cuatro escenarios considerados en la Tabla 6.7 comparando SPID, GA-SPID y NN-SPID. (a) Escenario 1 (b) Escenario 2 (c) Escenario 3 (d) Escenario 4.	243
6.12. Evolución en tiempo real del ancho de banda máximo cuando el entrenamiento de la red neuronal se realiza con ejemplos correctos y erróneos. (a) SLA ₁ (b) SLA ₂	245
6.13. Evolución en tiempo real de los parámetros de sintonización K_p , T_i y T_d cuando la etapa de entrenamiento se realiza con ejemplos erróneos.	246
6.14. Evolución en tiempo real del ancho de banda demandado para una ONU de cada SLA (azul) y variación en tiempo real del ancho de banda medio asignado (negro) considerando una frecuencia de cambio de 1800 s. (a) SLA ₀ (b) SLA ₁ (c) SLA ₂	247
6.15. Evolución en tiempo real del ancho de banda demandado por cada ONU (azul) y variación en tiempo real de los parámetros de sintonización K_p , T_i y T_d considerando una frecuencia de cambio de 1800 s.	248
6.16. Evolución en tiempo real del ancho de banda demandado para una ONU de cada SLA (azul) y variación en tiempo real del ancho de banda medio asignado (negro) considerando una frecuencia de cambio de 900 s. (a) SLA ₀ (b) SLA ₁ (c) SLA ₂	249
6.17. Ancho de banda demandado por cada ONU (azul) y variación en tiempo real de los parámetros de sintonización K_p , T_i y T_d considerando una frecuencia de cambio de 900 s.	250
7.1. Ciclo para la resolución de problemas del razonamiento basado en casos. .	255
7.2. Ejemplo de una red con encaminamiento por longitud de onda donde se han establecido tres <i>lightpaths</i>	258
7.3. Factor Q de los <i>lightpaths</i> en función de su longitud para la red GÉANT2 con 32 longitudes de onda considerando distintas cargas de red.	260
7.4. Porcentaje de aciertos en la clasificación de la QoT de los <i>lightpaths</i> para los métodos R-CBR y FixE-CBR para la red DT.	268
7.5. Porcentaje de aciertos en la clasificación de la QoT de los <i>lightpaths</i> para los métodos R-CBR y FixE-CBR para la red GÉANT2.	269
7.6. Porcentaje de aciertos del CBR en la clasificación de la QoT en la zona de incertidumbre para las redes DT y GÉANT2 con 32 longitudes de onda y aciertos apostando siempre a la clase más probable.	270

7.7. Tiempo de computación para la clasificación de un <i>lightpath</i> para la red DT con 32 longitudes de onda utilizando la <i>Q-Tool</i> , R-CBR y FixE-CBR.	273
7.8. Tiempo de computación para la clasificación de un <i>lightpath</i> para la red GÉANT2 con 32 longitudes de onda utilizando la <i>Q-Tool</i> , R-CBR y FixE-CBR.	274
7.9. Porcentaje de aciertos para tamaños pequeños de la base de conocimiento con y sin optimización previa para la red DT con 64 longitudes de onda.	275
7.10. Porcentaje de aciertos y tamaño de la base de conocimiento para diferente número de configuraciones de red testeadas en el proceso de construcción de la base de conocimiento para la red DT con 64 longitudes de onda.	276
7.11. Esquema del banco de pruebas desarrollado por la Universidad de Dinamarca para la validación experimental del estimador cognitivo.	278
7.12. Banco de pruebas del proyecto CHRON.	279
7.13. Diagrama de bloques propuesto para la sintonización adaptativa del controlador PID en SPID mediante un sistema CBR.	281
7.14. Variación en tiempo real de la desviación del ancho de banda medio asignado sobre el ancho de banda garantizado para las condiciones de red del Escenario 1. (a) SLA ₀ (b) SLA ₁ (c) SLA ₂ .	284
7.15. Evolución en tiempo real de los parámetros de sintonización en los algoritmos CBR-SPID y NN-SPID para el Escenario 1.	286
7.16. Variación en tiempo real de la desviación del ancho de banda medio asignado sobre el ancho de banda garantizado para las condiciones de red del Escenario 2. (a) SLA ₀ (b) SLA ₁ (c) SLA ₂ .	287
7.17. Evolución en tiempo real de los parámetros de sintonización en CBR-SPID y NN-SPID para las condiciones del Escenario 2.	288
7.18. Ancho de banda demandado para una ONU de cada SLA (azul) y variación en tiempo real del ancho de banda medio asignado (negro) considerando una frecuencia de cambio de 1800 s. (a) SLA ₀ (b) SLA ₁ (c) SLA ₂ .	289
7.19. Ancho de banda demandado por cada ONU (azul) y variación en tiempo real de los parámetros de sintonización K_p , T_i y T_d considerando una frecuencia de cambio de 1800 s.	290
7.20. Ancho de banda demandado para una ONU de cada SLA (azul) y variación en tiempo real del ancho de banda medio asignado (negro) considerando una frecuencia de cambio de 900 s. (a) SLA ₀ (b) SLA ₁ (c) SLA ₂ .	291

Índice de tablas

2.1. Principales diferencias entre 1G-EPON y 10G-EPON.	19
2.2. Principales diferencias entre GPON y XG-PON.	19
3.1. Acciones a realizar en DySGAB para la adaptación del ancho de banda máximo.	54
3.2. Parámetros de red considerados en el entorno inicial de simulación.	62
3.3. Valores de los parámetros óptimos relacionados con el diseño del algoritmo DySGAB.	68
3.4. Pesos utilizados en los escenarios considerados para establecer los anchos de banda iniciales para los tres perfiles de usuario.	70
3.5. Niveles de ancho de banda garantizado para cada SLA en los dos escenarios considerados.	75
3.6. Niveles de ancho de banda garantizado considerados para cada SLA a lo largo del tiempo.	76
3.7. Distribución de ONUs en los nuevos escenarios de red considerados.	78
3.8. Contribución de cada SLA a la carga total de red para los distintos escenarios considerados con tráfico asimétrico.	81
4.1. Diferentes estrategias de asignación de ancho de banda soportadas por SPID para garantizar los niveles mínimos de ancho de banda estipulados para cada perfil.	103
4.2. Parámetros de red considerados en el escenario de simulación de SPID.	105
4.3. Parámetros de los controladores P, PI y PID para el método de la respuesta en frecuencia de Ziegler-Nichols.	105
4.4. Valores de sintonización de los controladores P, PI y PID en SPID.	107
4.5. Pesos utilizados en los escenarios considerados para establecer los anchos de banda máximos iniciales para los tres perfiles de usuario.	116

4.6. Niveles de ancho de banda garantizado considerados para cada SLA a lo largo del tiempo de simulación.	118
4.7. Número de ONUs asociadas a cada perfil de abonado en los escenarios considerados.	122
4.8. Contribución de cada SLA a la carga total de red para los distintos escenarios considerados con tráfico asimétrico en SPID.	124
4.9. Retardos máximos considerados en el escenario de simulación en DaSPID para las distintas clases de servicio.	140
4.10. Valores de sintonización de los controladores P, PI y PID en DaSPID. . . .	142
4.11. Pesos utilizados en los escenarios considerados para establecer los anchos de banda iniciales en DaSPID y DySLa para los tres perfiles de usuario considerados.	155
4.12. Cotas máximas de retardo consideradas en el nuevo escenario de simulación.	158
4.13. Contribución de cada SLA a la carga total de red para los distintos escenarios considerados con tráfico asimétrico en DaSPID.	160
 5.1. Parámetros de red en el escenario de simulación para GA-SPID.	180
5.2. Parámetros finales del algoritmo genético diseñado para GA-SPID.	187
5.3. Valores de los parámetros K_p , T_i y T_d seleccionados para cada tipo de sintonización.	194
5.4. Niveles de ancho de banda garantizado estipulados para cada SLA en los escenarios considerados.	198
5.5. Parámetros de red en el escenario de simulación para GA-DaSPID. . . .	202
5.6. Parámetros finales del algoritmo genético diseñado para sintonizar el controlador P en GA-DaSPID.	206
5.7. Valores del parámetro K_p seleccionado para cada tipo de sintonización. .	209
5.8. Cotas máximas de retardo consideradas para las diferentes clases de servicio en el nuevo escenario de simulación.	212
 6.1. Parámetros de red considerados en el entorno de simulación de NN-SPID. .	231
6.2. Comparativa de la desviación estándar del término $\overline{B_{alloc}^{onu_i}}$ sobre el término $B_{guarantee}^{sla \in onu_i}$ considerando diferente número de iteraciones del PID. . . .	233
6.3. Comparativa de la desviación estándar del término $\overline{B_{alloc}^{onu_i}}$ sobre el término $B_{guarantee}^{sla \in onu_i}$ considerando diferente número de iteraciones en la etapa de entrenamiento.	233
6.4. Comparativa de la desviación estándar del término $\overline{B_{alloc}^{onu_i}}$ sobre el término $B_{guarantee}^{sla \in onu_i}$ considerando diferente número de neuronas en la capa oculta. .	234

6.5.	Parámetros finales escogidos para el diseño de la red neuronal en NN-SPID.	235
6.6.	Valores de los parámetros K_p , T_i y T_d seleccionados para cada tipo de sintonización comparada.	236
6.7.	Niveles de ancho de banda garantizado estipulados para cada SLA en los cuatro escenarios considerados.	240
6.8.	Comparativa de la desviación estándar del término $\overline{B_{alloc}^{onu_i}}$ sobre el término $B_{guarantee}^{sla \in onu_i}$ considerando diferente número de ejemplos en el entrenamiento de la red neuronal.	244
7.1.	Límite superior e inferior (en km) de la zona de incertidumbre para las redes DT y GÉANT2.	266
7.2.	Porcentaje de acierto en la clasificación de los <i>lightpaths</i> para la pareja origen-destino con peor porcentaje de aciertos en la clasificación.	272
7.3.	Parámetros de red en el escenario de simulación de CBR-SPID.	283
7.4.	Niveles de ancho de banda garantizado considerados para cada SLA a lo largo del tiempo en los dos escenarios analizados.	284
7.5.	Desviación del ancho de banda medio asignado sobre el ancho de banda garantizado en 300 s, 600 s, 900 s y 1200 s para las condiciones del Escenario 1 comparando SPID, CBR-SPID y NN-SPID.	285
7.6.	Desviación del ancho de banda medio asignado sobre el ancho de banda garantizado en 300 s, 600 s, 900 s y 1200 s para las condiciones del Escenario 2 comparando SPID, CBR-SPID y NN-SPID.	287
8.1.	Características principales de los algoritmos propuestos en esta Tesis para redes LR-EAPON.	300

Capítulo 1

Introducción

En los últimos años la red de telecomunicaciones ha experimentado grandes e importantes cambios. El crecimiento y la llegada de nuevas aplicaciones y servicios, junto con un mercado más competitivo, dan lugar a un escenario complejo en el que la calidad de servicio (QoS, *Quality of Service*) juega un papel muy importante. En este sentido, los usuarios finales cada vez demandan un mayor nivel de calidad de servicio, debido al creciente interés en las aplicaciones multimedia y de tiempo real. En concreto, según los datos presentados en [1], el tráfico de Internet se ha cuadriplicado en los últimos cinco años y se espera que se triplique en los próximos tres, con una tasa de crecimiento anual del 23 % desde 2012 hasta 2017. Además, se estima que de todo el tráfico de Internet que consuma cada usuario en 2017, el 69 % se corresponderá con tráfico de vídeo. Así mismo, se prevé que los juegos en línea tengan una tasa de crecimiento anual del 22 % en el tramo de 2012 a 2017. Por su parte, los proveedores de servicio necesitan adoptar e implementar nuevas políticas para la gestión de la calidad de servicio con el objetivo de garantizar la satisfacción de los usuarios. Esto supone un reto significativo, puesto que las redes de telecomunicaciones deben acomodar tráfico con unos requisitos muy diversos, todo ello, además, maximizando los recursos disponibles y minimizando los costes asociados al despliegue.

Por tanto, la red de comunicaciones transporta una multitud de datos y aplicaciones. En este sentido, los distintos servicios soportados presentan restricciones diversas, no solo en términos de ancho de banda, sino también en otros parámetros como el retardo medio, la tasa de pérdida de paquetes o la variación del retardo de los paquetes. Por tanto, es necesaria una gestión eficiente de los recursos disponibles en todos los tramos de la red, tanto en la red troncal como en la red de acceso, para así poder garantizar una calidad de servicio extremo a extremo. Esto supone además, que las redes de hoy en día tienen que reaccionar ante cambios que se produzcan en las condiciones de red,

especialmente en tiempo real, tales como cambios en las demandas de tráfico, fallos en los enlaces, cambios en los requisitos de calidad establecidos, etc. En este sentido, las redes de comunicaciones necesitan incluir hasta cierto nivel, una auto-adaptación a las nuevas condiciones, para intentar ser eficientes e independientes de un operador manual y así asegurar una respuesta rápida y una mayor eficiencia en la gestión de los recursos disponibles. Así pues, la incorporación de técnicas adaptativas en la gestión y control de la red resulta de gran interés. De la misma forma, la integración de cognición en las redes permite, aparte de una auto-adaptación, una gestión proactiva de la red, ya que las técnicas cognitivas permiten predecir cambios en las condiciones a partir del aprendizaje y del análisis de experiencias pasadas.

Hoy en día, para poder ofrecer con ciertas garantías todos los servicios y aplicaciones emergentes, haciendo frente a la gran demanda de ancho de banda y gestionando de forma efectiva el tráfico generado, las redes de telecomunicaciones necesitan una gran capacidad. En las redes troncales o redes de transporte, el medio de transmisión más común es la fibra óptica. Además de ofrecer un gran ancho de banda, la fibra óptica proporciona una infraestructura común sobre la que se pueden transportar multitud de servicios. Para explotar toda la capacidad que la fibra es capaz de ofrecer de una forma eficiente y económica, las redes de transporte actuales utilizan técnicas de multiplexación por longitud de onda (WDM, *Wavelength Division Multiplexing*). Con esta técnica, se pueden transmitir diferentes canales de forma simultánea por la fibra utilizando distintas longitudes de onda. No obstante, a medida que la señal óptica se propaga por estos canales, ésta se ve afectada por ciertos efectos físicos que atenúan y distorsionan la señal. De esta forma, la calidad de transmisión (QoT, *Quality of Transmission*) se puede reducir e incluso puede no llegar a cumplir con los requisitos estipulados por los proveedores de servicio. En la actualidad, existe un creciente interés en el desarrollo de algoritmos que llevan a cabo el enrutamiento de las conexiones ópticas y la asignación de longitudes de onda teniendo en cuenta estas limitaciones físicas. En concreto, resulta muy interesante determinar si una conexión tendrá la calidad suficiente antes de su establecimiento, o si su establecimiento influirá en la calidad de otras conexiones ya establecidas. Todo esto, con el objetivo de hacer un uso eficiente de los recursos existentes en la red. Por ese motivo, una parte de esta Tesis se centrará en el diseño de técnicas cognitivas en redes troncales con encaminamiento por longitud de onda, que sean capaces de estimar y de predecir con cierta exactitud si una conexión cumplirá con la calidad de transmisión necesaria.

Por otro lado, aunque la red troncal ha experimentado grandes avances en las últimas décadas, en las redes de acceso los cambios han comenzado a producirse en los últimos

años. En este sentido, las redes de acceso ópticas pasivas o redes PON (*Passive Optical Networks*) están mostrando un papel fundamental en el despliegue del segmento de acceso, ya que se han perfilado como la mejor solución para eliminar el “cuello de botella” en el que se había convertido la red de acceso, gracias al gran ancho de banda que son capaces de ofrecer. Además, existe un gran interés por las arquitecturas de redes PON de largo alcance (LR-PON, *Long-Reach PON*), que extienden el alcance de 20 km de las redes PON comunes hasta los 100 km. Esta nueva infraestructura resulta ser muy económica, ya que no solo elimina parte de los elementos electrónicos y simplifica la arquitectura de la red, integrando la red metropolitana y de acceso en una sola, sino que incorporan a un mayor número de abonados de forma que los costes compartidos se reducen. Tanto las redes PON como las redes LR-PON están basadas, en su forma más simplificada, en una arquitectura típicamente punto-multipunto, donde todos los usuarios comparten la transmisión en el canal ascendente, esto es, desde la ubicación del usuario hasta la oficina central. Por tanto, es necesario un mecanismo de control de acceso al medio (MAC, *Medium Access Control*) que impida que varios usuarios transmitan simultáneamente. El diseño de estos mecanismos es uno de los principales retos en el segmento de acceso, ya que la eficiencia de estas redes depende en gran medida de que los algoritmos repartan de forma justa y óptima el ancho de banda disponible entre los usuarios sin que se produzcan colisiones y se garanticen ciertos niveles de calidad de servicio.

En este sentido, gran parte de la investigación en estas redes se centra en el desarrollo de algoritmos de asignación de ancho banda dinámicos (DBA, *Dynamic Bandwidth Allocation*) que hagan un reparto eficiente del ancho de banda disponible, ofreciendo además la calidad de servicio estipulada por el proveedor de servicios. De este modo, es muy conveniente que los algoritmos contemplen diferenciación de servicios, para poder tratar por separado cada tipo de tráfico según sus requisitos, y que incorporen también una diferenciación de usuarios, para que así, dependiendo del nivel de prioridad del abonado, se ofrezcan una serie de garantías sobre los diversos servicios soportados. En el caso de las redes LR-PON, estos algoritmos no solo tienen que ofrecer calidad de servicio, sino que además deben tener en cuenta los inconvenientes de aumentar la cobertura extremo a extremo hasta los 100 km o más. Así mismo, también es primordial que se adapten a diferentes condiciones y escenarios de red. Por ese motivo, el eje central de esta Tesis está enfocado al desarrollo de algoritmos de asignación dinámica de ancho de banda que gestionen y repartan los recursos disponibles, ofreciendo los requisitos de calidad de servicio estipulados en redes LR-PON y que además ofrezcan características de adaptabilidad en tiempo real bajo diferentes condiciones y escenarios de red.

1.1. Objetivos de la Tesis

El objetivo global de esta Tesis Doctoral es el **diseño e implementación de mecanismos de control adaptativos y cognitivos que gestionen de forma eficiente los recursos en redes ópticas** para optimizar la calidad de servicio y la calidad de transmisión. Este objetivo global, tal y como se ha comentado anteriormente, se divide en dos objetivos, atendiendo al segmento de red en el que se van a implementar.

Por tanto, un primer objetivo, que abarca gran parte de la investigación de esta Tesis, está enfocado a la red de acceso y consiste en el **desarrollo de algoritmos de asignación dinámica de ancho de banda en redes de largo alcance LR-EPON**, que controlen en tiempo real los recursos de la red con el fin de garantizar la calidad de servicio acorde a los requisitos contratados por los usuarios finales y que se adapten a las diferentes condiciones de red. Algunos de los objetivos parciales para la consecución de este primer objetivo son los siguientes:

- **Diseño, implementación y evaluación de algoritmos DBA en redes LR-EPON para controlar dinámicamente el ancho de banda asignado en función de la prioridad de los usuarios.** Se pretenden implementar algoritmos DBA capaces de ofrecer calidad de servicio mediante la diferenciación de perfiles de abonado, asegurando de forma adaptativa unos niveles mínimos de ancho de banda garantizados por el proveedor de servicios. Además, para lograr una mayor aplicabilidad, este algoritmo se podrá emplear en redes EPON y LR-EPON con diferente rango de cobertura.
- **Incorporación de sistemas de control de procesos en algoritmos DBA para optimizar la asignación de recursos y la calidad de servicio.** Con el fin de mejorar y optimizar las prestaciones de los algoritmos que controlan dinámicamente la calidad de servicio, se van a incluir técnicas de control de procesos, utilizadas en otros campos, con el fin de controlar de forma automática la calidad de servicio ofrecida a los abonados de redes LR-EPON. Con el uso de estas técnicas se pretende aumentar la estabilidad del sistema, reduciendo los errores cometidos en la consecución de los requisitos de calidad estipulados. En concreto, se ha optado por incluir controladores Proporcional-Integral-Derivativo (PID, *Proportional-Integral-Derivative*) debido a su simplicidad, robustez, aplicabilidad y buen funcionamiento en el control de parámetros en distintos tipos de procesos. Para abordar este objetivo, se llevará a cabo el diseño e implementación de los siguientes algoritmos de gestión de recursos basados en controladores PID:

- Diseño, implementación y evaluación de algoritmos para asegurar dinámicamente que el ancho de banda asignado se ajuste a los niveles mínimos garantizados establecidos por los proveedores de servicio, según la prioridad del perfil de abonado contratado por cada usuario.
- Diseño, implementación y evaluación de algoritmos para controlar el retardo medio de servicios y aplicaciones con restricciones en este parámetro. Este tipo de estrategias serán diseñadas con el objetivo de proporcionar una calidad de servicio más global y acorde tanto a la prioridad de los abonados como a la prioridad de los servicios ofertados.

Por otro lado, para otorgar una mayor aplicabilidad de los algoritmos desarrollados, éstos serán diseñados para ser eficientes en redes EPON y LR-EAPON de diferente longitud extremo a extremo.

- **Diseño de métodos de sintonización automática y/o adaptativa de controladores PID para optimizar la calidad de servicio en redes LR-EAPON.** Para automatizar y optimizar el proceso de sintonización del que depende el funcionamiento de los controladores PID, se utilizarán y adaptarán a nuestro escenario de red LR-EAPON tres tipos de técnicas de inteligencia artificial diferentes:

- Diseño e implementación de controladores PID con sintonización basada en algoritmos genéticos para la gestión de recursos en redes LR-EAPON. En este sentido, los algoritmos genéticos son métodos que han demostrado en numerosos campos un gran potencial en la resolución de problemas de búsqueda y optimización. Con su uso, se pretenden encontrar los parámetros óptimos de sintonización para cada estado de la red de acceso de un modo automático, rápido y preciso.
- Diseño e implementación de controladores PID con sintonización basada en redes neuronales para la gestión de recursos en redes LR-EAPON. Se pretende que la red neuronal sintonice en tiempo real el controlador de forma adaptativa, dependiendo de la situación concreta de la red LR-EAPON en cada momento.
- Diseño e implementación de controladores PID con sintonización basada en técnicas cognitivas, en concreto usando la técnica de Razonamiento Basado en Casos, para la gestión de recursos en redes LR-EAPON. A partir de experiencias de sintonización previas, se pretende sintonizar de forma adaptativa y en tiempo real el controlador PID según las condiciones o estado de la red actuales.

- **Caracterización de patrones de tráfico dinámico y asimétrico en redes LR-EPON.** En un escenario de red real, el patrón de tráfico no va a ser constante a lo largo de las horas del día. Esto se debe a que algunos usuarios generarán más tráfico a ciertas horas, mientras que otros lo harán a otras (tráfico asimétrico), y además un mismo usuario no generará el mismo tráfico de forma homogénea a lo largo del día (tráfico dinámico). Por lo tanto, en primer lugar será necesario caracterizar un patrón de tráfico que cumpla estas premisas en una red de acceso y después realizar pruebas de algunos de los algoritmos de asignación dinámica de recursos desarrollados previamente bajo dichos patrones más realistas en un contexto de red.

El segundo objetivo está enfocado al segmento de red troncal. En concreto, consiste en el **diseño e implementación de un estimador cognitivo de calidad de transmisión de conexiones ópticas en redes troncales**. De esta forma, se pretende estimar la calidad de transmisión antes del establecimiento de una nueva conexión y ver si su establecimiento afecta al resto de conexiones existentes para así utilizar de forma eficiente los recursos disponibles.

1.2. Estructura de la Tesis

La memoria de la Tesis Doctoral se encuentra dividida en ocho capítulos. En este primer capítulo se ha descrito la motivación y los principales objetivos a cumplir en la investigación llevada a cabo. En el Capítulo 2 se realiza una introducción a las redes de acceso ópticas pasivas o redes PON, presentando las principales arquitecturas y estándares para el despliegue de dichas redes en el segmento de acceso. En concreto, esta Tesis está enfocada en la tecnología EPON basada en el estándar Ethernet debido a su gran nivel de despliegue. Posteriormente, el capítulo se centra en el principio de funcionamiento de las redes PON de largo alcance (LR-PON). En primer lugar, se presenta una revisión de las tecnologías utilizadas para el despliegue de redes LR-PON, así como de los mecanismos utilizados para proporcionar calidad de servicio en estas arquitecturas de red. El capítulo finaliza presentando un resumen de los principales algoritmos de asignación de ancho de banda existentes en la literatura, diseñados para la gestión de recursos y de calidad de servicio en este tipo de redes, analizando las ventajas y desventajas de cada uno de ellos.

El Capítulo 3 está centrado en el diseño e implementación de un algoritmo de asignación de ancho de banda en redes LR-EPON que lleva a cabo diferenciación de usuarios. En concreto, el algoritmo está enfocado a garantizar dinámicamente unos niveles mínimos

de ancho de banda estipulados por el proveedor de servicios de acuerdo a la prioridad de los perfiles de abonado que soporte. Tras describir su funcionamiento y hallar los parámetros óptimos de su diseño, el capítulo finaliza con la evaluación de sus prestaciones bajo diferentes escenarios de red y distintos patrones de tráfico para demostrar sus funcionalidades, comparándolo así mismo con otros algoritmos existentes en la literatura con el mismo objetivo.

En el Capítulo 4 se presentan otros dos algoritmos de asignación de ancho de banda dinámico desarrollados en esta Tesis. Estos dos nuevos algoritmos incorporan un sistema de control PID con la finalidad de ofrecer un control del ancho de banda más robusto y estable, ya que tienen en cuenta errores presentes, pasados y predicciones de errores futuros a la hora de realizar el ajuste dinámico del ancho de banda disponible para ofrecer la calidad de servicio necesaria. En la primera parte del capítulo, se presenta un algoritmo que lleva a cabo diferenciación de usuarios mediante el diseño de un controlador PID que garantiza niveles mínimos de ancho de banda estipulados por el proveedor de servicios en función de la prioridad del perfil contratado por el abonado. Posteriormente, se realiza el estudio de simulación donde se analizan las prestaciones y comportamiento de dicho algoritmo bajo distintos escenarios de red y patrones de tráfico, comparándolo con otros algoritmos existentes. La segunda parte del capítulo se centra en el diseño e implementación de un algoritmo para gestionar de forma adaptativa el retardo del tráfico y que lleva a cabo una doble diferenciación, de usuarios y de servicios. En concreto, el algoritmo desarrollado mantiene, a través de un controlador PID, el retardo de las diferentes clases de servicio prioritario bajo unos límites estipulados por el proveedor de servicios dependiendo de la prioridad del usuario. Finalmente, se evalúan las prestaciones de dicho algoritmo bajo diferentes escenarios de red y patrones de tráfico, comparándolo con propuestas previas existentes en la literatura.

En este sentido, uno de los aspectos fundamentales del uso de controladores PID es su proceso de sintonización, esto es, hallar las constantes de las que depende su funcionamiento, para que el controlador PID cumpla con el objetivo para el cual ha sido diseñado. Por este motivo, los siguientes capítulos de la Tesis se centran en el diseño de técnicas de sintonización optimizadas para los controladores PID. En concreto, el Capítulo 5 presenta el diseño de un nuevo método de sintonización basado en el uso de algoritmos genéticos. Este nuevo algoritmo permite automatizar el proceso de sintonización de los controladores PID, con el fin de reducir el tiempo de sintonización que necesitan otros métodos manuales que realizan las mismas funciones, y para obtener una mayor precisión en el cálculo de dichos parámetros. En la primera parte del capítulo, el nuevo método de sintonización se aplica al algoritmo diseñado en el Capítulo 4 que ga-

rantiza niveles de ancho de banda mínimos a diferentes perfiles de abonado y se analizan sus prestaciones. En la segunda parte del capítulo, el método de sintonización diseñado basado en algoritmos genéticos se incorpora al controlador PID que gestiona el retardo y, finalmente, se realiza un estudio de simulación donde se evalúa su comportamiento.

En el Capítulo 6, siguiendo con el desarrollo de métodos de sintonización adaptativos y automáticos, se lleva a cabo el diseño e implementación de un método de sintonización adaptativo basado en redes neuronales para ofrecer una sintonización óptima que se adapte a diferentes cambios en tiempo real en las condiciones de una red de acceso LR-EAPON.

El Capítulo 7 de la Tesis está centrado en la incorporación de técnicas cognitivas en redes ópticas de acceso y troncales, con el fin de realizar una gestión eficiente de los recursos disponibles. En primer lugar, se introduce la técnica cognitiva que se va a utilizar para llevar a cabo la optimización. Tras ello, la primera parte del capítulo está dedicada a las redes ópticas troncales. Por ello, en primer lugar se realiza una breve introducción a las redes troncales actuales y a su evolución en los últimos años. A continuación, se presenta el diseño y descripción de un estimador de calidad de transmisión que se ha desarrollado para redes troncales ópticas mediante el uso de cognición, así como el análisis de sus prestaciones a través de su correspondiente estudio de simulación. En la segunda parte del capítulo esta técnica cognitiva se introduce en el proceso de sintonización de controladores PID, para lograr una técnica de sintonización adaptativa y automática en tiempo real. El capítulo finaliza con un análisis de simulación de este nuevo método de sintonización desarrollado bajo una arquitectura de red de acceso LR-EAPON.

En el último capítulo de la Tesis, el Capítulo 8, se recogen las conclusiones más relevantes que se extraen de los resultados obtenidos en este trabajo de investigación desarrollado, así como las posibles líneas futuras y nuevas líneas de investigación que podrían abrirse tras la consecución de esta Tesis Doctoral.

Capítulo 2

Redes de acceso ópticas pasivas

2.1. Introducción

Este capítulo comienza presentando una visión general de las redes de acceso ópticas pasivas basadas en fibra óptica o redes PON (*Passive Optical Network*), ya que están consideradas como la tecnología idónea para cubrir las necesidades actuales del segmento de acceso, hecho corroborado por su actual despliegue masivo. En primer lugar, se explican los conceptos básicos de estas redes y se presentan las características principales de los estándares utilizados en el despliegue de las mismas, así como datos actuales de su grado de implantación a nivel mundial. A continuación, se introducen los conceptos fundamentales de las redes ópticas pasivas de largo alcance (LR-PON, *Long-Reach Passive Optical Network*), por ser una tecnología que está emergiendo con mucha fuerza gracias a la reducción de costes que supone su implantación. En este sentido, se presentan en primer lugar las diferentes tecnologías existentes en la literatura para su despliegue, así como un resumen de los diferentes mecanismos de asignación de recursos y calidad de servicio desarrollados para este tipo de redes.

2.2. Redes de acceso PON

La red de acceso se puede definir como el conjunto de elementos que permiten conectar a cada abonado con la central local de la que es dependiente. En los últimos años, a esta red se le ha asociado el nombre de “primera milla” (*first mile*) o “última milla” (*last mile*) [2].

La red de acceso, tal y como se observa en la Figura 2.1, se compone principalmente de la oficina central (CO, *Central Office*) y, en el otro extremo de la red, del usuario final conectado. El tipo de enlace que une ambos extremos de la red dependerá de la tecnología

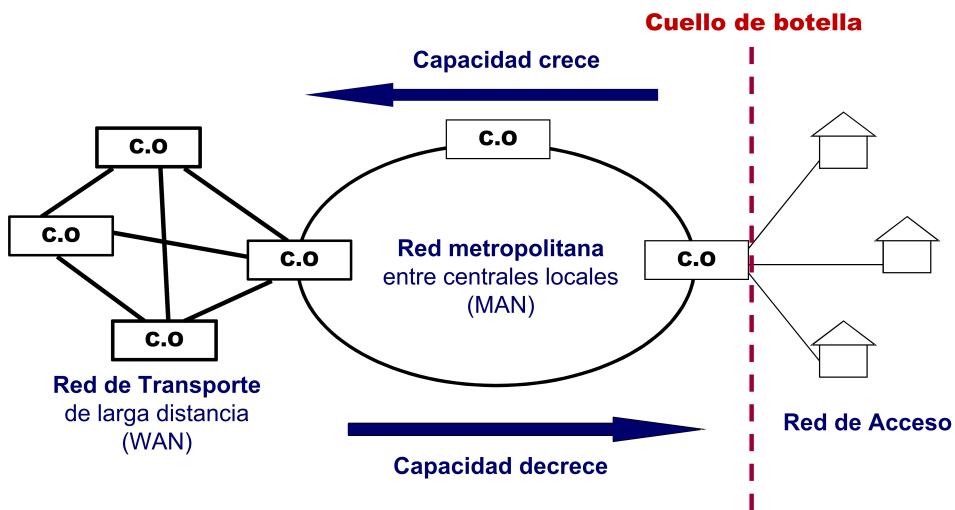


Figura 2.1: Esquema de una red de comunicaciones.

desplegada, pudiendo ser cobre, cable coaxial, fibra óptica o tecnología inalámbrica.

Actualmente, el tráfico en las redes se ha incrementado a unas tasas sin precedentes, propiciando que en las últimas décadas las redes de transporte hayan evolucionado y experimentado un sustancial crecimiento. Sin embargo, los cambios significativos en la red de acceso se están produciendo en estos últimos años. Las tecnologías desplegadas hasta hace poco en la red de acceso, tales como la línea digital de abonado (*DSL, Digital Subscriber Line*) y cable coaxial, no eran capaces de cubrir todo el ancho de banda necesario para abastecer la creciente demanda de los usuarios finales de la red. La perceptible carencia de estas tecnologías y los nuevos requisitos de los servicios emergentes propiciaron el incremento del “cuello de botella” en el acceso. Por este motivo, fue necesario un cambio en la infraestructura de acceso y la fibra óptica se planteó como una buena solución debido a su enorme capacidad para transportar información, ofreciendo en este sentido un gran ancho de banda de modo eficiente, relativamente económico y flexible. Además, la fibra óptica presenta una mayor robustez frente a otros medios de conexión tales como el coaxial o el par trenzado, así como una mayor capacidad de transmisión, menor atenuación y tasa de error e inmunidad ante interferencias electromagnéticas.

En este sentido se desarrollaron las infraestructuras de banda ancha o de fibra hasta un punto *x*, *FTTx* (*Fiber to the x*) que sustituyen el cobre o cable coaxial por fibra óptica hasta un punto cercano al usuario final [3]. Dependiendo de la cercanía de la fibra hasta el usuario, se pueden distinguir las siguientes variantes de redes FTTx (representadas en la Figura 2.2):

- **FTTN** (*Fiber To The Neighborhood* o *Fiber To The Node*, fibra hasta el vecindario

o fibra hasta el nodo). La fibra se coloca hasta puntos alejados del usuario, típicamente distancias de 1 km, siendo la parte final de la conexión hasta el usuario de cobre. A partir de este punto, se puede utilizar tecnología VDSL (*Very high speed Digital Subscriber Line*) para alcanzar velocidades de bajada alrededor de los 25-50 Mbit/s y velocidades de subida de 2-10 Mbit/s.

- **FTTC** (*Fiber To The Curb*, fibra hasta el bordillo o acera). En este caso, la fibra llega hasta distancias de unos 300 m del usuario. Después se puede colocar un multiplexor VDSL2 para proporcionar velocidades de bajada alrededor de los 50-100 Mbit/s y de 5-15 Mbit/s de subida [4].
- **FTTB** (*Fiber To The Building*, fibra hasta el edificio). La fibra llega a una distancia mayor que en los dos casos anteriores, hasta el edificio del usuario. La distribución final de las conexiones es de cobre. En este caso, se puede colocar un multiplexor VDSL2 o un commutador Ethernet en la base del edificio proporcionando velocidades de bajada de 100 Mbit/s y velocidades de subida en torno a los 10-20 Mbit/s. Las distancias desde los edificios a la oficina central del operador pueden ser de hasta decenas de kilómetros [4].
- **FTTH** (*Fiber To The Home*, fibra hasta el hogar). La fibra llega hasta el propio domicilio del usuario. Dependiendo de la tecnología usada, las distancias desde las casas hasta la oficina central pueden llegar a ser desde 20 km hasta 100 km. En el caso de FTTH se puede usar un commutador Ethernet o una unidad de red óptica (ONU, *Optical Network Unit*) para proveer una velocidad simétrica de bajada y subida de más de 100 Mbit/s (con una capacidad superior a 1 Gbit/s en ciertas tecnologías).

De los distintos tipos de tecnologías FTTx, los dos más desplegados hoy en día son FTTB y FTTH. En concreto, a mediados del 2013, FTTH/B representaba el 66 % de todas las suscripciones FTTx [5]. En principio, ambas infraestructuras ofrecen capacidades de 100 Mbit/s, pero FTTH es más idóneo para satisfacer las necesidades futuras en términos de capacidad [6]. En la actualidad, muchos operadores a nivel mundial están ya desplegando la fibra óptica hasta la ubicación del usuario final, lo que hace que FTTH se esté consolidando como la tecnología líder en el acceso. Tal y como reflejan los datos de mediados de 2013, Asia sigue siendo la región dominante en cuanto a su despliegue. En concreto, Japón y Corea del Sur con 24 y 11 millones de usuarios a mediados de 2013 [5], respectivamente, han sido los pioneros globales del acceso basado en fibra óptica, optando por FTTH. No obstante, en 2013 China se convirtió en el país líder, superando

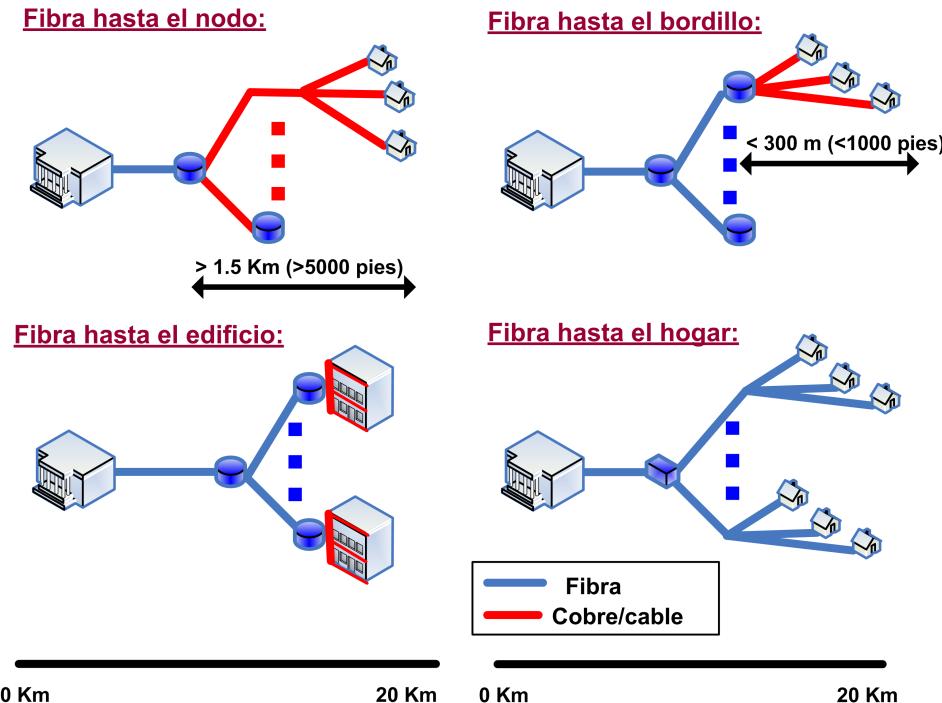


Figura 2.2: Escenarios de distintos tipos de despliegues FTTx.

a Japón, con 32 millones de usuarios [5]. Además, se prevé que para finales del año 2016 China cuente con 110 millones de usuarios FTTx [7]. Mientras tanto, Estados Unidos se sitúa en cuarta posición en la clasificación por países, con 9 millones de usuarios. Por otro lado, Europa ha conseguido un incremento del 33 % en el número de usuarios en el año 2013, gracias al despegue de los países de Centroeuropa y Europa del Este. En particular, Rusia es un mercado que ha experimentado un gran crecimiento con casi 9 millones de usuarios a finales de 2013 [8], así como Ucrania, con algo más de 1 millón de usuarios. En el oeste de Europa, en Diciembre de 2013, Francia contaba ya con 1.2 millones de usuarios, seguido de cerca por Suecia, con un número de usuarios similar [8]. Por último, África y Oriente Medio cuentan con el 0.7 % de los usuarios de FTTH/B a nivel mundial, mientras que Latinoamérica constituye el 0.5 % [9].

Por otro lado, para el despliegue de fibra óptica en la red de acceso se han propuesto diversas topologías (Figura 2.3) [10–12]. Una de ellas es la topología punto a punto (P2P, *Point to Point*), donde existe una fibra dedicada para conectar a cada usuario con la oficina central. Otro ejemplo es la topología punto-multipunto activa donde existe un tramo con una única fibra que llega hasta un punto de distribución, donde se encuentra un conmutador, y a partir de ese punto se utilizan fibras para cada usuario. Por último, las topologías punto-multipunto pasivas tienen la misma filosofía que las topologías

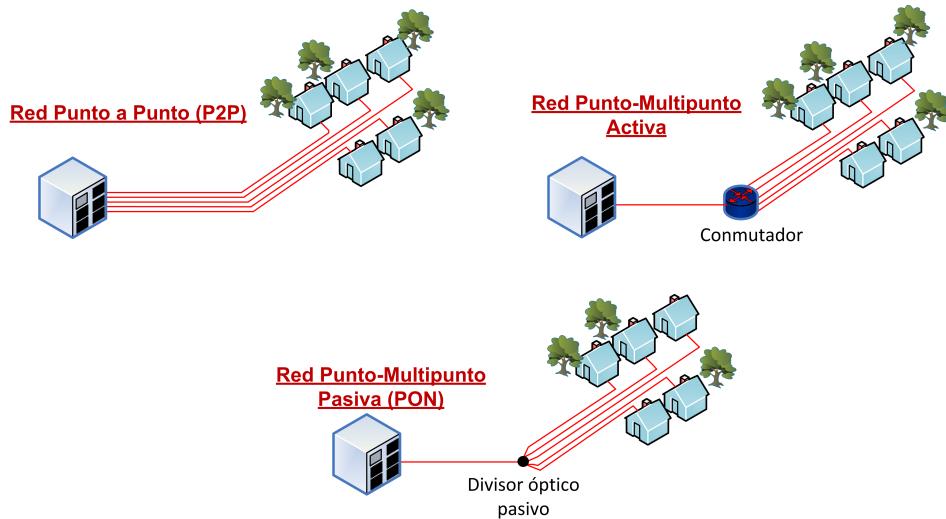


Figura 2.3: Topologías propuestas en el despliegue de tecnologías FTTx.

punto-multipunto activas, solo que utilizan un divisor óptico pasivo que, a diferencia del conmutador, no necesita alimentación, resultando por tanto, más económico y fácil de mantener [10, 12]. Esta última topología es conocida como red óptica pasiva o PON (*Passive Optical Network*).

De estas tres topologías propuestas, las dos más desplegadas son la punto a punto (P2P) y la red óptica pasiva (PON). La primera de ellas tiene la ventaja de que cada usuario puede disfrutar de toda la capacidad de la fibra. Sin embargo, los costes son elevados ya que no se comparte infraestructura con el resto de usuarios, tal y como ocurre en las redes PON. Además en P2P puede existir un gran número de fibras, lo que también hace que el mantenimiento de la red sea más complejo. Por tanto, la mayor parte de los modelos FTTx se basan en redes PON debido a su buena rentabilidad y su bajo consumo de energía por bit [13], convirtiéndose actualmente en la arquitectura dominante. De hecho, a nivel mundial el despliegue de la tecnología punto a punto en 2008 ya comprendía el 14 % de los despliegues en fibra mientras que el restante 86 % se correspondía con los de las redes PON [14]. En la actualidad esta tendencia persiste, ya que todos los despliegues de redes de acceso ópticas que se están produciendo en la actualidad obedecen a esta arquitectura PON.

2.2.1. Arquitectura básica de una red PON

En la configuración más básica, una red PON se compone de un Terminal de Línea Óptico (OLT, *Optical Line Terminal*) y varias estaciones o Unidades de Red Ópticas (ONU, *Optical Network Units*) conectadas mediante fibra óptica, y a través de un di-

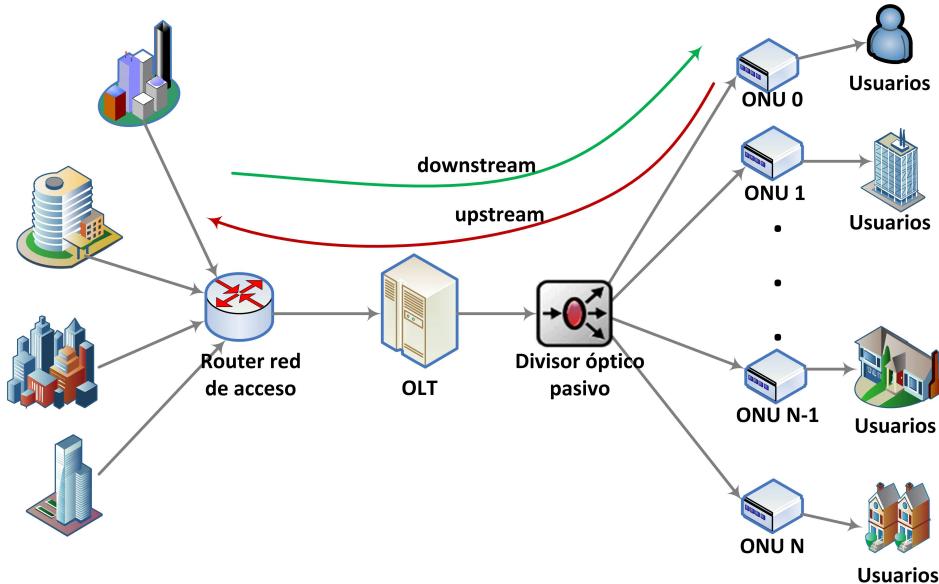


Figura 2.4: Arquitectura básica de una red PON con topología en árbol.

visor óptico pasivo o *splitter* siguiendo un despliegue en árbol como el mostrado en la Figura 2.4. El OLT se encuentra en la oficina central, conectando la red de acceso óptica a la red troncal, mientras que las ONUs están dentro o cerca de las dependencias del abonado final [10, 15].

Aunque existen otras topologías para el despliegue de redes PON (bus, anillo o estrella), el despliegue en árbol es el más frecuente ya que tiene como principal ventaja el hecho de que gran parte de la infraestructura es compartida por todos los usuarios, lo que permite una importante reducción en los costes de implantación y mantenimiento.

En el canal descendente o *downstream* (desde el OLT hasta la ONU), una PON es una red punto multipunto, donde el OLT envía a todas las ONUs la misma información a través del divisor óptico en una longitud de onda determinada. Junto con la información se envía además un identificador. De esta manera, cada ONU solo procesa el tráfico que le corresponde o para el que tiene acceso por parte del operador. Así, cuando el paquete de información llega a la ONU, si el identificador que trae dicho paquete corresponde con el de la estación, se acepta la información. Por el contrario, si esos identificadores no corresponden a la ONU, ésta elimina la información sin procesarla.

En el canal ascendente o *upstream* (desde las ONUs al OLT), la red PON es una red punto-a-punto donde múltiples ONUs transmiten hacia el OLT. En este caso, el divisor pasivo funciona como combinador y envía todo lo que le llega de las ONUs hacia el OLT por un único canal. El OLT acepta todo el tráfico procedente de las

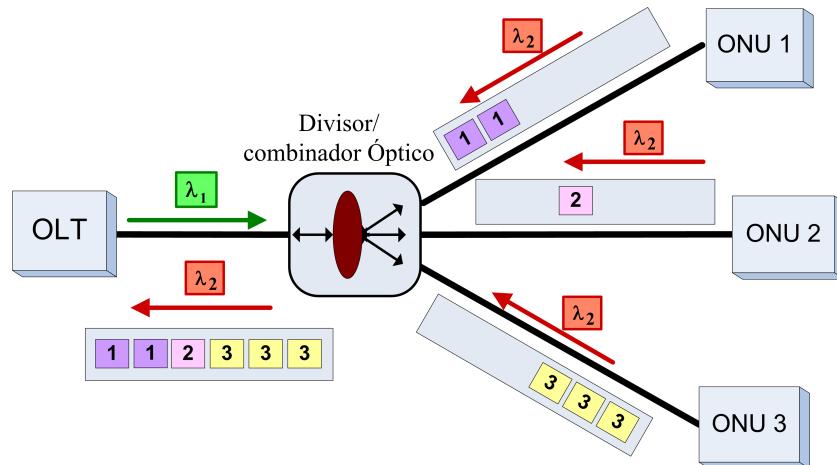


Figura 2.5: Ejemplo de arquitectura de una red PON con protocolo TDMA.

ONUs y es el encargado de procesarlo y enviarlo fuera de la red de acceso hacia la red troncal. Al ser el canal de subida un canal compartido por todas las ONUs, es necesario implementar un protocolo de control de acceso al medio (MAC, *Medium Access Control*) para evitar colisiones si varias ONUs transmiten simultáneamente, y para que además realice un reparto justo de la capacidad de la fibra entre todos los usuarios conectados. Los mecanismos que se utilizan con mayor frecuencia en las redes PON están basados en el protocolo de Acceso Múltiple por División en el Tiempo (TDMA, *Time Division Multiple Access*). Este protocolo utiliza una única longitud de onda para todas las transmisiones en el canal de acceso compartido y asigna un cierto intervalo de tiempo a cada usuario a lo largo de un ciclo. En la Figura 2.5 se observa de forma esquemática el principio de funcionamiento del protocolo TDMA en una red PON, en concreto con tres ONUs conectadas a la red. Otro de los posibles mecanismos más populares que se pueden aplicar en redes PON es el de Acceso Múltiple por División en Longitud de Onda (WDMA, *Wavelength Division Multiple Access*). Este protocolo asigna una longitud de onda a cada usuario. De este modo se consigue una conectividad punto-a-punto sobre una arquitectura de red física punto-multipunto. Sin embargo, es necesario que las ONUs estén equipadas con láseres a una longitud de onda determinada, lo cual aumenta su coste y complejidad [16]. No obstante, esta solución se presenta como la más predominante en un futuro inmediato.

Respecto al protocolo TDMA, que es el más utilizado actualmente, existe un esquema de acceso múltiple por división en el tiempo puro, donde la capacidad se presenta en forma de ciclos de tiempo de longitud fija e igual para todas las ONUs. Por tanto, cada ciclo se divide en tantos intervalos de tiempo como número de ONUs compartan

el canal de transmisión de la red, siendo cada intervalo de tiempo dedicado a una única ONU. Así, cada ciclo está organizado de tal manera que periódicamente un intervalo de tiempo transporta los paquetes de una misma ONU. Sin embargo, aunque el esquema TDMA puro es fácil de implementar, resulta ineficiente bajo condiciones de tráfico real (no homogéneo), ya que siempre asigna el mismo ancho de banda independientemente de la demanda actual de cada ONU. Por este motivo, la aplicación de TDMA puro en redes PON no es idónea puesto que el tráfico presenta una naturaleza rafagosa, obteniéndose un mal comportamiento en algunos parámetros de red como el retardo medio o la probabilidad de pérdida de paquetes [17]. Para subsanar esta deficiencia, se han propuesto en la literatura algoritmos basados en TDMA pero que distribuyen de forma dinámica el ancho de banda disponible en cada ciclo. Estos algoritmos, denominados algoritmos de asignación dinámica de ancho de banda (DBA, *Dynamic Bandwidth Allocation*), asignan ancho de banda a cada ONU en función de su demanda actual, con lo que los recursos disponibles en la red se utilizan de forma óptima.

Por otro lado, existe un interés cada vez mayor en el desarrollo de redes ópticas pasivas de largo alcance (LR-PON, *Long-Reach PON*). Estas arquitecturas ofrecen un coste menor ya que son capaces de simplificar la red, combinando la parte de acceso y la red metropolitana en una sola utilizando un tramo de 100 km de fibra en vez de los 20 km de una red PON típica [18, 19]. De esta forma, se pueden eliminar los costes asociados a las interfaces electrónicas entre ambas partes de la red y dar cobertura a un mayor número de usuarios, resultando por tanto más económicas. Sin embargo, es necesario desarrollar nuevos algoritmos DBA que sean capaces de superar las inefficiencias de los algoritmos DBA desarrollados para redes PON típicas de 20 km, al incrementarse en gran medida la distancia extremo a extremo.

Independientemente de la arquitectura de red implementada, uno de los mayores retos en las redes de comunicaciones ópticas desplegadas en la actualidad, es abordar todas las cuestiones importantes relativas a la Calidad de Servicio (QoS, *Quality of Service*) y proporcionar la Calidad de Experiencia (QoE, *Quality of Experience*) esperada por los usuarios finales. En este sentido, las redes actuales tienen que ser capaces de transportar tráfico multimedia de servicios con distintos requisitos de calidad de servicio y calidad de experiencia, cada vez más restrictivos. Todo ello, maximizando los recursos disponibles y minimizando los costes. Así, cada aplicación presenta una serie de restricciones en términos de ancho de banda, retardos o tasa de pérdida de paquetes, que se deben cumplir a lo largo de todo el camino (calidad extremo a extremo). Por ello es importante que se ofrezca calidad de servicio en todos los tramos de la red, tanto en la red troncal como en la red de acceso. En este sentido, los algoritmos DBA para redes PON y LR-PON

tienen que implementar mecanismos para poder ofrecer diferenciación de servicios y de usuarios, y cumplir así con los requisitos de QoS estipulados por los proveedores de servicio y requisitos de QoE adecuados para los usuarios finales.

Por tanto, el eje central de esta Tesis Doctoral es el diseño e implementación de algoritmos DBA eficientes en redes LR-PON que sean capaces de ofrecer y controlar dinámicamente la calidad de servicio de distintas aplicaciones y tráfico. Además, para aumentar la aplicabilidad de los algoritmos, éstos se diseñarán para que puedan utilizarse también en redes PON de diferente cobertura.

2.2.2. Estándares para el despliegue de redes PON

La falta de una estandarización inicial global en las redes de acceso PON provocó la aparición de distintas alternativas para su diseño. En este sentido, existen dos grandes grupos dedicados a la estandarización de las redes PON. Por un lado, la Unión Internacional de Telecomunicaciones (ITU-T, *International Telecommunication Union*), que con el grupo FSAN (*Full Service Access Network*) llevan trabajando en este campo desde 1995 y han propuesto diversos estándares, entre ellos APON (*Asynchronous PON*) que pasaría más tarde a denominarse BPON (*Broadband PON*) [20], GPON (*Gigabit PON*) [21] y en 2009 XG-PON (10 *Gigabit Capable PON*) [22], conocido también como el estándar para la primera fase de la próxima generación de redes PON (NG-PON1, *Next Generation PON1*) [12]. Además, desde 2011, el grupo FSAN está trabajando en el estándar para la segunda fase de la próxima generación de redes PON (NG-PON2) [23]. Por otra parte, el Instituto de Ingenieros Eléctricos y Electrónicos (IEEE, *Institute of Electrical and Electronic Engineers*) ha desarrollado dos estándares basados en la tecnología Ethernet, EPON (Ethernet PON) [24] también llamado 1G-EPON y 10G-EPON (10 *Gigabit Ethernet PON*) [25].

El estándar BPON utiliza un protocolo en modo de transferencia asíncrono (ATM, *Asynchronous Transfer Mode*) y permite una tasa de transmisión simétrica de subida y bajada a 622 Mbit/s. Sin embargo, el uso de ATM crea ineficiencias debido a la segmentación y reensamblado que hay que hacer en los puntos finales de la red, lo que se traduce en un aumento significativo del retardo [16]. Por este motivo, GPON fue desarrollado para mejorar la eficiencia en la fragmentación y transportar diferentes tipos de trama y células ATM con tasas de bit superiores a 1 Gbit/s. En concreto, GPON soporta tasas de bit asimétricas de 2.5 Gbit/s de bajada y 1.2 Gbit/s de subida.

Por otro lado, EPON, gracias al uso del estándar Ethernet, se perfila como la tecnología dominante en las redes PON. Esto es así debido a que el estándar Ethernet es relativamente simple, económico, fácil de implementar y sus componentes han reducido

su coste en las últimas décadas [10]. Además, teniendo en cuenta que el 90 % del tráfico se origina y termina en tramas Ethernet, usar una red EPON puede reducir la adaptación necesaria para mover datos entre las redes de área local (LAN, *Local Area Network*) y la red de acceso [16].

De entre estas dos tecnologías, EPON es actualmente la más popular para implementar tecnologías FTTH/B a nivel mundial. De hecho, según los datos publicados en 2013 [9], a finales de 2011 el 80 % de los usuarios de los 10 operadores más fuertes en FTTH/B a nivel mundial utilizaban tecnología EPON. Los mayores despliegues de redes EPON se encuentran en Japón, China y Corea del Sur. A mediados del 2013, existían en Japón más de 24 millones de usuarios de FTTH utilizando EPON. En Corea del Sur, donde los despliegues de redes EPON comenzaron en 2006, contaban en Abril de 2012 con más de cuatro millones de usuarios [26]. China, que tal y como se mencionó anteriormente ha eclipsado el mercado de Japón y de Corea del Sur, está desplegando ambas tecnologías. Sin embargo, el despliegue de redes EPON en China continúa siendo mayor que GPON, ya que el 60 % del equipamiento que se compra es para redes EPON [27]. En concreto, en 2012 existían más de 20 millones de usuarios de banda ancha utilizando EPON, convirtiendo a *China Telecom* en el operador de red más grande y con un mayor crecimiento [26].

Sin embargo, tanto GPON como EPON presentan una serie de restricciones que hacen difícil un incremento en el número de usuarios, alcance o tasas de datos. En ese respecto, las redes PON de 10 Gbit/s han sido recientemente estandarizadas para soportar servicios con una mayor demanda de ancho de banda. Estos estándares, definidos tanto por el IEEE como por la ITU-T, permiten compatibilidad hacia atrás y coexistencia con las generaciones de PON existentes, permitiendo actualizaciones progresivas con una inversión mínima y un impacto operacional también mínimo en los usuarios existentes [13]. Por un lado, 10G-EPO es el sucesor de EPON y extiende sus tasas de datos hasta 10 Gbit/s. Este estándar, aprobado en Septiembre de 2009, soporta tasas simétricas de 10 Gbit/s de subida y bajada, y asimétricas de 10 Gbit/s de bajada y 1 Gbit/s de subida. También se especifica una nueva alimentación para soportar 32 usuarios con distancias de al menos 20 km y la implementación de un código de corrección de errores hacia delante (FEC, *Forward Error Correction*), que antes era optativo, pero que con la inclusión de la nueva alimentación pasa a ser obligatorio [28]. En concreto, se utiliza el protocolo *Reed Solomon*. Además, la longitud de onda asignada para la dirección de bajada se sitúa entre 1575 y 1580 nm, las cuales están fuera de la banda de distribución de video RF (Radio Frecuencia). Para la dirección de subida se utilizan longitudes entre 1260 y 1280 nm. Las principales diferencias entre 1G-EPO y 10G-EPO se muestran

	EPON (1G-EPON)	10G-EPON
Tasa de transmisión	1G/1G simétrico	10G/10G simétrico 10G/1G asimétrico
Relación de división (<i>split ratio</i>)	1:16	1:16/1:32
Longitudes de onda	1480-1500 nm (bajada) 1260-1360 nm (subida)	1575-1580 nm (bajada) 1260-1280 nm (subida)
FEC	<i>Reed Solomon</i> (255,239) (optativo)	<i>Reed Solomon</i> (255,223) (obligatorio)

Tabla 2.1: Principales diferencias entre 1G-EPON y 10G-EPON.

en la Tabla 2.1. Algunas pruebas experimentales de este estándar con tasa simétrica se han realizado ya en Japón [29] y en China ya se ha empezado su comercialización [30].

Por otro lado, el grupo FSAN ha estudiado soluciones para facilitar una mayor provisión, conseguir un mayor número de usuarios y una mayor longitud extremo a extremo. En concreto, se ha propuesto un estándar para la NG-PON1. Este estándar es el XG-PON, que se centra en desarrollar PONs que sean compatibles con redes GPON operacionales. En este sentido, la variante asimétrica de XG-PON (conocida como XG-PON1) soporta una capacidad de ancho de banda asimétrico de 10 Gbit/s y 2.5 Gbit/s en la dirección de bajada y subida, respectivamente. También introduce mejoras en cuanto a la seguridad mediante sistemas de autenticación y técnicas de minimización del consumo de energía. En la variante simétrica de XG-PON, denominada XG-PON2, se propone una especificación con una capacidad de ancho de banda simétrica de 10 Gbit/s [13]. También se especifica en ambos estándares la corrección de errores hacia delante. En cuanto a las longitudes de onda utilizadas, XG-PON utiliza las mismas que el estándar 10G-EPON. Las principales diferencias entre GPON y XG-PON se recogen en la Tabla 2.2 [31].

La estandarización para NG-PON2, que proporcionará tasas superiores a 10 Gbit/s,

	GPON	XG-PON
Tasa de transmisión	2.5G/1.2G simétrico	10G/2.5G XG-PON1 10G/10G XG-PON2
Relación de división (<i>split ratio</i>)	1:32, 1:64	1:64/1:256 (capa lógica)
Longitudes de onda	1480-1500 nm (bajada) 1260-1360 nm (subida)	1575-1580 nm (bajada) 1260-1280 nm (subida)
FEC		<i>Reed Solomon</i> (248,216) (bajada) <i>Reed Solomon</i> (248,232) (subida)

Tabla 2.2: Principales diferencias entre GPON y XG-PON.

todavía se está llevando a cabo a día de hoy. Los principales requisitos para estas nuevas arquitecturas son al menos 40 Gbit/s, 40 km de longitud, una relación de división 1:64, un alcance diferencial de 40 km y al menos una tasa de acceso de 1 Gbit/s por cada ONU [23]. Se han propuesto diversas tecnologías para proporcionar un ancho de banda mayor a 10 Gbit/s, tales como redes PON con multiplexación por división en tiempo a 40 Gbit/s (40G TDM PON, 40G *Time Division Multiplexing* PON, también denominada XLG-PON), redes PON con multiplexación por división en frecuencia ortogonal (OFDM-PON, *Orthogonal Frequency Division Multiplexing* PON), PON con división por multiplexación en longitud de onda (WDM-PON, *Wavelength Division Multiplexing* PON) y redes PON con multiplexación en tiempo y longitud de onda (TWDM-PON, *Time Wavelength Division Multiplexing* PON) [32]. No obstante, de entre todas estas tecnologías, TWDM-PON es la que más respaldo tiene actualmente. En concreto, en Abril de 2012, FSAN la adoptó como la solución primaria para la NG-PON2 [23]. La arquitectura básica de TWDM-PON está formada por cuatro redes XG-PON multiplexadas usando cuatro pares de longitudes de onda (ocho longitudes de onda en total), y con láseres y receptores sintonizables en las ONUs.

No obstante, debido a las ventajas que ofrecen EPON y su sucesor 10G-EAPON sobre el resto de estándares y por ser una de las tecnologías con un mayor despliegue a nivel mundial en la actualidad, en esta Tesis se ha elegido EPON como la tecnología PON objeto de estudio.

2.2.3. Calidad de servicio en redes PON

Las redes PON se diseñan para transportar todo tipo de servicios y aplicaciones tales como comunicaciones de voz, televisión estándar (SDTV, *Standard Definition Television*) y de alta definición (HDTV, *High Definition Television*), vídeo conferencia (vídeo interactivo), transacciones en tiempo real o semi-real, tráfico de datos, juegos *online*, etc. Para soportar esta multitud de aplicaciones, cada una con diferentes requisitos, las redes PON deben exhibir una actuación adecuada dependiendo del tipo de aplicación soportada. Por ello, las redes de acceso tienen que ofrecer una calidad de servicio adecuada para usar eficientemente los recursos disponibles y ofrecer un tratamiento preferencial a ciertos tipos de servicios.

Para proporcionar QoS existen principalmente dos tecnologías [33]. La primera se basa en realizar diferenciación a nivel de usuario. En este sentido, los usuarios finales contratan acuerdos de servicio (SLA, *Service Level Agreement*) con el proveedor de servicios. En estos SLAs se definen fundamentalmente los servicios y los requisitos de calidad de los mismos, los cuales están generalmente relacionados con un nivel de ancho

de banda mínimo que es necesario garantizar. Sin embargo, estos requisitos se pueden extender de forma inherente a otros parámetros de la red. En este sentido, un segundo tipo de calidad de servicio se caracteriza por la priorización de un determinado tipo de tráfico. Para ello, el tráfico se segmenta y se diferencia para que la red lo trate de forma distinta. Esto se conoce como diferenciación de servicios y para ello se definen las clases de servicio (CoS, *Class of Service*). Estas CoS están definidas en una serie de estándares [34, 35] donde aparecen los requisitos necesarios de las distintas aplicaciones. Estos requisitos se refieren principalmente al retardo extremo a extremo de los paquetes, tasa de pérdida de paquetes y en algunos casos a la variación del retardo o *jitter*.

Por tanto, otra de las cuestiones importantes a tener en cuenta en el diseño de los algoritmos de asignación de ancho de banda en redes PON, es la implementación de mecanismos eficientes que permitan realizar diferenciación de usuarios y/o diferenciación de servicios. De este modo, se podrá asegurar una buena adecuación de la red a los distintos tipos de tráfico y a la variedad en los perfiles de usuario que existen en la actualidad.

2.3. Redes de acceso LR-PON

Las redes de acceso pasivas de largo alcance (LR-PON, *Long-Reach Passive Optical Network*) se han propuesto como una solución más económica a las redes de acceso ópticas de banda ancha. Estas redes extienden el alcance de los tradicionales 20 km hasta los 100 km e incluso mayores distancias, mediante el uso de amplificadores ópticos y tecnologías de multiplexación por longitud de onda (WDM). Gracias a este aumento en la cobertura, las redes LR-PON son capaces de combinar la red de acceso y metropolitana sustituyendo el equipamiento de la jerarquía digital síncrona (SDH, *Synchronous Digital Hierarchy*) por una fibra compartida [18]. Como resultado, la jerarquía de la red de telecomunicaciones se simplifica (Figura 2.6) reduciendo de manera significativa tanto la inversión inicial necesaria para desplegar la red, es decir el CapEx (*Capital Expenditure*), como los costes derivados del mantenimiento y funcionamiento de la misma, es decir, el OpEx (*Operational Expenditure*) [19].

Además, en los estudios realizados sobre arquitecturas de redes reales (Irlanda y Reino Unido) en [36–38] se comprobó que, no solo el uso de redes LR-PON incrementaría la cobertura, de forma que tanto la población de zonas urbanas como de zonas rurales podría recibir servicios de banda ancha con alta velocidad, sino que también se reduciría el número de nodos activos en una proporción de 50. En concreto para la red de Irlanda implementada con tecnología DSL, el número de nodos activos pasaría de 1100

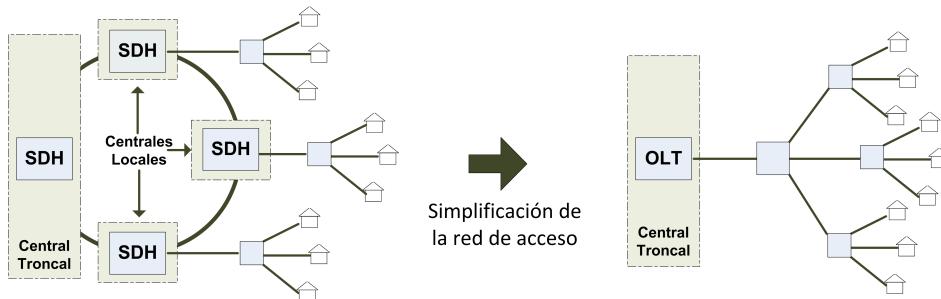


Figura 2.6: Simplificación de la red de acceso mediante redes LR-PON.

a únicamente 20 considerando una arquitectura de red LR-PON. Para el caso de Reino Unido, con tecnología DSL se necesitarían 5600 nodos activos frente a los 75 que serían necesarios con redes LR-PON. Incluso, si para la red en Reino Unido se utilizan despliegues GPON, serían necesarios 2000 nodos activos, mientras que con las arquitecturas LR-PON se precisarían únicamente 75, lo que conlleva una reducción del 96.25 %.

2.3.1. Arquitecturas para el despliegue de redes LR-PON

Para el desarrollo de este tipo de redes se han propuesto diversas arquitecturas. En función de la topología que presentan, se pueden dividir en redes con un despliegue en árbol y redes con un despliegue en anillo. A continuación se realiza un estudio del estado del arte, donde se recogen las principales características de cada arquitectura.

2.3.1.1. LR-PON con topología en árbol

2.3.1.1.1. Arquitectura de red SuperPON

Este proyecto [39] surgió con la idea de aumentar la distancia, el número de usuarios conectados y las tasas de transmisión. En concreto, el prototipo de red SuperPON implementado soporta un total de 2048 ONUs (mediante una división en dos etapas) y alcanza una distancia de 100 km (donde los últimos 10 km pertenecen a la sección de distribución). Además, se consiguen tasas de 2.5 Gbit/s en el sentido de bajada (entre el OLT y las ONUs) y de 311 Mbit/s para la transmisión de subida (entre las ONUs y el OLT). Para compensar la atenuación debida al aumento de la distancia y al número de usuarios, SuperPON introduce amplificadores ópticos tanto para la transmisión en sentido de bajada como para la transmisión en el sentido de subida (Figura 2.7). Como la comunicación en el sentido descendente utiliza un modo de transmisión de onda continua, se usaron amplificadores de fibra dopados con Erbio (*Erbium Doped Fiber Amplifier*), que se caracterizan por tener una figura de ruido baja y por proporcionar una alta eficiencia energética.

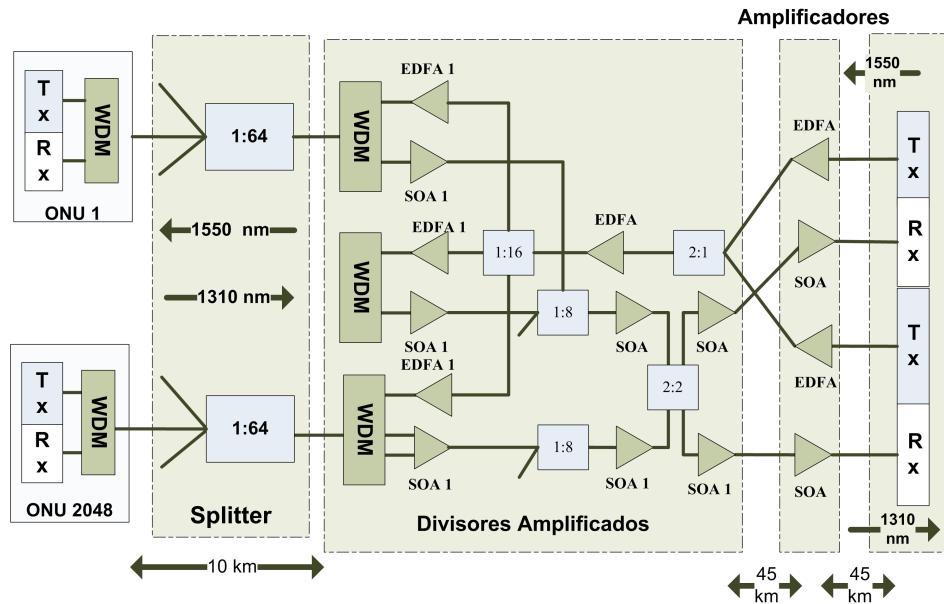


Figura 2.7: Arquitectura de una red SuperPON de 100 km de distancia [39].

nar ganancias altas. Sin embargo, mantienen una ganancia constante durante un tiempo largo, por lo que no es posible su uso en el sentido de subida debido a la naturaleza rafagosa del tráfico originado en las ONUs. Además, en esta dirección el divisor óptico funciona como un combinador, por lo que el ruido de emisión espontánea amplificada (ASE, *Amplified Spontaneous Emission*) introducido por cada amplificador, se suma a las contribuciones ASE del resto de amplificadores debido a su ubicación en paralelo. Para evitar este efecto no deseado, los amplificadores solo deben amplificar cuando se envíen datos a través de ellos. Por ello, para la dirección de subida se optó por la utilización de amplificadores ópticos de semiconductor (SOA, *Semiconductor Optical Amplifier*), que tienen la ventaja de que el tiempo de encendido/apagado es del orden de nanosegundos. Sin embargo, para hacer más escalable la red y reducir los costes de los componentes ópticos, la arquitectura de red SuperPON evolucionó hacia las arquitecturas PON de largo alcance que se explican a continuación.

2.3.1.1.2. Arquitectura LR-PON TDMA de 10 Gbit/s

La operadora inglesa *British Telecom* presenta en [40] una arquitectura que es capaz de transmitir a distancias de 100 km y a tasas de 10 Gbit/s. El tamaño del divisor óptico es de 1024, la mitad que en SuperPON, pero solo requiere seis amplificadores ópticos para la transmisión en el sentido de subida y de bajada, a diferencia de los 39 necesarios en la arquitectura SuperPON [18]. Se coloca una etapa intermedia de amplificadores

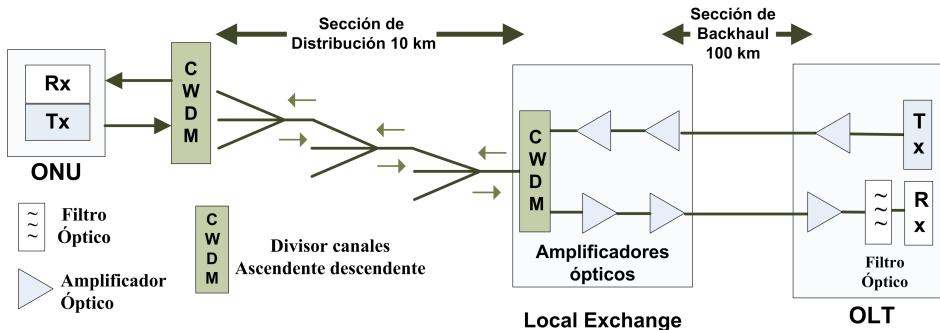


Figura 2.8: Arquitectura de una red LR-PON TDMA de 10 Gbit/s [40].

justo antes de dividir la señal (antes del divisor óptico) de forma que no existen amplificadores ópticos en paralelo (Figura 2.8). Por tanto, la combinación del ruido ASE no es un problema y no se necesitan sistemas ópticos de puertas como en SuperPON. Además, para la transmisión en sentido ascendente, se utiliza una amplificación dual. La primera etapa, está formada por un preamplificador de bajo ruido, y la segunda etapa de amplificación proporciona a la señal la potencia suficiente para superar la atenuación hasta llegar al OLT.

Otras tecnologías usadas en esta nueva arquitectura de red son los mecanismos para la corrección de errores, tales como el corrector de errores hacia delante FEC. Gracias a su uso, los errores de transmisión pueden ser detectados y corregidos y, utilizando un filtro óptico, se aumenta la relación señal a ruido (SNR, *Signal to Noise Ratio*) de la señal recibida reduciendo así el ruido ASE. Además, como las transmisiones sobre 100 km a 10 Gbit/s provocan penalizaciones debido a la dispersión cromática de la fibra monomodo, son necesarias técnicas para contrarrestar la dispersión. Por lo tanto, usando las tecnologías ópticas apropiadas, es posible conseguir una transmisión de 10 Gbit/s en canales tanto de subida como de bajada soportando hasta 1024 usuarios en distancias de 100 km. Aunque estas arquitecturas incrementan la distancia y el número de usuarios, no resultan del todo escalables y flexibles, por lo que la tendencia ha sido evolucionar hacia arquitecturas híbridas que combinan WDM con tecnologías de multiplexación por división en tiempo (TDM).

2.3.1.1.3. Arquitectura LR-PON híbrida DWDM-TDM

Se trata de una arquitectura que combina el uso de WDM denso (DWDM, *Dense Wavelength Division Multiplexing*) con TDMA [41]. De esta forma coexisten un número de TDM PONs, cada una trabajando a diferente longitud de onda, compartiendo la misma planta de amplificadores y fibra (Figura 2.9). Esta arquitectura híbrida DWDM-TDM

implica una mejora significativa en las redes de larga distancia desplegadas hasta ese momento. Esto es así porque el uso de WDM permite una mayor cantidad de ancho de banda a compartir, aprovechada solo si se incluye un mayor número de usuarios en la red, lo que puede hacerse utilizando redes de largo alcance.

Uno de los principales retos en términos de componentes optoelectrónicos de una arquitectura híbrida de este tipo es el transmisor del usuario, que necesita tener un alineamiento muy preciso con la longitud de onda asignada en DWDM. La opción convencional de usar láseres con una longitud de onda específica es demasiado cara para las redes de acceso. Esto es así debido al alto coste de estos componentes y a los costes de mantenimiento asociados al despliegue de un gran número de láseres con diferentes longitudes de onda a lo largo de la base del usuario. Una solución más económica sería emplear los mismos componentes en cada transmisor del usuario, es decir, que sean independientes de la longitud de onda asignada por la red, dando lugar a ONUs denominadas *colorless*. Por tanto, esta arquitectura utiliza un modulador con una generación centralizada de portadora óptica. El modulador está basado en un semiconductor de electro-absorción (EAM, *Electro Absorption Modulator*) integrado con dos amplificadores ópticos semiconductores (SOA). El primer SOA se utiliza para amplificar la portadora óptica a la entrada de la ONU, y el segundo se usa para amplificar la señal después de ser modulada por el EAM. Sin embargo, la portadora óptica que se envía a cada usuario es generada en la oficina central, de manera que el operador de red tiene un control completo sobre la longitud de onda asignada a cada transmisor. Por tanto, esta arquitectura de red va a estar formada por varias redes PON de largo alcance independientes, cada una trabajando a distinta longitud de onda pero sobre la misma infraestructura de fibra, gracias al uso de multiplexadores/demultiplexadores DWDM.

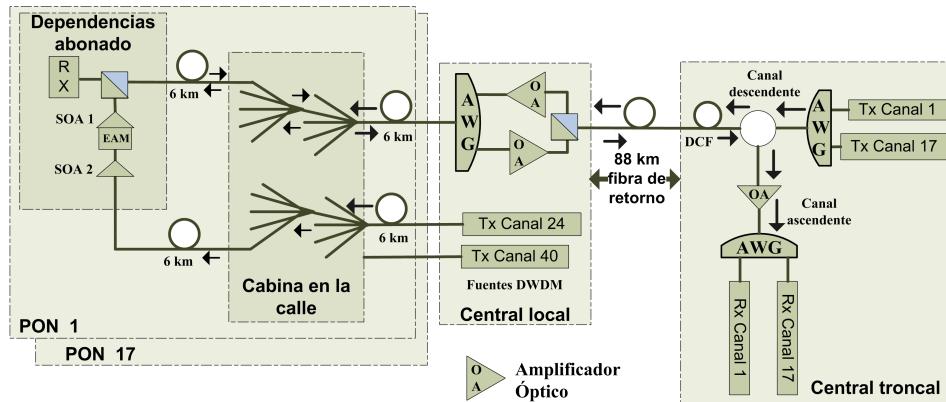


Figura 2.9: Arquitectura general de una red híbrida DWDM-TDM [41].

Existen otros trabajos como el presentado por Davey et al. [42] que analizan el comportamiento del protocolo GPON en redes DWDM LR-PON sobre distancias de 135 km. Los resultados experimentales demuestran que el estándar GPON puede ser desplegado fácilmente para operar sobre redes de largo alcance cuando se utiliza regeneración optoelectrónica. Por otro lado, también se han planteado arquitecturas DWDM-TDM con conversores de longitud de onda (WC, *Wavelength Converter*) que utilizan esta conversión en el canal ascendente para pasar los datos transmitidos de la longitud de onda de la ONU a la longitud de onda DWDM [43, 44]. De este modo, un conjunto de PONs podrían ser agrupadas bajo la misma fibra de retorno, siendo cada PON convertida a una longitud de onda distinta. Esto reduce el coste de la fibra de retorno ya que su utilización y eficiencia se incrementa y además los costes son compartidos por un mayor número de usuarios.

2.3.1.1.4. Arquitectura híbrida WDM/TDMA PIEMAN

El proyecto PIEMAN (*Photonic Integrated Extended Metro and Access Network*) propone una red LR-PON con un alcance de 100 km y con 32 longitudes de onda DWDM [45]. Cada longitud de onda, a una tasa de 10 Gbit/s, se asigna únicamente a una PON con un divisor de 512, permitiendo a la red soportar un máximo de 16384 usuarios (32×512) con una media de ancho de banda de 20 Mbit/s para cada usuario. En la transmisión en sentido ascendente se utilizarán ONUs independientes de la longitud de onda (*colorless*) para así evitar la necesidad de que cada usuario compre una ONU con una longitud de onda específica (Figura 2.10). Para el diseño de los transmisores *colorless* de las ONUs, este proyecto ha investigado dos alternativas diferentes. La primera de ellas es una ONU con un transmisor sintonizable de bajo coste que pueda seleccionar 32 longitudes de onda con un espaciado de 50 GHz. Una segunda aproximación bajo consideración es el uso de SOAs reflectivos (RSOA, *Reflective Semiconductor Optical Amplifier*). Esta aproximación es similar a la utilizada en la configuración híbrida DWDM-TDM del Apartado 2.3.1.1.3, donde la red proporciona la longitud de onda de la portadora óptica sobre la que la ONU modula su información en el canal ascendente.

Por otro lado, y tal y como se explicó con anterioridad, la naturaleza de ráfagas del tráfico TDMA en la dirección de subida de las redes LR-PON hace que los EDFA no puedan ser usados debido a la lenta ganancia dinámica. Esto es, la ganancia del EDFA no puede cambiar lo suficientemente rápido para asegurar que el paquete entero reciba una ganancia constante. La ganancia cambia a medida que el paquete se propaga a través del EDFA causando que la amplitud del paquete sea no uniforme. La solución seleccionada por esta arquitectura consiste en utilizar una longitud de onda auxiliar que se ajusta de forma relativa al paquete en sentido ascendente de forma que la potencia óptica a través

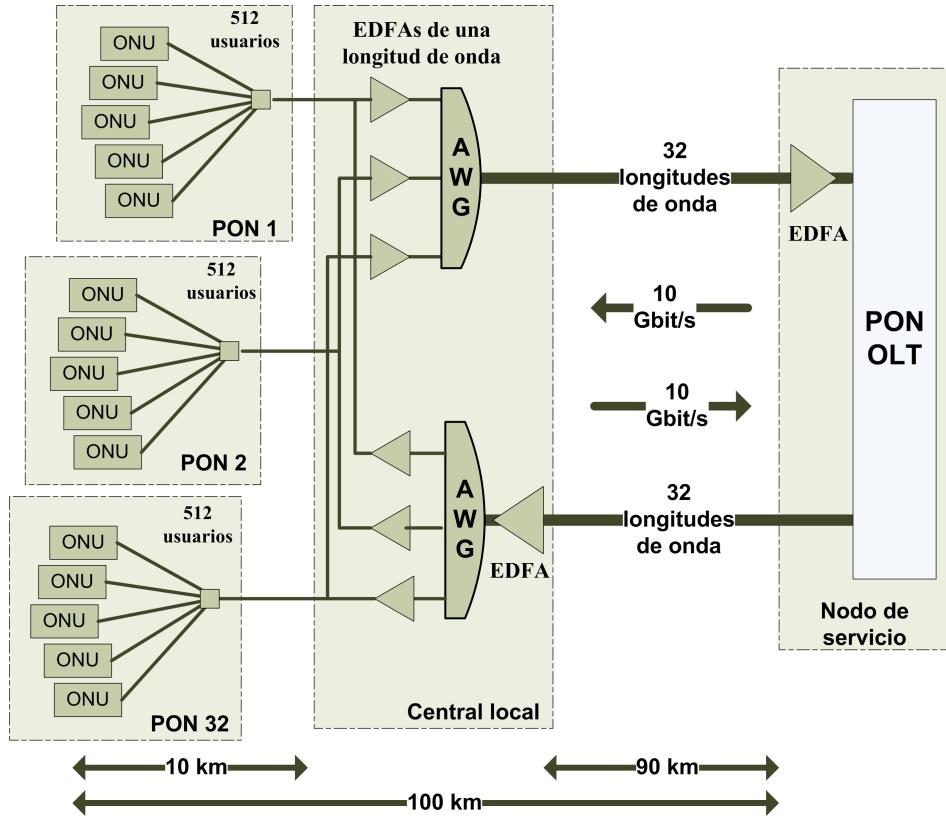


Figura 2.10: Arquitectura general de la red híbrida WDM/TDMA del proyecto PIEMAN [45].

del EDFA permanezca constante.

2.3.1.1.5. Arquitectura PON con Ultra-DWDM coherente

Esta arquitectura tiene un alcance de 100 km, donde los últimos 20 km se corresponden con la sección de distribución [46]. Esta solución utiliza la multiplexación por división de longitud de onda ultra-densa (UDWDM, *Ultra Dense Wavelength Division Multiplexing*) como sistema de asignación de ancho de banda, permitiendo la multiplexación en la misma fibra de un alto número de longitudes de onda, de forma que un gran número de usuarios pueden ser servidos, asignando una longitud de onda a cada usuario a una tasa de 1 Gbit/s. Para asignar la longitud de onda a cada usuario se utiliza un filtro paso banda en cada receptor coherente, lo que permite una gran selectividad en el plano de longitudes de onda soportadas. Esto, junto con el uso de UDWDM permite incrementar el número de longitudes de onda incluso a más de mil.

2.3.1.1.6. Arquitectura PON proyecto DISCUS

En el proyecto Europeo DISCUS (*DIStributed Core for unlimited bandwidth supply for all Users and Services*) [38] se presenta la arquitectura LR-PON de la Figura 2.11. Esta arquitectura presenta una etapa con divisores alimentados y después etapas con divisores pasivos, para compartir el OLT y la fibra de distribución entre un gran número de usuarios (hasta 1024). La principal razón de esta elección se debe a que los divisores pasivos evitan el uso de dispositivos de longitud de onda selectiva, lo que reduce la flexibilidad para proveer servicios que puedan surgir en el futuro.

A diferencia de otras arquitecturas donde las fibras de conexión más largas se encuentran en la sección de retorno y tienen una sección de distribución de menor tamaño (unos 10 km), en DISCUS se están investigando también diseños que permitan una mayor longitud en la sección de distribución (incluso de 80 km) para así dar cobertura en zonas rurales. Por otro lado, en la sección de retorno se utilizan dos fibras una de ellas para protección o para usos futuros.

Inicialmente, la tasa de transmisión soportada es de 10 Gbit/s aunque se están considerando actualizaciones para soportar hasta 40 Gbit/s en la dirección de bajada.

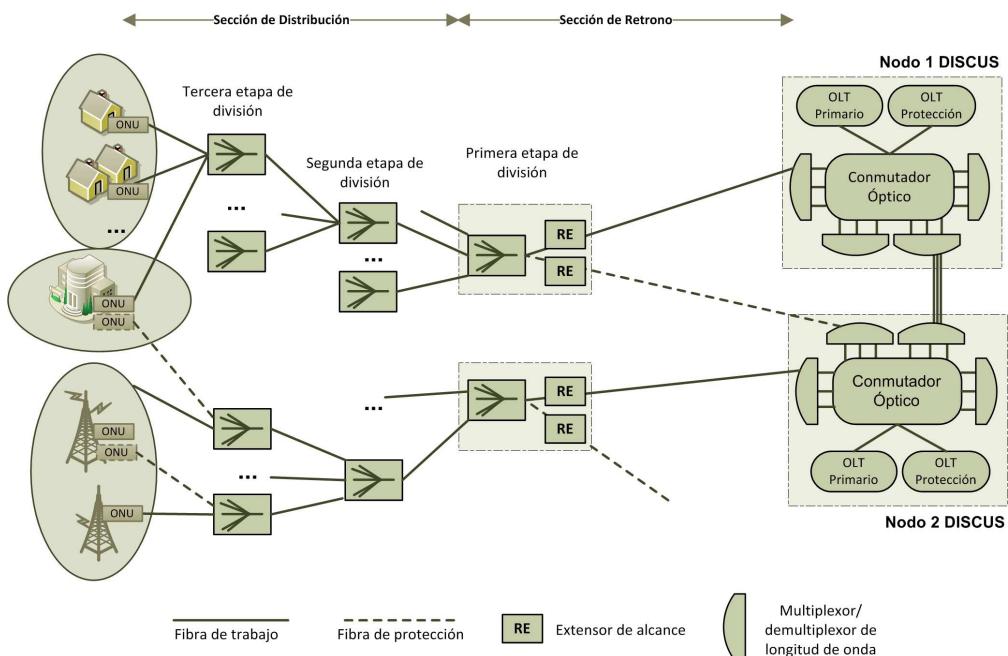


Figura 2.11: Arquitectura general de la red LR-PON del proyecto DISCUS [38].

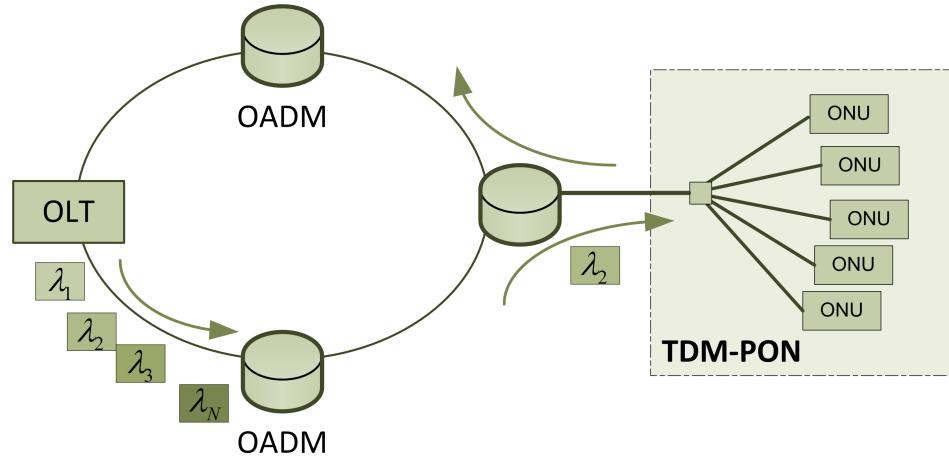


Figura 2.12: Arquitectura básica de una red LR-PON en anillo.

2.3.1.2. LR-PON con topología en anillo

Una variante de las redes LR-PON con topología en árbol son las que implementan una topología en anillo, lo que resulta perfecto para la transmisión bidireccional y para las recuperaciones ante posibles fallos de los enlaces [19, 47]. En la configuración más básica, el elemento principal es un multiplexador óptico de inserción/extracción (OADM, *Optical Add-Drop Multiplexer*). Éste estará situado en cada nodo, como se muestra en la Figura 2.12 y extraerá o insertará la longitud de onda dirigida hacia los usuarios finales. Estos usuarios finales están conectados a cada OADM mediante una configuración PON en árbol soportando una longitud de onda, y por lo tanto conformando una TDM PON independiente. Aunque cada OADM extrae una longitud de onda y la asocia a cada red TDM PON independiente, todas las longitudes de onda son combinadas y viajan por la misma fibra a lo largo del anillo.

A continuación se presentan diferentes ejemplos de redes LR-PON desplegadas con esta topología en anillo, así como una descripción de sus características fundamentales.

2.3.1.2.1. Arquitectura WDM/TDM PON Híbrida SUCCESS

La arquitectura híbrida WDM/TDM PON SUCCESS también denominada SUCCESS-HPON [48, 49], está compuesta por un anillo colector con una serie de acopladore en estrella WDM conectados al anillo por medio de un Nodo Remoto (RN, *Remote Node*), que es el centro de los despliegues en árbol de las subredes PON. Esta arquitectura se ha presentado como una opción para la migración de las redes TDM-PON actuales a los futuros despliegues de redes WDM-PONs. Por ello, la arquitectura considera que las subredes PON adjuntas al anillo pueden ser tanto TDM-PONs como WDM-PONs

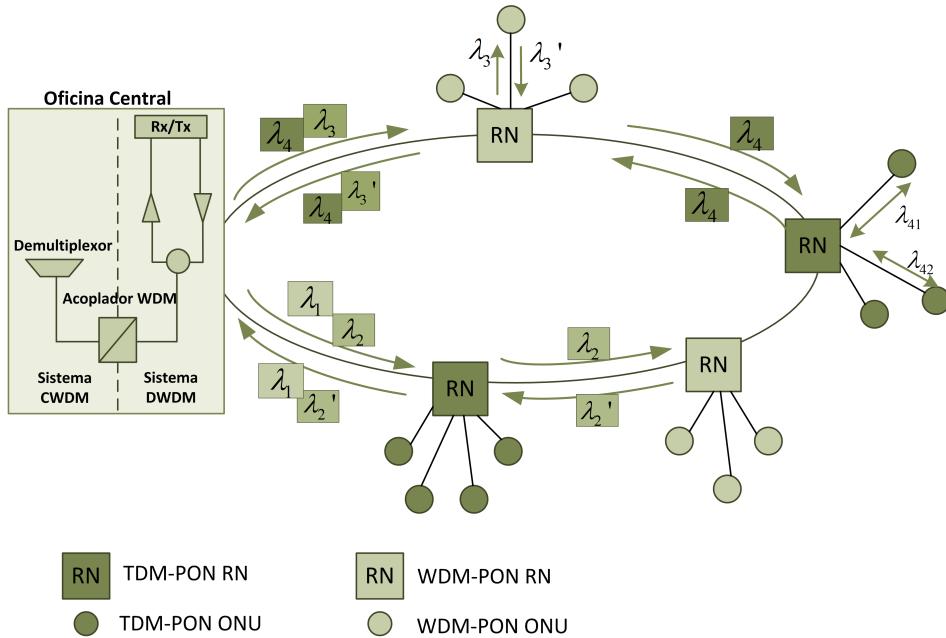


Figura 2.13: Arquitectura general del prototipo de red SUCCESS [48, 49].

(Figura 2.13). Por otro lado, en el anillo se utiliza un sistema WDM. Normalmente, cada nodo remoto une un número entre 16 y 64 ONUs y se consiguen tasas de transmisión de 1.25 Gbit/s.

2.3.1.2.2. Arquitectura XL-PON

La red PON extra-larga (XL-PON, *Extra-Large PON*) presentada en [50] tiene una extensión de 100 km, donde 70 km se corresponden con el alcance del anillo y los 30 km restantes a la longitud de las redes PON en árbol conectadas al anillo. Las redes en árbol tienen un tamaño de división de 512, aunque también se han realizado pruebas con 1024. Para conectar las redes en árbol al anillo se utilizan puntos de acceso a la red metropolitana (MAP, *Metro Access Points*). Para conseguir el factor de división de 512 se utilizan *splitters* 1:8 en el MAP y 1:64 en el nodo remoto. Los MAPs están formados por un OADM seguido de un preamplificador. Después se utiliza un amplificador EDFA bidireccional, para amplificar la señal de subida y de bajada. También existe regeneración 3R, por lo que existe un traspondedor seguido del divisor 1:8. El último componente se utiliza para dividir/combinar la señal dirigida a las ONUs o procedente de mismas (Figura 2.14). Esta arquitectura presenta tasas de transmisión de 10 Gbit/s en el sentido de bajada y de 2.5 Gbit/s en el de subida, utilizando el algoritmo de detección de errores FEC, en concreto *Reed Solomon*.

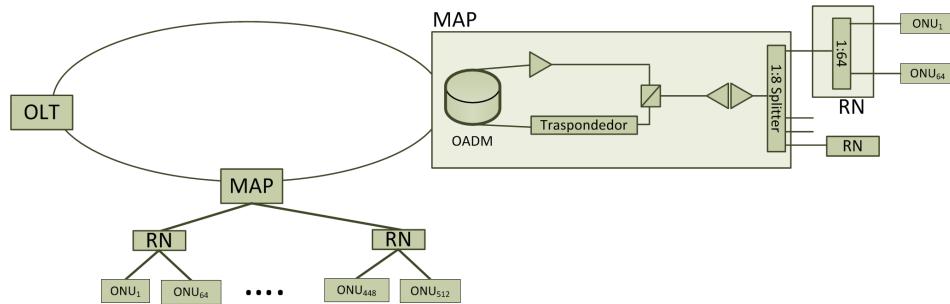


Figura 2.14: Arquitectura general del prototipo de red XL-PON [50].

2.3.1.2.3. Arquitectura WDM-Ethernet PON

La arquitectura WDM-Ethernet PON (WE-PON) [51] utiliza un anillo central mediante tecnología WDM donde se conectan 16 redes TDMA-PON. En el anillo, se transmiten 16 longitudes de onda diferentes, una para cada red TDMA-PON, en las cuales los factores de división van desde 8 hasta 32. Esta es la configuración típica para un despliegue FTTC. Para una configuración FTTH, el anillo se compone de un conjunto de fibras, que alcanzan diferentes AWGs, donde las longitudes de onda transportadas se llevan hasta las ONUs. El esquema de este tipo de arquitectura para la configuración FTTH se muestra en la Figura 2.15.

Para reducir el coste de las ONUs, se utiliza un mecanismo de reutilización de longitud de onda, de forma que la señal de subida se consigue remodulando la señal de bajada.

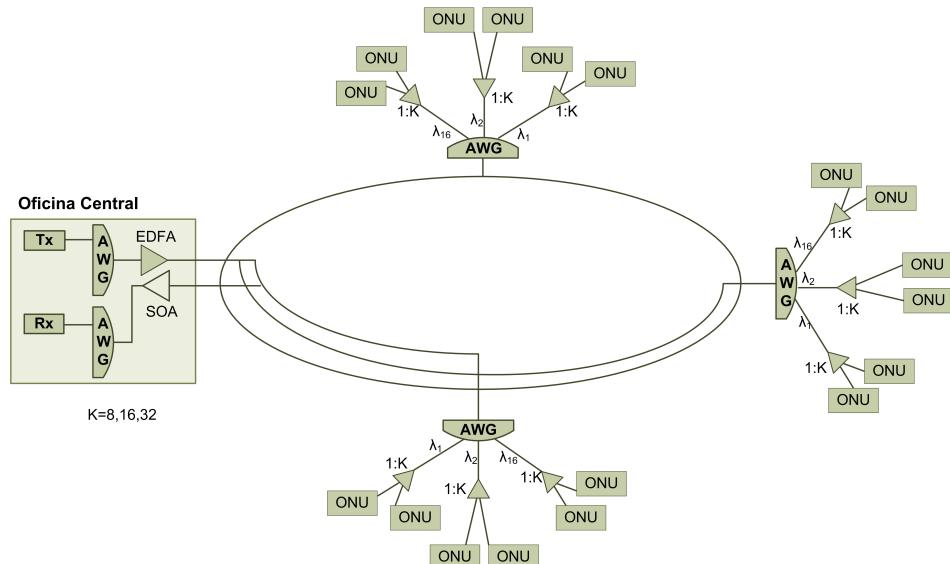


Figura 2.15: Arquitectura del prototipo de red WE-PON para FTTH [51].

En este sentido, no es necesario usar un láser en una longitud de onda concreta, sino solo un componente reflectivo. El componente utilizado por esta arquitectura es un RSOA, que modula la señal de información en función de la longitud de onda de la portadora recibida. Los multiplexores y demultiplexores WDM se encuentran tanto en la oficina central como en los nodos remotos.

2.3.1.2.4. Arquitectura de red SARDANA

La arquitectura de red SARDANA (*Scalable Advanced Ring Dense Access Network Architecture*) [52], implementa un doble anillo WDM con una extensión de 100 km, a una tasa de transmisión de 10 Gbit/s en el canal de bajada y de 1.25/2.5/5 Gbit/s para el canal de subida usando RSOAs (Figura 2.16). El anillo WDM transporta un mínimo de 32 longitudes de onda, una para cada TDM-PON. Con esta configuración, la red de acceso puede soportar más de 1000 usuarios. En cada red TDM-PON, también es posible transmitir varias longitudes de onda, habilitando a un número de operadores a compartir la planta de fibra y permitiendo a los usuarios elegir el operador mediante unos filtros intercambiables en la ONU. También propone un diseño de protección, donde la estructura del anillo permite dos caminos desde la oficina central hacia el nodo remoto. De esta forma, si uno de los caminos se ve envuelto en un fallo de la fibra, el segundo camino se puede usar para enrutar el tráfico a través de la red.

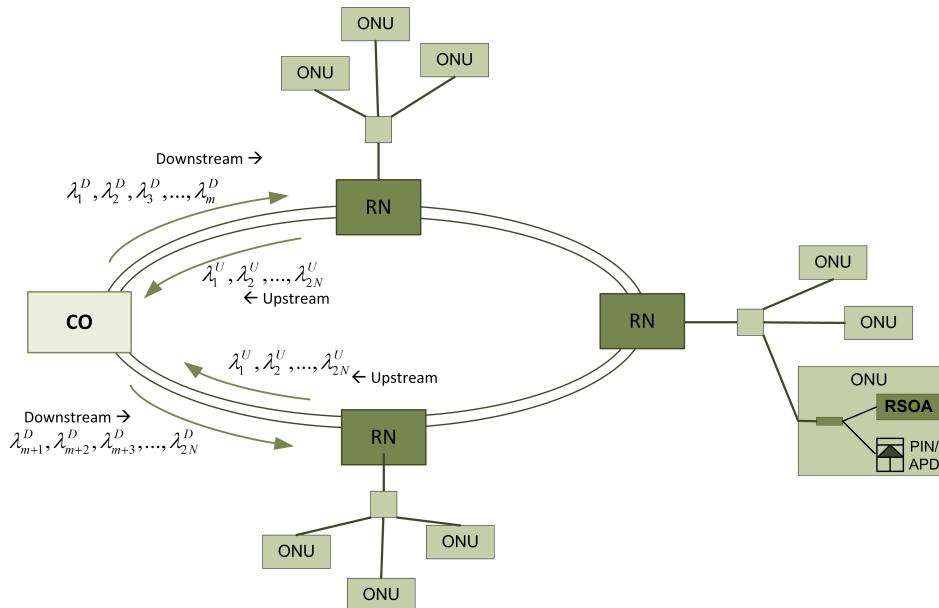


Figura 2.16: Arquitectura de red SARDANA [52].

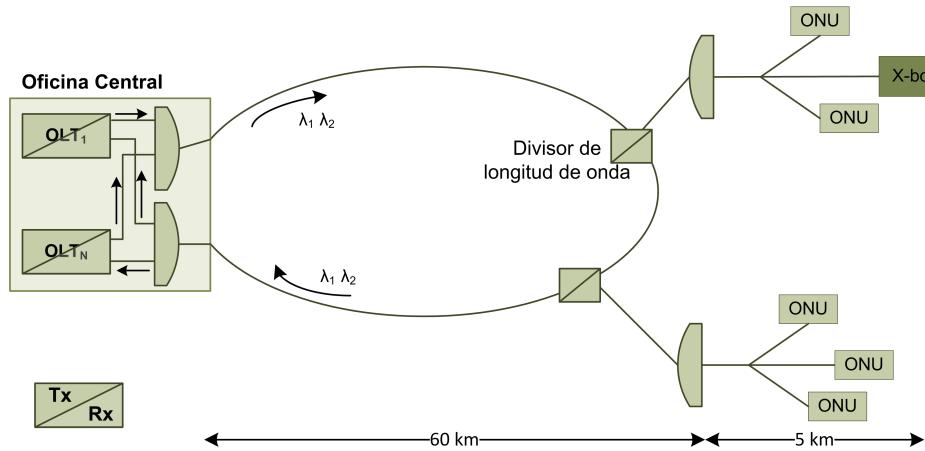


Figura 2.17: Arquitectura del prototipo de red Wx-PON [53].

2.3.1.2.5. Arquitectura Wx-PON

La arquitectura Wx-PON es una red óptica pasiva híbrida donde un conjunto de redes WDM-PON y redes TDMA-PON existentes se integran en una única red PON. Esta arquitectura basada en anillo [53], proporciona una cobertura de unos 60 km en el anillo y 5 km en las redes TDM-PON convencionales (Figura 2.17). En el anillo, el ancho de banda está compartido a través de WDM con 32 longitudes de onda, cada una soportando una tasa de transmisión 1.25 Gbit/s.

Aparte de las ONUs TDMA-PON convencionales, existe una ONU especial, denominada *X-box*, que además de realizar las funciones de una ONU TDMA-PON convencional, lleva a cabo la función de conexión entre el anillo WDM y las secciones TDMA [53]. Para ello, la *X-box* convierte la señal DWDM en banda C a la señal de bajada de TDMA, y a su vez la señal TDMA de subida se convertirá a la señal DWDM en banda C. La *X-box* funciona a una tasa 1.25 Gbit/s y utiliza RSOAs para modular los datos del usuario.

Hasta este punto de la Tesis se han presentado las arquitecturas más típicas para el despliegue de redes LR-PON a nivel físico, así como sus principales características. En el siguiente apartado se explicará la problemática de las redes de acceso pasivas de largo alcance a nivel de acceso, eje central del desarrollo de esta Tesis.

2.3.2. Control de acceso al medio en redes LR-PON

Tal y como se ha visto anteriormente, en las distintas arquitecturas para el despliegue de redes de acceso de largo alcance, independientemente de si utilizan una arquitectura en árbol o en anillo, se multiplexan/demultiplexan un conjunto de redes LR-PON, cada

una utilizando una longitud de onda determinada. En este sentido, las redes híbridas WDM-TDM PON, pueden considerarse, a nivel de protocolo de acceso, como un conjunto de redes TDMA-PON de largo alcance independientes.

Por tanto, las redes LR-PON, al igual que las redes PON tradicionales, son redes punto-multipunto con comunicación bidireccional entre el OLT y las ONUs. Al igual que ocurre en las redes PON comunes, todas las ONUs comparten el mismo canal para la transmisión de subida, de forma que es necesario un mecanismo de control de acceso al medio. En este sentido, los algoritmos de asignación dinámica de ancho de banda (DBA) basados en el protocolo TDMA resultan, al igual que para las redes PON de corto alcance, muy convenientes y útiles para la gestión de los recursos de la red en este tipo de redes de largo alcance.

Estos algoritmos de asignación dinámica de ancho de banda, tal y como se comentó anteriormente, son capaces de adaptar la capacidad de la red a las condiciones de tráfico actuales modificando la distribución del ancho de banda asignado a cada ONU en función de la demanda actual o en función de los requisitos de QoS. Tanto para las redes PON comunes como para las PON de largo alcance, la asignación dinámica de ancho de banda requiere la comunicación entre el OLT y las ONUs. De este modo, el OLT distribuye el ancho de banda del canal ascendente entre todas las ONUs, controlando así la transmisión de datos desde las ONUs al OLT ciclo tras ciclo. El tiempo de ciclo es el tiempo total en el que todas las ONUs transmiten una vez siguiendo un orden consecutivo. En el caso de las redes basadas en el estándar Ethernet, se utiliza el protocolo de control multipunto (MPCP, *Multi-Point Control Protocol*) para establecer dicha comunicación entre el OLT y las ONUs. Para ello, el protocolo MPCP utiliza dos tipos de mensajes, el mensaje *Gate* y el mensaje *Report*. Cada ONU envía un mensaje *Report* al final de su tiempo de transmisión en cada ciclo para informar al OLT del estado de sus colas y donde se incluye la demanda de ancho de banda para el siguiente ciclo de transmisión. Por otro lado, el OLT asigna ancho de banda a cada ONU en función del estado de sus colas y de los requisitos de calidad de servicio estipulados y le envía un mensaje *Gate* para informar del ancho de banda asignado y del instante en el que debe comenzar a transmitir, esto es, para asignarle un tiempo de transmisión. En el estándar GPON la comunicación se realiza de forma similar. En concreto, el OLT utiliza el campo PCBd (*Physical Control Block downstream*) de la trama descendente para informar a cada ONU del comienzo y final de su transmisión a través del mapa de ancho de banda ascendente (*Upstream bandwidth map*). Por otro lado, la ONU informa al OLT sobre su demanda para el ciclo siguiente utilizando el campo DBRu (*Dynamic Bandwidth Report upstream*) de la cabecera de la trama GPON en sentido ascendente [54].

Por otro lado, los algoritmos DBA desarrollados para redes PON de 20 km pueden no resultar eficientes al ser aplicados en redes de largo alcance, debido al incremento en el tiempo de ida y vuelta (RTT, *Round Trip Time*) asociado al notable aumento de la distancia (100 km). Por ello, uno de los principales retos a nivel de acceso en las redes LR-PON es el diseño de algoritmos de asignación de ancho de banda que resulten eficientes en distancias de 100 km o superiores y que además ofrezcan una calidad de servicio adecuada a los requisitos establecidos por cada proveedor de servicios. Aunque las redes LR-PON son una arquitectura muy prometedora que está emergiendo con mucha fuerza, los estudios centrados en el desarrollo de algoritmos de asignación de ancho de banda para estas redes son menores que para las redes PON tradicionales [55]. Sin embargo, un diseño adecuado de los mismos es crucial para el despliegue real y eficiente de este tipo de tecnología de acceso. En este sentido, en el siguiente apartado se presenta un breve resumen del estado del arte actual de algoritmos DBA en redes LR-PON.

2.3.3. Algoritmos de asignación dinámica de ancho de banda en redes LR-PON

En los algoritmos de asignación dinámica de ancho de banda (DBA), el ancho de banda asignado a cada ONU en cada ciclo no es fijo, sino que depende de su demanda actual y de los requisitos de calidad de servicio estipulados por los proveedores. De esta forma, se aprovechan de manera óptima los recursos de la red, ya que su reparto se realiza en función de los requisitos de cada ONU. Gracias a su buen funcionamiento, los algoritmos DBA son los más desarrollados para el control de acceso al medio, tanto en redes PON de 20 km [16], como en redes de larga distancia LR-PON [55]. En este tipo de algoritmos, el OLT es el elemento central que controla dicha asignación ciclo tras ciclo. Según la política de asignación utilizada a la hora de distribuir el ancho de banda, los algoritmos DBA pueden dividirse en dos tipos, algoritmos centralizados u *offline* y algoritmos de testeo continuo (*polling* u *online*). En los siguientes subapartados se explican en detalle las características más importantes de cada política.

2.3.3.1. Algoritmos DBA centralizados en redes LR-PON

En los algoritmos centralizados, también llamados *offline*, el OLT asigna ancho de banda a cada ONU para el ciclo siguiente una vez ha recibido la demanda actualizada de todas las ONUs en el ciclo actual. Por lo tanto, las ONUs tienen que esperar un tiempo igual al RTT para poder transmitir en el siguiente ciclo, tal y como puede verse en la Figura 2.18. Como en las redes LR-PON este tiempo se incrementa de manera

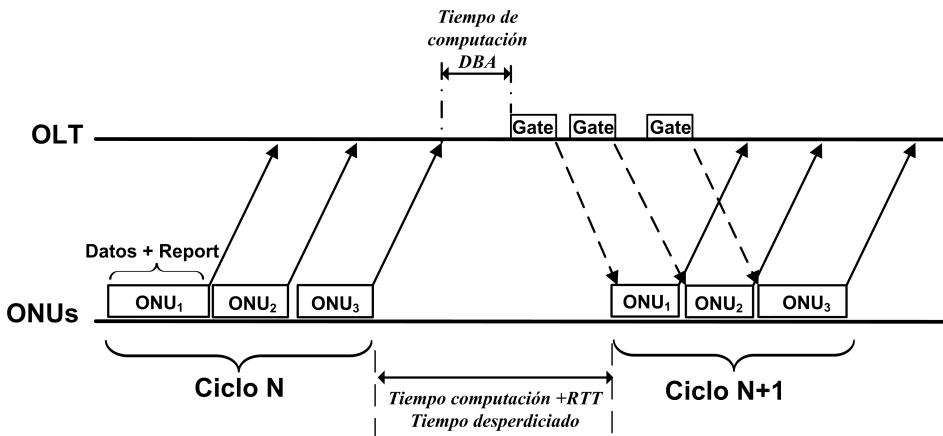


Figura 2.18: Esquema de funcionamiento de un algoritmo centralizado en una red LR-PON.

considerable al aumentar la distancia a 100 km, los algoritmos centralizados tienen que implementar algún procedimiento que permita a las ONUs transmitir durante ese tiempo y que no se desperdicie tanto ancho de banda.

En concreto, el algoritmo centralizado desarrollado en [56, 57] y denominado TSD (*Two-State DMB*) considera este tiempo de ida y vuelta como un ciclo virtual de tamaño fijo donde las ONUs pueden transmitir (Figura 2.19). Puesto que en el instante en el que habría que planificar estas nuevas transmisiones la demanda de todas las ONUs no es conocida, TSD realiza una predicción de tráfico basada en la demanda del ciclo anterior, lo que conlleva un incremento de la complejidad del algoritmo.

En el algoritmo centralizado *Multi-Thread polling* (MT) presentado en [19, 47] se utilizan varios procesos de asignación, también llamados hilos, corriendo en paralelo. De

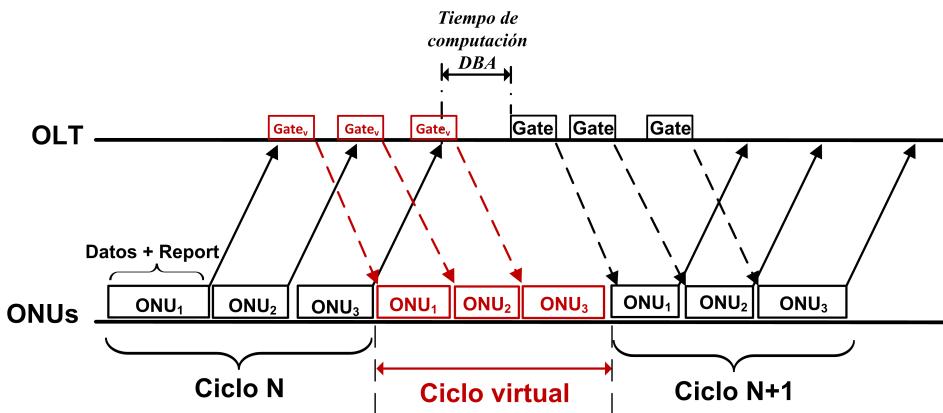


Figura 2.19: Esquema de funcionamiento del algoritmo centralizado TSD [56, 57].

esta forma, el tiempo de espera se elimina ya que cuando las ONUs están esperando a recibir el mensaje de control del OLT para transmitir en el ciclo siguiente, pueden transmitir paquetes que han sido planificados en otro proceso de asignación paralelo. Un ejemplo de la ejecución de este algoritmo puede verse en la Figura 2.20, donde por simplificación solo se han considerado dos ONUs en la red. Sin embargo, este algoritmo presenta una serie de inconvenientes. En primer lugar, este algoritmo no soporta calidad de servicio. El otro inconveniente viene dado por el sistema de planificación entre hilos. En este sentido, en los algoritmos con múltiples hilos, es muy importante que exista una coordinación entre los mismos, de forma que se pueda intercambiar información del ancho de banda demandado y asignado entre los hilos existentes y así evitar que un hilo pueda asignar ancho de banda para paquetes que ya han sido asignados por otro hilo [58, 59]. No obstante, la coordinación entre hilos en el algoritmo MT se basa en la presunción de que la actualización de la información de las colas de las ONUs, que llegan a través de mensajes *Report* procedentes de otros hilos, están disponibles al realizar la asignación de ancho de banda en el hilo actual [58], lo cual no siempre ocurre. Además, tal y como se demuestra en [59], el esquema de planificación que utiliza MT no impide que se duplique la asignación de paquetes.

Por tanto, a partir de la idea de este algoritmo con múltiples hilos, han surgido algunas variaciones que tratan de mejorar algunos de sus aspectos negativos. Por un lado, se han desarrollado algoritmos centrados en la optimización del proceso de asignación de ancho de banda del algoritmo MT, mientras que, por otro lado, han surgido otras propuestas que se centran en mejorar la planificación entre hilos. Con respecto a los primeros, en [60] se propone una mejora en la distribución del ancho de banda en exceso entre aquellas ONUs que demandan más en el ciclo, de tal manera que dicho ancho de exceso se reparte de forma equitativa en función de la demanda, pero sin que la asignación final sobrepase su ancho de banda solicitado. De esta forma, la asignación es más

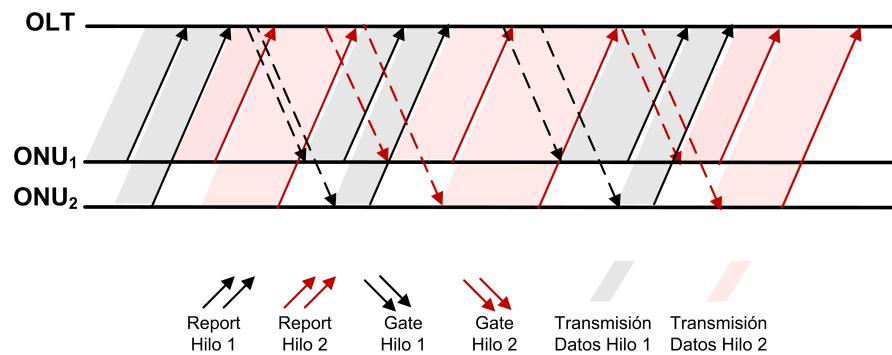


Figura 2.20: Esquema de funcionamiento del algoritmo *Multi-Thread polling* (MT) [19, 47].

eficiente y justa. El algoritmo PGO (*Periodic GATE Optimization*) presentado en [61], mejora también las prestaciones del algoritmo MT. En concreto, este algoritmo añade formulación ILP (*Integer Linear Programming*) en el OLT para optimizar la distribución del ancho de banda en exceso en cada ciclo a las ONUs más sobrecargadas, esto es, a las ONUs que demandan más que su ancho de banda mínimo estipulado. En este sentido, se estima el ancho de banda extra que se le va a asignar a dichas ONUs en el siguiente periodo a partir de la formulación ILP y utilizando la información de los mensajes *Report* enviados en cada hilo. Mediante esta optimización se consigue decrementar el retardo medio, sobre todo para altas cargas de red. No obstante, la incorporación de la formulación ILP aumenta la complejidad del algoritmo. Por otro lado, en [62] cambian el número de hilos dependiendo de la carga de la red, consiguiendo mejoras el retardo para escenarios tanto con cargas altas, como con cargas bajas.

Entre los algoritmos que mejoran la planificación entre los distintos hilos existentes, en [58] se propone un nuevo esquema, denominado *Newly Arrived Frames Plus* (NA+), que actúa como coordinador entre los procesos DBA corriendo en paralelo. Este esquema elimina la sobreasignación de ancho de banda a las ONUs, evitando que el OLT asigne el ancho de banda destinado al mismo paquete en varios hilos. Del mismo modo, en [59] se presenta el algoritmo *Enhanced Inter-Thread Scheduling* (EIS), que combina las ideas presentadas tanto en [58] para evitar la sobreasignación, como en [60] para la asignación de ancho de banda. Sin embargo en estos algoritmos la complejidad aumenta, ya que es necesario un control del ancho de banda asignado en los diferentes procesos de asignación.

Por otro lado, el algoritmo presentado en [63] es un algoritmo centralizado aplicable a redes LR-EPON donde las ONUs están situadas a distintas distancias del OLT. Para las ONUs cercanas, este algoritmo tiene el comportamiento de un algoritmo centralizado común. Sin embargo, para las ONUs situadas a mayores distancias, el ancho de banda a asignar se calcula realizando una predicción de la demanda utilizando información del tráfico y de las demandas anteriores. Esta predicción, tal y como ocurre con TSD, aumenta la complejidad del algoritmo.

Por último, cabe destacar que los algoritmos anteriores no proporcionan calidad de servicio, algo que resulta crucial para poder soportar los distintos tipos de tráfico existentes ofreciendo las garantías acordadas por los proveedores de servicios.

2.3.3.2. Algoritmos de *polling* o de testeo continuo

En los algoritmos que utilizan un esquema de *polling* o de testeo continuo (Figura 2.21), también denominados algoritmos *online*, el OLT asigna ancho de banda a cada ONU para el ciclo siguiente en el instante que recibe su mensaje *Report* con la demanda

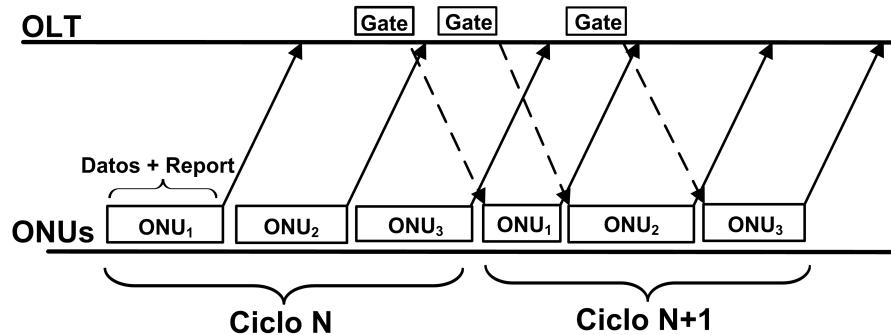


Figura 2.21: Esquema de funcionamiento de los algoritmos de testeo continuo o algoritmos de *polling* en una red LR-PON.

actualizada en el presente ciclo. De esta forma, las ONUs pueden transmitir justo cuando la transmisión de la ONU anterior termina, lo que da lugar a una utilización muy eficiente del ancho de banda. Además no es necesario realizar ningún tipo de predicción por lo que son algoritmos más sencillos. En este sentido, en [60] se demuestra que los algoritmos de testeo continuo son capaces de mantener un buen comportamiento con respecto al retardo en redes LR-PON sin necesidad de llevar a cabo ningún tipo de mejora en el algoritmo (a diferencia de los algoritmos centralizados). Así mismo, obtienen una mayor eficiencia en la utilización del canal de transmisión comparados con el algoritmo centralizado MT.

Un algoritmo desarrollado para redes LR-EPON que sigue este esquema es LIPSA (*Long-Reach Interleaved Polling algorithm with Service level Agreement*) [64]. Este algoritmo, al igual que otros algoritmos de testeo continuo, obtiene una alta eficiencia ya que, para cargas de red bajas y medias, donde no hay mucho tráfico, el aumento en el RTT no influye demasiado en las prestaciones, ya que ciclo tras ciclo se puede enviar todo el tráfico entrante sin apenas retardo en las colas de las ONUs. Además, para cargas de red altas, el tamaño de ciclo máximo es mayor que el RTT, esto es, 2 ms de ciclo máximo definido por el estándar EPON frente a 1 ms del RTT por el incremento a 100 km, por lo que no existe tiempo desperdiciado entre transmisiones consecutivas. Sin embargo, para optimizar las prestaciones de este algoritmo para cargas de red medias y bajas, en [65] se diseñó el algoritmo LOHEDA (*LOng reach Highly Efficient Dynamic bandwidth Assignment*). Este algoritmo reduce el retardo medio de los paquetes para cargas de red bajas y medias aprovechando el tiempo desperdiciado entre ciclos que existe en algunos casos para dichas cargas. Además, la optimización en estas cargas resulta de gran interés ya que son cargas a las que normalmente trabaja la red, puesto que los operadores de servicio prefieren trabajar en niveles de carga donde las pérdidas de paquetes sean nulas

o muy bajas.

Por otro lado, el algoritmo OEBD (*Online Excess Bandwidth Distribution*) [66], es un algoritmo de testeo continuo que utiliza un esquema de asignación de ancho de banda en exceso. A diferencia de los algoritmos centralizados que suelen aplicar estas técnicas, OEBD asigna anchos de banda en exceso a cada ONU en tiempo real, sin esperar a tener la demanda actualizada de todas las ONUs. Para calcular el ancho de banda en exceso de cada ONU, se tienen en cuenta unos pesos asociados a cada ONU (w_i) y una ventana de ancho de banda extra que se incrementa cuando las ONUs piden menos ancho de banda que un mínimo establecido y que se decrementa dinámicamente por un factor (γ).

Por último, se ha adaptado recientemente el uso de varios hilos en la asignación de ancho de banda a esquemas de testeo continuo. En este sentido, en [67] se presenta el algoritmo denominado *online MTP* (*online Multi-Thread Polling*). Sin embargo, aunque este algoritmo obtiene mejores retardos que el algoritmo MT centralizado, no obtiene mejoras significativas al comparar con otros algoritmos de testeo continuo que no introducen procesamiento con hilos paralelos.

2.3.4. Calidad de servicio en redes LR-PON

Al igual que en las redes PON comunes, para que las redes LR-PON puedan soportar los distintos tipos de servicios y aplicaciones existentes, con una calidad de servicio y de experiencia adecuada, es necesario implementar técnicas de diferenciación de usuarios y de servicios. Por ello, en este apartado se describen en primer lugar los métodos de diferenciación de usuarios hasta ahora aplicados en redes LR-PON, para después presentar los métodos que permiten priorizar el tráfico.

2.3.4.1. Integración de métodos de diferenciación de usuarios

Dado que en las redes de acceso actuales los proveedores de servicios ofrecen una serie de perfiles con diferentes requisitos, los algoritmos DBA tienen que ser capaces de proporcionar esta diferenciación de usuarios demandada. Para llevarla a cabo, es necesario cumplir con los requisitos de calidad de servicio dependiendo del SLA contratado por el usuario. Estos requisitos pueden referirse tanto a retardo garantizado, probabilidad de pérdida de paquetes y de forma muy común, garantizando unos niveles mínimos de ancho de banda.

A continuación, se presentan algunos métodos para realizar esta diferenciación de usuarios que han sido utilizados en redes PON y LR-PON. En [68], Banerjee *et al.* proponen un algoritmo que realiza un reparto justo del ancho de banda con SLAs duales. Aunque en principio está diseñado para asignar el ancho de banda disponible en sentido

descendente para distintos proveedores de servicios, también puede aplicarse en el canal ascendente de la red. El SLA primario se define como aquel cuyos requisitos mínimos especificados se tienen que cumplir con la máxima prioridad. Después de que el SLA primario cumpla sus requisitos, la siguiente prioridad es que se cumplan los del SLA secundario. Si no hay suficiente ancho de banda para acomodar los servicios del SLA secundario, el algoritmo adopta una política de máximos-mínimos para asignar el ancho de banda de forma justa. Esta política va a dar a la entidad que ha recibido la menor asignación de ancho de banda, una mayor cantidad del ancho de banda remanente. Sin embargo, este algoritmo solo distingue entre dos tipos de perfiles de abonado. Algo similar ocurre con el algoritmo BGP (*Bandwidth Guaranteed Polling*) [69], donde las ONUs se dividen en dos grupos, ONUs a las que se garantiza un ancho de banda mínimo y ONUs a las que no se garantiza un cierto nivel de ancho de banda. En primer lugar se garantiza el ancho de banda a las ONUs del primer grupo. Después, el ancho de banda restante se reparte entre las ONUs que no tienen garantías de ancho de banda.

Una forma común para ofrecer diferenciación de usuarios con más de dos perfiles de abonado, se basa en la utilización de unos pesos fijos que se asignan a cada ONU dependiendo del SLA contratado. Así, la distribución de ancho de banda se realiza de acuerdo a esos pesos. Un ejemplo de algoritmo que garantiza el ancho de banda mediante el uso de pesos fijos es el desarrollado en [70]. Cabe destacar que todos estos algoritmos que se han presentado anteriormente solo se han aplicado para redes PON comunes de 20 km y no en redes de largo alcance. Sin embargo, estas técnicas podrían implementarse de forma directa a redes LR-PON. De los algoritmos desarrollados hasta el momento para arquitecturas LR-PON, TSD [56, 57], LIPSA [64] y LOHEDA [65] soportan diferenciación de usuarios y los tres utilizan este esquema de pesos fijos para la asignación del ancho de banda. Los valores de estos pesos dependen del ancho de banda mínimo que es necesario garantizar a cada usuario dependiendo del SLA que hayan contratado. De esta forma el OLT utiliza estos pesos fijos para repartir el ancho de banda disponible y garantizar así los niveles de ancho de banda estipulados por el proveedor. La diferencia entre ellos es que el primero es centralizado, mientras que los otros dos son algoritmos de *polling*. Por otro lado, el uso de pesos fijos no garantiza una adaptabilidad en tiempo real a todas las condiciones de tráfico, ya que el ancho de banda máximo que se va a poder asignar a cada ONU en cada ciclo de transmisión será el mismo independientemente del escenario. Del mismo modo, si el proveedor de servicios comete algún error en el cálculo de los pesos, los anchos de banda máximos permitidos no serán los adecuados. Por ello, resulta muy interesante, en primer lugar, lograr una adaptabilidad en tiempo real a diferentes escenarios con distintos tipos de tráfico y en segundo lugar, hacer que la red

sea independiente de sus condiciones iniciales, tales como el ancho de banda asignado inicialmente.

2.3.4.2. Integración de métodos de diferenciación de tráfico

Para poder garantizar requisitos de retardo, *jitter* y pérdida de paquetes de los distintos tipos de aplicaciones existentes, es necesario separar el tráfico en un número limitado de clases y proporcionar una diferenciación de servicios (CoS). Para llevar a cabo esta diferenciación se pueden aplicar métodos de planificación dentro de las ONUs (*intra-ONU scheduling*) y planificación entre ONUs (*inter-ONU scheduling*). El primero de ellos se refiere a la gestión de las prioridades del tráfico dentro de las ONUs, mientras que el segundo regula la transmisión de paquetes fuera de las diferentes ONUs.

Para realizar la planificación dentro de las ONUs (*intra-ONU scheduling*), éstas están equipadas con tantas colas como número de clases de servicio soportadas. De esta forma, cuando la ONU recibe un paquete procedente de un usuario, se clasifica teniendo en cuenta su prioridad. Después los paquetes se extraen de las colas siguiendo un determinado orden y en una proporción fijada por la política escogida. Por otro lado, en los métodos de planificación entre ONUs (*inter-ONU scheduling*), el OLT es el encargado de realizar la planificación, que consiste en realizar la asignación de ancho de banda a cada ONU en cada ciclo de transmisión. Para llevar a cabo esta asignación, puede tener en cuenta otros aspectos, como el tamaño de las colas más prioritarias de las ONUs.

Así pues, el algoritmo [71] ofrece calidad de servicio en redes LR-PON utilizando ambos métodos de planificación (en las ONUs y entre ONUs). En este sentido, los autores han extendido el algoritmo PGO presentado en [61], incluyendo un módulo de planificación dentro de las ONUs que modifica la proporción en la que los paquetes salen de cada cola. Para ello, la ONU comprueba el tamaño de la cola de la clase más prioritaria. Si el tamaño de esta cola es mayor que el tamaño previo, la proporción para sacar paquetes de esta cola aumenta y disminuye la proporción de otra cola con menor prioridad. Además, estas proporciones se envían al OLT en el mensaje *Report*, que las utiliza en la formulación ILP para priorizar en el proceso de asignación de ancho de banda a aquellas ONUs donde las colas de mayor prioridad tienen mayor cantidad de paquetes almacenados. Con esto se reduce el retardo de las clases de servicio con mayor prioridad.

Otros algoritmos que ofrecen garantías en el retardo (aunque no son específicos para redes de largo alcance), son los presentados en [72, 73]. Sin embargo, estos algoritmos solo ofrecen garantías en el retardo para la clase de servicio más prioritaria. No obstante, con la aparición de nuevas aplicaciones, es necesario proporcionar también garantías de retardo para clases de tráfico de prioridad media, como en [74]. En este sentido, los

autores de [75, 76] proponen un algoritmo para redes LR-PON con restricciones de retardo para varias clases de tráfico. Para ello utiliza un control de admisiones y se estima en la ONU el retardo con el que el paquete va a llegar al OLT, de forma que si no cumple con las restricciones de retardo de su clase de servicio, el paquete se elimina.

Por otro lado, para ofrecer de forma más completa la calidad de servicio estipulada, es necesaria una combinación adecuada tanto de la diferenciación de clases de servicio como de usuarios. En este sentido, ninguno de los algoritmos descritos anteriormente lleva a cabo esta doble diferenciación, esto es, ninguno ofrece diferentes garantías de retardo dependiendo de la prioridad del SLA. En [77], se desarrolla un algoritmo que considera diferenciación de SLAs y que además ofrece garantías de retardo, pero solo para la clase de servicio de mayor prioridad. Sin embargo, tal y como se ha mencionado anteriormente, es necesario ofrecer garantía de retardo también para otros servicios de carácter prioritario.

Así pues, en la literatura, solo existen dos algoritmos que realizan un control del retardo para tráfico de prioridad alta y media y que además llevan a cabo diferenciación de usuarios. Esto es, dependiendo de la prioridad del SLA mantienen el nivel del retardo de las clases de servicio prioritarias por debajo de su cota máxima garantizada. En [78] se presenta un algoritmo donde para conseguir asegurar los retardos se utilizan en la asignación del ancho de banda dos algoritmos diferentes, uno para la clase de tráfico más prioritaria y otro para la de prioridad media, lo cual aumenta la complejidad. Así mismo, se permite el envío de los paquetes de la clase de servicio más restrictiva antes de que el mensaje *Gate* llegue a la ONU con su asignación, ya que considera que estos paquetes se generan a una tasa constante. Además, para llevar a cabo la diferenciación de usuarios, a las ONUs que pertenecen a los SLAs más prioritarios se les asigna ancho de banda con una mayor frecuencia. Esto es, mientras que a las prioritarias siempre se les asigna ancho de banda en cada ciclo de transmisión, no todas las ONUs con baja prioridad transmiten en todos los ciclos. De esta forma, el retardo que obtienen estas ONUs prioritarias es menor que el de las ONUs de menor prioridad. El segundo algoritmo que realiza también la doble diferenciación de usuarios y clientes se denomina DySLa (*Dynamic Service Level agreement*) [79]. Este algoritmo controla dinámicamente que todos los perfiles de abonado satisfagan los requisitos de retardo de todas las clases de servicio soportadas asegurando que los servicios sensibles mantengan el retardo medio por debajo de un límite máximo permitido para cada SLA dependiendo de su prioridad. Para llevar a cabo esta tarea, DySLa distribuye el ancho de banda disponible a cada SLA dependiendo de su prioridad a partir de un conjunto de pesos que se van a ir modificando en tiempo real para alcanzar el mejor comportamiento y adaptarse a posibles cambios

en la red. Sin embargo, esta evolución de los pesos se realiza restando una cantidad fija del peso asociado a los abonados que cumplen con los requisitos y sumando esa misma cantidad al peso de los abonados que no cumplen con las restricciones. Esta cantidad se mantiene siempre fija, lo cual puede no resultar óptimo para acotar de forma dinámica los retardos bajo diferentes circunstancias y condiciones de red.

Por tanto, tal y como se ha visto en este apartado, existen diversas técnicas para ofrecer calidad de servicio de acuerdo a diferentes perfiles de abonado y clases de servicio soportadas. Sin embargo, para realizar esta doble diferenciación, los métodos utilizados resultan no ser adaptativos, bien por utilizar esquemas de pesos fijos, o porque los pesos siempre se modifican una misma cantidad, lo que puede no ser óptimo. En este sentido, a lo largo de la Tesis se desarrollarán algoritmos que gestionen de forma adaptativa y en tiempo real la asignación de ancho de banda con el objetivo de que todos los usuarios cumplan con los requisitos de calidad estipulados por los proveedores de servicio para diferentes parámetros de red, tales como niveles mínimos de ancho de banda garantizados o cotas en el retardo según la prioridad del servicio. Además, ya que los recursos se reparten de forma óptima dinámicamente, se consigue una independencia de las condiciones iniciales de la red. Así mismo, aunque estos algoritmos se han desarrollado para redes de largo alcance, son escalables de forma que pueden ser perfectamente aplicables a redes de menor alcance.

2.4. Conclusiones

En este capítulo de introducción, se han presentado los conceptos básicos de las redes PON en general y de las redes LR-PON de forma particular. En la primera parte del capítulo, se ha visto como la fibra óptica es la solución óptima para solventar el “cuello de botella” en la red de acceso actual. Así mismo, se han descrito las características básicas de las redes FTTx. Para el despliegue de estas tecnologías, la arquitectura más económica es la punto-multipunto, ya que utiliza una única fibra desde la oficina central hasta el punto de distribución, donde se utiliza un divisor óptico para dividir la señal y llevarla a los abonados. Esta infraestructura de red es conocida como red de acceso óptica pasiva (PON). Para el despliegue de las redes PON existen en la actualidad estándares que pertenecen a dos grupos distintos. Por un lado, GPON y XG-PON propuestos por la ITU-T, y por otro lado EPON y 10G-EPON, propuestos por IEEE y basados en el estándar Ethernet. Los estándares más recientes son XG-PON y 10G-EPON que permiten desplegar redes PON de 10 Gbit/s así como aumentar el número de usuarios y el alcance de las redes.

En la segunda parte del capítulo se estudiaron las redes de acceso ópticas pasivas de largo alcance (LR-PON). Estas redes están emergiendo con fuerza ya que son capaces de reducir los costes de las redes PON tradicionales extendiendo el alcance desde los 20 km hasta los 100 km. Gracias a este aumento de la distancia, las redes LR-PON son capaces de combinar la red de acceso y metropolitana simplificando la jerarquía de red e incluyendo un mayor número de usuarios finales, de forma que se reducen los gastos a compartir entre los usuarios. En este capítulo se han presentado diversas arquitecturas para el despliegue de este tipo de redes. De entre ellas, las más estudiadas son las redes híbridas WDM-TDM PON, donde se multiplexan/demultiplexan un conjunto de redes TDM-PON de larga distancia, cada una de ellas trabajando a una longitud de onda dedicada. Por tanto, a nivel de control de acceso, cada una de las redes LR-PON multiplexadas puede tratarse de forma independiente. De esta forma, los retos a nivel de acceso que se plantean en las redes PON de largo alcance son los mismos que en una red PON tradicional, con el inconveniente añadido de que al extender la distancia el retardo extremo a extremo también aumenta, lo que puede reducir en gran medida la eficiencia de estas redes. En este sentido, en la última parte del capítulo, se presentan diversos algoritmos existentes en la literatura y orientados a resolver dicho inconveniente en redes LR-PON. Finalmente, se realiza un breve resumen donde se recogen las principales técnicas para ofrecer calidad de servicio en arquitecturas LR-PON.

Capítulo 3

Algoritmo DBA para garantizar dinámicamente niveles de ancho de banda en LR-EPONs

3.1. Introducción

Tal y como se ha comentado en el capítulo anterior, todos los algoritmos que soportan diferenciación de usuarios en redes LR-PON se basan en la utilización de pesos fijos. Sin embargo, eso no garantiza una adaptabilidad dinámica independiente de las condiciones del tráfico y de la red. Esto es así porque con el uso de pesos fijos se mantiene un ancho de banda máximo fijo de forma que, si las condiciones del tráfico cambian, estos anchos de banda no se adaptarán en tiempo real al nuevo escenario. En este sentido, los requisitos establecidos por los proveedores para cada uno de los perfiles de abonado podrían cambiar en tiempo real, y sería necesario recalcular unos nuevos pesos para poder garantizar dichos nuevos requisitos. Por otro lado, resulta también muy interesante hacer que el algoritmo sea independiente de las condiciones iniciales de la red, algo imposible de conseguir mediante el uso de pesos fijos, y también de los diferentes patrones de tráfico, tales como tráfico asimétrico y dinámico.

Por este conjunto de razones, se ha desarrollado un algoritmo que no utiliza pesos fijos para llevar a cabo la diferenciación de usuarios. En concreto, se ha supuesto un escenario donde los proveedores de servicio garantizan a cada perfil de abonado (SLA) unos requisitos de ancho de banda garantizado. Bajo estas condiciones, el algoritmo desarrollado reajusta constantemente el ancho de banda asignado a cada usuario con el objetivo de que todos los usuarios cumplan con sus requisitos de nivel mínimo de

ancho de banda y de que el comportamiento del algoritmo sea independiente del ancho de banda asignado inicialmente. Para ello lleva a cabo una modificación dinámica en tiempo real del ancho de banda máximo permitido a cada ONU. Así pues, en este capítulo, se presenta una descripción del algoritmo DySGAB (*Dynamic Service levels with GuAranteed Bandwidth*), que ha sido diseñado siguiendo esta filosofía, así como un estudio de simulación en el que se analizan sus prestaciones y comportamiento bajo diferentes escenarios y condiciones de red.

3.2. Algoritmo DySGAB (*Dynamic Service levels with GuAranteed Bandwidth*)

El primer algoritmo desarrollado en esta Tesis, denominado DySGAB (*Dynamic Service levels with GuAranteed Bandwidth*, algoritmo dinámico con niveles de servicio de ancho de banda garantizado) [80–82], tiene el objetivo de realizar un control dinámico de la asignación de ancho de banda de forma que todos los usuarios tengan asegurado los requisitos mínimos de ancho de banda según el perfil de abonado contratado.

Se trata de un algoritmo de asignación dinámica de ancho de banda que implementa una política de asignación de testeo continuo o *polling*. Tal y como se comentó en el Capítulo 2, en los algoritmos de testeo continuo, el OLT asigna el ancho de banda a cada ONU justo después de recibir su demanda actualizada, independientemente del estado del resto de las ONUs. De esta forma, el OLT no necesita esperar para recibir la información de las colas de cada ONU y por tanto, la utilización del ancho de banda es más eficiente en el canal compartido y además son algoritmos muy sencillos computacionalmente [64].

En las siguientes secciones se explicará el proceso de asignación de ancho de banda en DySGAB, así como el proceso de adaptación dinámico que lleva a cabo en tiempo real para garantizar los niveles mínimos de ancho de banda estipulado.

3.2.1. Asignación de ancho de banda en DySGAB

Para asignar ancho de banda a cada ONU en cada ciclo, DySGAB utiliza un esquema limitado, por ser el que presenta un mejor comportamiento y una mayor eficiencia en redes EPON [83] y LR-EAPON [64, 65]. En este esquema el OLT dará a cada ONU el ancho de banda demandado, siempre y cuando dicha demanda esté por debajo de un ancho de banda máximo considerado. Por otro lado, si la demanda es superior a ese límite establecido, el OLT asignará un ancho de banda a la ONU igual a dicho límite. De esta forma, se evita que en un ciclo una ONU monopolice todo el ancho de banda disponible si su demanda es muy elevada, o que el ciclo sea de duración infinita (y por tanto existan

retardos muy elevados) si la demanda de las ONUs crece considerablemente.

Por tanto la asignación de ancho de banda en un ciclo se realiza según la ecuación (3.1), donde el término $B_{alloc}^{onu_i}$ es el ancho de banda asignado a la ONU i , $B_{demand}^{onu_i}$ es el ancho de banda demandado por dicha ONU y $B_{max}^{onu_i}$ es el ancho de banda máximo permitido para la ONU i (todos ellos en bits).

$$B_{alloc}^{onu_i} = \begin{cases} B_{demand}^{onu_i} & \text{si } B_{demand}^{onu_i} \leq B_{max}^{onu_i} \\ B_{max}^{onu_i} & \text{si } B_{demand}^{onu_i} > B_{max}^{onu_i} \end{cases} \quad (3.1)$$

Al contrario que otros algoritmos propuestos para estas arquitecturas de red, DySGAB constantemente reajusta el ancho de banda máximo permitido para cada ONU en cada ciclo, $B_{max}^{onu_i}$, dependiendo de la estimación actual del ancho de banda medio asignado y de los requisitos de ancho de banda garantizado estipulados. Por tanto, y tal como se explica en la siguiente sección, el algoritmo incrementa o decrementa estos anchos de banda máximos para usar eficientemente y de forma adaptativa los recursos de la red.

Esta característica es una gran fortaleza de DySGAB, ya que los anchos de banda máximos se actualizan constantemente con el propósito de que los usuarios cumplan con los niveles mínimos de ancho de banda garantizados de acuerdo con la prioridad de su SLA. También es importante recalcar que esta auto-adaptación presenta interesantes ventajas para los proveedores de servicio. En concreto, si necesitan cambiar los anchos de banda garantizados de los SLAs en tiempo real, DySGAB es capaz de reajustar los anchos de banda asignados automáticamente y de acuerdo a los nuevos requisitos de calidad de servicio estipulados, sin tener que calcular nuevos pesos o interrumpir el servicio de la red.

Aunque DySGAB actualiza los valores de $B_{max}^{onu_i}$ de forma periódica, para establecer inicialmente un ancho de banda máximo asignado a cada ONU se ha utilizado un esquema de pesos fijos relacionados con la prioridad del SLA contratado por cada uno de los usuarios, al igual que en [56, 64, 84]. De esta forma, las ONUs que pertenecen al mismo SLA comienzan con un mismo ancho de banda inicial que será proporcional al ancho de banda que se le tiene que garantizar.

Los anchos de banda máximos iniciales vienen dados según la ecuación (3.2), donde $B_{cycle_available}$ es el ancho de banda contenido en el tiempo de ciclo máximo considerado (por ejemplo, el ciclo máximo del estándar EPON es de 2 ms). El término W^{sla_j} , es el peso asignado al SLA j mientras que $W^{sla_k/onu_i \in sla_k}$ se corresponde con el peso asignado al SLA k que ha contratado la ONU i . Cada SLA existente en la red tiene asignado un peso diferente de acuerdo con los niveles de ancho de banda que se le tiene que garantizar. De esta forma, el SLA con mayor prioridad y que tendrá un mayor ancho de banda mínimo

a garantizar, tendrá mayor peso. Por último, N^{sla_j} se refiere al número de ONUs que pertenecen a cada SLA j .

$$B_{max}^{onu_i} = \frac{B_{cycle_available} \cdot W^{sla_k / onu_i \in sla_k}}{\sum_j W^{sla_j} \cdot N_{onus}^{sla_j}} \quad (3.2)$$

3.2.2. Adaptación dinámica del ancho de banda en DySGAB

Tal y como se ha comentado anteriormente, DySGAB adapta los anchos de banda máximos de cada una de las ONUs a lo largo del tiempo dependiendo de la estimación en tiempo real de su ancho de banda medio asignado, $\overline{B}_{alloc}^{onu_i}$. Para calcular este valor, cada ONU tiene una ventana temporal deslizante de duración fija de T_{window} segundos. En esta ventana temporal se almacenan muestras que contienen el ancho de banda asignado a cada una de las ONUs ($B_{alloc}^{onu_i}$) y el tiempo en el cual ese ancho de banda se asignó ($t_N, t_{N+1}, t_{N+2} \dots$). En la Figura 3.1 se muestra el funcionamiento de este tipo de ventana. Cuando llega una nueva muestra, se comprueba si la diferencia de tiempos entre la primera muestra y la entrante es menor que el tamaño fijo de la ventana (T_{window}). En caso de que la diferencia de tiempos no supere el valor de T_{window} la muestra entrante se añade al final de la ventana. Por el contrario, si la diferencia de tiempos supera dicho valor, se añade la muestra entrante pero se borra la primera muestra de la ventana. Una vez que se ha añadido la nueva muestra se actualiza el nuevo valor medio ($\overline{B}_{alloc}^{onu_i}$) con

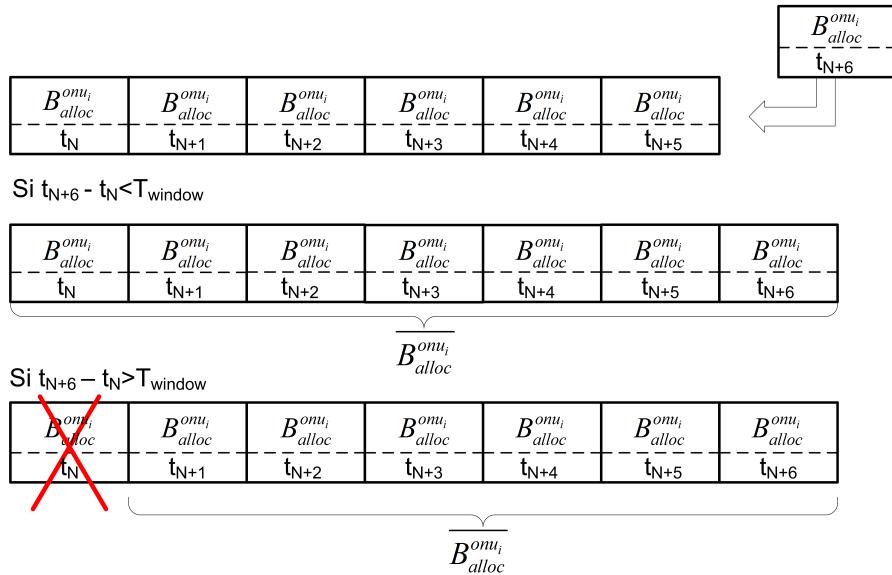


Figura 3.1: Funcionamiento de la ventana deslizante de la ONU i para el cálculo del ancho de banda medio asignado a dicha ONU.

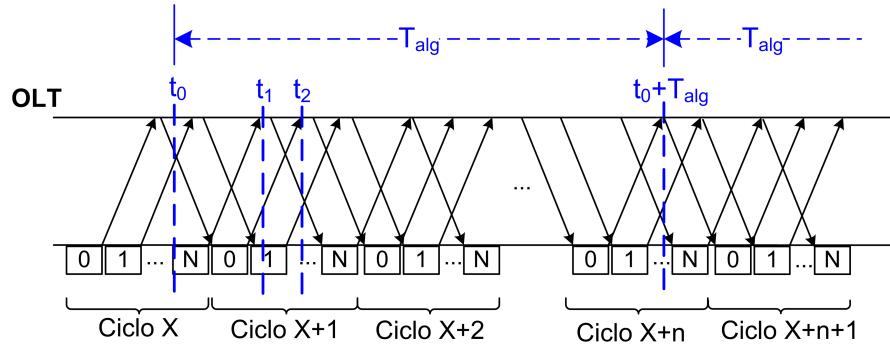


Figura 3.2: Esquema explicativo de la adaptación de máximos realizada por el algoritmo DySGAB.

todas las muestras contenidas en la ventana (Figura 3.1). Las muestras se insertan cada vez que se asigna un nuevo ancho de banda a cada ONU, esto es, en cada ciclo. Se ha optado por este tipo de ventana porque proporciona datos actualizados al ir eliminando las muestras menos recientes, por lo que en el proceso de asignación solo se tiene en cuenta el estado más reciente de la red.

Por otro lado, la adaptación de los máximos se va a ejecutar periódicamente, cada cierto tiempo denominado T_{alg} (medido en segundos). Un ejemplo del funcionamiento del algoritmo aparece en la Figura 3.2. En todo momento, se va a ir guardando un registro de las ONUs que en cada ciclo demandan un mayor ancho de banda que su máximo y de las que han pedido menos. En el ejemplo de la Figura 3.2, en t_0 ya ha transcurrido un tiempo igual a T_{alg} desde la última adaptación de los máximos. Entonces, el OLT espera a recibir el ancho de banda demandado por todas las ONUs en el ciclo actual (*Ciclo X*) y, en t_1 (cuando recibe la demanda de ancho de banda de la última ONU en ese ciclo) ejecutará la adaptación de los máximos. De esta forma, se tienen todas las demandas de ancho de banda actualizadas antes de realizar la adaptación. Por tanto, en t_1 se lleva a cabo la adaptación de los máximos, que tendrá efecto en las asignaciones realizadas a partir de este momento. En el caso de la Figura 3.2 esto ocurre a partir del instante t_2 , que es cuando se envía el mensaje *Gate* a la primera ONU en el *Ciclo X+2*. Por tanto, los nuevos máximos calculados por DySGAB ya se utilizan en el *Ciclo X+2* y en los ciclos consecutivos hasta que se realice la siguiente adaptación. Así, en el tiempo $t_0 + T_{alg}$, el proceso se repite, esperando a recibir todas las demandas de ancho de banda en el *Ciclo X+n* y realizando de nuevo la actualización de los máximos. Es conveniente enfatizar que DySGAB actualiza los máximos permitidos de todas las ONUs ($B_{max}^{onu_i}$) cada tiempo T_{alg} , mientras que la asignación de ancho de banda a cada ONU se realiza cada ciclo siguiendo una política de *polling*.

El principio básico de operación de DySGAB es la redistribución de ancho de banda entre todas las ONUs de forma dinámica y continua. Esto es, el ancho de banda máximo en un ciclo ($B_{max}^{onu_i}$) de algunas ONUs se reducirá, mientras que otras ONUs incrementarán el suyo dependiendo del ancho de banda total reducido. Para hacer esto, el primer paso de la adaptación consiste en comprobar la demanda de todas las ONUs. Entonces, DySGAB diferencia entre tres grupos diferentes de ONUs, un grupo de m ONUs (representadas a partir de ahora por $ONUs^-$) cuyos anchos de banda máximos serán reducidos, un grupo de n ONUs (representadas por $ONUs^+$), cuyos máximos serán incrementados, y en algunos casos, ONUs cuyos máximos no se verán modificados. Dependiendo de la demanda de todas las ONUs, el número de miembros de cada grupo es diferente. Por tanto, para hacer la clasificación, DySGAB analiza si hay ONUs cuya demanda es menor que su ancho de banda máximo permitido y si hay ONUs que están demandando más que sus máximos y no están cumpliendo con su ancho de banda garantizado, que son las ONUs que están siendo en ese momento más perjudicadas. El resultado de este análisis indica cuál de los cuatro casos considerados por DySGAB se va a aplicar para organizar las ONUs en los diferentes grupos. Estos cuatro casos se explicarán más adelante y recogerán los casos posibles que se pueden dar en la red LR-EPON.

Una vez que las ONUs están clasificadas, los valores de los anchos de banda máximos permitidos se modifican. Entonces, las m ONUs que pertenecen al grupo $ONUs^-$, reducen su ancho de banda máximo un número de *bytes* igual a $B_{reduction}$ siguiendo la ecuación (3.3). Esta cantidad es igual para todas las ONUs que pertenezcan al conjunto de $ONUs^-$ y su valor se discutirá en la siguiente sección.

$$\forall ONUi \in ONUs^- \Rightarrow B_{max}^{onu_i} = B_{max}^{onu_i} - B_{reduction} \quad (3.3)$$

Por otro lado, cada ONU i que pertenece al grupo $ONUs^+$, incrementa su ancho de banda máximo un número de *bytes* igual a $B_+^{onu_i}$ (ecuación (3.4)). Este término se calcula siguiendo la ecuación (3.5). En esta ecuación (3.5), el término $\sum_m B_{reduction}$ representa el ancho de banda total reducido de todas las $ONUs^-$ en una adaptación, y es igual al ancho de banda que se va a distribuir entre las ONUs pertenecientes al grupo $ONUs^+$. Además, el término $B_{guarantee}^{sla \in onu_i}$ es el ancho de banda estipulado para el SLA que ha contratado la ONU i , esto es, el ancho de banda que hay que garantizar a dicha ONU.

$$\forall ONUi \in ONUs^+ \Rightarrow B_{max}^{onu_i} = B_{max}^{onu_i} + B_+^{onu_i} \quad (3.4)$$

$$B_{+}^{onu_i} = \frac{B_{guarantee}^{sla \in onu_i}}{\sum_n B_{guarantee}^{sla \in onu_n}} \cdot \sum_m B_{reduction} \quad (3.5)$$

Con la distribución realizada mediante la ecuación (3.5), las ONUs que pertenecen al grupo $ONUs^+$ incrementan sus máximos dependiendo del nivel de ancho de banda que se les tiene que garantizar. Esto es, aquellas ONUs del grupo $ONUs^+$ que tengan que cumplir con un nivel de ancho de banda garantizado mayor, incrementarán su máximo más que aquellas que tienen que cumplir con un ancho de banda garantizado menor.

Los diferentes casos y las acciones que realiza DySGAB según el estado y la demanda de las diferentes ONUs están recogidos en la Tabla 3.1.

A continuación, se explican los pasos seguidos por DySGAB para realizar la adaptación de los anchos de banda máximos y los distintos casos considerados junto con los correspondientes pseudocódigos. Para ello, en primer lugar se explica el proceso principal donde se determina el caso en el que se encuentra la red. Después se presentan los cuatro casos considerados por DySGAB para realizar el reajuste de los anchos de banda máximos permitidos a cada ONU ($B_{max}^{onu_i}$).

Proceso principal:

En el proceso principal, DySGAB lleva a cabo un análisis para determinar el estado de las demandas de cada una de las ONUs. En la primera línea del pseudocódigo, todas las variables utilizadas en la adaptación se inicializan. Después, se comprueba la demanda de cada una de las ONUs de la red (donde la variable *num_ONUs_red* es el número de ONUs en la red). Si la ONU ha demandado menos ancho de banda que su máximo, se incrementa un contador que guarda el número de ONUs que han demandado menos ancho de banda (*onus_piden_menos*). Por tanto, la variable *onus_piden_menos* representa el número total de ONUs en la red cuya demanda es menor que su máximo permitido. Por el contrario, si la ONU demanda más ancho de banda que su máximo actual, DySGAB comprueba si su ancho de banda medio asignado es menor que el ancho de banda garantizado según su SLA. Si esto ocurre, la variable *onus_dar* se incrementa. Por tanto, esta variable representa el número de ONUs que además de demandar un ancho de banda mayor que su máximo, su ancho de banda medio no alcanza el nivel garantizado. Además, como este último grupo de ONUs es el más probable para incrementar su máximo permitido, DySGAB también acumula la suma de sus niveles garantizados ($\sum B_{guarantee}^{sla \in onu}$) para calcular más tarde la cantidad de ancho de banda que se va a aumentar, esto es, el término $B_{+}^{onu_i}$.

	ONUs que reducen su ancho de banda máximo (ONUs del grupo $ONUs^-$)	ONUs que aumentan su ancho de banda máximo (ONUs del grupo $ONUs^+$)
CASO 1: Hay algunas ONUs que demandan menos ancho de banda que su máximo y otras que no cumplen con su nivel garantizado	$\forall ONUi / B_{demand}^{onu_i} < B_{max}^{onu_i}$ $B_{max}^{onu_i} = B_{max}^{onu_i} - B_{reduction}$	$\forall ONUi / (B_{demand}^{onu_i} > B_{max}^{onu_i})$ y $(\overline{B}_{alloc}^{onu_i} < B_{guarantee}^{onu_i})$ $B_{max}^{onu_i} = B_{max}^{onu_i} + B_+^{onu_i}$ donde $B_+^{onu_i}$ depende de $\sum_m B_{reduction}$ y de $B_{guarantee}^{onu_i}$
CASO 2: Todas las ONUs demandan más ancho de banda que su máximo pero hay ONUs que sí cumplen con sus niveles de ancho de banda garantizados.	$\forall ONUi / \overline{B}_{alloc}^{onu_i} > B_{guarantee}^{onu_i}$ $B_{max}^{onu_i} = B_{max}^{onu_i} - B_{reduction}$	$\forall ONUi / \overline{B}_{alloc}^{onu_i} < B_{guarantee}^{onu_i}$ $B_{max}^{onu_i} = B_{max}^{onu_i} + B_+^{onu_i}$ donde $B_+^{onu_i}$ depende de $\sum_m B_{reduction}$ y de $B_{guarantee}^{onu_i}$
CASO 3: Todas las ONUs demandan más ancho de banda que su máximo pero no hay ONUs que cumplan con sus niveles garantizados.	$\forall ONUi / \in SLA$ de menor prioridad $B_{max}^{onu_i} = B_{max}^{onu_i} - B_{reduction}$	$\forall ONUi / \notin SLA$ de menor prioridad $B_{max}^{onu_i} = B_{max}^{onu_i} + B_+^{onu_i}$ donde $B_+^{onu_i}$ depende de $\sum_m B_{reduction}$ y de $B_{guarantee}^{onu_i}$
CASO 4: Todas las ONUs cumplen con sus niveles garantizados o todas demandan menos ancho de banda.	$B_{max}^{onu_i}$ no se modifica	$B_{max}^{onu_i}$ no se modifica

Tabla 3.1: Acciones a realizar en DySGAB para la adaptación del ancho de banda máximo.

Dependiendo de los valores de las variables $onus_piden_menos$ y $onus_dar$, DySGAB distingue cuatro casos que se explicarán seguidamente. El pseudocódigo correspondiente al proceso principal es el que se muestra a continuación:

```

 $onus\_piden\_menos \leftarrow 0; onus\_dar \leftarrow 0; \sum B_{reduction} \leftarrow 0; \sum B_{guarantee}^{sla \in onu_i} \leftarrow 0$ 
para  $i \leftarrow 0$  hasta  $num\_ONUs\_red$  hacer
    //si las ONUs demandan menos ancho de banda que su máximo
    si  $B_{demand}^{onu_i} < B_{max}^{onu_i}$  entonces
        //contador que guarda el número de ONUs que han demandado
        //menos ancho de banda que su máximo
         $onus\_piden\_menos \leftarrow onus\_piden\_menos + 1$ 
    fin si
    si  $((B_{demand}^{onu_i} > B_{max}^{onu_i}) \text{ y } (\overline{B}_{alloc}^{onu_i} < B_{guarantee}^{sla \in onu_i}))$  entonces
        //contador que guarda el número de ONUs que demandan más que su
        //máximo permitido y no cumplen con los requisitos de ancho de banda
         $onus\_dar \leftarrow onus\_dar + 1$ 
        //sumatorio de todos los niveles
        //garantizados de las ONUs que recibirán ancho de banda.
         $\sum B_{guarantee}^{sla \in onu} \leftarrow \sum B_{guarantee}^{sla \in onu} + B_{guarantee}^{sla \in onu_i}$ 
    fin si
fin para

```

A partir de aquí se explica cada uno de los cuatro casos contemplados por DySGAB.

CASO 1:

En este primer caso hay ONUs que demandan menos ancho de banda que su máximo ($onus_piden_menos > 0$) y hay ONUs que demandan más ancho de banda que su máximo y que además no cumplen con su nivel garantizado ($onus_dar > 0$). Entonces DySGAB considera que el primer grupo de ONUs no necesita máximos tan altos, así que los reduce una cantidad igual a $B_{reduction}$. Esto es, se considera $ONUs^-$ a las ONUs cuya demanda es menor que su máximo. A continuación, DySGAB acumula el ancho de banda que se ha reducido en esas ONUs ($\sum B_{reduction}$). Después, se analizan de nuevo todas las ONUs de la red. Aquellas cuya demanda es mayor que su máximo y no cumplen con su nivel de ancho de banda garantizado se consideran del grupo $ONUs^+$, es decir, son las ONUs a las que se incrementará su máximo porque lo necesitan para cumplir con sus niveles mínimos estipulados. A continuación, como el ancho de banda total reducido y la suma de los anchos de banda garantizados de las $ONUs^+$ es conocido (este último término coincide con $\sum B_{guarantee}^{sla \in onu}$ calculado en el proceso principal), el ancho de banda máximo en estas ONUs ($ONUs^+$) se incrementa siguiendo las ecuaciones (3.4) y (3.5). Como se ha

indicado anteriormente, el término $\sum B_{guarantee}^{sla \in onu}$ calculado en el proceso principal sirve en este CASO 1 porque las $ONUs^+$ que van a incrementar su $B_{max}^{onu_i}$ son aquellas ONUs que demandan más que su máximo y no cumplen con su nivel garantizado. Finalmente, las ONUs que demandan más ancho de banda pero satisfacen el nivel garantizado acorde al SLA contratado no modifican sus máximos, ya que sus requisitos se cumplen. El pseudocódigo para el CASO 1 es el que se muestra a continuación:

```

si ((onus_piden_menos > 0) y (onus_dar > 0)) entonces
    para i ← 0 hasta num_ONUs_red hacer
        si ( $B_{demand}^{onu_i} < B_{max}^{onu_i}$ ) entonces
            //Estas son  $ONUs^-$ .
             $B_{max}^{onu_i} \leftarrow B_{max}^{onu_i} - B_{reduction}$ 
             $\sum B_{reduction} \leftarrow \sum B_{reduction} + B_{reduction}$ 
        fin si
    fin para
    para i ← 0 hasta num_ONUs_red hacer
        si ( $B_{demand}^{onu_i} > B_{max}^{onu_i}$ ) y ( $\overline{B}_{alloc}^{onu_i} < B_{guarantee}^{sla \in onu_i}$ ) entonces
            //Estas son  $ONUs^+$ .
             $B_{max}^{onu_i} = B_{max}^{onu_i} + B_{+}^{onu_i}$ 
        fin si
    fin para

```

CASO 2:

Este caso se produce cuando todas las ONUs en la red piden más ancho de banda que su máximo (*onus_piden_menos* = 0). Además, existen ONUs que no cumplen con su nivel garantizado y ONUs cuyo ancho de banda medio asignado es superior al garantizado y que por tanto cumplen con sus requisitos (*onus_dar* ≠ *num_ONUs_red*). Por eso, DySGAB analiza cada ONU y reduce el ancho de banda máximo de todas aquellas ONUs cuyo ancho de banda medio sea mayor que su ancho garantizado, es decir, reduce el máximo de las ONUs que sí cumplen. El objetivo de esta reducción consiste en que estas ONUs sigan cumpliendo con su ancho de banda garantizado, pero que el resto de ONUs que no lo cumplen puedan evolucionar hasta él. Por tanto, en este caso las ONUs que pertenecen al grupo $ONUs^-$ son aquellas que cumplen con su nivel de ancho de banda garantizado de acuerdo al SLA contratado pero demandan más que su máximo permitido. Además, DySGAB calcula el ancho de banda total reducido. A continuación, todas las ONUs se analizan de nuevo y aquellas que no cumplen con su nivel de ancho de banda garantizado pertenecerán al grupo $ONUs^+$. De nuevo, como el ancho de banda total reducido y la suma de los anchos de banda garantizados de las ONUs que pertenecen

al grupo $ONUs^+$ son conocidos (este último término ya ha sido calculado en el proceso principal), los máximos se incrementan siguiendo las ecuaciones (3.4) y (3.5). Al igual que ocurría en el CASO 1, el término $\sum B_{guarantee}^{sla \in onu}$ calculado en el proceso principal sirve en el CASO 2 porque las $ONUs^+$ que incrementan su $B_{max}^{onu_i}$ son aquellas ONUs que demandan más que su máximo y no cumplen con su nivel garantizado. El pseudocódigo para este CASO 2 es el que se muestra a continuación:

```

si ((onus_piden_menos == 0) y (onus_dar ≠ num_onus_red)) entonces
  para  $i \leftarrow 0$  hasta num_ONUs_red hacer
    si ( $\overline{B_{alloc}^{onu_i}} > B_{guarantee}^{sla \in onu_i}$ ) entonces
      //Estas son  $ONUs^-$ .
       $B_{max}^{onu_i} \leftarrow B_{max}^{onu_i} - B_{reduction}$ 
       $\sum B_{reduction} \leftarrow \sum B_{reduction} + B_{reduction}$ 
    fin si
  fin para
  para  $i \leftarrow 0$  hasta num_ONUs_red hacer
    si ( $\overline{B_{alloc}^{onu_i}} < B_{guarantee}^{sla \in onu_i}$ ) entonces
      //Estas son  $ONUs^+$ .
       $B_{max}^{onu_i} = B_{max}^{onu_i} + B_{+}^{onu_i}$ 
    fin si
  fin para

```

CASO 3:

Al igual que en el CASO 2, todas las ONUs demandan más ancho de banda que sus máximos ($onus_piden_menos = 0$). Sin embargo, a diferencia del CASO 2, no hay ONUs que cumplan con el nivel de ancho de banda estipulado según su SLA ($onus_dar = num_onus_red$). Bajo esta situación, DySGAB opta por una política que beneficia a aquellos usuarios con SLAs más prioritarios a expensas de reducir el ancho de banda máximo de aquellos usuarios que han contratado el SLA de menor prioridad. Aunque se podrían haber elegido otras políticas, ésta ofrece un trato preferente a los usuarios con mayor prioridad. Por tanto, se considera que las ONUs que han contratado el SLA de menor prioridad pertenecen al grupo $ONUs^-$, mientras que el resto de ONUs pertenecen al grupo $ONUs^+$. Entonces, DySGAB analiza el SLA de todas las ONUs de la red. Si la ONU ha contratado el SLA menos prioritario, entonces se reducen sus máximos y la variable con el ancho de banda total reducido se actualiza. Si por el contrario, la ONU no pertenece al SLA de menor prioridad, DySGAB acumula la suma de su nivel garantizado para recalcular el término $\sum B_{guarantee}^{sla \in onu}$. Esto ocurre porque el término calculado en el procedimiento principal no sirve para este caso, a diferencia de lo que ocurría para el

CASO 1 y el CASO 2. Esto es así porque ahora el grupo de $ONUs^+$ son todas las ONUs que pertenecen a los SLAs prioritarios, y no todas aquellas que demandan más ancho de banda y no cumplen con su nivel garantizado, que es lo que se considera en el proceso principal para calcular este término. Una vez que se conoce el ancho de banda total reducido y la suma de los anchos de banda garantizados de las $ONUs^+$, DySGAB incrementa el ancho de banda máximo permitido de las ONUs que no pertenecen al SLA de menor prioridad siguiendo las ecuaciones (3.4) y (3.5). El pseudocódigo para este CASO 3 es el que se muestra a continuación:

```

si ((onus_piden_menos == 0) y (onus_dar == num_onus_red)) entonces
    para i  $\leftarrow 0$  hasta num_ONUs_red hacer
        si ( $ONU_i \in SLA_{menor\_prioridad}$ ) entonces
            //Estas son  $ONUs^-$ .
             $B_{max}^{onu_i} \leftarrow B_{max}^{onu_i} - B_{reduction}$ 
             $\sum B_{reduction} \leftarrow \sum B_{reduction} + B_{reduction}$ 
        fin si
        sino
             $\sum B_{guarantee}^{sla \in onu} \leftarrow \sum B_{guarantee}^{sla \in onu} + B_{guarantee}^{sla \in onu_i}$ 
        fin sino
    fin para
    para i  $\leftarrow 0$  hasta num_ONUs_red hacer
        si ( $ONU_i \notin SLA_{menor\_prioridad}$ ) entonces
            //Estas son  $ONUs^+$ .
             $B_{max}^{onu_i} = B_{max}^{onu_i} + B_{+}^{onu_i}$ 
        fin si
    fin para

```

CASO 4:

Este último caso se da cuando todas las ONUs demandan menos ancho de banda que su máximo ($onus_piden_menos = num_onus_red$) o cuando todas las ONUs cumplen con su ancho de banda garantizado ($onus_dar = 0$). Bajo estas condiciones, los máximos no se modifican. Por este motivo, para este caso no existe pseudocódigo.

Por último, cabe destacar que el término $B_{reduction}$, tal y como se comentó anteriormente, se corresponde con la cantidad en *bytes* que se va a reducir el ancho de banda máximo de cada ONU del grupo $ONUs^-$. Este valor influye considerablemente en el comportamiento del algoritmo. Por tanto, en el Apartado 3.3.2 de este capítulo se realizará un estudio para la elección de su valor óptimo. Por otro lado, la Figura 3.3 muestra

el diagrama de flujo donde aparecen representadas la secuencia de acciones que realiza DySGAB, según los casos recogidos en la Tabla 3.1, y las explicaciones realizadas sobre el pseudocódigo de cada uno de los casos.

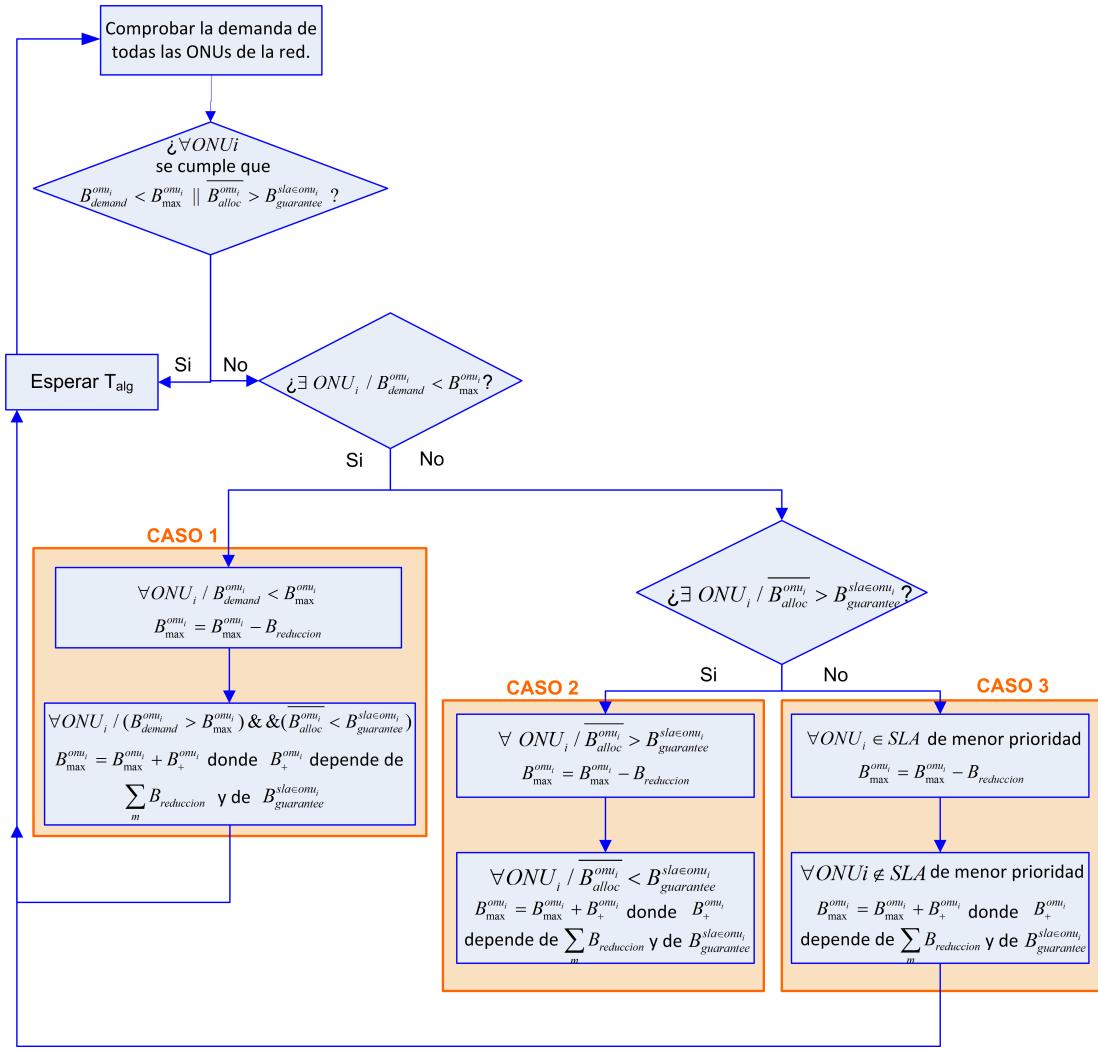


Figura 3.3: Esquema con los pasos de la adaptación que realiza DySGAB.

3.3. Estudio de simulación del algoritmo DySGAB

Para comprobar el funcionamiento y las prestaciones del algoritmo, se ha procedido en primer lugar a la elección de los parámetros propios del diseño de DySGAB para obtener un buen comportamiento del mismo. En concreto, se llevarán a cabo simulaciones para obtener un valor óptimo para el tamaño de ventana (T_{window}), para el término $B_{reduction}$ y para el intervalo entre ejecuciones consecutivas de la adaptación de los máximos (T_{alg}).

Una vez fijados estos valores, se realizarán diversas pruebas para comprobar las prestaciones del algoritmo bajo distintos escenarios de red, tales como variando los niveles de ancho de banda garantizados, las condiciones de ancho de banda iniciales, el número de ONUs asociadas a cada perfil de abonado y la longitud de la red LR-EPON.

3.3.1. Escenario de simulación

Para analizar el comportamiento de DySGAB se han realizado simulaciones utilizando la herramienta OPNET Modeler v.16.0 [85]. Se ha considerado una red con 16 ONUs conectadas mediante un divisor óptico pasivo a un OLT y un único usuario conectado a cada ONU. Esta es una arquitectura típica de fibra hasta el hogar (FTTH), similar a la considerada en la mayor parte de algoritmos DBA propuestos [17, 56, 64, 83, 84]. La tasa de transmisión en sentido ascendente entre las ONUs y el OLT se ha fijado a 1 Gbit/s (estándar EPON) y los enlaces entre las ONUs y los usuarios soportan tasas de transmisión de 100 Mbit/s [47, 56, 64, 83]. El principal análisis de este algoritmo se realiza para redes LR-EPON ya que son una tecnología muy prometedora debido a la reducción de costes y a su gran cobertura. Por ello, se ha establecido una distancia de red (entre el OLT y las ONUs) de 100 km, típica en las arquitecturas EPON de largo alcance [18, 19, 43], lo que supone un RTT de 1 ms (ecuación (3.6)). Sin embargo, para demostrar la aplicabilidad de DySGAB en redes EPON y LR-EPON de distinto alcance, también se han incluido en el análisis de prestaciones simulaciones para diferente rango de distancias.

$$\frac{RTT}{2} = \frac{Distancia_{OLT_ONU}}{V_{fibra_optica}} = \frac{100 \cdot 10^3 m}{2 \cdot 10^8 m/s} = 5 \cdot 10^{-4} s \Rightarrow RTT = 1ms \quad (3.6)$$

Se ha asumido un escenario de tráfico simétrico, como el utilizado en la mayoría de los algoritmos de asignación de ancho de banda desarrollados para redes LR-PON [47, 56, 64]. En este tipo de escenarios todas las ONUs contribuyen por igual en la carga de red, es decir, todas las ONUs generan datos a la misma tasa de transmisión. Por tanto, la relación entre la carga total de red ($\rho_{network}$) y la carga de las ONUs (ρ_{onu})

se muestra en la ecuación (3.7), donde N_{onus} representa el número de ONUs de la red, $EPON_Line_rate$ es la tasa de transmisión del canal de subida entre las ONUs y el OLT (1 Gbit/s) y $User_Line_rate$ es la tasa en el enlace entre las ONUs y los usuarios (100 Mbit/s).

$$\rho_{onu} = \frac{EPON_Line_rate}{User_Line_rate} \cdot \frac{\rho_{network}}{N_{onus}} \quad (3.7)$$

Se ha considerado también que cada ONU tiene una única memoria con una sola cola y una capacidad de 6 Mbit, donde se encolarán los paquetes según el orden de llegada. Los paquetes generados por los usuarios finales siguen una distribución de Pareto con un parámetro de *Hurst* (*H*), igual a 0.8, y con un tamaño de paquete variable entre 64 y 1500 bytes más 38 bytes de cabeceras siguiendo el estándar Ethernet 802.3 [86]. Para evitar colisiones entre las transmisiones de ONUs adyacentes, se ha escogido un tiempo de guarda de 1 μ s y un tiempo máximo de ciclo de 2 ms (estándar EPON [86]).

Por otro lado, para comprobar los beneficios de la dinamicidad implementada por DySGAB, su comportamiento se ha comparado con LIPSA [64], un algoritmo de *polling* diseñado para redes LR-EAPON que también garantiza niveles mínimos de ancho de banda según la prioridad del SLA contratado, pero que se basa en la utilización de un esquema de pesos fijos, siguiendo la ecuación (3.2). Tal y como se comentó en el capítulo anterior, los valores de estos pesos se fijan de tal manera que a cada SLA se le asigna un ancho de banda acorde al nivel que se le debe garantizar. Para poder llevar a cabo esta comparación, se ha considerado la misma distribución de perfiles de usuario que LIPSA [64]. Por tanto en la red existen tres SLAs diferentes donde una única ONU contrata el servicio más prioritario (SLA_0), cinco contratan el de prioridad intermedia (SLA_1) y diez ONUs contratan el menos prioritario (SLA_2). Así mismo, los niveles de ancho de banda garantizados que tienen que cumplir las distintas ONUs se han fijado a 100/75/50 Mbit/s para el SLA_0 , SLA_1 y SLA_2 , respectivamente, siguiendo los planes de servicio de NTT (*Nippon Telegraph and Telephone*) [87]. Además, para llevar a cabo la comparación de ambos algoritmos bajo las mismas condiciones, se han fijado los pesos para calcular los anchos de banda iniciales (ecuación (3.2)) iguales a 1 ($W^{sla0} = 1, W^{sla1} = 1, W^{sla2} = 1$). Sin embargo, para demostrar la flexibilidad e independencia que ofrece DySGAB respecto a las condiciones iniciales de ancho de banda y a los niveles de ancho de banda garantizados, se han considerado a lo largo de este estudio otros escenarios con diferentes valores tanto en los pesos iniciales, como en el ancho de banda mínimo a garantizar. En la Tabla 3.2 se recoge un resumen de los parámetros de red considerados en el entorno de simulación inicial general.

Parámetros de simulación	Valores
Número de ONUs	16 ONUs (N_{onu})
Tasa de transmisión canal ascendente	1 Gbit/s ($EPON_Line_rate$)
Tasa de transmisión ONU-abonado	100 Mbit/s ($User_Line_rate$)
RTT	1 ms
Tiempo de guarda	1 μ s
Tiempo de ciclo máximo	2 ms
Distribución de tráfico	Fuente de Pareto H=0.8
Tamaño de paquete	Variable entre 64-1500 bytes (más 38 bytes de cabeceras)
<i>Service Level Agreement (SLA)</i>	SLA ₀ : 1 ONU Ancho de banda garantizado: 100 Mbit/s
	SLA ₁ : 5 ONUs Ancho de banda garantizado: 75 Mbit/s
	SLA ₂ : 10 ONUs Ancho de banda garantizado: 50 Mbit/s

Tabla 3.2: Parámetros de red considerados en el entorno inicial de simulación.

3.3.2. Establecimiento de los parámetros asociados al diseño de DySGAB

El funcionamiento de DySGAB depende de varios parámetros involucrados en la ejecución del algoritmo. De hecho, existen tres parámetros fundamentales relacionados con su diseño que caracterizan su funcionamiento, y hallar los valores óptimos es esencial para garantizar una distribución eficiente en tiempo real del ancho de banda disponible.

El primer factor es el tiempo de ventana (T_{window}), durante el cual las muestras asociadas al ancho de banda de cada ONU son almacenadas. Éste es un parámetro significativo, ya que la adaptación dinámica de los anchos de banda máximos permitidos a cada ONU ($B_{max}^{onu_i}$) depende directamente de su ancho de banda medio asignado ($\overline{B}_{alloc}^{onu_i}$) y se calcula a través de las muestras contenidas en su ventana asociada. Por lo tanto, si la duración de la ventana es muy grande, el algoritmo reacciona con una mayor latencia y los valores no se actualizarán con suficiente frecuencia, por lo que la adaptación no se realizará de forma adecuada, sobre todo si existen cambios abruptos en el patrón de tráfico existente. Por otro lado, si el tamaño de la ventana es demasiado pequeño, la estimación media del ancho de banda no representa de manera apropiada el ancho de banda medio asignado, ya que el número de muestras en la ventana no es suficientemente grande. Por tanto, se debe seleccionar un valor intermedio que permita obtener una estimación real del ancho de banda medio para conseguir una adaptación adecuada al nivel garantizado de cada ONU. Para hallar los valores más adecuados, se han llevado

a cabo simulaciones bajo el escenario de red de la Tabla 3.2 pero considerando distintos tiempos del parámetro T_{window} .

La Figura 3.4 (a), (b) y (c) representa la evolución en tiempo real del ancho de banda medio ($\overline{B}_{alloc}^{onu_i}$ en la Figura 3.1) asignado a una ONU del SLA₀, SLA₁ y SLA₂ respectivamente, considerando distintas duraciones de la ventana (120 s, 60 s, 30 s y 7.5 s). Se ha considerado solo una ONU de cada SLA puesto que al generar todas el mismo tráfico, el comportamiento de una ONU es idéntico para el resto de ONUs del mismo SLA. Como puede observarse en las gráficas, para las ventanas con un tamaño menor (7.5 s y 30 s), la variación del ancho de banda medio asignado es menor, especialmente en los dos SLAs menos prioritarios, consiguiendo una situación más estable y una evolución a los valores estipulados más rápida (Figura 3.4 (b) y (c)). Por el contrario, para mayores tiempos de ventana (120 s y 60 s), el rango de variación es mayor y la evolución es también

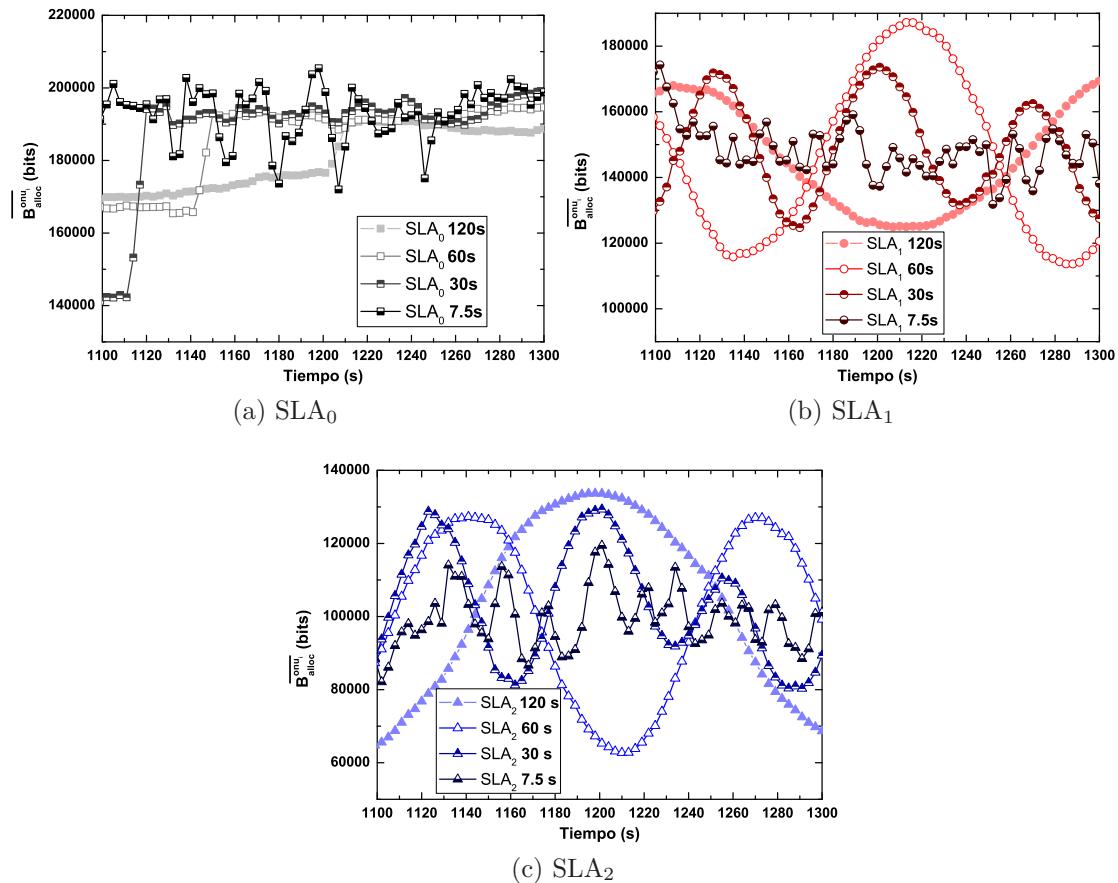


Figura 3.4: Variación en tiempo real del ancho de banda medio para tiempos de ventana relativamente altos. (a) SLA₀ (b) SLA₁ (c) SLA₂.

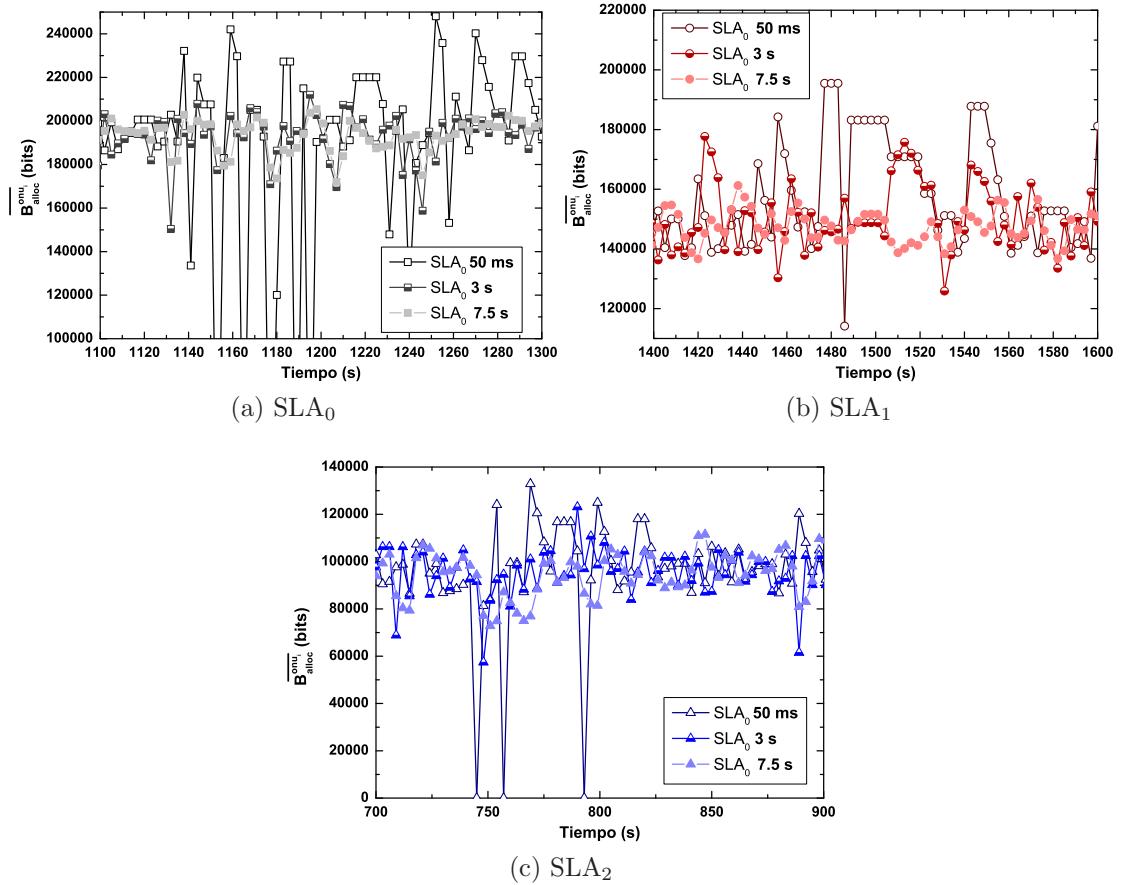


Figura 3.5: Variación en tiempo real del ancho de banda medio para tiempos de ventana relativamente bajos. (a) SLA₀ (b) SLA₁ (c) SLA₂.

más lenta, por lo que este comportamiento es no deseado. Para el SLA de mayor prioridad se observa la misma tendencia, aunque las diferencias entre las diferentes ventanas son menores que para los dos SLAs de menor prioridad (Figura 3.4 (a)).

Por otro lado, la Figura 3.5 (a), (b) y (c) muestra la evolución del ancho de banda medio asignado a una ONU del SLA₀, SLA₁ y SLA₂, respectivamente, para tiempos de ventana mucho más pequeños (50 ms, 3 s y 7.5 s). Tal y como se aprecia, para los tamaños más pequeños (50 ms y 3 s), se observan grandes oscilaciones del ancho de banda en los tres SLAs, en especial para la ventana más pequeña de 50 ms. Por tanto, analizando la Figura 3.4 y la Figura 3.5, el valor escogido para el tamaño de la ventana será de 7.5 segundos, ya que este valor intermedio permite simultáneamente una evolución en tiempo real más rápida hacia los valores de ancho de banda garantizados y presenta una respuesta más estable del ancho de banda asignado.

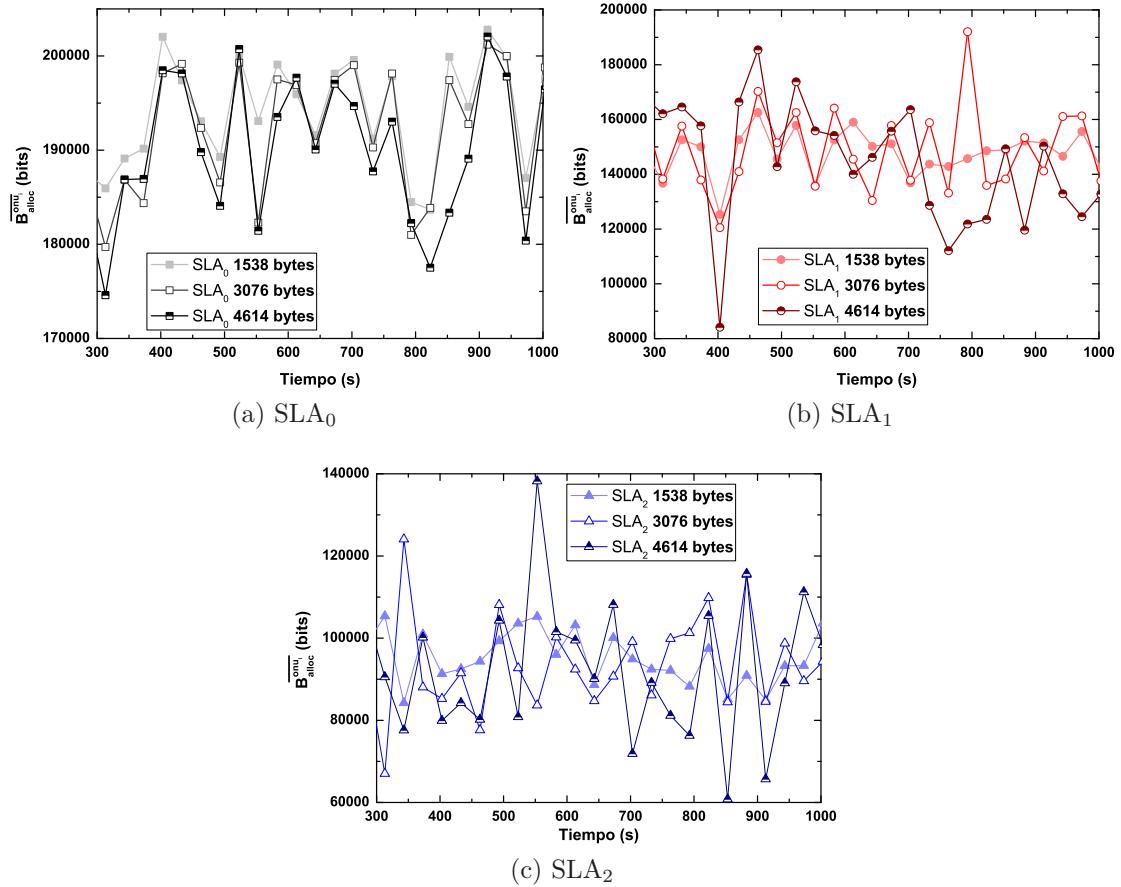


Figura 3.6: Evolución en tiempo real del ancho de banda medio asignado a una ONU de cada SLA para valores elevados del término $B_{\text{reduction}}$ (en bytes). (a) SLA₀ (b) SLA₁ (c) SLA₂.

Otro parámetro muy importante a tener en cuenta es el ancho de banda involucrado en el reajuste, esto es, el término $B_{\text{reduction}}$ de la ecuación (3.3). Si el valor escogido es demasiado pequeño, la evolución del ancho de banda medio hacia los niveles de ancho de banda garantizados puede llegar a ser excesivamente lenta. Por el contrario, si este valor es demasiado alto, los cambios en los anchos de banda máximos permitidos serán muy abruptos y por tanto no convergerán o la respuesta en tiempo real no será demasiado estable. Por ello, se han realizado simulaciones considerando diferentes tamaños (en bytes) para dicho término involucrado en la adaptación.

En concreto, en la Figura 3.6 (a), (b) y (c) se representa la variación en tiempo real del ancho de banda medio asignado a una ONU de cada SLA cuando el ancho de banda reducido es igual a 1538 bytes, 3076 bytes y 4614 bytes. Como se observa, las variaciones más pequeñas del ancho de banda medio asignado se obtienen para el caso

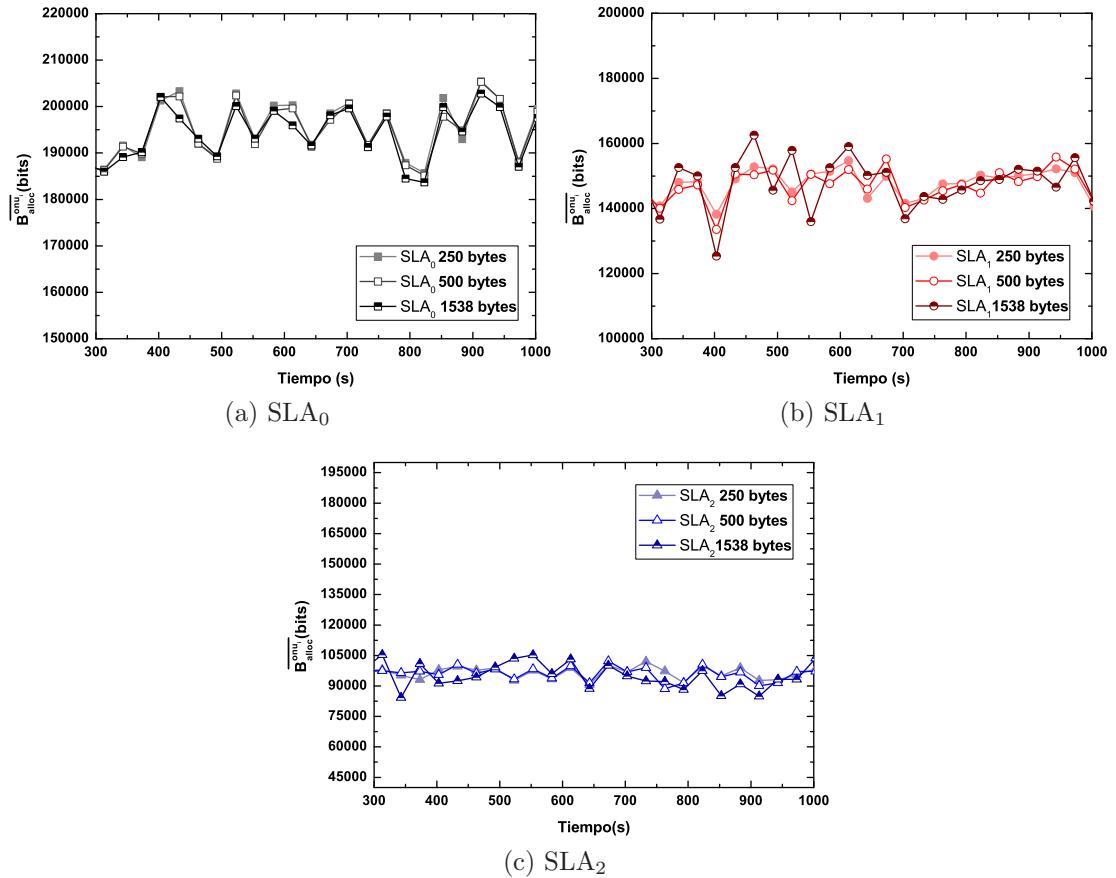


Figura 3.7: Evolución en tiempo real del ancho de banda medio asignado a una ONU de cada SLA para valores pequeños del término $B_{reduction}$ (en bytes). (a) SLA₀ (b) SLA₁ (c) SLA₂.

en el que el ancho de banda se reduce una cantidad igual a 1538 *bytes* (en especial para el SLA₁ y SLA₂). Esto es así porque para reducciones de ancho de banda muy altas, se puede quitar y/o añadir demasiado ancho de banda de forma repentina, por lo que existirían demasiadas fluctuaciones en el ancho de banda medio asignado. Por el contrario, reducciones muy pequeñas pueden implicar una evolución de los anchos de banda demasiado lenta. Para analizar este comportamiento del algoritmo cuando la reducción del ancho de banda máximo es igual a 250, 500 y 1538 *bytes*. Como puede observarse, la fluctuación obtenida es menor a medida que los valores de $B_{reduction}$ son menores, aunque las diferencias no son demasiado grandes entre ellos.

No obstante y tal como se ha comentado anteriormente, para valores excesivamente

pequeños del término $B_{reduction}$, la velocidad de convergencia hacia los niveles garantizados estipulados es mucho más lenta. Para comprobar este comportamiento, la Figura 3.8 (a), (b) y (c) muestra la evolución instantánea del ancho de banda medio para los tres SLAs cuando se consideran reducciones de ancho de banda iguales e inferiores 1538 bytes para los 100 primeros segundos. Como se observa, la convergencia hacia los niveles estipulados es más lenta a medida que se reduce el valor de $B_{reduction}$. Por ejemplo, para el caso del SLA_0 , si dicho parámetro es igual a 1538 bytes, la convergencia de los valores hacia su nivel garantizado (100 Mbit/s), se realiza en unos 50 s aproximadamente. Por el contrario, para reducciones de 250 bytes, los valores no convergen hasta los 75 s aproximadamente. Para el SLA_2 , cuando $B_{reduction}$ es igual a 1538 bytes el ancho de banda medio converge rápidamente, en torno a 20 s, mientras que para 250 bytes este tiempo

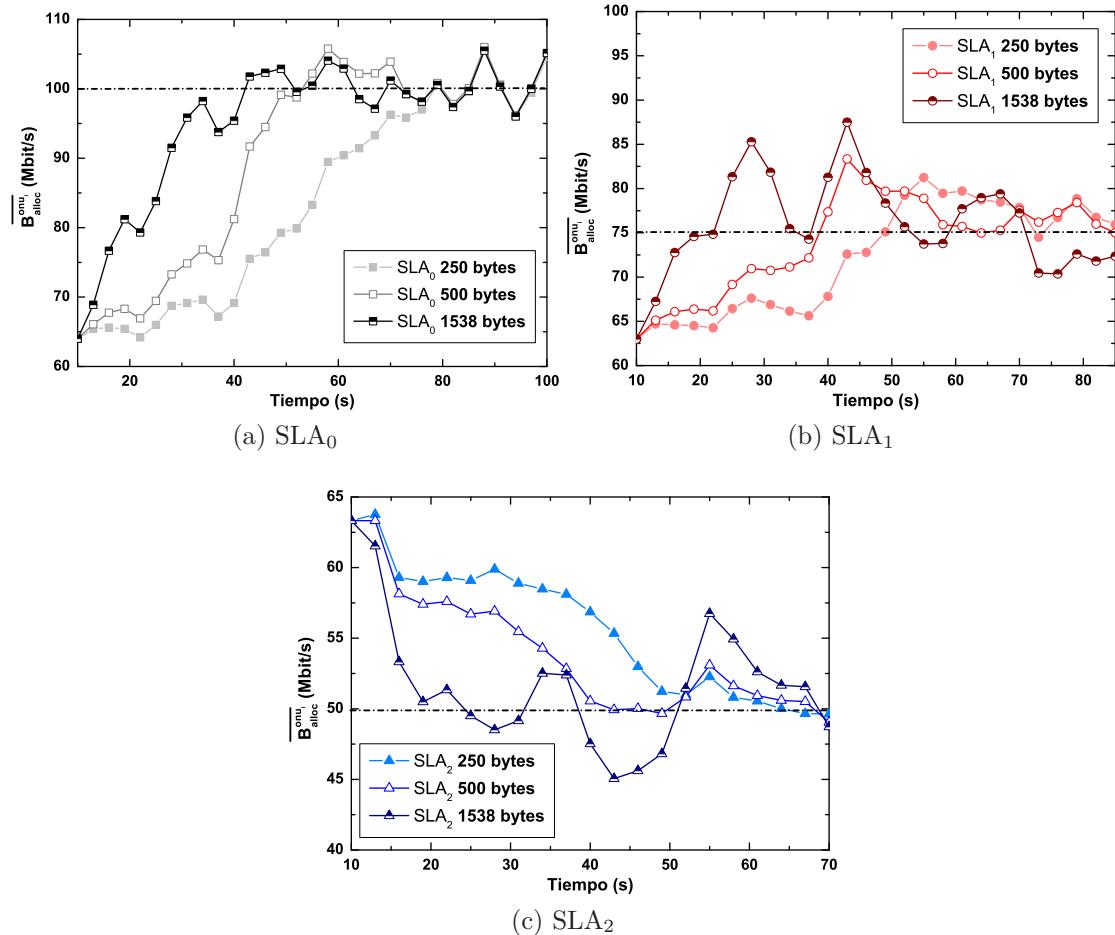


Figura 3.8: Tiempo de convergencia hacia los niveles de ancho de banda garantizados considerando valores pequeños del término $B_{reduction}$. (a) SLA_0 (b) SLA_1 (c) SLA_2 .

asciende hasta los 60 s. Esta diferencia de tiempos puede ser significativa en entornos de tráfico dinámico, escenarios en los que se pretende que el algoritmo reaccione de forma rápida a los cambios de tráfico. Por tanto, el término $B_{reduction}$ debe tener un valor que marque un compromiso entre la velocidad de convergencia y la estabilidad, por lo que se ha escogido el valor intermedio de 1538 *bytes*, que cumple ambas premisas deseadas.

El último parámetro importante involucrado en el diseño de DySGAB es el tiempo entre ejecuciones consecutivas de la adaptación de los máximos, T_{alg} . Este valor no puede ser muy elevado porque la adaptación se haría de forma muy lenta, por lo que durante un gran número de ciclos el ancho de banda asignado no se ajustaría a los niveles garantizados y además tardaría mucho en converger hacia estos valores estipulados. Por otro lado, si la adaptación se realiza con demasiada frecuencia sería más difícil el control de los anchos de banda máximos, ya que se actualizarían con demasiada frecuencia y casi sin obtener resultados fiables de la adaptación realizada en el paso o pasos previos.

Por tanto, se han realizado simulaciones con diferentes valores del término T_{alg} para analizar el comportamiento de DySGAB. La Figura 3.9 (a), (b) y (c) recoge los resultados obtenidos para el SLA₀, SLA₁ y SLA₂, respectivamente, suponiendo cuatro valores diferentes. Como se observa, para los dos tiempos más pequeños (50 ms y 2 s) la fluctuación del ancho de banda es demasiado grande, por lo que el comportamiento para estos casos no es el óptimo. Para 3 s y 10 s, se observan unos niveles de oscilación muy parecidos, por lo que se ha optado por escoger el menor valor, es decir 3 s, para asegurar una adaptación más rápida.

A modo de resumen, la Tabla 3.3 recoge los valores finales escogidos tras este estudio de simulación realizado para determinar los distintos parámetros de DySGAB con el fin de obtener un funcionamiento relativamente óptimo.

Parámetro	Valor
Tamaño ventana (T_{window})	7.5 segundos
Tiempo entre ejecuciones consecutivas del algoritmo (T_{alg})	3 segundos
Ancho de banda a reducir en los reajustes del algoritmo ($B_{reduction}$)	1538 <i>bytes</i>

Tabla 3.3: Valores de los parámetros óptimos relacionados con el diseño del algoritmo DySGAB.

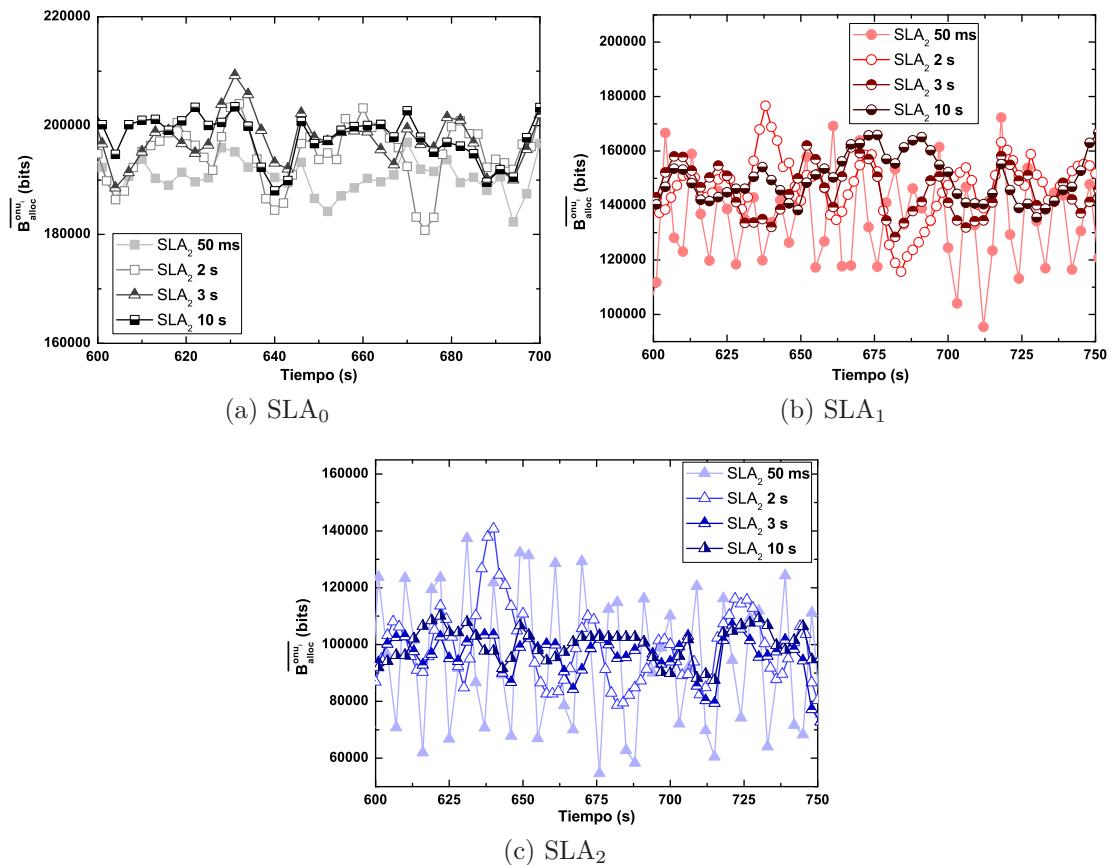


Figura 3.9: Variación en tiempo real del ancho de banda medio para una ONU de cada SLA con distintos valores de T_{alg} . (a) SLA₀ (b) SLA₁ (c) SLA₂.

3.3.3. Análisis de prestaciones de DySGAB

En el siguiente apartado del capítulo, se realiza un estudio del comportamiento de DySGAB bajo distintos escenarios de red, comparándolo con el algoritmo LIPSA. En primer lugar, se van a realizar simulaciones variando las condiciones iniciales de ancho de banda. Con esto se pretende demostrar que DySGAB evoluciona el ancho de banda medio asignado hacia los diferentes niveles garantizados independientemente de los anchos de banda máximos asignados inicialmente. En segundo lugar, se han considerado niveles garantizados distintos a los fijados en la Tabla 3.2 para cada perfil de abonado, con el fin de comprobar que DySGAB es capaz de adaptar los anchos de banda medios a distintos valores. Además, se ha considerado un escenario de red donde estos valores se van a modificar a lo largo del tiempo para analizar así la auto-adaptación que realiza DySGAB cuando los cambios se producen en tiempo real. Para seguir demostrando la independencia de DySGAB en distintos escenarios, también se ha evaluado el comportamiento

miento cuando la distribución del número de ONUs asociadas a cada perfil de abonado es diferente. Por último, se han realizado simulaciones de DySGAB en redes EPON y LR-EAPON de distinta longitud para demostrar que este algoritmo puede ser aplicado en redes con distinta distancia manteniendo sus buenas funcionalidades intactas.

3.3.3.1. Independencia frente a las condiciones iniciales de ancho de banda

Una de las principales características y bondades de DySGAB, que le diferencia de otros algoritmos en LR-EPONs, es la independencia de su comportamiento frente a las condiciones iniciales de ancho de banda asignado, debido al ajuste adaptativo que realiza dinámicamente en tiempo real. Mientras que DySGAB evoluciona hacia los anchos de banda garantizados automáticamente, el resto de algoritmos que realizan diferenciación de clientes mediante el uso de pesos fijos, no están diseñados para ello. Para demostrarlo, se han llevado a cabo simulaciones comparando DySGAB y LIPSA [64] bajo tres escenarios diferentes, donde ambos algoritmos parten de los mismos valores de ancho de banda asignados inicialmente. Por tanto, para establecer los anchos de banda iniciales se ha utilizado la ecuación (3.2) asumiendo diferentes pesos en cada escenario de simulación. En concreto, los pesos asociados a cada uno de los SLAs (W^{sla_j}) en cada escenario están recogidos en la Tabla 3.4.

	W^{sla_0}	W^{sla_1}	W^{sla_2}
Escenario 1	1	1	1
Escenario 2	1	2	3
Escenario 3	2	3	1

Tabla 3.4: Pesos utilizados en los escenarios considerados para establecer los anchos de banda iniciales para los tres perfiles de usuario.

Para comparar el funcionamiento de DySGAB frente a LIPSA, la Figura 3.10 (a), (b) y (c) muestra el ancho de banda asignado a una ONU del SLA_0 , SLA_1 y SLA_2 , respectivamente, para los tres escenarios de la Tabla 3.4 y para todas las cargas de red. En estos tres gráficos, el término $B_{demandado}$ representa el ancho de banda demandado por una ONU de cada SLA, que es una función lineal de 0 a 100 Mbit/s dependiendo de la carga de red. Además, como todas las ONUs tienen la misma distribución de tráfico (tráfico simétrico), todas las ONUs demandan el mismo ancho de banda. Para cada perfil de usuario (SLA), los algoritmos deben ofrecer al menos el mínimo nivel garantizado en caso de que la capacidad en sentido ascendente no cubra la demanda total de ancho de banda. Por el contrario, si se demanda un menor ancho de banda que el nivel garantizado, los algoritmos asignan el ancho de banda demandado.

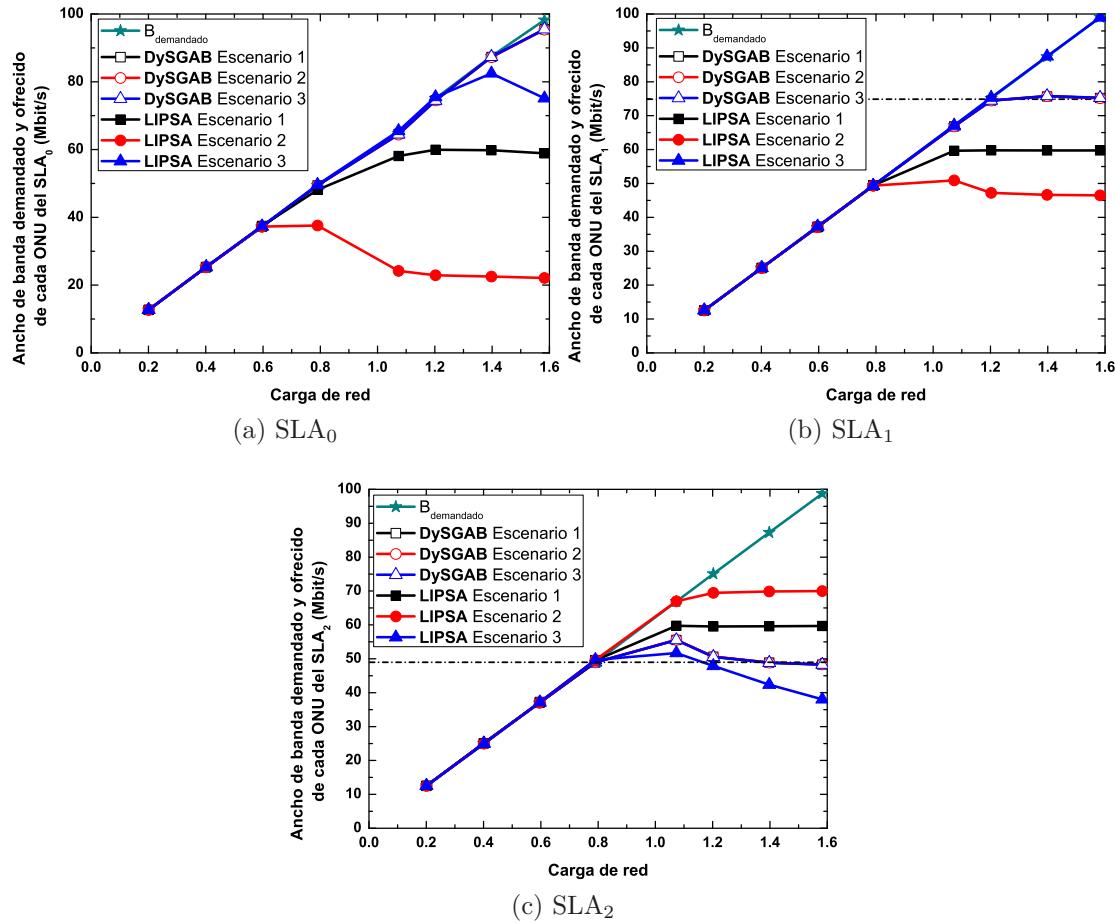


Figura 3.10: Ancho de banda demandado y ofrecido para una ONU de cada SLA bajo los tres escenarios considerados en la Tabla 3.4. (a) SLA₀ (b) SLA₁ (c) SLA₂.

Para el perfil más prioritario (SLA₀), se puede observar en la Figura 3.10 (a) que DySGAB, independientemente de las condiciones iniciales de ancho de banda, ofrece para todas las cargas de red el ancho de banda total demandado para todas las ONUs (≤ 100 Mbit/s). Por el contrario, como en LIPSA los pesos de los tres escenarios no están ajustados de forma adecuada a los valores de ancho de banda estipulados, estos no se garantizan y al ser fijos, no se ajustan dinámicamente a dichos valores. Para el SLA de prioridad media (SLA₁), se observa el mismo comportamiento. Mientras DySGAB converge automáticamente al nivel de ancho de banda garantizado para este perfil (75 Mbit/s) independientemente del escenario considerado, LIPSA muestra una gran dependencia de los pesos iniciales, los cuales no están adecuadamente ajustados en todos los escenarios a los anchos de banda estipulados. De hecho, LIPSA no es capaz de garantizar el ancho de banda para el Escenario 1 y el Escenario 2 para cargas altas

(cuando las ONUs transmiten a más de 50 Mbit/s). Por el contrario, DySGAB garantiza los niveles mínimos estipulados para todos los escenarios en cargas de red donde no se puede dar el ancho de banda demandado a cada ONU; esto es, cargas de red superiores a 1.2 (ONUs transmitiendo a 75 Mbit/s), que conllevan una demanda total de 1.2 Gbit/s ($75 \text{ Mbit/s} \cdot 16 \text{ ONUs}$), un valor superior a la capacidad de 1 Gbit/s del canal ascendente. Por el contrario, cuando las ONUs demandan menos ancho de banda, ambos algoritmos siempre ofrecen dicho ancho de banda. Como puede observarse en la Figura 3.10 (c), los mismos resultados se consiguen para el SLA de menor prioridad (SLA_2). Mientras DySGAB adapta automáticamente el ancho de banda para garantizar el mínimo ancho de banda estipulado para este perfil (50 Mbit/s), LIPSA muestra de nuevo esa gran dependencia de los pesos. Es destacable que LIPSA no es capaz de garantizar el ancho de banda estipulado para el Escenario 3 en cargas de red superiores a 1.2. Por el contrario, DySGAB garantiza el ancho de banda en los tres escenarios cuando los recursos de la red no son suficientes para garantizar la demanda total.

Para ver como se realiza la adaptación instantánea a lo largo del tiempo, las Figuras 3.11, 3.12 y 3.13 muestran la variación en tiempo real del ancho de banda medio asignado a una ONU del SLA_0 , SLA_1 y SLA_2 respectivamente, comparando LIPSA y DySGAB en los tres escenarios considerados. Se ha asumido una carga de red de 1.6 por ser el caso más restrictivo, ya que todas las ONUs están transmitiendo a su tasa máxima de transmisión (100 Mbit/s). Por tanto, si la adaptación de DySGAB se consigue en esta situación límite de la red, también se conseguirá para el resto de cargas menores. Para los usuarios del SLA de mayor prioridad (SLA_0), se puede observar en la Figura 3.11

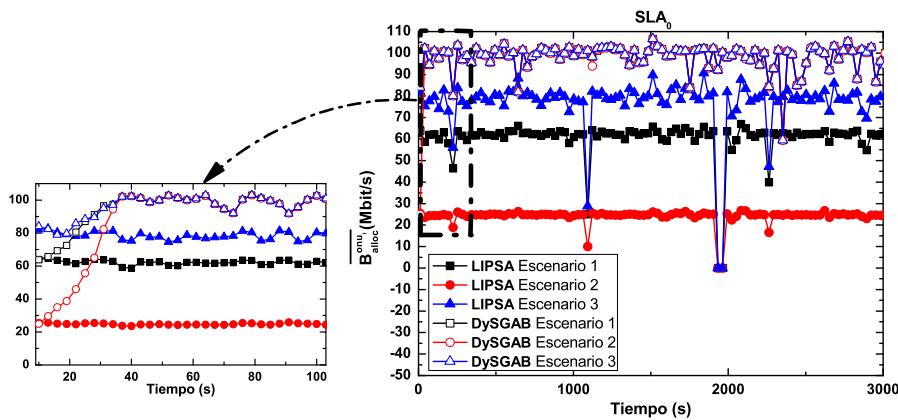


Figura 3.11: Comparación de la evolución en tiempo real del ancho de banda medio en LIPSA y DySGAB considerando diferentes condiciones iniciales de ancho de banda para el SLA_0 .

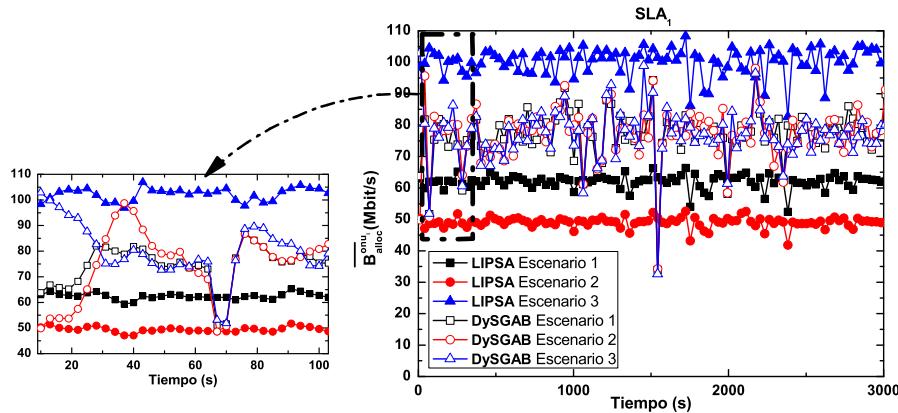


Figura 3.12: Comparación de la evolución en tiempo real del ancho de banda medio en LIPSA y DySGAB considerando diferentes condiciones iniciales de ancho de banda para el SLA₁.

que los valores del ancho de banda medio asignado por DySGAB, con independencia de las condiciones de ancho de banda iniciales, convergen hacia el nivel garantizado del SLA₀, esto es 100 Mbit/s, y luego fluctúan alrededor de él durante el resto del tiempo de simulación. Por el contrario, los valores medios en LIPSA oscilan sobre otros valores que dependen fuertemente de los pesos seleccionados para cada perfil de abonado y que no se corresponden con los niveles que se deben garantizar. Para observar de forma más detallada la adaptación que realiza DySGAB en el tiempo, a la izquierda de esta figura se muestra un zoom de los 100 primeros segundos. Puede verse que la adaptación se realiza con bastante rapidez, ya que a partir de los 40 segundos aproximadamente, los valores en DySGAB han convergido al nivel deseado en los tres escenarios considerados.

Respecto a los usuarios de prioridad intermedia, la Figura 3.12 muestra la evolución en tiempo real del ancho de banda medio asignado por DySGAB y LIPSA. Como se observa, al igual que el caso anterior, DySGAB consigue valores medios que fluctúan alrededor del nivel garantizado para este perfil (75 Mbit/s) con independencia de los pesos iniciales asignados. Por otro lado, los valores en LIPSA oscilan alrededor de otros niveles que dependen de los pesos considerados en el escenario. En el zoom de esta figura (a la izquierda), se representa la adaptación que realiza DySGAB en los primeros segundos de simulación para todos los escenarios considerados. Se puede ver que ambos algoritmos comienzan con los mismos niveles de ancho de banda pero, mientras que los valores en LIPSA no se adaptan, DySGAB consigue la convergencia al nivel deseado en aproximadamente 60 segundos.

Por último, la Figura 3.13, que representa la evolución del ancho de banda medio en tiempo real para los usuarios de menor prioridad, refuerza lo comentado anteriormente.

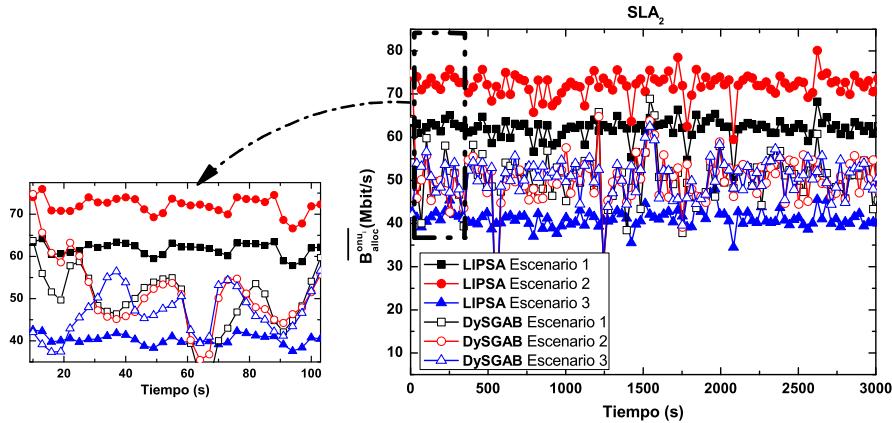


Figura 3.13: Comparación de la evolución en tiempo real del ancho de banda medio en LIPSA y DySGAB considerando diferentes condiciones iniciales de ancho de banda para el SLA₂.

Para el SLA₂, DySGAB, al contrario que LIPSA, también consigue adaptar los anchos de banda asignados al nivel requerido (50 Mbit/s) para todos los escenarios considerados. En el zoom de los 100 primeros segundos representado a la izquierda de dicha figura, se comprueba que la adaptación también se produce para estos usuarios, aunque en este caso los valores convergen algo más despacio que en los dos casos anteriores.

Por tanto, mientras en DySGAB los pesos iniciales no son significativos para una adaptación adecuada en tiempo real a los requisitos de calidad de servicio establecidos, si el proveedor de servicios está utilizando LIPSA y no escoge unos pesos adecuados a los niveles estipulados, éstos no se garantizarán nunca, por lo que se deberá interrumpir el servicio y calcular de nuevo unos pesos adecuados al nuevo escenario de red considerado.

3.3.3.2. Independencia frente a los niveles de ancho de banda garantizados para cada perfil de abonado

Puesto que DySGAB ha sido diseñado para reajustarse a cualquier nivel de ancho de banda garantizado, se ha ampliado el estudio de simulación modificando los valores estipulados para cada SLA, y así evaluar de forma más completa sus prestaciones. En concreto, se han llevado a cabo simulaciones con los parámetros del escenario inicial (Tabla 3.2), con pesos iniciales de $W^{sla_0} = 1$, $W^{sla_1} = 1$ y $W^{sla_2} = 1$, pero considerando dos escenarios cuyos niveles de ancho de banda garantizado para cada SLA aparecen recogidos en la Tabla 3.5.

Bajo estos dos escenarios, se ha estudiado la evolución en tiempo real del ancho de banda de una ONU de cada SLA para la carga de red más elevada (ONUs transmitiendo

	SLA₀	SLA₁	SLA₂
Escenario 1	90 Mbit/s	70 Mbit/s	55 Mbit/s
Escenario 2	100 Mbit/s	80 Mbit/s	40 Mbit/s

Tabla 3.5: Niveles de ancho de banda garantizado para cada SLA en los dos escenarios considerados.

a 100 Mbit/s). La Figura 3.14 (a) y (b) representa dicha evolución para los dos escenarios considerados. Como lo que se quiere demostrar es la adaptación de DySGAB a diferentes niveles garantizados, en este caso no aparece representado LIPSA, al no realizar este algoritmo ninguna adaptación en tiempo real.

Tal y como se aprecia en la Figura 3.14 (a), para el Escenario 1 los valores del ancho de banda medio asignado a cada ONU oscilan sobre sus niveles garantizados. Para este escenario no existe ancho de banda en exceso disponible, puesto que para asegurar el ancho de banda garantizado a todas las ONUs son necesarios 990 Mbit/s ($90 \text{ Mbit/s} + 5 \cdot 70 \text{ Mbit/s} + 10 \cdot 55 \text{ Mbit/s} = 990 \text{ Mbit/s}$), valor que se aproxima mucho a la capacidad total del canal (1 Gbit/s). Por tanto, para el SLA₂ los valores en algunos momentos son ligeramente inferiores a los 55 Mbit/s.

A diferencia del Escenario 1, para el Escenario 2 (Figura 3.14 (b)) existe ancho de banda extra disponible para repartir a mayores del ancho de banda garantizado, ya que la suma del ancho total a garantizar de todas las ONUs es 900 Mbit/s ($100 \text{ Mbit/s} + 5 \cdot 80 \text{ Mbit/s} + 10 \cdot 40 \text{ Mbit/s} = 900 \text{ Mbit/s}$), cantidad menor que la capacidad del canal (1 Gbit/s). Por ello, a las ONUs del SLA₀ se les asigna todo el ancho de banda demandado

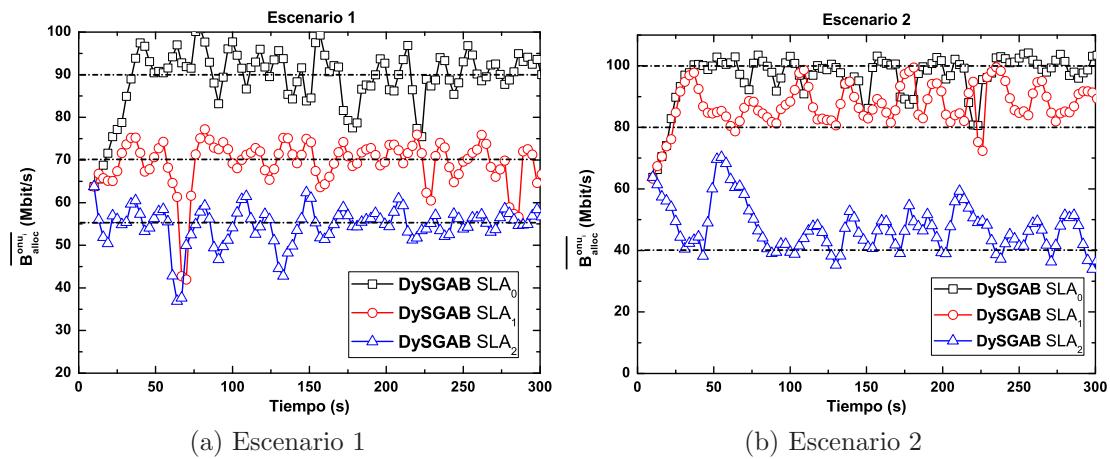


Figura 3.14: Evolución en tiempo real del ancho de banda medio asignado por DySGAB para una ONU de cada SLA y para carga de red de 1.6 considerando los dos escenarios de la Tabla 3.5. (a) Escenario 1 (b) Escenario 2.

(que coincide con el garantizado), mientras que las ONUs del SLA₁ y SLA₂ reciben un ancho de banda mayor que su nivel garantizado, ya que como se advierte, los valores fluctúan por encima de estos niveles (80 y 40 Mbit/s respectivamente).

3.3.3.3. Adaptación a niveles de ancho de banda garantizados que varían en tiempo real

Por otro lado, la auto-adaptabilidad que presenta DySGAB es muy adecuada cuando los proveedores de servicio requieran cambios en tiempo real de los anchos de banda garantizados de algunos o todos los perfiles. La adaptación automática que realiza DySGAB permite un reajuste en tiempo real sin interrupciones en el servicio de red. Esto supone una gran ventaja frente a otros algoritmos que utilizan pesos fijos, ya que no sería necesario interrumpir el servicio para recalcular nuevos pesos que se ajusten a los nuevos niveles garantizados. Para demostrarlo, se han realizado simulaciones asumiendo que cada 1000 segundos el proveedor cambia los niveles de ancho de banda garantizado a cada perfil de abonado. La Tabla 3.6 recoge los valores considerados para los anchos de banda garantizados de los tres SLAs a lo largo del tiempo. En la Figura 3.15 (a), (b) y (c) está representada la evolución en tiempo real del ancho de banda medio asignado a una de las ONUs del SLA₀, SLA₁ y SLA₂, respectivamente, y para la carga máxima de red (1.6). En este caso, no está representado el comportamiento del algoritmo LIPSA, ya que como no realiza ninguna adaptación mantendría siempre un nivel constante. Sin embargo, tal y como se puede ver, en DySGAB el ancho de banda medio de cada uno de los SLAs se adapta de forma rápida y eficiente a los nuevos niveles de ancho de banda garantizados en cada intervalo temporal.

Tiempo (s)	Nivel de ancho de banda garantizado a cada SLA		
	SLA ₀	SLA ₁	SLA ₂
0-1000	100 Mbit/s	75 Mbit/s	50 Mbit/s
1000-2000	70 Mbit/s	100 Mbit/s	40 Mbit/s
2000-3000	90 Mbit/s	70 Mbit/s	55 Mbit/s
3000-4000	60 Mbit/s	90 Mbit/s	45 Mbit/s

Tabla 3.6: Niveles de ancho de banda garantizado considerados para cada SLA a lo largo del tiempo.

Por tanto, DySGAB es capaz de adaptar automáticamente el ancho de banda de forma que los valores evolucionan hasta los niveles de ancho de banda garantizado de forma rápida y estable. Por consiguiente, la adaptación llevada a cabo por DySGAB es muy útil para posibles cambios hechos por los proveedores de servicio, tanto modificaciones iniciales como variaciones en tiempo real. Por tanto, el algoritmo diseñado DySGAB, es

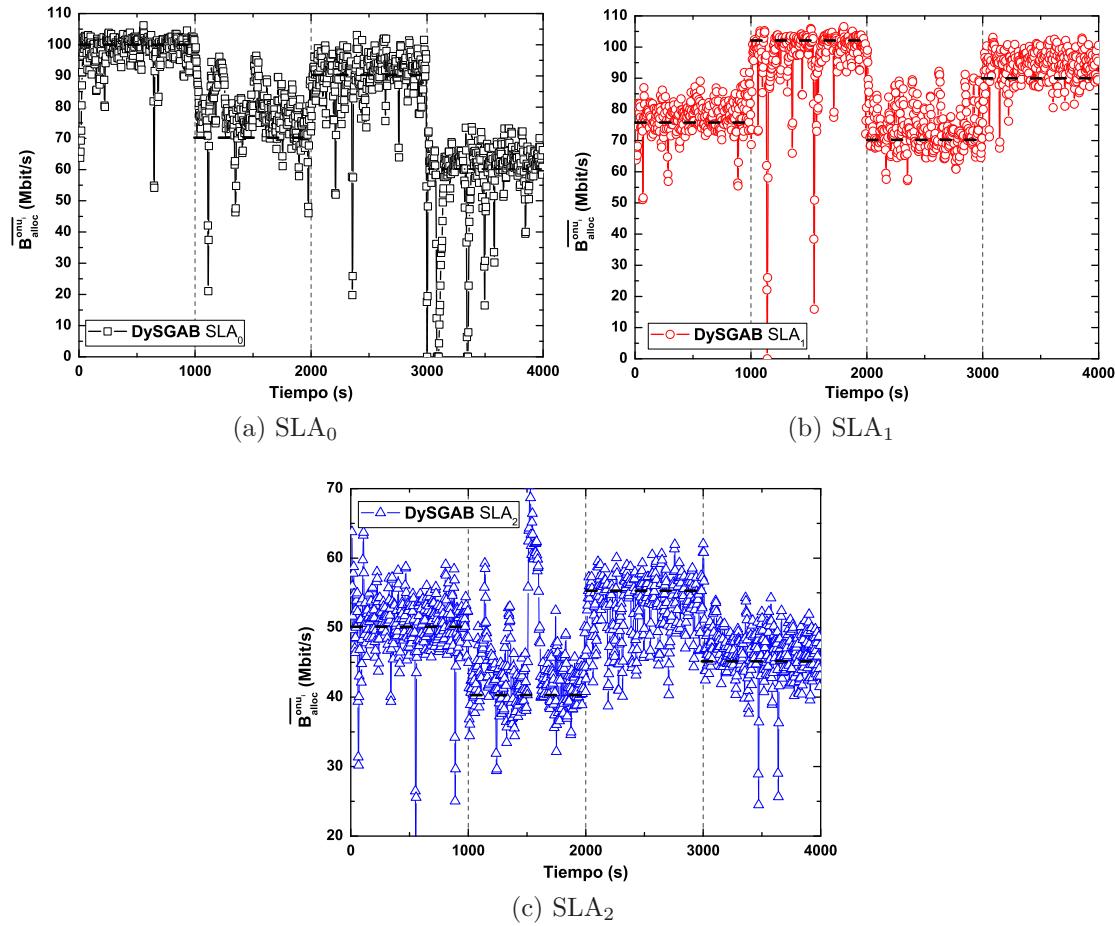


Figura 3.15: Evolución a lo largo del tiempo del ancho de banda medio asignado a una ONU de cada SLA en DySGAB cuando varían los anchos de banda garantizados en tiempo real. (a) SLA₀ (b) SLA₁ (c) SLA₂.

el primer algoritmo en la literatura que consigue esta funcionalidad gracias a la auto-adaptación en tiempo real que éste realiza de una forma relativamente óptima, rápida y estable.

3.3.3.4. Análisis de DySGAB considerando distintas distribuciones de ONUs en cada perfil de abonado

Desde el punto de vista de un proveedor de servicios, los cambios en la distribución de las ONUs asociadas a cada perfil de abonado también puede ser algo que sea necesario realizar. Por ello, la auto-adaptación que realiza DySGAB también es muy conveniente para estas situaciones, puesto que no se necesitan recalcular los pesos ni interrumpir el

servicio para reajustar los recursos al nuevo estado de la red.

Para analizar el comportamiento de DySGAB se han considerado dos nuevos escenarios siguiendo las características del escenario inicial (Tabla 3.2) pero modificando el número de ONUs que contratan cada uno de los SLAs según la Tabla 3.7.

	ONUs del SLA ₀	ONUs del SLA ₁	ONUs del SLA ₂
Escenario 1	2	6	8
Escenario 2	2	3	11

Tabla 3.7: Distribución de ONUs en los nuevos escenarios de red considerados.

La Figura 3.16 (a) y (b) muestra el ancho de banda asignado a una ONU de cada SLA para todas las cargas red, comparando LIPSA y DySGAB en los dos escenarios. Para LIPSA, los pesos asociados a cada uno de los SLAs (W^{sla_j}) para establecer los anchos de banda iniciales son los usados en [64] para ajustar adecuadamente a los niveles garantizados de 100, 75 y 50 Mbit/s. En concreto estos valores son, $W^{sla_0} = 4$, $W^{sla_1} = 3$ y $W^{sla_2} = 2$.

En el caso del Escenario 1 (Figura 3.16 (a)), se observa que para cargas de red por encima de 1.2, ambos algoritmos no son capaces de asegurar los niveles de ancho de banda garantizados para todos los usuarios. Esto es así porque por ejemplo para la carga máxima de red de 1.6, el ancho de banda total que habría que garantizar sería de 1.05 Gbit/s (2·100 Mbit/s + 6·75 Mbit/s + 8·50 Mbit/s = 1050 Mbit/s), que es un valor superior a la tasa del canal (1 Gbit/s). En el caso de DySGAB, como ha sido diseñado para seguir una política agresiva cuando no se puede garantizar el ancho de banda de

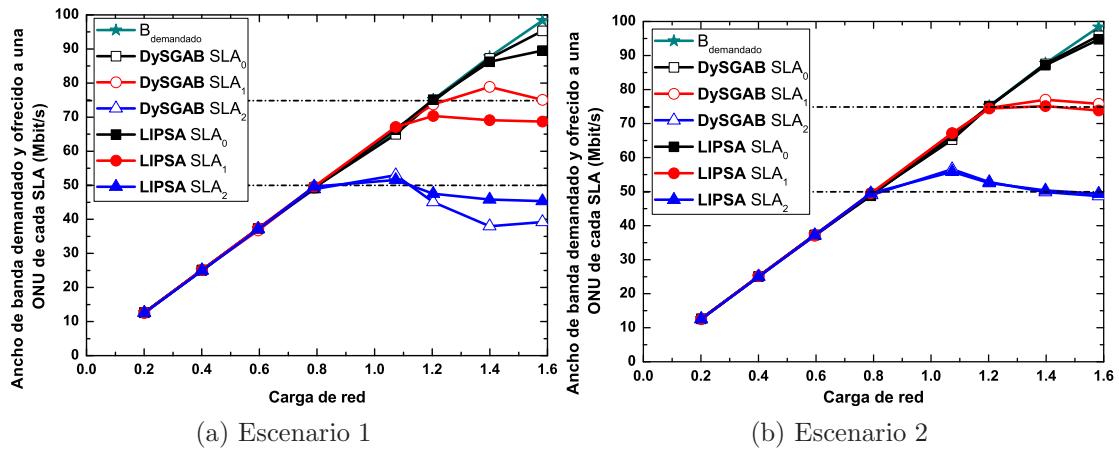


Figura 3.16: Ancho de banda demandado y ofrecido a una ONU de cada SLA comparando DySGAB y LIPSA frente la carga de red para los escenarios de la Tabla 3.7. (a) Escenario 1 (b) Escenario 2.

todos los usuarios, se reducen los anchos de banda máximos de las ONUs del SLA de menor prioridad (SLA_2) para que los SLAs más prioritarios cumplan con los requisitos estipulados (SLA_0 , SLA_1). De esta forma, las ONUs del SLA_0 y SLA_1 cumplen con sus niveles mientras que las del SLA_2 están unos 10 Mbit/s por debajo de los 50 Mbit/s que hay que garantizar. Por otro lado LIPSA, que no sigue esta política, no consigue garantizar el ancho de banda a ningún tipo de usuario, aunque mantiene los niveles de ancho de banda por encima de DySGAB para el perfil de menor prioridad (SLA_2).

A diferencia del caso anterior, en el Escenario 2 (Figura 3.16 (b)), se observa como sí que es posible garantizar los niveles de ancho de banda estipulados por el proveedor para todos los usuarios, por lo que ambos algoritmos cumplen con las especificaciones de QoS requeridas.

3.3.3.5. Análisis de DySGAB considerando redes de distinta longitud

Por último, es importante destacar que DySGAB es perfectamente aplicable a redes EPON y LR-EPON con diferente rango de cobertura. Por tanto, ha sido diseñado para ofrecer características de adaptabilidad e independencia bajo distintos escenarios en redes LR-EPON con diferentes distancias. Para demostrar esto, la Figura 3.17 representa el ancho de banda demandado y ofrecido para una ONU de cada SLA cuando se consideran distancias desde 20 km hasta 120 km en saltos de 20 km. Los anchos de banda garantizados se han establecido a 100/75/50 Mbit/s para los perfiles SLA_0 , SLA_1 y SLA_2 , respectivamente. Además se han considerado pesos iguales a 1 para establecer

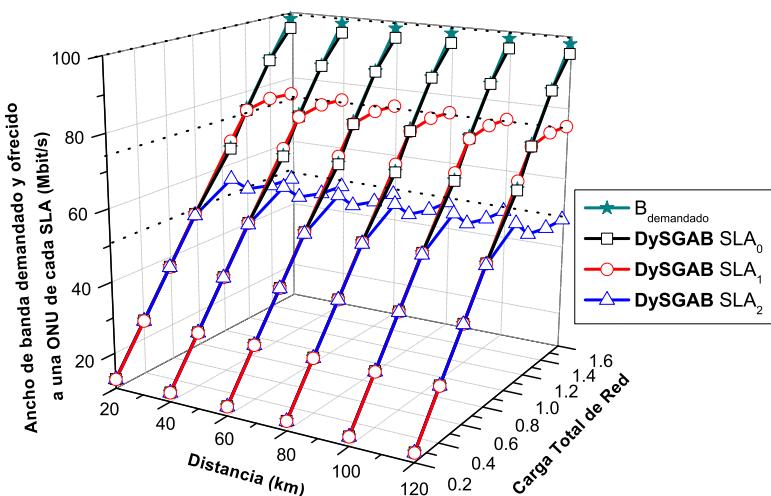


Figura 3.17: Ancho de banda demandado y ofrecido a una ONU de cada SLA para redes EPON y LR-EPON de distinta distancia.

las condiciones iniciales. Tal y como se observa, para todas las distancias consideradas, los resultados demuestran un comportamiento idéntico de DySGAB independientemente de la longitud de la red LR-EPON y EPON. Por tanto, los anchos de banda garantizados de todos los perfiles se aseguran para todas las distancias cuando la capacidad del canal en sentido ascendente no es suficiente para cubrir la demanda total en la red. Además para cargas bajas de red, donde la capacidad cubre la demanda total, DySGAB asigna todo el ancho de banda demandado en todas las distancias consideradas.

Por lo tanto, todas las ventajas que ofrece el algoritmo desarrollado inicialmente en redes LR-EPON de 100 km, gracias a su adaptación, pueden ser aprovechadas en redes LR-EPON y EPON con distinta cobertura.

3.4. Análisis de prestaciones de DySGAB bajo patrones de tráfico asimétrico y dinámico

En el siguiente apartado del capítulo se van a analizar las prestaciones del algoritmo DySGAB bajo condiciones de tráfico más reales, tales como tráfico asimétrico y tráfico dinámico. De esta forma se pretende demostrar la robustez de DySGAB ante diferentes patrones de tráfico más acordes al contexto de una red real.

3.4.1. Escenarios con un patrón de tráfico asimétrico

En una red de acceso real existen distintos tipos de usuarios que demandan diferentes servicios y en diferentes proporciones. Por ejemplo, la tasa de información generada por una gran empresa donde se haga un uso extensivo de videoconferencias, teletrabajo o de redes privadas virtuales (VPN, *Virtual Private Networks*), no es comparable con la que puede generar un usuario convencional en su hogar.

Por tanto, la consideración de un escenario simétrico donde todos los usuarios generan paquetes a la misma tasa supone una simplificación del modelo de red de acceso. El hecho de considerar tráfico asimétrico o con carga no balanceada, donde los usuarios no contribuyen por igual a la carga total de red, es más realista. Por ello, y para demostrar la robustez de DySGAB ante otros patrones de tráfico más reales, se han analizado sus prestaciones considerando en primer lugar, un patrón de tráfico asimétrico.

Para ello, se han realizado simulaciones con las características del escenario inicial de la Tabla 3.2, pero considerando que cada ONU transmite a una tasa aleatoria entre 0 y 100 Mbit/s (máxima tasa de transmisión) y limitando la contribución de cada SLA a la carga total de red. En concreto, se han considerado dos escenarios con carga de red

distinta, donde la contribución de cada SLA a la carga total de red aparece recogida en la Tabla 3.8.

	Carga de red	SLA ₀	SLA ₁	SLA ₂
Escenario 1	1.0	10 %	40 %	50 %
Escenario 2	1.2	4 %	30 %	66 %

Tabla 3.8: Contribución de cada SLA a la carga total de red para los distintos escenarios considerados con tráfico asimétrico.

Así pues, la Figura 3.18 muestra el ancho de banda medio asignado a cada ONU para las condiciones de red del Escenario 1. En este escenario, la carga de total es de 1.0, lo que implica que el ancho de banda demandado por todas las ONUs es de 1 Gbit/s. Además, bajo este escenario se considera que el 10 % del tráfico es generado por la ONU del SLA₀, el 40 % lo generan de forma aleatoria las cinco ONUs del SLA₁ y el 50 % restante se reparte aleatoriamente entre las diez ONUs del SLA₂. Tal y como puede observarse en la Figura 3.18, el ancho de banda demandado por cada ONU (representado en la figura por una estrella) es diferente, mostrando esta asimetría en el tráfico generado por las diferentes ONUs. Sin embargo, independientemente del ancho de banda demandado por las ONUs, DySGAB asegura el mínimo ancho de banda garantizado estipulado por el proveedor de servicios para aquellas ONUs cuya demanda supera dicho valor. Incluso, en algunos casos DySGAB ha asignado a estas ONUs un valor mayor de su nivel garantizado, ya que existía ancho de banda en exceso. Finalmente, se observa que DySGAB asigna

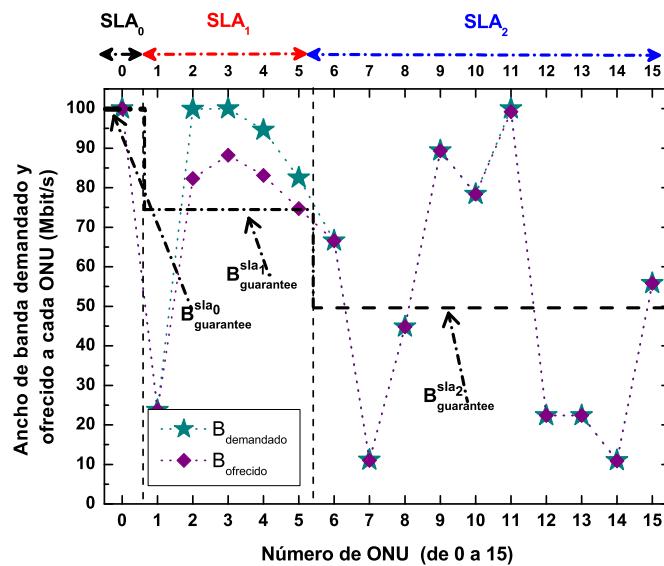


Figura 3.18: Ancho de banda demandado y ofrecido a cada ONU de la red LR-EPON bajo las condiciones del Escenario 1 de la Tabla 3.8.

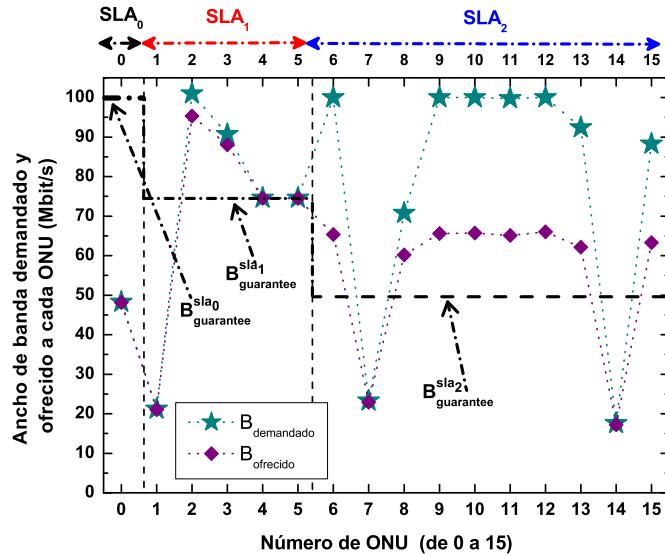


Figura 3.19: Ancho de banda demandado y ofrecido a cada ONU de la red LR-EPON bajo las condiciones del Escenario 2 de la Tabla 3.8.

todo el ancho de banda demandado a aquellas ONUs que presentan una demanda menor de su nivel mínimo garantizado.

De forma similar, la Figura 3.19, muestra el ancho de banda demandado y ofrecido por DySGAB a cada ONU en las condiciones de tráfico del Escenario 2. En este caso, la carga de red total del escenario es de 1.2, lo que implica que la demanda total de las ONUs es de 1.20 Gbit/s. Además, en este escenario el 4% del tráfico es generado por la ONU del SLA₀, el 30% se genera aleatoriamente por las cinco ONUs del SLA₁ y el 66% restante entre las diez ONUs del SLA₂. Al igual que para el Escenario 1, en este escenario DySGAB es capaz de garantizar el ancho de banda mínimo para aquellas ONUs cuya demanda es mayor que dicho nivel mínimo y de forma simultánea asigna el ancho de banda demandado para aquellas ONUs cuya demanda es menor que su nivel mínimo garantizado.

Por tanto, se demuestra que DySGAB mantiene un buen comportamiento y ofrece buenas prestaciones bajo un patrón de tráfico asimétrico, lo cual demuestra su robustez ante escenarios de red con una distribución de tráfico más realista.

3.4.2. Escenarios con un patrón de tráfico dinámico

Otro escenario de red más realista en una red LR-EPON es el que considera tráfico dinámico, ya que los usuarios no demandan la misma cantidad de tráfico siempre, sino que ésta varía dependiendo de las horas del día. Para modelar este tipo de tráfico se

ha seguido el modelo de tráfico dinámico propuesto por Gençata y Mukherjee [88]. Este modelo considera que el tráfico entre dos nodos va a fluctuar con el tiempo. Estas fluctuaciones pueden ser variaciones de periodo corto, de periodo largo o abruptas. Las variaciones abruptas son las menos comunes y aparecen cuando la demanda cambia de forma repentina (por ejemplo ante un fallo de red). Las variaciones de periodo corto aparecen de forma continua, son del orden de segundos y están causadas por la naturaleza rafagosa del tráfico. Por último, las de periodo largo cambian el tráfico de forma suave a lo largo del tiempo, dependiendo de la hora del día en la que nos situemos [89]. Para modelar una carga de red que varía con el tiempo según este modelo de tráfico se va a seguir la ecuación (3.8) [89].

$$\lambda^{s,d}(t) = \Lambda^{s,d} \cdot \beta(t) \cdot \left[1 + \frac{1}{2} \sin \left(2\pi \left(\frac{t + \delta_s \cdot 3600}{86400} \right) \right) \right] \quad (3.8)$$

En esta ecuación, el término $\Lambda^{s,d}$ se corresponde con la media de tráfico entre dos nodos en el tiempo de variación de ciclo largo (un día, 86400 s). El término $\beta(t)$ es una función que se emplea para crear efecto rafagoso. En general, el valor de dicha función se genera de forma aleatoria cada segundo en el intervalo $[1 - \epsilon, 1 + \epsilon]$, donde ϵ es una constante que varía entre 0 y 1 y que determina el grado de rafagosidad del tráfico. El término δ_s es un número que representa la zona horaria del nodo origen, para conseguir que nodos en diferentes zonas horarias generen sus máximos valores de tráfico en diferentes tiempos. Por último, la función seno es la que genera las variaciones de periodo largo [89].

Para llevar a cabo las simulaciones en la red LR-EPON con un patrón de tráfico dinámico, se ha supuesto una carga media igual en todas las ONUs de 0.6 (ONUs transmitiendo a una tasa de 60 Mbit/s). En este caso, no se ha considerado ni rafagosidad en el tráfico ($\beta(t)$ siempre igual a 1.0) ni variaciones en la zona horaria ($\delta_s = 1$). Además, el tiempo de ciclo largo considerado no es un día completo, sino 6 horas (21600 s) para simplificar el estudio de simulación. Esto quiere decir, que en 6 horas la carga de cada ONU ha cumplido con un ciclo completo del seno, y por tanto ha fluctuado desde el nivel mínimo hasta el máximo. Por tanto, la ecuación utilizada para modelar la variación de carga en las ONUs de la red LR-EPON, ρ_{onu} , es la ecuación (3.9).

$$\rho_{onu}(t) = 0.6 \cdot \left[1 + \frac{1}{2} \sin \left(2\pi \left(\frac{t + 3600}{21600} \right) \right) \right] \quad (3.9)$$

Además, se han escogido dos valores distintos para simular la frecuencia con la que se cambia el valor de la carga dinámica, para así comprobar la reacción del algoritmo DySGAB ante cambios más rápidos y más lentos de dicho patrón de tráfico. En concreto,

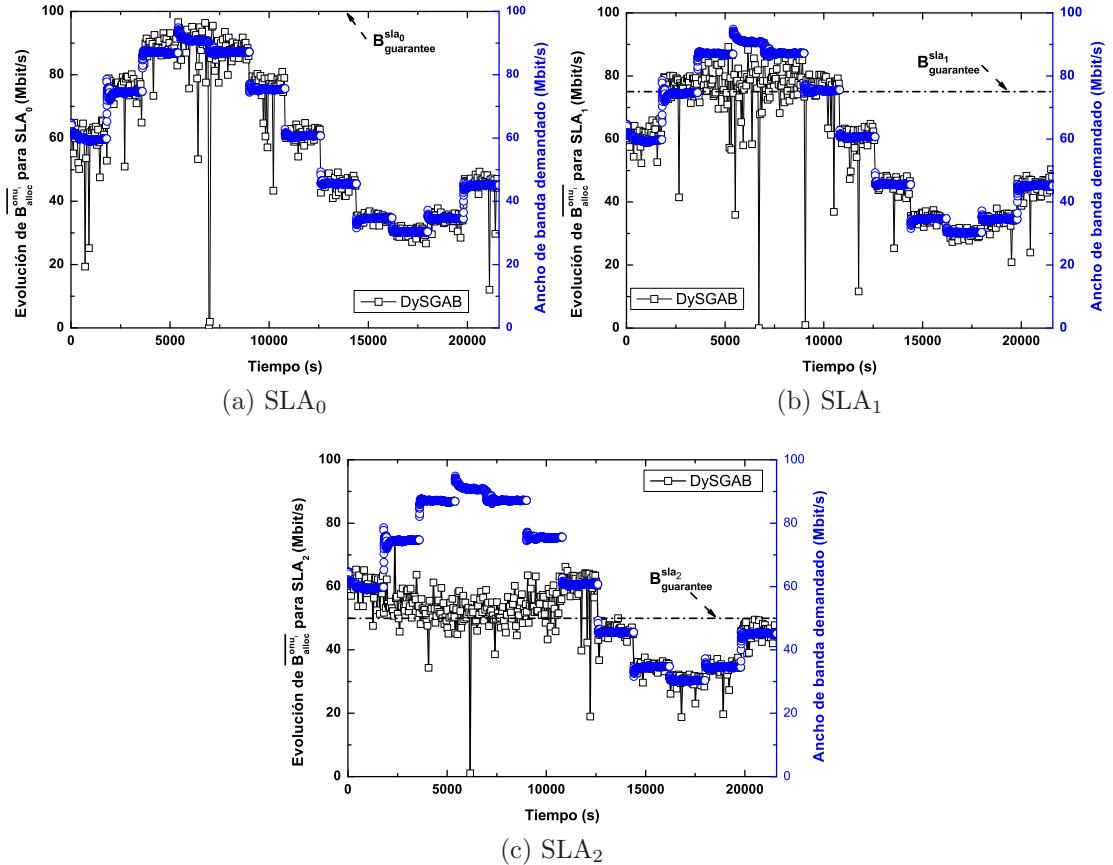


Figura 3.20: Ancho de banda demandado para una ONU de cada SLA (azul) y variación en tiempo real del ancho de banda medio asignado (negro) considerando una frecuencia de cambio de 1800 s. (a) SLA₀ (b) SLA₁ (c) SLA₂.

se han escogido los valores de 1800 s y de 900 s.

La Figura 3.20 (a), (b) y (c) ilustra la variación del ancho de banda medio asignado a lo largo del tiempo para una ONU del SLA₀, SLA₁ y SLA₂, respectivamente, cuando la frecuencia de los cambios en la carga es de 1800 s. En azul está representada la variación en la carga de ONU en Mbit/s, referida al eje de la derecha también en azul (de 0-100 Mbit/s). En negro aparece la variación en tiempo real del ancho de banda asignado por DySGAB (eje de la izquierda). Tal y como puede observarse, para los tres SLAs, el ancho de banda medio asignado se corresponde con el ancho de banda demandado, siempre que existan recursos para satisfacer la demanda total. En el caso de que la demanda total del ancho de banda sea mayor que la capacidad del canal, DySGAB asigna el nivel mínimo garantizado a cada SLA (100, 75 y 50 Mbit/s respectivamente). Esto ocurre cuando el ancho de banda demandado por cada ONU es mayor o igual a

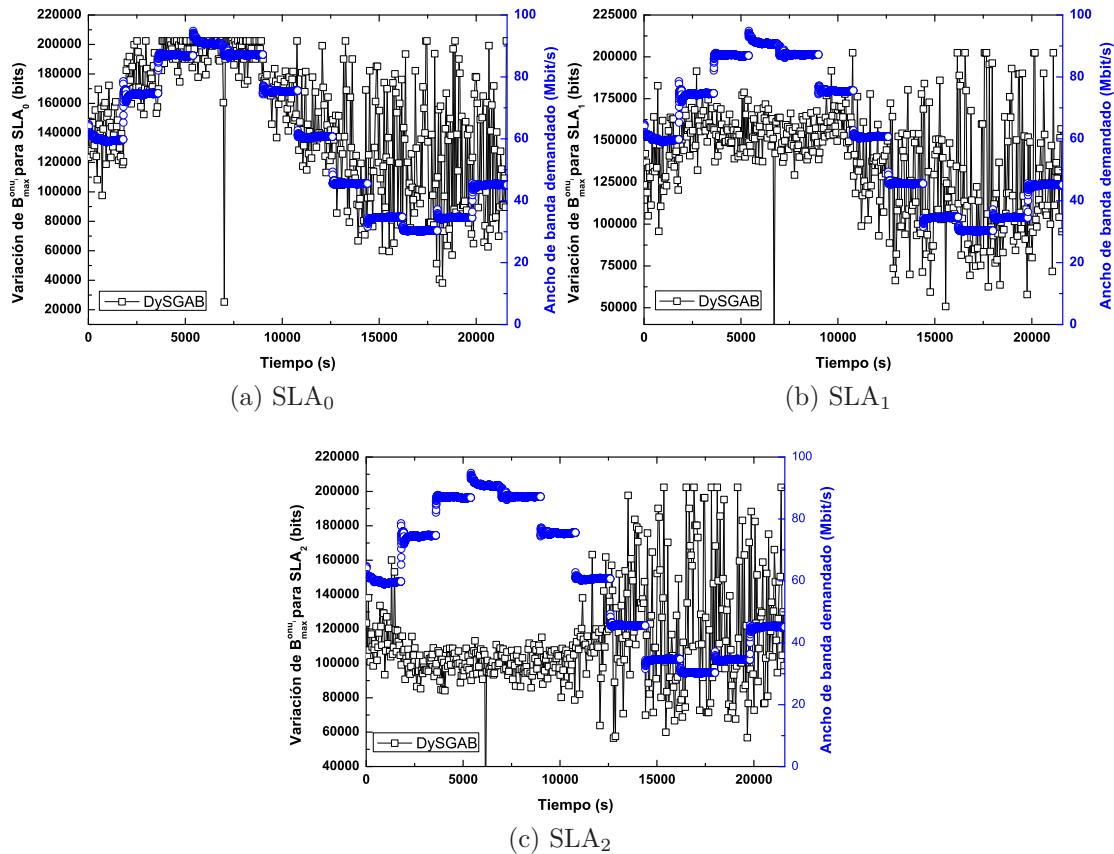


Figura 3.21: Ancho de banda demandado para una ONU de cada SLA (azul) y variación en tiempo real del ancho de banda máximo (negro) considerando una frecuencia de cambio de 1800 s. (a) SLA₀ (b) SLA₁ (c) SLA₂.

62.5 Mbit/s aproximadamente, en cuyo caso la demanda total ($62.5 \text{ Mbit/s} \cdot 16 = 1 \text{ Gbit/s}$) es igual o superior a la capacidad de 1 Gbit/s del canal ascendente.

La Figura 3.21 (a) (b) y (c) representa la variación del ancho de banda máximo (B_{max}^{onui}) para el SLA₀, SLA₁ y SLA₂, respectivamente, para una frecuencia de cambio de 1800 s. Al igual que en la figura anterior, en azul está representada la variación de la carga de ONU (referida al eje de la derecha también en azul). En negro aparece la variación dinámica del ancho de banda máximo (eje de la izquierda). Tal y como puede observarse, en la zona de mayor carga, la variación de los máximos es más estable, ya que es en esta zona donde al ancho de banda asignado se ajusta adecuadamente a los niveles que DySGAB debe garantizar a cada perfil de abonado. En la zona de cargas inferiores a 0.6 se observa una mayor fluctuación de los máximos. Esto es debido a que como existe ancho de banda en exceso para cubrir toda la demanda, los valores máximos

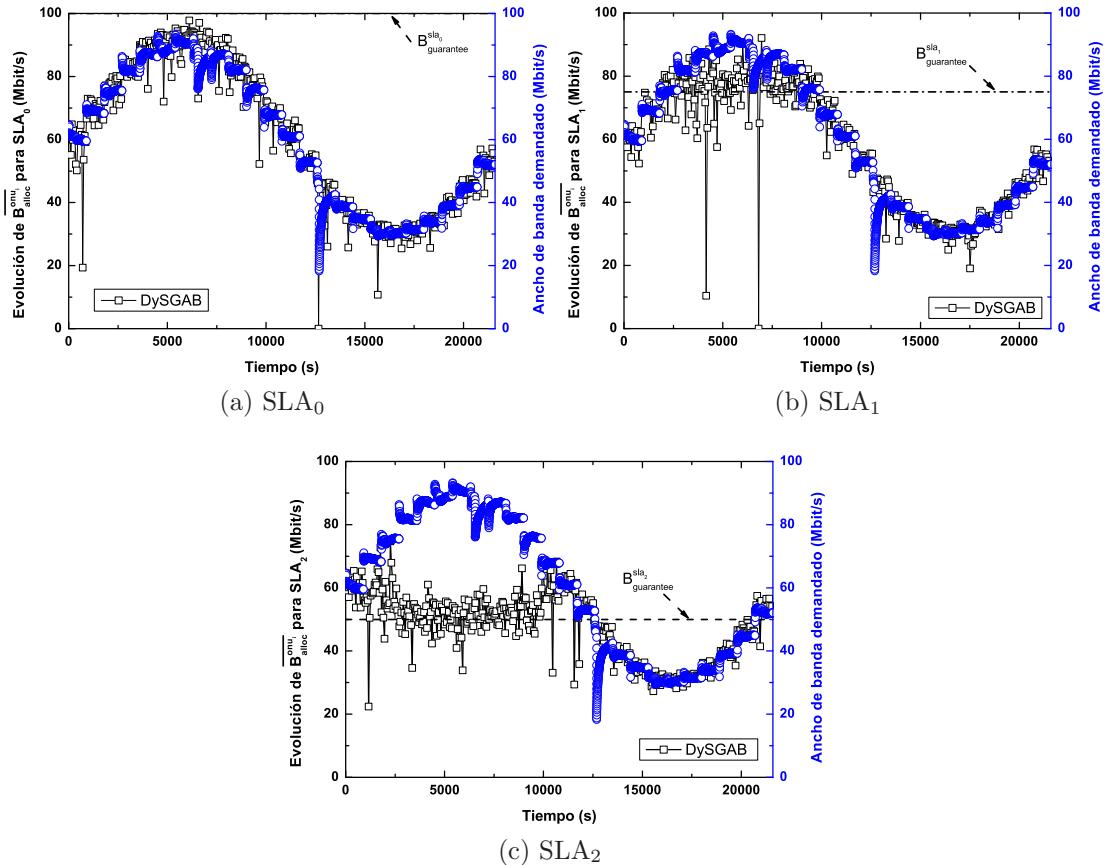


Figura 3.22: Ancho de banda demandado para una ONU de cada SLA (azul) y variación en tiempo real del ancho de banda medio asignado (negro) considerando una frecuencia de cambio de 900 s. (a) SLA $_0$ (b) SLA $_1$ (c) SLA $_2$.

se modifican en mayor grado. Además, aunque para estas cargas la media del tráfico de las ONUs es inferior a los niveles garantizados, en ciertos instantes las ONUs pueden tener una demanda mayor que dicho nivel debido a la naturaleza rafagosa del tráfico, lo que provoca el reajuste de los máximos. No obstante, esto no tiene ninguna implicación en las prestaciones de DySGAB, ya que como se observó en la Figura 3.20, para estas cargas se asigna todo el ancho de banda demandado por las ONUs.

Por otro lado, la Figura 3.22 y la Figura 3.23 representan la variación del ancho de banda medio y del ancho de banda máximo (representadas en negro) respectivamente, según la variación de la carga de ONU (representada en azul), considerando una frecuencia de cambio de 900 s. Al igual que en el caso anterior, DySGAB asigna todo el ancho de banda demandado siempre que existan recursos suficientes. Cuando la demanda total no permite asignar todo el ancho de banda demandado, DySGAB asegura adecuadamente

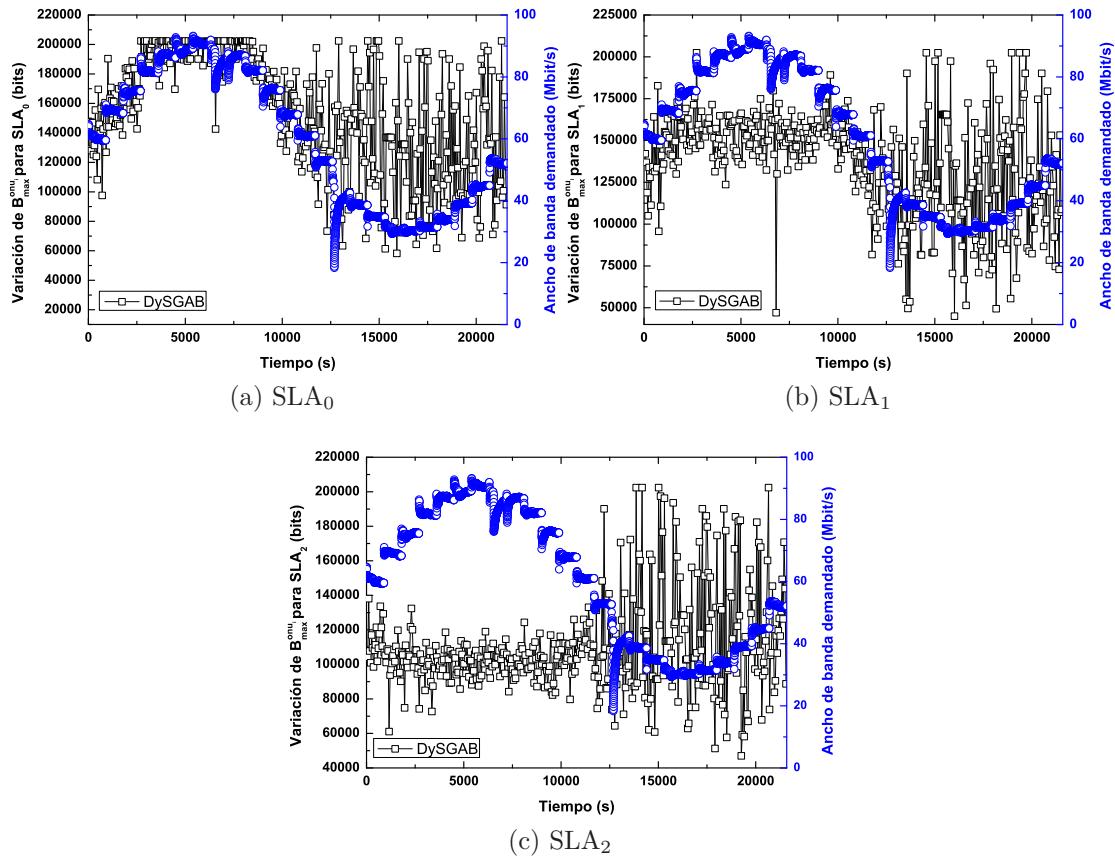


Figura 3.23: Ancho de banda demandado para una ONU de cada SLA (azul) y variación en tiempo real del ancho de banda máximo (negro) considerando una frecuencia de cambio de 900 s. (a) SLA₀ (b) SLA₁ (c) SLA₂.

el nivel garantizado establecido por el proveedor de servicios. Además, la variación del ancho de banda máximo es similar a la obtenida para una frecuencia de cambio de 1800 s.

Por tanto, puede concluirse que DySGAB es capaz de garantizar el ancho de banda estipulado por el proveedor de servicios bajo escenarios con condiciones de tráfico más reales, tales como tráfico dinámico, independientemente de la frecuencia de los cambios de la carga en el patrón de tráfico.

3.5. Conclusiones

En este capítulo de la Tesis se ha presentado un novedoso algoritmo de asignación de ancho de banda para redes LR-EAPON denominado DySGAB (*Dynamic Service levels with GuAranteed Bandwidth*). Este algoritmo realiza un control dinámico de la asignación de ancho de banda de forma que todos los usuarios tengan asegurado los niveles de ancho de banda mínimos estipulados por el proveedor de servicios según el perfil contratado. De esta forma, DySGAB realiza un autoajuste en tiempo real, incrementando o decrementando los anchos de banda máximos permitidos a cada ONU, para usar de forma eficiente y adaptativa los recursos de la red. Esta funcionalidad, que nunca antes había sido implementada en ningún algoritmo, supone una gran ventaja frente al resto de algoritmos hasta ahora existentes, ya que gracias a esta adaptabilidad en tiempo real, DySGAB hace en primer lugar que la red sea independiente de las condiciones iniciales de la misma y que además se auto-adapte eficientemente ante posibles cambios en tiempo real en las condiciones de red o en el patrón de tráfico.

Para demostrar las bondades y ventajas que ofrece el buen funcionamiento del algoritmo, se han realizado en primer lugar simulaciones con el propósito de fijar una serie de parámetros clave que marcan el punto de operación en DySGAB y que están relacionados con su diseño. Estos parámetros son el tiempo de la ventana con la que se calcula la media del ancho de banda asignado a cada ONU, el tiempo entre ejecuciones del algoritmo y el ancho de banda involucrado en la adaptación de los máximos. Una vez definidos estos parámetros de forma relativamente óptima, se han analizado las principales fortalezas y prestaciones de DySGAB llevando a cabo un estudio de simulación. Además, se ha comparado el comportamiento de DySGAB con el algoritmo LIPSA que también garantiza los niveles de ancho de banda estipulados, pero que lo hace mediante el tradicional y extendido método de pesos fijos.

Para todos los escenarios considerados, se ha comprobado que DySGAB, a diferencia de las propuestas con las que se ha comparado, consigue adaptar los anchos de banda asignados a diferentes niveles mínimos de ancho de banda garantizado estipulados por el proveedor de servicio. Incluso esta adaptación se consigue con niveles garantizados que varían en tiempo real, lo cual es una gran ventaja, ya que utilizando DySGAB no es necesario interrumpir el servicio para recalcular unos nuevos pesos acordes a la nueva situación. Esto supone una gran novedad frente al resto de algoritmos existentes hasta el momento. Además, DySGAB dota a la red de independencia frente a las condiciones iniciales de la misma. En concreto, esto quiere decir que no es necesario calcular nuevos pesos cuando cambian los niveles de ancho de banda a garantizar o la distribución de ONUs en cada perfil de abonado, lo que implica una gran desventaja en los algoritmos

basados en pesos fijos.

Por último se ha demostrado que DySGAB se puede emplear en redes EPON y LR-EAPON de distinta distancia sin perder su eficacia y funcionalidades principales. Así mismo, se ha mostrado la eficiencia del algoritmo bajo patrones de tráfico más reales, tales como tráfico asimétrico y tráfico dinámico, lo cual prueba su adaptabilidad en tiempo real a distintos escenarios de red más realistas.

Capítulo 4

Controladores PID (Proporcional-Integral-Derivativo) en la gestión de recursos en redes LR-EPON

4.1. Introducción

Este capítulo se centra en la incorporación de sistemas de control de procesos en algoritmos de asignación dinámica de ancho de banda para el control de recursos en redes LR-EPON, algo que hasta el momento no se había planteado en este campo de investigación. La incorporación de estos sistemas permite conseguir un aumento de la estabilidad del sistema, reduciendo así el grado de oscilación de los parámetros de red bajo control y el error cometido en la consecución de los requisitos de calidad de servicio estipulados. En concreto, se ha optado por el uso de controladores PID (Proporcional-Integral-Derivativo, *Proportional-Integral-Derivative*) debido a su simplicidad, robustez, funcionalidad, aplicabilidad y buen funcionamiento en numerosas aplicaciones en campos muy diferentes. Los controladores PID nunca antes habían sido utilizados en la gestión de recursos en redes ópticas de acceso. Sin embargo, el control robusto de parámetros de red tales como el ancho de banda, el retardo medio o la probabilidad de bloqueo, son una tarea esencial y crítica para cualquier operador o proveedor de servicios para ofrecer una calidad de servicio satisfactoria. Por lo tanto, el control de dichos parámetros de red a través de un sistema PID, otorgará a la red un rendimiento óptimo y adecuado a la hora de garantizar los requisitos estipulados a los abonados y a los servicios soportados en la red de acceso.

En la primera parte del capítulo se describe el principio de funcionamiento de los

controladores PID. Así mismo, se presentan algunos ejemplos del uso de controladores PID en redes de telecomunicaciones, y de forma más concreta en el campo de las redes ópticas. A continuación, se explica el primer algoritmo, denominado SPID (*Service level agreement PID*), desarrollado para el control del ancho de banda medio asignado en redes LR-EPON mediante un controlador PID. Este algoritmo adapta de forma dinámica, mediante el uso de un PID, los recursos disponibles en la red para garantizar los niveles de ancho de banda estipulados por el proveedor a cada perfil de abonado. Seguidamente se realiza un análisis exhaustivo del proceso de sintonización llevado a cabo en SPID. La sintonización es un proceso que consiste en hallar el valor de las constantes que gobiernan el comportamiento del PID, y cuya determinación es esencial para asegurar su buen funcionamiento dentro del sistema en el que se implementa.

En la segunda parte del capítulo, se presenta otro algoritmo desarrollado en esta Tesis, denominado DaSPID (*Delay aware Service level PID*), para controlar el retardo medio de ciertas clases de servicio de carácter prioritario mediante un PID. La idea principal de este algoritmo consiste en controlar que los servicios ofrecidos a los distintos perfiles de usuario (SLAs) cumplan con las restricciones de retardo impuestas por los estándares y por los proveedores de servicio dependiendo del tipo de tráfico al que pertenecen. De este modo, se considera el comportamiento final de cada uno de los servicios ofrecidos a cada abonado con distinta prioridad para garantizar así una calidad de servicio más completa e integral. Se ofrece por tanto una doble diferenciación (de servicios y usuarios), que la mayoría de los algoritmos desarrollados hasta el momento no considera, para así conseguir un mejor control de la calidad de servicio en la red LR-EPON. Este es un aspecto crítico que debe ser cubierto por los algoritmos DBA, ya que los proveedores de servicio tienen que garantizar que el tráfico más prioritario cumpla con las restricciones en parámetros tan importantes como el retardo medio o la probabilidad de pérdida de paquetes de acuerdo con la prioridad del SLA contratado por el usuario.

4.2. Controladores PID (Proporcional-Integral-Derivativo)

Un PID es un sistema de control ampliamente utilizado hoy en día en muchos campos. De hecho, se estima que el 90 % de los controladores industriales se implementan con sistemas PID [90, 91]. Su sencillez, robustez en una amplia gama de condiciones de operación, funcionalidad, aplicabilidad y buen funcionamiento han hecho que el control PID sea muy popular y extensamente aplicado en diversos ámbitos [92, 93]. Los controladores PID se han utilizado en una gran variedad de problemas, tales como el control de procesos, control de vuelo, instrumentación, etc. [91].

En el campo de las telecomunicaciones, el control PID ha sido utilizado en técnicas de gestión activa de colas (AQM, *Active Queue Management*) en el protocolo de control de transmisión (TCP, *Transmission Control Protocol*) y en el protocolo de Internet (IP, *Internet Protocol*) para controlar el nivel de congestión [94–96]. La base de los algoritmos AQM es descartar paquetes antes de que se produzca un desbordamiento de la memoria e informar acerca de la congestión que se espera, para que las fuentes puedan evitar grandes retardos y múltiples pérdidas de paquetes mediante el ajuste de su tasa de envío. Estos algoritmos presentan problemas para detectar y controlar la congestión cuando las condiciones de la red cambian de forma dinámica. El controlador PID, por el contrario, es capaz de detectar y controlar la congestión de forma adaptativa en estos entornos de red que cambian dinámicamente. En este sentido, el controlador PID calcula la tasa de entrada esperada en cada *router* y envía esta información hacia el *router* superior, de tal manera que el tamaño de la memoria puede ser estabilizado a un valor apropiado.

De forma más específica, en el campo de las redes ópticas, este tipo de procesos de control está emergiendo de forma potente en el establecimiento de *lightpaths* (caminos ópticos) en redes troncales WDM [97, 98]. En las redes WDM el número de longitudes de onda disponibles es limitado, por lo que es indispensable establecer y liberar los caminos ópticos de forma dinámica. El controlador PID implementado en [97, 98] determina la cantidad de datos que deben transmitirse de modo que en la memoria se almacene una cantidad constante de datos. Los *lightpaths* se establecen y liberan dinámicamente comparando la señal de salida del PID con la cantidad máxima de datos que pueden ser transmitidos a través de los caminos ópticos establecidos. Gracias al uso del PID se establece un menor número de *lightpaths*, lo que lleva a un uso más eficiente de la red comparado con la utilización de otros métodos.

Por otro lado, en las redes de acceso, la implementación de controladores PID en el control de recursos y de parámetros nunca antes había sido utilizada, aunque su aplicabilidad en este tipo de gestión podría hacerse de un modo sencillo y eficiente. Sin embargo, en la red de acceso, se han propuesto otras técnicas de control en la asignación de ancho de banda. En concreto, en [99] se propone un algoritmo que estima el ancho de banda máximo permitido para el siguiente ciclo dependiendo de la diferencia entre el máximo permitido actual y el tamaño de las colas de la ONU una vez que transmite sus datos, esto es, se modifica el máximo dependiendo del error cometido en la asignación del ancho de banda. En [100, 101], se presentan las directrices para el diseño de un compensador óptimo, cuya función es mantener la estabilidad del sistema bajo diferentes escenarios y que consigue mantener un tamaño de cola deseado.

Por tanto en este capítulo se propone el uso de controladores PID para la gestión

eficiente de recursos en redes LR-EPON. Por ello, en primer lugar se va a proceder a describir los fundamentos básicos de los controladores PID para después presentar los algoritmos diseñados e implementados.

4.2.1. Arquitectura general de un PID

Un PID es un controlador simple basado en un lazo de realimentación, cuyo propósito es mantener la variable que se controla lo más cerca posible del valor deseado [102]. El diagrama de bloques típico para el control de procesos con un PID se muestra en la Figura 4.1. Como se puede observar, está formado por dos bloques principales, el proceso bajo control y el controlador PID. El proceso tiene una entrada, la variable manipulada, también denominada variable de control (representada en la Figura 4.1 por u). La salida del proceso (y) es la variable que se quiere mantener cercana al valor deseado (y_0). Por tanto, el PID compara el valor de salida del proceso y con el valor de referencia y_0 . La diferencia entre estos dos valores es el error cometido (denotado por e). Entonces, el PID en función del error cometido, ajusta la variable de control u para mantener el valor de la variable controlada (y) lo más cercano posible al valor de referencia (y_0), de forma que se minimice el error cometido.

La ecuación que modela el comportamiento de un controlador PID en el dominio del tiempo es la que se muestra en la ecuación (4.1), donde $u(t)$ es la variable de control en el dominio del tiempo y $e(t)$ es el error, definido como $e(t) = y_0 - y(t)$.

$$u(t) = \underbrace{K_p \cdot e(t)}_{\text{Proporcional}(P)} + \underbrace{K_p \cdot \frac{1}{T_i} \int_0^t e(\tau) d\tau}_{\text{Integral}(I)} + \underbrace{K_p \cdot T_d \cdot \frac{de(t)}{dt}}_{\text{Derivativo}(D)} \quad (4.1)$$

Por tanto, se observa que la señal de control $u(t)$ es la suma de tres términos, el término P (proporcional al error), el término I (proporcional a la integral del error) y el término D (proporcional a la derivada del error). Además, el parámetro K_p se define como la ganancia proporcional, T_i como el tiempo integral y T_d como el tiempo derivativo [102]. El proceso de sintonización del PID mencionado anteriormente consiste en la determinación de los valores óptimos de estos tres parámetros.

Sin embargo, para aplicar el control de un PID en una simulación de eventos discretos,

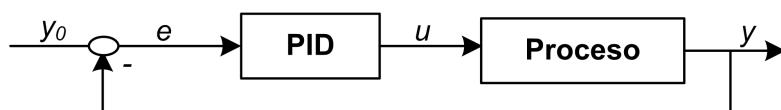


Figura 4.1: Diagrama de bloques de un proceso controlado por un PID.

como es el caso del sistema que nos ata e, es necesario aplicar un controlador PID discreto. Este tipo de PID lee el error y lo calcula en un determinado intervalo de tiempo, llamado tiempo de muestreo. As , una aproximaci n de la ecuaci n (4.1) en el dominio discreto se muestra en la ecuaci n (4.2) [103, 104], donde T se refiere a dicho tiempo de muestreo.

$$u[n] = \underbrace{K_p \cdot e[n]}_{\text{Proporcional}} + \underbrace{K_p \cdot \frac{T}{T_i} \sum_{m=0}^n e[m]}_{\text{Integral}} + \underbrace{K_p \cdot \frac{T_d}{T} (e[n] - e[n-1])}_{\text{Derivativo}} \quad (4.2)$$

Cada uno de los t rmidos de la se al de control tiene una finalidad distinta. As , la componente proporcional (P), que es proporcional al error cometido ($e[n]$), se refiere al error presente. La componente integral (I), que es proporcional a la integral del error, tiene como funci n principal asegurarse de que la salida del proceso se mantenga igual al valor de referencia en un estado estacionario, y est  relacionada con la acumulaci n de errores pasados ($\sum_{m=0}^n e[m]$). Finalmente, la componente derivativa (D), que es proporcional a la derivada del error, proporciona la estabilidad al sistema y ser  la encargada de realizar una predicci n de los errores futuros.

Un controlador PID completo, se compone de un compensador proporcional y derivativo (PD) en cascada con un compensador proporcional e integral (PI). El prop sito del compensador PD es mejorar la respuesta transitoria, al mismo tiempo que mantiene la estabilidad deseada; mientras que el prop sito del compensador PI es mejorar la exactitud del estado permanente del sistema sin degradar la estabilidad [105].

El controlador PD o controlador derivativo es esencialmente un controlador anticipativo, ya que $\frac{de(t)}{d(t)}$ representa la pendiente del error $e(t)$. Al conocer la pendiente, el controlador puede anticipar la direcci n del error y emplearla para controlar mejor el proceso. El control derivativo mide la pendiente instant nea del error, predice el sobre-paso adelante en el tiempo y hace un esfuerzo correctivo antes de que se produzca un sobre-paso excesivo. Este tipo de control afecta al error en estado estable de un sistema solo si el error en estado estable var a con el tiempo. Si el error en estado estable de un sistema es constante con respecto al tiempo, la derivada con respecto al tiempo de este error es cero y la porci n derivativa del controlador no proporciona ninguna entrada al proceso. Pero si el error en estado estable se incrementa con el tiempo, se genera otra vez la parte derivativa proporcional al t rmino $\frac{de(t)}{d(t)}$, lo cual reduce la magnitud del error. Este controlador mejora el amortiguamiento, lo que implica que se puede emplear una ganancia proporcional m s elevada para acelerar la respuesta transitoria y reducir el sobre-paso m ximo, pero no es efectivo para sistemas ligeramente amortiguados o

inicialmente inestables [106].

Por otro lado, la parte integral del controlador PID produce una señal proporcional a la integral con respecto al tiempo de la entrada del controlador. El controlador PI es el más común y además es adecuado para todos los procesos donde las dinámicas son esencialmente de primer orden. El error en estado estable del sistema original se mejora en un orden; es decir, si el error en estado estable a una entrada dada es constante, el control PI lo reduce a cero (considerando que el sistema compensado permanece estable). Además, si un proceso no requiere un control muy exigente, incluso si el proceso tiene dinámicas de orden superior, lo que se necesita es una acción para proporcionar una desviación nula en estado estacionario y una respuesta transitoria adecuada mediante una acción proporcional. El controlador PI mejora el error en estado estable a costa de la estabilidad. Sin embargo, si se seleccionan adecuadamente los parámetros K_p y T_i , tanto el amortiguamiento como el error en estado estable se pueden mejorar. El controlador PI mejora el amortiguamiento y reduce el sobrepaso máximo. Sin embargo, aumenta el tiempo de respuesta. Un controlador PI solo considera datos presentes y pasados [103, 106].

Como se ha visto anteriormente, el controlador PD puede añadir amortiguamiento a un sistema, pero no afecta la respuesta en estado estable. Por otro lado, el controlador PI puede mejorar la estabilidad relativa y el error en estado estable al mismo tiempo, pero el tiempo de respuesta se incrementa. Esto conduce a utilizar un controlador PID para que se empleen las mejores características de los controladores PI y PD. El controlador PID, así como sus componentes PD y PI, representan formas simples de controladores que emplean operaciones de derivación e integración en la compensación de sistemas de control. El control PID es suficiente para procesos donde las dinámicas dominantes son de segundo orden.

No obstante, aunque un sistema PID completo está formado por las tres componentes anteriores, éstas no son estrictamente necesarias para llevar a cabo un control adecuado de una variable en ciertos sistemas. Por otro lado, cuantas más componentes haya, mayor será la complejidad del controlador.

Finalmente, para un correcto funcionamiento de un PID, en cualquiera de sus variantes, es necesario determinar los parámetros que ajustan la señal de control, esto es, la ganancia proporcional K_p , el tiempo integral T_i y el tiempo derivativo T_d . A este proceso se le denomina sintonización del PID y es una etapa muy importante, ya que es necesario determinar un valor óptimo para estos parámetros en función de la aplicación, con el fin de que el sistema sea estable y se alcancen los objetivos para los cuales ha sido diseñado el controlador. En el siguiente apartado, se explicarán con mayor detalle

algunos de los métodos de sintonización empleados en controladores PID.

4.2.2. Métodos de sintonización de controladores PID

Tal y como se especifica en el apartado anterior, un controlador PID tiene que ser sintonizado adecuadamente a través de un proceso de sintonización. En este apartado, se van a analizar algunos de los métodos de sintonización existentes para sistemas PID, así como su implementación en la gestión y control de parámetros en redes de acceso LR-EAPON. A continuación, se presenta una descripción de los métodos de sintonización de sistemas PID más extendidos:

- **Método de Ziegler-Nichols de respuesta transitoria:** Este método consiste en someter al sistema a una entrada escalón y, a partir de la respuesta obtenida, ajustar los parámetros del controlador en función de ciertos parámetros de dicha respuesta, tales como el retardo o la pendiente máxima obtenida [102, 103]. Este método ha sido descartado ya que al trabajar con redes en las que la demanda de ancho de banda es en la práctica aleatoria, introducir una entrada escalón se alejaba por completo de nuestro entorno de trabajo.
- **Método de Ziegler-Nichols de respuesta en frecuencia o sensibilidad final:** Este método se basa también en una caracterización simple de la dinámica del proceso. En este método se ajustan los parámetros del controlador de forma que la acción del control es únicamente proporcional ($T_i = \infty$ y $T_d = 0$) y se incrementa el parámetro K_p desde cero hasta un valor crítico en donde la salida presente oscilaciones sostenidas. Una vez obtenido el parámetro K_p , los parámetros T_i y T_d se ajustan con unas fórmulas de aplicación directa [102, 103]. A pesar de que la elección del valor crítico es puramente experimental y puede dar lugar a inexactitudes o imprecisiones, este método resulta muy eficaz y relativamente sencillo. Por este motivo, es un método ampliamente utilizado, en especial en sistemas donde no existe un modelo matemático o éste es muy difícil de calcular.
- **Método de Cohen-Coon:** En este método, el criterio principal de diseño es el rechazo de las perturbaciones de carga. Por otro lado, la metodología viene a ser similar a la del método de Ziegler-Nichols de respuesta transitoria, planteando similares problemas en su ejecución. Cabe decir, además, que la calidad de la información que este método proporciona no es mejor que en los dos métodos anteriormente mencionados [102, 103]. Por ello, se ha descartado la utilización de este método en el control de parámetros de redes LR-EAPON.

- **Método de Chien, Hrones y Reswick:** Es una variante del método de Ziegler-Nichols de respuesta transitoria, y si bien proporciona una robustez algo mayor que éste (debido a que tiene en cuenta un parámetro más) [102, 103], plantea las mismas dificultades en su aplicación, por lo que su uso ha sido descartado.
- **Método de Harriot:** Este método considera únicamente la parte proporcional del controlador al igual que el método de Ziegler-Nichols de respuesta en frecuencia. Posteriormente, se aplican entradas escalón y se va ajustando la ganancia proporcional (K_p) hasta conseguir una razón de amortiguamiento de 1/4. Esto es, que la amplitud de la segunda oscilación sea cuatro veces menor que la amplitud de la primera. Este método plantea el mismo problema con la entrada escalón que se ha señalado en métodos anteriores. Por otro lado, su eficacia y robustez son similares al método de Ziegler-Nichols de respuesta en frecuencia o sensibilidad final.
- **Métodos algebraicos y asignación de polos:** Existen numerosos métodos algebraicos y métodos basados en la asignación de polos que se utilizan en la sintonización de los controladores. Sin embargo, para la aplicación de estos métodos es necesario el conocimiento de la función de transferencia. En nuestro caso, esta función no se puede obtener, por lo que no se ha realizado un estudio más profundo de estas técnicas.
- **Métodos de optimización:** Estos métodos pueden verse como una especie de control óptimo, donde los parámetros del PID se obtienen de forma específica para cada aplicación mediante una optimización previa, utilizando algún tipo de heurística o algoritmo evolutivo [92]. Un ejemplo de este tipo de sintonización es la realizada mediante algoritmos genéticos. El objetivo principal de este tipo de algoritmo, es realizar una sintonización automática previa para asegurar que los parámetros del PID sean los más óptimos posibles. El proceso consiste en originar un conjunto de posibles soluciones seleccionadas aleatoriamente, evaluarlas, seleccionar parte de esta población, cruzarla y obtener una nueva generación mejor que la anterior, hasta llegar a obtener la solución óptima que sea capaz de ajustar el proceso al valor deseado con las condiciones del sistema.
- **Métodos con características adaptativas y variables en tiempo real:** Estos métodos se caracterizan porque los parámetros del controlador PID varían dinámicamente con las características del sistema, de tal modo que en todo momento el controlador es el óptimo para el estado actual del sistema. En algunas ocasiones se pueden aplicar los métodos de optimización en tiempo real para llevar a cabo esta sintonización adaptativa.

Un ejemplo de este tipo de sintonización es la realizada mediante redes neuronales. Este método utiliza una arquitectura de red neuronal para determinar los parámetros de sintonización del controlador modificándolos a lo largo del tiempo con el fin de obtener una sintonización óptima en tiempo real para las distintas condiciones que se presenten en el sistema.

Por tanto, se puede concluir que al no disponer de una función de transferencia que caracterice el sistema de asignación de ancho de banda en una red LR-EAPON, se utilizarán métodos experimentales para llevar a cabo la sintonización. En primer lugar, y para llevar a cabo una primera aproximación del uso de controladores PID se ha optado por utilizar el método de Ziegler-Nichols de respuesta en frecuencia o sensibilidad final, que es un método ampliamente utilizado por su sencillez y buenos resultados. En capítulos posteriores de la Tesis, se implementarán otras técnicas de sintonización más complejas basadas en algoritmos genéticos y redes neuronales para realizar una sintonización más óptima y adaptativa que con otros métodos.

4.3. Controlador PID para garantizar niveles de ancho de banda. Algoritmo SPID (*Service level Agreement PID*)

El segundo algoritmo de asignación de ancho de banda desarrollado en esta Tesis, denominado SPID (*Service level agreement PID*) [107, 108], es un algoritmo que realiza un control automático de la asignación de ancho de banda mediante el uso de un controlador PID, con el fin de garantizar dinámicamente requisitos de nivel de ancho de banda según la prioridad del perfil contratado.

Para describir el funcionamiento de este nuevo algoritmo se va a seguir la misma estructura del capítulo anterior. Por tanto en primer lugar, se explicará el proceso de asignación de ancho de banda en SPID, así como la adaptación dinámica que se lleva a cabo para garantizar los niveles de ancho de banda. A continuación, se describe el proceso de sintonización del PID y por último se analizan las prestaciones del algoritmo diseñado e implementado en una red LR-EAPON.

4.3.1. Asignación de ancho de banda en SPID

Al igual que DySGAB, SPID es un algoritmo de asignación dinámica de ancho de banda que implementa una política de asignación de *polling* o de testeo continuo con un esquema limitado, por presentar éste un comportamiento muy eficiente. Tal y como se explicó en el Apartado 3.2.1 del Capítulo 3, la asignación de ancho de banda se realiza

según la ecuación (4.3), donde el término $B_{alloc}^{onu_i}$ es el ancho de banda asignado a la ONU i , $B_{demand}^{onu_i}$ es el ancho de banda demandado por dicha ONU y $B_{max}^{onu_i}$ es el ancho de banda máximo permitido para la ONU i (todos ellos en bits).

$$B_{alloc}^{onu_i} = \begin{cases} B_{demand}^{onu_i} & \text{si } B_{demand}^{onu_i} \leq B_{max}^{onu_i} \\ B_{max}^{onu_i} & \text{si } B_{demand}^{onu_i} > B_{max}^{onu_i} \end{cases} \quad (4.3)$$

La gran novedad que implementa SPID, es que la asignación de ancho de banda está controlada por un PID, según se explica en el siguiente apartado. Se prevé que este tipo de control sea más robusto y eficaz que otros métodos ya existentes hasta la fecha.

4.3.2. Adaptación dinámica del ancho de banda en SPID

El objetivo del algoritmo SPID es mantener el ancho de banda asignado a cada usuario lo más cerca posible de su ancho de banda garantizado, de modo que los requisitos de calidad de su SLA contratado puedan cumplirse. Por tanto, el objetivo de garantizar cierto nivel de ancho de banda a cada usuario, encaja en el tipo de procesos que pueden ser controlados por un PID. Tal y como se verá en posteriores secciones, otra de las grandes ventajas que ofrece un controlador PID, es que su diseño puede modificarse fácilmente para controlar otros parámetros esenciales en la red, tales como el retardo medio de los paquetes.

La estructura del sistema propuesto en SPID puede observarse en el diagrama de bloques de la Figura 4.2. Como se puede ver, la variable de salida del sistema (y) es el ancho de banda medio asignado a cada ONU, $\overline{B}_{alloc}^{onu_i}$. Este valor se calcula de igual forma que en DySGAB. Por tanto, cada ONU tiene una ventana deslizante de duración fija de T_{window} segundos. En esta ventana se introducen las últimas muestras que contienen el ancho de banda asignado a dicha ONU en cada ciclo ($B_{alloc}^{onu_i}$) y el tiempo en el que fue asignado ese ancho de banda. Cada vez que se inserta una muestra, la estimación del ancho de banda medio asignado a cada ONU se actualiza. Para calcular el término

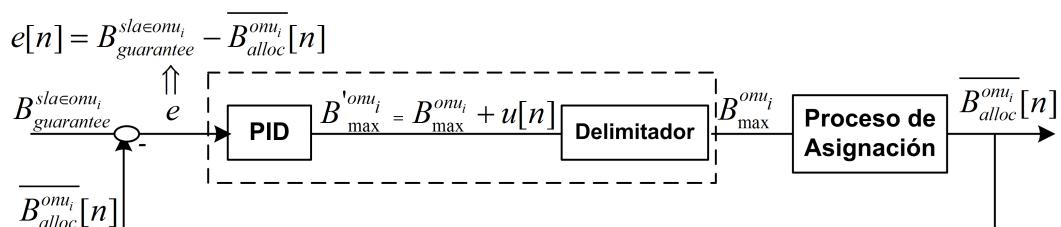


Figura 4.2: Diagrama de bloques del proceso controlado por un PID propuesto para el control del ancho de banda en una red LR-EPON.

$\overline{B_{alloc}^{onu_i}}[n]$ en un cierto instante n en el que se calcula el error, se hace la media aritmética con todos los valores $B_{alloc}^{onu_i}$ contenidos en la ventana en ese momento. Además, la primera muestra de la cola se descartará si la diferencia entre el tiempo de la última muestra y el tiempo de la primera muestra supera la duración máxima de la ventana.

Por otro lado, el valor de referencia (y_0) al que se quiere aproximar el ancho de banda medio asignado, es el nivel de ancho de banda garantizado a cada ONU, de acuerdo con su SLA contratado ($B_{guarantee}^{sla \in onu_i}$). Por lo tanto, el controlador PID, en función del error cometido en un momento n , definido en nuestro caso como $e[n] = B_{guarantee}^{sla \in onu_i} - \overline{B_{alloc}^{onu_i}}[n]$, cambia de forma dinámica el ancho de banda máximo permitido para cada ONU en un ciclo. Para realizar dicha modificación, la variable de control $u[n]$, calculada según la ecuación (4.2), se añade al término $B_{max}^{onu_i}$ (Figura 4.2), para obtener un nuevo máximo para esta ONU, $B'_{max}^{onu_i}$. De esta manera, por ejemplo, si el ancho de banda medio asignado a la ONU i ($\overline{B_{alloc}^{onu_i}}[n]$) es muy inferior a su ancho de banda garantizado ($B_{guarantee}^{sla \in onu_i}$), el error cometido será alto, de modo que el PID incrementa dinámicamente el término $B_{max}^{onu_i}$, para facilitar que la ONU i cumpla con los requisitos de su ancho de banda garantizado. Por el contrario, si el ancho de banda medio asignado es mayor que su ancho de banda garantizado, el error cometido será negativo y el PID decrementa su ancho de banda máximo asignado.

Sin embargo, en nuestro sistema, el ancho de banda máximo calculado por el PID tiene que estar delimitado en cada ciclo. Esto se basa en el hecho de que el estándar EPON define un tiempo de ciclo máximo de 2 ms. Por lo tanto, la suma total del ancho de banda máximo permitido para todas las ONUs en un ciclo no puede exceder el ancho de banda contenido en un tiempo de ciclo de 2 ms. Por eso, el sistema incluye un delimitador (Figura 4.2). Si la suma total de los anchos de banda máximos permitidos de todas las ONUs propuestos por el PID supera el ancho de banda total máximo, el delimitador reduce los máximos en proporción a los calculados por el PID, pero que se ajusten a un ciclo de 2 ms. De lo contrario, el delimitador no cambia los valores máximos. Finalmente, si el ancho de banda demandado por una ONU es inferior a su máximo actual, el PID permanece inactivo, lo que significa que no calcula un nuevo máximo para dicha ONU. En ese caso, siguiendo la ecuación (4.3), el OLT asigna a dicha ONU el ancho de banda demandado.

Por tanto, el proceso completo de asignación de ancho de banda se puede resumir de la siguiente forma. Cuando un mensaje *Report* de una ONU llega al OLT, éste aplica la ecuación (4.3) para asignar el ancho de banda. Aparte de este proceso, que es el funcionamiento general de un algoritmo de *polling* con un esquema limitado, cada cierto tiempo de muestreo (T segundos), el PID se ejecuta en paralelo y se aplica la ecuación

(4.2) de manera que se actualizan los anchos de banda máximos de cada ONU. Como esta ecuación solo consta de sumas, productos y restas, el ajuste es muy rápido y no incrementa la complejidad del algoritmo dentro de la red. Una vez calculada la señal de control para cada ONU y aplicado el delimitador, se actualizan los anchos de banda máximos permitidos para cada ONU ($B_{max}^{onu_i}$). Éste será el ancho de banda máximo permitido para cada ONU hasta que transcurran otros T segundos y el PID se ejecute de nuevo, obteniendo nuevos valores de $B_{max}^{onu_i}$.

Por otro lado, aunque el PID adapta constantemente los valores de ancho de banda máximo, éstos se inicializan a un cierto valor, al igual que ocurría en DySGAB, con un peso asociado al SLA contratado (W^{sla_j}). Por lo tanto, las condiciones iniciales de ancho de banda se calculan siguiendo la ecuación (4.4), donde $B_{cycle_available}$ es el ancho de banda contenido en el tiempo de ciclo máximo considerado (el ciclo máximo del estándar EPON es de 2 ms). El término W^{sla_j} es el peso asignado al SLA j , mientras que $W^{sla_k / onu_i \in sla_k}$ es, en concreto, el peso asignado al SLA k que ha contratado la ONU i . Por último, $N_{onus}^{sla_j}$ se refiere al número de ONUs que pertenecen a cada SLA j .

$$B_{max}^{onu_i} = \frac{B_{cycle_available} \cdot W^{sla_k / onu_i \in sla_k}}{\sum_j W^{sla_j} \cdot N_{onus}^{sla_j}} \quad (4.4)$$

El control automático y robusto que ofrece un PID resulta muy ventajoso, ya que el ancho de banda asignado se ajusta constantemente, para asegurar los niveles garantizados a los usuarios de acuerdo a los requisitos de calidad de su perfil. Sin embargo, esta auto-adaptación tiene otros beneficios para los proveedores de servicios. Al igual que ocurría con DySGAB, si se necesita cambiar el nivel garantizado de algunos perfiles, el PID reajusta automáticamente el ancho de banda en tiempo real. Por el contrario, otros algoritmos basados en pesos fijos, tienen que interrumpir el servicio de la red para volver a calcular los nuevos pesos. Por otro lado, tal y como se vio en el Capítulo 3, DySGAB que también adapta dinámicamente el ancho de banda, actualiza estos valores reduciendo o aumentando una cantidad fija el ancho de banda máximo permitido en cada iteración del algoritmo. Sin embargo, este comportamiento, tal y como se demostró en el Capítulo 3, denota una gran dependencia de esta cantidad fija para garantizar de forma más o menos óptima los niveles de calidad de servicio estipulados. Por el contrario, el controlador PID propuesto cambia los anchos de banda máximos permitidos de acuerdo con el error cometido, por lo que la capacidad de adaptación a cambios en el patrón de tráfico o en las condiciones de la red será más robusta, más independiente y más rápida que en DySGAB. Por todo ello, en este capítulo de la Tesis se analizará y comparará el comportamiento de ambos algoritmos.

4.3.3. Diseño de estrategias de asignación de ancho de banda de SPID

Otra ventaja que presenta SPID es que ha sido diseñado para adoptar dos políticas diferentes de asignación de ancho de banda cuando el ancho de banda disponible no es suficiente para garantizar todos los anchos de banda a cada ONU. Bajo esta situación, los proveedores de servicio y operadores pueden escoger entre una política conservativa o una política agresiva, dependiendo de sus preferencias. En la primera de ellas, todos los usuarios se ven afectados de igual forma por la escasez de recursos cuando el ancho de banda disponible no es suficiente para garantizar los anchos de banda mínimos estipulados. En este caso, SPID no distingue entre la prioridad de los usuarios y trata a todos por igual. Por tanto, ningún SLA cumplirá con su nivel garantizado, pero ningún SLA se verá más penalizado que otro.

Por otro lado, si los proveedores de servicio u operadores seleccionan la política agresiva, se tendrá en cuenta el grado de prioridad de los usuarios, tratando de beneficiar a aquellos con una prioridad más alta. Por lo tanto, si el ancho de banda disponible no es suficiente para garantizar los anchos de banda estipulados de todos los usuarios, aquellos que hayan contratado el SLA de menor prioridad ceden parte de su ancho de banda hasta que el resto de mayor prioridad satisfagan sus anchos de banda garantizados. De esta forma, todos los SLAs cumplen con sus niveles de ancho de banda excepto el SLA de menor prioridad. Por ejemplo, si en la red se consideran tres SLAs, SLA₀ de prioridad alta, SLA₁ de prioridad media, y SLA₂ de prioridad baja y se aplica esta política agresiva, cuando no exista ancho de banda suficiente para garantizar todos los niveles de ancho de banda, el SLA₂ cederá ancho de banda para que los usuarios del SLA₀ y del SLA₁ cumplan con sus niveles mínimos garantizados.

En la Tabla 4.1 se presentan las características de las políticas de asignación que se han implementado en SPID.

Estrategia	Descripción
Conservativa	<ul style="list-style-type: none"> - Ningún SLA cumple con su nivel garantizado - Ningún SLA es más penalizado que otro
Agresiva	<ul style="list-style-type: none"> - Los SLAs de alta prioridad cumplen con sus niveles garantizados gracias al ancho de banda que cede el SLA de menor prioridad - El SLA de menor prioridad es penalizado a favor del resto de SLAs más prioritarios

Tabla 4.1: Diferentes estrategias de asignación de ancho de banda soportadas por SPID para garantizar los niveles mínimos de ancho de banda estipulados para cada perfil.

4.4. Estudio de simulación del algoritmo SPID

Para analizar el comportamiento de SPID se va a llevar a cabo en primer lugar el proceso de sintonización del controlador y a seleccionar el tipo de controlador más adecuado en este sistema (P, PI o PID). A continuación, se va a proceder a determinar el valor los parámetros relacionados con el diseño del controlador PID. En concreto, una vez sintonizado el PID, solo es necesario determinar el tiempo de muestreo. Tras ello se realizará un análisis de simulación de las prestaciones de SPID comparándolo con el algoritmo DySGAB, considerando distintos escenarios y condiciones de red.

4.4.1. Escenario de simulación

El escenario de simulación considerado es el mismo que en DySGAB (Apartado 3.3.1 del Capítulo 3). La Tabla 4.2 recoge un resumen de los principales parámetros de red y los valores utilizados. Para llevar a cabo el análisis de prestaciones, el algoritmo SPID se compara con DySGAB, ya que los dos son algoritmos que siguen una política de *polling* con esquema limitado, y son los únicos algoritmos existentes que han sido diseñados para garantizar dinámicamente el ancho de banda estipulado en escenarios con diversos perfiles de usuario. Tal y como se vio en el Capítulo 3, DySGAB modifica los anchos de banda máximos permitidos incrementando y decrementando una cantidad fija de ancho de banda, lo que puede dar lugar a inefficiencias si el valor no se escoge de forma correcta o si las condiciones de la red cambian. Sin embargo, en SPID la adaptación se realiza a través de un robusto controlador PID, que modifica los máximos permitidos dependiendo del error cometido. Además, para permitir una mejor comparación de ambos algoritmos, el tamaño de la ventana con la que se calcula el ancho de banda medio asignado (T_{window}), se ha fijado a 7.5 s al igual que en el caso de DySGAB, valor que da lugar a un comportamiento adecuado del sistema.

4.4.2. Proceso de sintonización y comparativa de estrategias P-PI-PID

El PID tiene que ser sintonizado de forma adecuada para obtener los valores óptimos de la ganancia proporcional K_p , el tiempo integral T_i y el tiempo derivativo T_d de la ecuación (4.2). Así mismo, si el controlador escogido es un P (solo parte proporcional) o PI (parte proporcional e integral), solo sería necesario determinar los valores óptimos de K_p o de K_p y T_i , respectivamente.

Para sintonizar el controlador se ha utilizado el método de Ziegler-Nichols de respuesta en frecuencia. Esta elección viene determinada por dos motivos. El primero de ellos es que, al no disponer de un modelo matemático o función de transferencia, es

Parámetros de simulación	Valores
Número de ONUs	16 ONUs (N_{onu})
Tasa de transmisión canal ascendente	1 Gbit/s ($EPON_Line_rate$)
Tasa de transmisión ONU-abonado	100 Mbit/s ($User_Line_rate$)
RTT	1 ms
Tiempo de guarda	1 μ s
Tiempo de ciclo máximo	2 ms
Distribución de tráfico	Fuente de Pareto H=0.8
Tamaño de paquete	Variable entre 64-1500 bytes (más 38 bytes de cabeceras)
<i>Service Level Agreement (SLA)</i>	SLA ₀ : 1 ONU Ancho de banda garantizado: 100 Mbit/s
	SLA ₁ : 5 ONUs Ancho de banda garantizado: 75 Mbit/s
	SLA ₂ : 10 ONUs Ancho de banda garantizado: 50 Mbit/s

Tabla 4.2: Parámetros de red considerados en el escenario de simulación de SPID.

necesario usar un método experimental. El segundo motivo radica en el hecho de que el método de Ziegler-Nichols es un método muy extendido y fácil de aplicar [92, 109, 110]. Esta técnica de sintonización ofrece reglas simples que se caracterizan por dos parámetros, llamados T_u (periodo último) y K_u (ganancia última), determinados por medio de experimentos sencillos sobre el sistema [90–92, 102, 103, 109, 110]. El primer paso en el método de Ziegler-Nichols de respuesta en frecuencia consiste en determinar los valores de estos dos parámetros (T_u y K_u). Para ello, solo se utiliza la respuesta proporcional ($T_i = \infty$, $T_d = 0$). Entonces, la ganancia se incrementa lentamente hasta que el proceso comienza a oscilar. La ganancia cuando esto ocurre es K_u y el período de oscilación es T_u . Una vez que se determinan estos valores, se pueden obtener de forma directa los parámetros del controlador P (K_p), los del controlador PI (K_p y T_i) y los del controlador PID (K_p , T_i y T_d) a partir de la Tabla 4.3, donde se recogen las relaciones establecidas por Ziegler-Nichols para dichos controladores [102, 103].

Controlador	K_p	T_i	T_d
P	$0.5K_u$		
PI	$0.4K_u$	$0.8T_u$	
PID	$0.6K_u$	$0.5T_u$	$0.125T_u$

Tabla 4.3: Parámetros de los controladores P, PI y PID para el método de la respuesta en frecuencia de Ziegler-Nichols.

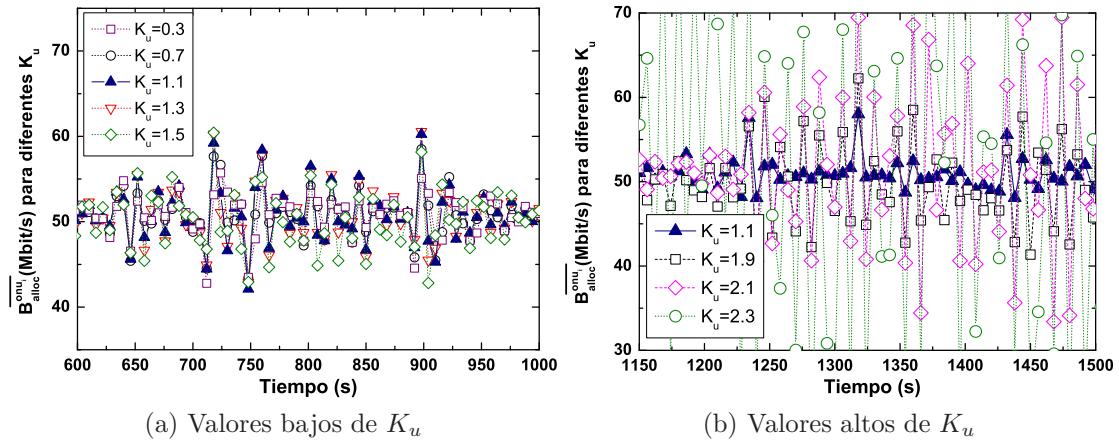


Figura 4.3: Variación del ancho de banda medio asignado para distintos valores de K_u . (a) Valores bajos de K_u (b) Valores altos de K_u .

Los resultados experimentales del proceso de sintonización en SPID se muestran en la Figura 4.3. En concreto, la Figura 4.3 (a) representa la variación del ancho de banda medio asignado a cada ONU ($\overline{B}_{alloc}^{onu_i}$) a lo largo del tiempo (puesto que es este valor el que se desea controlar mediante el PID) para diferentes valores de K_u (de 0.3 a 1.5). Como se puede observar, la oscilación se mantiene más o menos estable y similar para dicho rango de valores. En consecuencia, se puede elegir cualquier valor contenido en este rango de valores. Sin embargo, a medida que K_u continúa aumentando, el ancho de banda medio asignado sufre una mayor variación y el proceso comienza a oscilar en exceso. Para demostrarlo, en la Figura 4.3 (b), se compara el rendimiento de un valor bajo de K_u (1.1) frente a valores más altos de este parámetro. Como se puede observar en la gráfica, la oscilación no es constante y fluctúa demasiado para valores altos de K_u . Este comportamiento se puede apreciar tanto en el periodo como en la amplitud del ancho de banda medio asignado.

En conclusión, ya que el ancho de banda medio asignado es muy similar para valores bajos de K_u , se eligió un valor de $K_u = 1.1$. Para este valor, el periodo medio de oscilación T_u es igual a 22 segundos. Por lo tanto, aplicando las relaciones de la Tabla 4.3, se obtienen los valores de sintonización para cada uno de los controladores, recogidos en la Tabla 4.4.

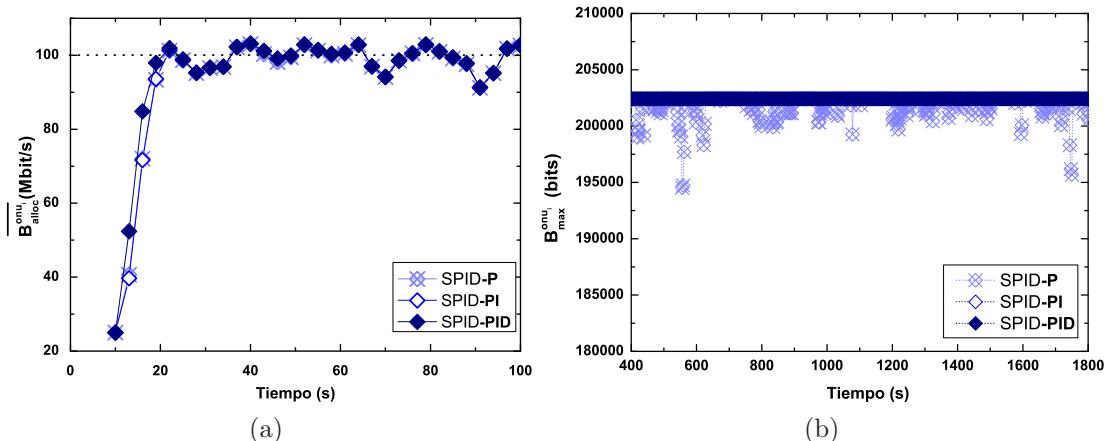
A continuación, se va a realizar un análisis del comportamiento de los distintos tipos de controladores (P-PI-PID) para elegir el que mejor se adapta al proceso de ajuste del ancho de banda máximo de SPID. Es necesario realizar este análisis porque para cada aplicación un tipo de controlador puede funcionar mejor que otros. Por tanto, se han

Controlador	K_p	T_i	T_d
P	0.55		
PI	0.44	17.6	
PID	0.66	11	2.75

Tabla 4.4: Valores de sintonización de los controladores P, PI y PID en SPID.

llevado a cabo simulaciones siguiendo el escenario de simulación genérico de la Tabla 4.2 y considerando los siguientes pesos asociados a cada SLA: $W^{sla_0} = 1$, $W^{sla_1} = 2$ y $W^{sla_2} = 3$. Estos pesos no están ajustados a los niveles garantizados, de forma que se podrá observar la evolución en tiempo real del ancho de banda medio asignado a los niveles estipulados por el proveedor de servicios. Además la carga total de red considerada es de 1.6, esto es, ONUs transmitiendo a su tasa de transmisión máxima, 100 Mbit/s.

La Figura 4.4 (a) representa la evolución del ancho de banda medio a lo largo del tiempo para una ONU del SLA₀ hacia su nivel de ancho de banda garantizado (100 Mbit/s) realizado por SPID con las tres implementaciones del controlador (P, PI y PID). Tal y como se puede observar, SPID utilizando un PID es ligeramente más rápido que con un controlador P y PI. Sin embargo, las diferencias entre los tres controladores son mínimas. Para completar la comparación, la Figura 4.4 (b) ilustra la variación instantánea del ancho de banda máximo calculado para el SLA₀ en SPID usando los tres controladores. Se puede observar como SPID es más estable con los controladores PI y PID, obteniendo ambos unos resultados muy similares. El uso de un controlador P produce variaciones ligeramente mayores que con los otros dos, aunque estas diferencias no son muy elevadas.

Figura 4.4: Evolución en tiempo real del (a) ancho de banda medio hacia nivel garantizado del SLA₀ (b) ancho de banda máximo permitido para el SLA₀.

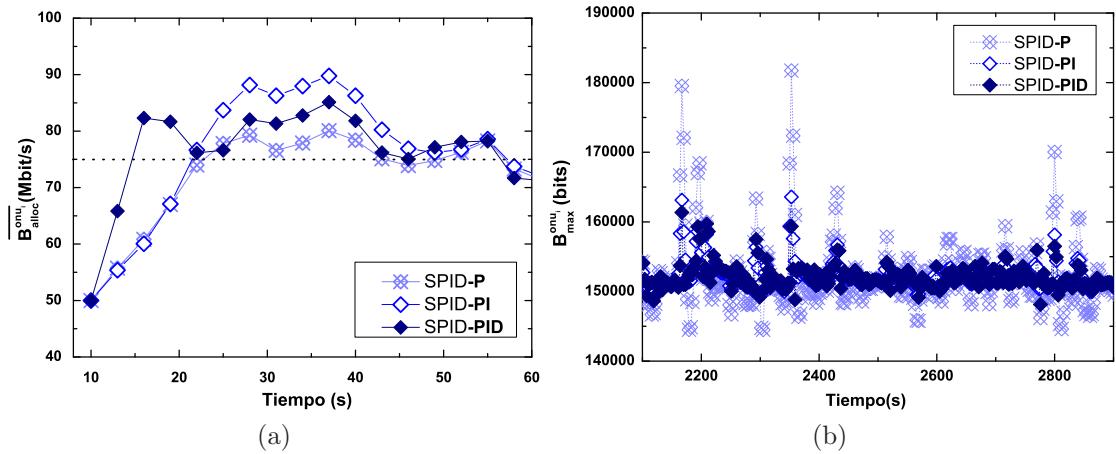


Figura 4.5: Evolución en tiempo real del (a) ancho de banda medio hacia nivel garantizado del SLA₁ (b) ancho de banda máximo permitido para el SLA₁.

El mismo comportamiento puede observarse en el SLA de prioridad intermedia (SLA_1). En la Figura 4.5 (a) está representada la evolución del ancho de banda medio en los primeros segundos de simulación para el SLA_1 cuando SPID considera los tres tipos de controlador (P, PI y PID). Se observa cómo en este caso las diferencias entre el uso de los tres controladores es también muy pequeña en cuanto al tiempo de convergencia al nivel de ancho de banda garantizado (75 Mbit/s). Por otro lado, en la Figura 4.5 (b) se muestra la variación instantánea del ancho de banda máximo permitido. En este caso, se observa que las menores fluctuaciones se consiguen en SPID cuando utiliza un controlador PI o PID. El uso del controlador P, al igual que en el SLA de mayor prioridad, produce alguna variación ligeramente superior.

Por último, la Figura 4.6 (a) muestra la evolución en tiempo real del ancho de banda medio asignado al SLA de menor prioridad (SLA_2). Tal y como puede verse, el comportamiento es similar al de los SLAs anteriores. SPID con cualquiera de los tres controladores alcanza el ancho de banda garantizado para este SLA (50 Mbit/s) en un tiempo muy reducido. Además, en la Figura 4.6 (b), se muestra la variación instantánea del ancho de banda máximo. Como puede observarse, el uso del controlador P en SPID, incrementa ligeramente las fluctuaciones obtenidas a lo largo del tiempo en el ancho de banda máximo permitido.

Por tanto, las gráficas anteriores demuestran que SPID consigue garantizar los niveles de ancho de banda estipulados utilizando cualquiera de los tres controladores (P, PI o PID). Sin embargo, el controlador P es ligeramente menos estable que los controladores PI y PID. Además, el uso de un controlador PID no supone un gran incremento en la

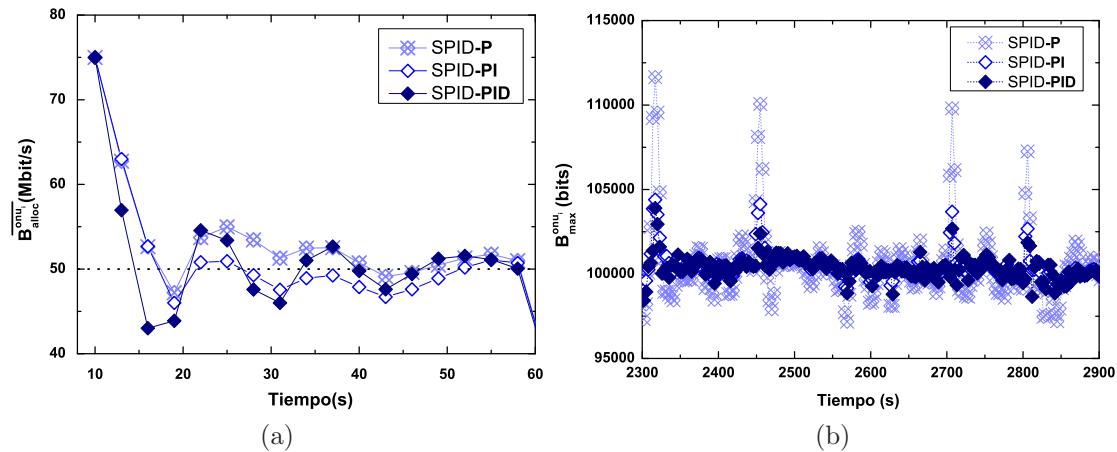


Figura 4.6: Evolución en tiempo real del (a) ancho de banda medio hacia nivel garantizado del SLA₂ (b) ancho de banda máximo permitido para el SLA₂.

complejidad de la red comparado con el controlador PI y además el PID es más completo. Por este motivo, para realizar el análisis de las prestaciones del algoritmo SPID se va a utilizar un controlador PID.

4.4.3. Establecimiento de los parámetros asociados al diseño del controlador en SPID

Además de los parámetros de sintonización del controlador PID, en el algoritmo SPID es necesario determinar el valor óptimo del tiempo de muestreo para obtener un comportamiento adecuado del sistema. El tiempo de muestreo (T en la ecuación (4.2)) marca el tiempo en el que el PID actualiza el error cometido y es un parámetro importante, ya que determina la frecuencia de ejecución del PID. En este sentido, si el valor de T es muy elevado, los máximos se actualizarán con poca frecuencia, por lo que el tiempo de convergencia hacia los niveles estipulados será mayor. Por el contrario, si el tiempo de muestreo es muy pequeño, la adaptación será más difícil ya que las ejecuciones del PID son muy frecuentes y los máximos se actualizarán sin tener en cuenta la influencia de las adaptaciones previas. Por tanto, es necesario llegar a un compromiso en el valor de este parámetro. Además, tal y como se observa en la ecuación (4.2), este parámetro también afecta a la señal de control $u[n]$.

Para seleccionar el valor más adecuado, se han realizado distintas simulaciones. La Figura 4.7 (a) y (b) representa la evolución en tiempo real del ancho de banda medio asignado considerando distintos valores de muestreo T , para el SLA de menor prioridad (SLA₂) y para la carga de red más alta (1.6) en los primeros 100 s. El comportamiento

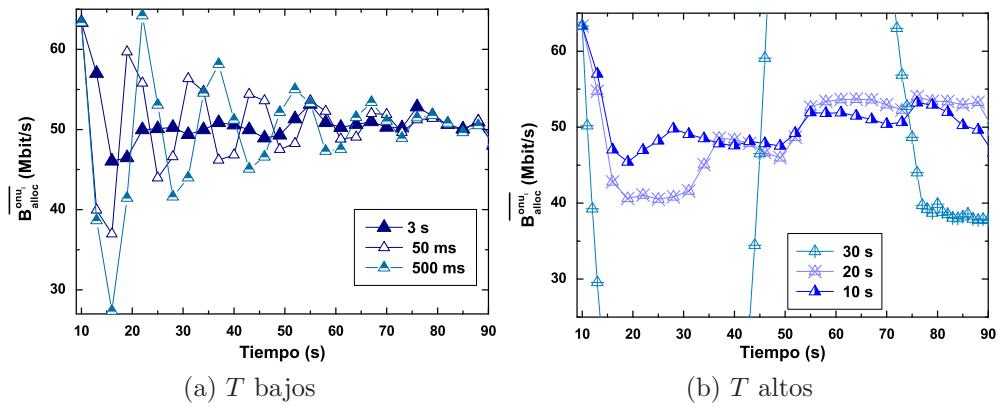


Figura 4.7: Variación en tiempo real del ancho de banda medio asignado del SLA₂ considerando (a) valores bajos de T (b) valores altos de T .

para el resto de SLAs y cargas será similar. En concreto, en la Figura 4.7 (a) se puede ver como para los valores más pequeños que se han considerado (50 y 500 ms) la fluctuación de dicho parámetro es bastante elevada. Sin embargo, en la Figura 4.7 (b) se observa el mismo comportamiento para valores altos de este parámetro (30 y 20 s), ya que valores altos de T generan un incremento sustancial en el término integral (según la ecuación (4.2)) y la señal de control $u[n]$ se ve modificada en gran medida. Este comportamiento puede dar lugar a cambios más abruptos en el ancho de banda, lo cual no es deseable para lograr una estabilidad en el sistema. Por tanto, hay que escoger un valor que permita una evolución estable y al mismo tiempo rápida hacia los niveles de ancho de banda garantizados. Un valor de 3 s (Figura 4.7(a)) presenta el comportamiento deseado y converge con rapidez al ancho de banda garantizado sin mostrar fluctuaciones significativas. Por lo tanto, este será el valor seleccionado.

4.4.4. Análisis de prestaciones de SPID

En los siguientes apartados del capítulo, se realiza un análisis de las prestaciones de SPID comparado con DySGAB bajo distintos escenarios de red. En primer lugar, se va a estudiar la evolución del ancho de banda medio a los niveles estipulados por el proveedor de servicios, comprobando no solo la velocidad de convergencia a los mismos, sino también la estabilidad de los algoritmos. Despu s se demostrar  la independencia de SPID frente a diferentes condiciones iniciales de la red. Seguidamente, se realizar  el an lisis modificando los niveles de ancho de banda garantizados en tiempo real con el fin de demostrar la auto-adaptaci n de SPID y su robustez ante cambios en tiempo real. As  mismo, se analizar  la robustez de ambos algoritmos frente a diferentes valores en

los parámetros de su diseño, tales como el tiempo de muestreo y el tiempo de la ventana deslizante. Además, se comprobará el comportamiento de la estrategia conservadora y agresiva que se ha desarrollado en SPID. Por último, se han realizado simulaciones de SPID en redes EPON y LR-EPON de distinta longitud para demostrar su aplicabilidad en redes de acceso de diferente cobertura.

4.4.4.1. Evolución del ancho de banda a los niveles garantizados por el proveedor de servicios

La principal característica de SPID frente a otros algoritmos ya existentes es que de forma dinámica y adaptativa controla el ancho de banda medio asignado a través de un PID para mantenerlo cercano al nivel de ancho de banda garantizado con gran robustez y estabilidad. Por tanto, SPID reajusta continuamente los anchos de banda máximos permitidos para asignar de forma adaptativa los recursos disponibles, y consecuentemente hacer que el algoritmo funcione independientemente de las condiciones iniciales de la red. Como los anchos de banda máximos de cada ONU están controlados por un PID, estos cambian de acuerdo al error presente, al pasado y a una predicción futura del mismo. Por tanto, la evolución del ancho de banda medio asignado hacia los niveles garantizados es más rápida y adecuada que si los máximos se incrementan y decrecen una cantidad fija (como ocurre con DySGAB). Para analizar las diferencias en el comportamiento entre estos dos algoritmos, la Figura 4.8, 4.9 y 4.10 compara el funcionamiento de SPID frente al de DySGAB cuando este último considera diferentes valores del ancho de banda involucrado en la adaptación del ancho de banda máximo permitido ($B_{reduction}$ definido en el Apartado 3.2.2 del Capítulo 3). Para llevar a cabo esta demostración se ha escogido el escenario genérico de la Tabla 4.2 con los siguientes pesos asociados a cada SLA: $W^{sla_0} = 1$, $W^{sla_1} = 2$ y $W^{sla_2} = 3$. Estos pesos no están ajustados a los anchos de banda garantizados que se han considerado (100, 75 y 50 Mbit/s para el SLA₀, SLA₁ y SLA₂, respectivamente) de forma que se observa la evolución y adaptabilidad de ambos algoritmos hacia los anchos de banda estipulados. Sin embargo, a lo largo de este apartado también se presentarán otros escenarios con diferentes condiciones iniciales para probar la flexibilidad en el comportamiento de SPID frente a DySGAB.

En concreto, la Figura 4.8 (a) representa la evolución del ancho de banda medio ($\overline{B}_{alloc}^{onu_i}$) a lo largo del tiempo para una ONU del SLA₀ hacia su nivel de ancho de banda garantizado (100 Mbit/s) cuando DySGAB considera diferentes cantidades de $B_{reduction}$ (250, 500, 1538 y 3076 bytes). Tal y como se puede observar, la velocidad de convergencia en DySGAB depende fuertemente de este parámetro. De hecho, los valores más bajos

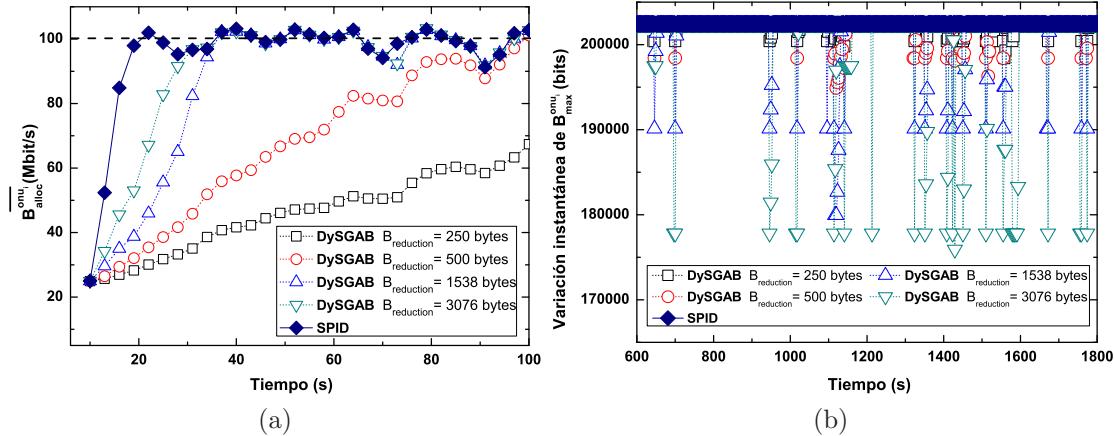


Figura 4.8: Evolución en tiempo real del (a) ancho de banda medio hacia nivel garantizado del SLA₀ (b) ancho de banda máximo permitido para el SLA₀.

implican una convergencia más lenta. Por el contrario, es reseñable destacar que SPID siempre converge de forma más rápida que DySGAB, ya que los máximos se modifican dependiendo del error cometido, esto es, de la diferencia entre el ancho de banda medio asignado y el valor a garantizar. Para completar la comparación, la Figura 4.8 (b) ilustra la variación instantánea del ancho de banda máximo permitido para el perfil SLA₀ en ambos algoritmos y para todos los valores anteriores de $B_{reduction}$. En esta gráfica, se puede observar como el comportamiento más estable del B_{max}^{onu} en DySGAB se obtiene cuando se consideran valores pequeños del término $B_{reduction}$. Sin embargo, SPID siempre presenta un comportamiento más estable que DySGAB independientemente del valor de $B_{reduction}$ seleccionado. De hecho, los valores más grandes de $B_{reduction}$ en DySGAB implican una respuesta más rápida pero también una mayor inestabilidad, que no es deseada. Por tanto, ambas figuras demuestran que la asignación de ancho de banda controlada con un PID proporciona de forma simultánea una respuesta más rápida y estable que la que se obtiene con DySGAB. Por el contrario, en DySGAB se necesita encontrar un compromiso entre rapidez y estabilidad, por lo que su respuesta no es totalmente óptima.

El mismo comportamiento puede observarse en el SLA de prioridad intermedia (SLA₁). En la Figura 4.9 (a) está representada la evolución del ancho de banda medio ($\overline{B}_{alloc}^{onu}$) para los primeros segundos de simulación cuando se compara SPID y DySGAB con diferentes valores de $B_{reduction}$. Como se aprecia, SPID siempre converge antes que DySGAB independientemente de $B_{reduction}$. Además, estas diferencias son aún mayores si se compara SPID con DySGAB utilizando valores bajos del término $B_{reduction}$. En la Figura 4.9 (b) se muestra la variación instantánea del ancho de banda máximo per-

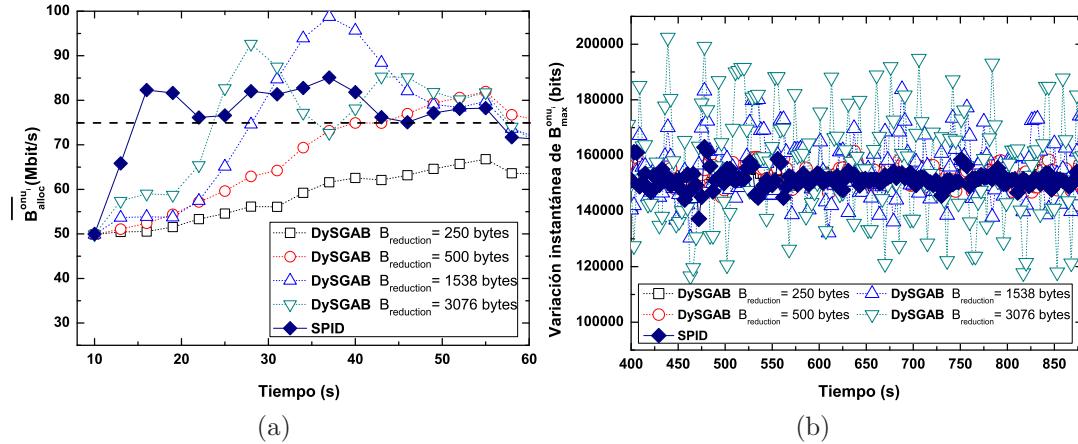


Figura 4.9: Evolución en tiempo real del (a) ancho de banda medio hacia nivel garantizado del SLA₁ (b) ancho de banda máximo permitido para el SLA₁.

mitido. En este caso, se observa que SPID tiene una menor fluctuación que DySGAB independientemente del ancho de banda considerado en la reducción. Así mismo, las fluctuaciones que presenta DySGAB para este SLA son aún mayores que para el SLA de mayor prioridad (SLA₀).

Por último, la Figura 4.10 (a) muestra la evolución del ancho de banda medio asignado al SLA de menor prioridad (SLA₂). Puede verse como SPID alcanza con mayor rapidez el ancho de banda garantizado para este SLA (50 Mbit/s) que DySGAB. Sin embargo, las diferencias entre ambos algoritmos son menores que las observadas en el SLA₀, donde las condiciones de ancho de banda iniciales asignadas en ambos algoritmos

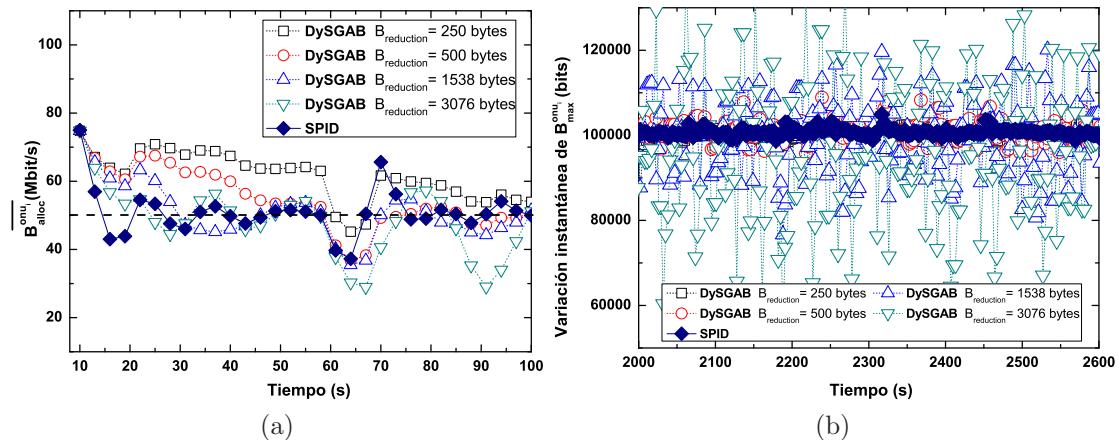


Figura 4.10: Evolución en tiempo real del (a) ancho de banda medio hacia nivel garantizado del SLA₂ (b) ancho de banda máximo permitido a lo largo del tiempo para el SLA₂.

para ese SLA distaban mucho de su nivel de ancho de banda estipulado. Por lo tanto, se aprecia también la adaptación más rápida que obtiene SPID frente a DySGAB en condiciones adversas. Además la Figura 4.10 (b), muestra la variación instantánea del ancho de banda máximo. Como puede observarse, SPID obtiene menores fluctuaciones que DySGAB para todos los valores del término $B_{reduction}$ considerados.

Por tanto, las gráficas anteriores demuestran que para valores pequeños del término $B_{reduction}$ en DySGAB, las fluctuaciones en el ancho de banda máximo permitido ($B_{max}^{onu_i}$) son menores, pero el tiempo de convergencia hacia los valores garantizados es mayor. Por el contrario, valores grandes de $B_{reduction}$ suponen tiempos de convergencia menores pero mayores fluctuaciones de los valores máximos. Sin embargo SPID, para todos los SLAs consigue tiempos de convergencia y fluctuaciones menores para dicho parámetro de red, mostrando, gracias al controlador PID, una mayor rapidez y un comportamiento mucho más estable.

Finalmente, para analizar con mayor detenimiento la estabilidad de ambos algoritmos, se ha representado la desviación estándar del ancho de banda medio asignado con respecto al valor garantizado, considerando diferentes anchos de banda garantizados para cada perfil de usuario. Esta medida proporciona un valor aproximado en Mbit/s de la exactitud con la que cada algoritmo aproxima el ancho de banda medio asignado a cada ONU a su correspondiente nivel de ancho de banda garantizado. Así, la Figura 4.11 (a), (b) y (c) representa la desviación del ancho de banda medio asignado con respecto a diferentes niveles de ancho de banda garantizado para el SLA_0 , SLA_1 y SLA_2 , respectivamente. En todas las gráficas, es importante destacar que SPID es el algoritmo más estable, ya que la desviación es menor que en DySGAB independientemente del valor de $B_{reduction}$ considerado. Además, se puede observar que mayores tamaños de ancho de banda involucrados en la adaptación en DySGAB, dan lugar a mayores niveles de desviación. Este resultado, concuerda con las conclusiones de la Figura 4.8 (b), Figura 4.9 (b) y Figura 4.10 (b), donde se observaba que tamaños mayores en $B_{reduction}$ dan lugar a una mayor inestabilidad en el proceso de asignación de ancho de banda. Por tanto, se ha demostrado que SPID presenta un mejor comportamiento que DySGAB, ya que además de ofrecer una mayor estabilidad, converge más rápido hacia los valores estipulados por el proveedor de servicios.

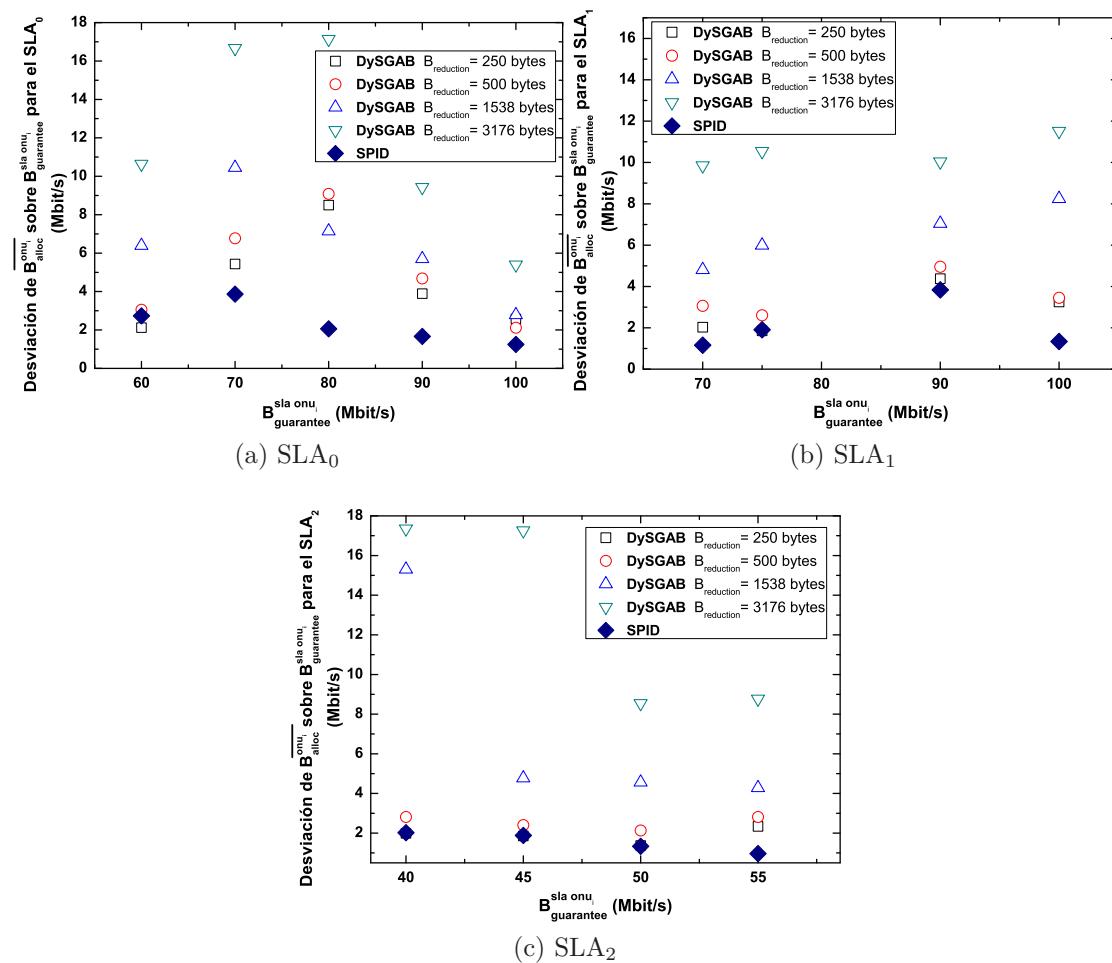


Figura 4.11: Desviación estándar del ancho de banda medio asignado sobre el ancho de banda garantizado. (a) SLA₀ (b) SLA₁ (c) SLA₂.

4.4.4.2. Independencia de las condiciones iniciales de ancho de banda

Por otro lado, tanto DySGAB como SPID han sido diseñados para ser independientes de las condiciones iniciales de ancho de banda máximo asignado, ya que evolucionan en tiempo real hacia los niveles estipulados. Para demostrar esta independencia, se han llevado a cabo simulaciones comparando DySGAB y SPID bajo tres escenarios diferentes, donde ambos algoritmos tienen los mismos valores de ancho de banda máximo inicial, pero éstos difieren de un escenario a otro. Para establecer los anchos de banda iniciales se ha utilizado la ecuación (4.4) asumiendo en cada escenario los pesos de cada SLA (W^{sla_j}) recogidos en la Tabla 4.5.

	W^{sla_0}	W^{sla_1}	W^{sla_2}
Escenario 1	1	1	1
Escenario 2	1	2	3
Escenario 3	2	3	1

Tabla 4.5: Pesos utilizados en los escenarios considerados para establecer los anchos de banda máximos iniciales para los tres perfiles de usuario.

Para ver cómo se comportan tanto DySGAB como SPID en los diferentes escenarios, la Figura 4.12 (a), (b) y (c) muestra la evolución en los primeros segundos de simulación del ancho de banda medio para los tres escenarios y para cada uno de los SLAs cuando en DySGAB se considera un valor de $B_{reduction}$ igual a 1538 bytes. Este valor óptimo de $B_{reduction}$ se escogió en el Capítulo 3, ya que presenta un valor de compromiso entre estabilidad y velocidad de convergencia. Estas gráficas demuestran que aunque ambos algoritmos son capaces de garantizar el ancho de banda estipulado independientemente del escenario considerado, SPID lo hace de forma más rápida para todos los SLAs y escenarios.

4.4.4.3. Adaptación a niveles de ancho de banda garantizados que varían con el tiempo

La adaptabilidad realizada tanto por DySGAB como por SPID es muy adecuada cuando los proveedores de servicio requieren cambios en tiempo real de los anchos de banda garantizados a cada SLA, ya que permiten la adaptación en tiempo real a esos nuevos niveles sin interrupciones en el servicio. Esto resulta muy útil, ya que los proveedores de servicio pueden desear actualizar los requisitos de calidad de ciertos SLAs o incluso el número de SLAs debido a diferentes necesidades, condiciones o demandas de la red. Para demostrar esta ventaja, se han llevado a cabo simulaciones cambiando en tiempo real los niveles de ancho de banda garantizados a los diferentes perfiles de

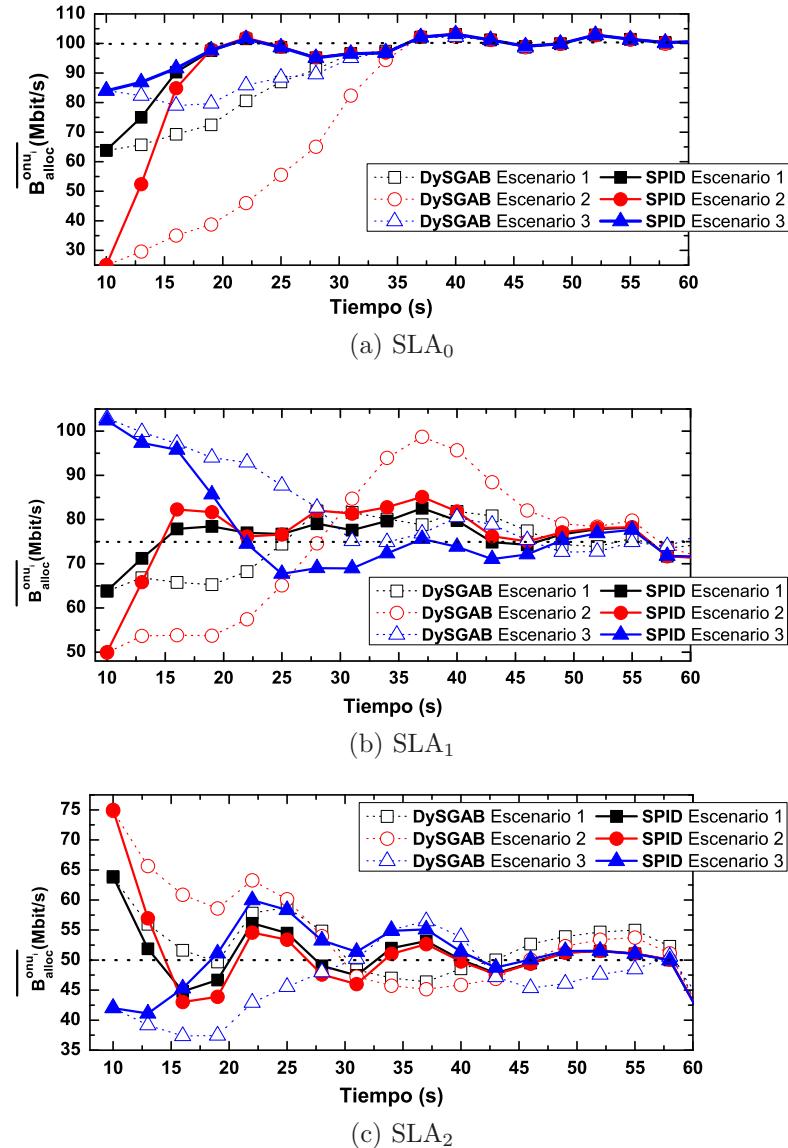


Figura 4.12: Evolución en tiempo real del ancho de banda medio asignado para los tres perfiles de abonado y los tres escenarios considerados en la Tabla 4.5. (a) SLA₀ (b) SLA₁ (c) SLA₂.

abonado. En concreto, los valores cambian cada 300 s para poder observar la adaptación que llevan a cabo ambos algoritmos ante cambios frecuentes de los niveles garantizados. Los valores en los diferentes tramos se muestran en la Tabla 4.6.

Tiempo (s)	Nivel de ancho de banda garantizado a cada SLA (Mbit/s)		
	SLA ₀	SLA ₁	SLA ₂
0-300	100	75	50
300-600	70	100	40
600-900	90	70	55
900-1200	60	90	45

Tabla 4.6: Niveles de ancho de banda garantizado considerados para cada SLA a lo largo del tiempo de simulación.

La Figura 4.13 (a), (b) y (c) caracteriza la evolución en tiempo real del ancho de banda medio asignado en función de los distintos niveles garantizados para el SLA₀, SLA₁ y SLA₂, respectivamente, y considerando distintos valores de $B_{reduction}$ en DySGAB. Tal y como se puede observar en las tres figuras, ambos algoritmos son capaces de adaptar el ancho de banda máximo de tal manera que el ancho de banda medio evolucione a los niveles garantizados. Sin embargo, la velocidad de convergencia en los dos algoritmos difiere mucho. En concreto, tal y como se ha visto a lo largo de este apartado, DySGAB es muy dependiente de la cantidad de ancho de banda involucrada en la adaptación de los máximos ($B_{reduction}$). Por ejemplo, para valores pequeños (250 bytes) la convergencia es tan lenta que, si los cambios de nivel garantizado son muy frecuentes, el ancho de banda puede no adaptarse al nivel estipulado antes del siguiente cambio. Por otro lado, para valores altos (3076 bytes) se mejora la velocidad pero a costa de empeorar la estabilidad del ancho de banda medio, tal y como se observa en cualquiera de las tres figuras. Por el contrario, puede verse como SPID consigue converger hacia los niveles de ancho de banda garantizados de forma más rápida que DySGAB y presenta una respuesta mucho más estable.

En concreto, la Figura 4.13 (a), muestra para el SLA₀ cómo la adaptación cuando el nivel garantizado baja desde 90 Mbit/s a 60 Mbit/s a los 900 s, se realiza casi de forma instantánea en SPID, mientras que para DySGAB, con 250 bytes, se produce aproximadamente en el segundo 1100 (esto es 200 s más tarde del cambio) y para 500 bytes en torno a 1000 s (100 s más tarde del cambio). Otro ejemplo de este comportamiento puede verse para el SLA₂ (Figura 4.13 (c)), donde para pasar a un ancho de banda garantizado mayor (de 40 a 55 Mbit/s en el instante 600 s), DySGAB con 250 bytes emplea casi 100 s y 50 s para 500 bytes. Por el contrario, SPID consigue alcanzar el nuevo nivel garantizado en aproximadamente 15 s.

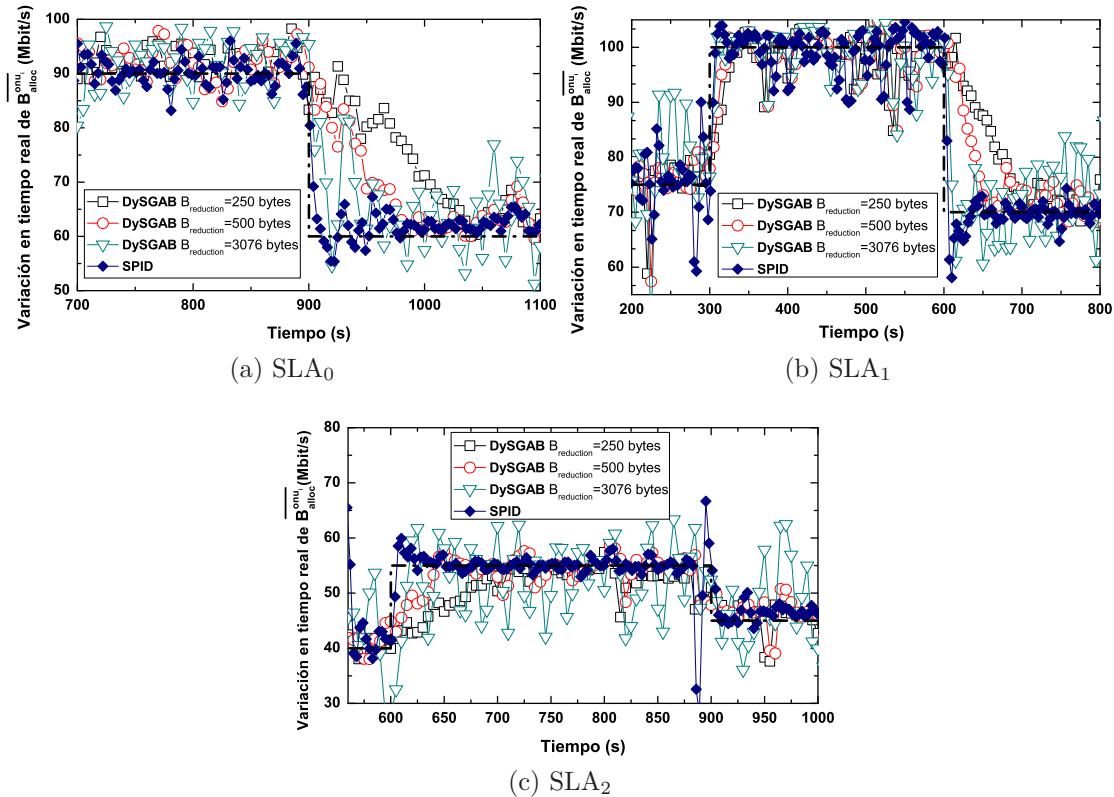


Figura 4.13: Variación en tiempo real del ancho de banda medio asignado a una ONU cuando los niveles garantizados varían a lo largo del tiempo. (a) SLA₀ (b) SLA₁ (c) SLA₂.

Por tanto, se demuestra que aunque ambos algoritmos son capaces de adaptarse a cambios en tiempo real del nivel de ancho de banda garantizado, SPID evoluciona los anchos de banda de forma más rápida y estable a los nuevos valores. Por el contrario, en DySGAB el término $B_{reduction}$ tiene una gran influencia en la rapidez y estabilidad del algoritmo, esto es, valores pequeños hacen que presente una respuesta más estable pero más lenta. Por tanto, en DySGAB es necesario un compromiso entre estabilidad y rapidez, mientras que en SPID, gracias al uso del controlador PID, se tienen simultáneamente ambas características.

4.4.4.4. Robustez de SPID y DySGAB ante parámetros comunes relacionados con el diseño de los algoritmos

En este apartado se va a analizar la robustez de SPID y DySGAB frente a los parámetros propios de los algoritmos y que caracterizan su funcionamiento, esto es, el tiempo de muestreo (T) y el tamaño de la ventana (T_{window}) que se utiliza para calcular

el ancho de banda medio asignado ($\overline{B}_{alloc}^{onu_i}$).

En cuanto al primer parámetro, el tiempo de muestreo (T en la ecuación (4.2)) determina la frecuencia de actualización de los máximos en SPID. De forma similar, para el algoritmo DySGAB este tiempo se definió en el Apartado 3.2.2 del Capítulo 3 como T_{alg} . Sin embargo, ambos tiempos son equivalentes en ambos algoritmos, ya que los dos fijan la frecuencia con la que se adaptan los máximos. Por lo tanto, se va a analizar la diferencia existente entre el comportamiento de DySGAB y de SPID con respecto a este parámetro. En concreto, la Figura 4.14 (a) representa la evolución del ancho de banda medio de DySGAB frente al tiempo cuando se asume un tiempo T_{alg} igual a 500 ms, 2 s, 3 s, 10 s, 20 s para el SLA₂. Así mismo, la Figura 4.14 (b) representa esa misma variación para el algoritmo SPID. Aunque solo se muestran los resultados del SLA₂, el resto de perfiles de abonado presentan el mismo comportamiento. Como se puede observar, SPID es más robusto que DySGAB, ya que SPID presenta una mayor estabilidad del ancho de banda medio asignado para todos los tiempos de muestreo considerados. Esto es, SPID, independientemente del valor escogido para el tiempo de muestreo, consigue asegurar el ancho de banda garantizado de forma más estable que DySGAB, donde se aprecia que para tiempos pequeños las variaciones son aún mucho mayores.

En cuanto al segundo parámetro, el tamaño de la ventana deslizante (T_{window}), DySGAB también presenta una mayor dependencia que SPID. Para ello, la Figura 4.15 representa la evolución del ancho de banda medio asignado para el SLA₂ cuando se comparan DySGAB (con 1538 bytes) y SPID, utilizando diferentes tiempos de ventana (T_{window}), 7.5 s en la Figura 4.15 (a) y 30 s en la Figura 4.15 (b). En ambas figuras puede verse la naturaleza de un controlador PID, alcanzando la situación estable independientemente de la duración de la ventana. Incluso cuando la ventana no tiene un valor óptimo (30 s) SPID consigue estabilizar el ancho de banda medio asignado, mientras que

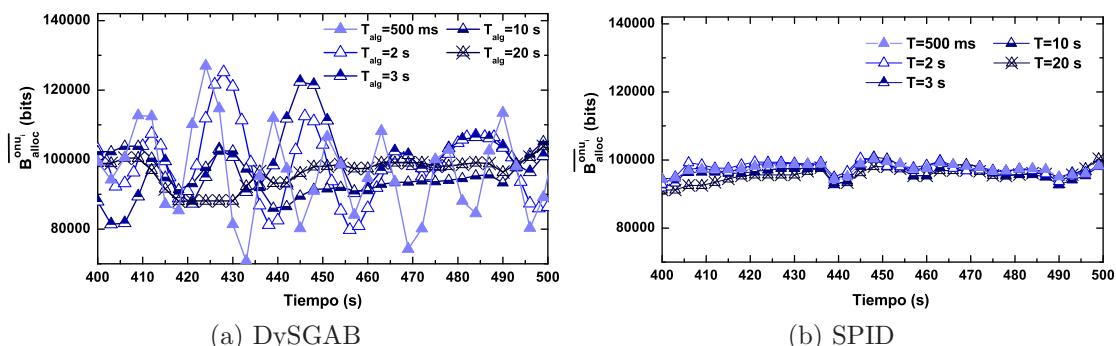


Figura 4.14: Variación en tiempo real del ancho de banda medio asignado a una ONU para diferentes tiempos de ejecución en los algoritmos (T_{alg}, T). (a) DySGAB (b) SPID.

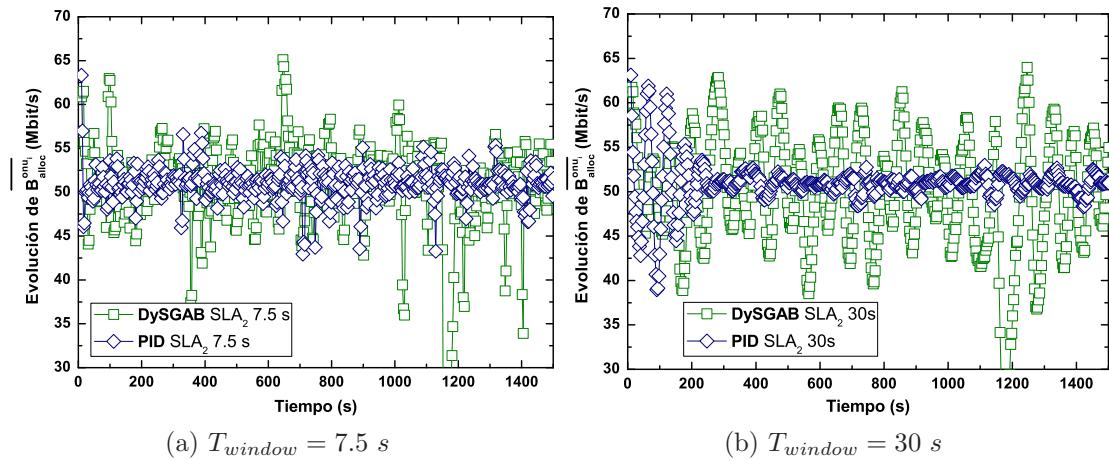


Figura 4.15: Evolución en tiempo real del ancho de banda medio asignado al SLA₂ para DySGAB y SPID considerando distintos tamaños de ventana (T_{window}). (a) 7.5 s (b) 30 s.

DySGAB presenta grandes fluctuaciones en la evolución de dicho parámetro.

Por tanto, se demuestra así que SPID presenta una mayor robustez que el algoritmo DySGAB frente a los parámetros comunes que definen el diseño de ambos algoritmos. Por el contrario, DySGAB presenta una fuerte dependencia con respecto a estos parámetros, ya que necesita una selección óptima de los mismos para asegurar un buen funcionamiento.

4.4.4.5. Comparativa de las estrategias de asignación de ancho de banda en SPID

Tal y como se mencionó en el Apartado 4.3.3 de este capítulo, SPID puede adoptar diferentes políticas para cumplir con los anchos de banda estipulados para cada SLA. En concreto, se han diseñado dos estrategias, conservativa y agresiva, para asignar el ancho de banda cuando no se pueden garantizar los niveles estipulados de todos los SLAs. En la estrategia conservativa, todos los SLAs son tratados de igual forma, así que ningún SLA cumplirá con su ancho de banda garantizado y tampoco se penalizará un SLA frente a los otros. Por el contrario, la política agresiva trata que los usuarios de mayor prioridad siempre cumplan con sus niveles garantizados mientras que el SLA de menos prioridad se verá penalizado.

Para comprobar el funcionamiento de ambas estrategias, en primer lugar se ha considerado el escenario básico de la Tabla 4.2. En este escenario, denominado Escenario 1, existen recursos suficientes para garantizar el ancho de banda para todas las ONUs, por lo que ambas estrategias deberían presentar un comportamiento similar. Por otro lado,

se han considerado otros dos nuevos escenarios donde no existen recursos suficientes para garantizar el ancho de banda de todos los usuarios. Para ello se ha considerado el escenario genérico de la Tabla 4.2 pero modificando el número de ONUs asociadas a cada SLA. La nueva configuración de los tres escenarios aparece recogida en la Tabla 4.7.

	Número de ONUs asociadas a cada SLA		
	SLA ₀	SLA ₁	SLA ₂
Escenario 1	1	5	10
Escenario 2	4	4	8
Escenario 3	2	6	8

Tabla 4.7: Número de ONUs asociadas a cada perfil de abonado en los escenarios considerados.

La Figura 4.16 muestra el ancho de banda ofrecido a cada ONU de cada SLA para los escenarios considerados en la Tabla 4.7 y para cargas de red elevadas (ONUs generando desde 75 Mbit/s hasta 100 Mbit/s). Como puede observarse, para el Escenario 1 donde el ancho de banda disponible es suficiente para ofrecer los anchos de banda garantizados de todas las ONUs, ambas estrategias presentan un comportamiento análogo. Para el Escenario 2 y el Escenario 3, sí se observan diferencias en el comportamiento de ambas

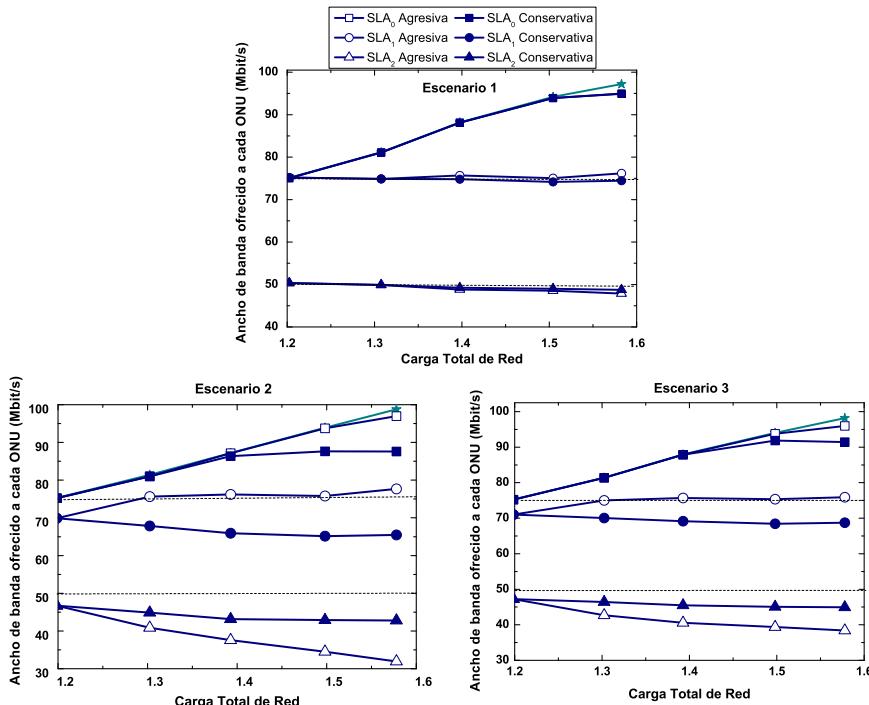


Figura 4.16: Ancho de banda ofrecido a cada ONU para las políticas conservativa y agresiva bajo las condiciones de los tres escenarios de la Tabla 4.7.

estrategias. La estrategia agresiva consigue cumplir con los niveles de ancho de banda estipulados para los usuarios de alta prioridad (SLA_0 y SLA_1) en ambos escenarios de red, en detrimento del SLA de menor prioridad (SLA_2). Por el contrario, cuando se aplica la estrategia conservativa, se puede observar que en ninguno de los dos escenarios se cumple con el ancho de banda garantizado para cada SLA. Sin embargo, el ancho de banda asignado al SLA de menor prioridad (SLA_2) es mayor que en el de la estrategia agresiva, ya en este caso, ningún SLA es más penalizado que otro.

4.4.4.6. Análisis de SPID considerando redes de distinta longitud

Por último, es importante destacar que SPID puede aplicarse a redes EPON y LR-EAPON con diferentes distancias. Para demostrar esto, la Figura 4.17 representa el ancho de banda demandado y ofrecido para una ONU de cada SLA cuando se consideran distancias desde 20 km hasta 120 km en saltos de 20 km. Los anchos de banda garantizados se han fijado a 100/75/50 Mbit/s para el SLA_0 , SLA_1 y SLA_2 , respectivamente. Tal y como se observa, para todas las distancias consideradas, los resultados demuestran que el comportamiento de SPID no depende de la distancia. Por tanto, las ventajas del algoritmo SPID se obtienen también en redes EPON y LR-EAPON de diferente cobertura, y por lo tanto su aplicabilidad en estas redes puede ser directa.

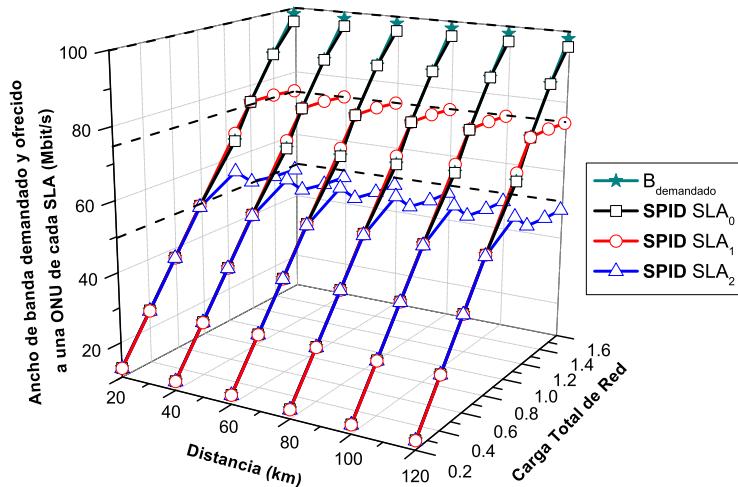


Figura 4.17: Ancho de banda demandado y ofrecido a una ONU de cada SLA para redes EPON y LR-EAPON de distinta distancia.

4.5. Análisis de prestaciones de SPID bajo patrones de tráfico asimétrico y dinámico

En este apartado se van a analizar las prestaciones del algoritmo SPID bajo condiciones de tráfico más reales, tales como tráfico asimétrico y tráfico dinámico. De esta forma se pretende demostrar la robustez de SPID ante diferentes patrones de tráfico.

4.5.1. Escenarios con un patrón de tráfico asimétrico

Para realizar el análisis se han llevado a cabo simulaciones en el escenario básico de la Tabla 4.2, pero considerando que cada ONU transmite a una tasa aleatoria entre 0 y 100 Mbit/s (máxima tasa de transmisión) y limitando la contribución de cada SLA a la carga total de red, al igual que se hizo para DySGAB en el Apartado 3.4.1 del Capítulo 3. En concreto, la contribución de cada SLA en la carga total de los dos nuevos escenarios considerados aparece recogida en la Tabla 4.8.

	Carga de red	SLA ₀	SLA ₁	SLA ₂
Escenario 1	1.0	10 %	40 %	50 %
Escenario 2	1.2	4 %	30 %	66 %

Tabla 4.8: Contribución de cada SLA a la carga total de red para los distintos escenarios considerados con tráfico asimétrico en SPID.

La Figura 4.18 muestra el ancho de banda medio asignado a cada ONU por SPID y DySGAB bajo las condiciones del Escenario 1. En este escenario, la carga de red es de 1.0 (el ancho de banda demandado por todas las ONUs es de 1 Gbit/s), donde el 10 % es generado por la ONU del SLA₀, el 40 % se genera aleatoriamente entre las cinco ONUs del SLA₁ y el 50 % restante entre las diez ONUs del SLA₂. Por tanto, el ancho de banda demandado por cada ONU (representado en la figura por una estrella) es diferente. Como se puede apreciar, el comportamiento de ambos algoritmos es similar. En este sentido, tanto SPID como DySGAB son capaces de garantizar el ancho de banda estipulado por el proveedor de servicios para aquellas ONUs cuya demanda supera dicho valor. Incluso, ambos algoritmos asignan un valor mayor de su nivel garantizado ya que existe ancho de banda en exceso. Por otro lado, ambos algoritmos asignan todo el ancho de banda demandado a aquellas ONUs con una demanda menor a su nivel de ancho de banda garantizado.

De forma similar, la Figura 4.19, muestra el ancho de banda demandado y ofrecido por SPID y DySGAB para cada ONU en las condiciones del Escenario 2. En este caso, la carga de red total del escenario es de 1.2 (la demanda total de las ONUs es de 1.20 Gbit/s)

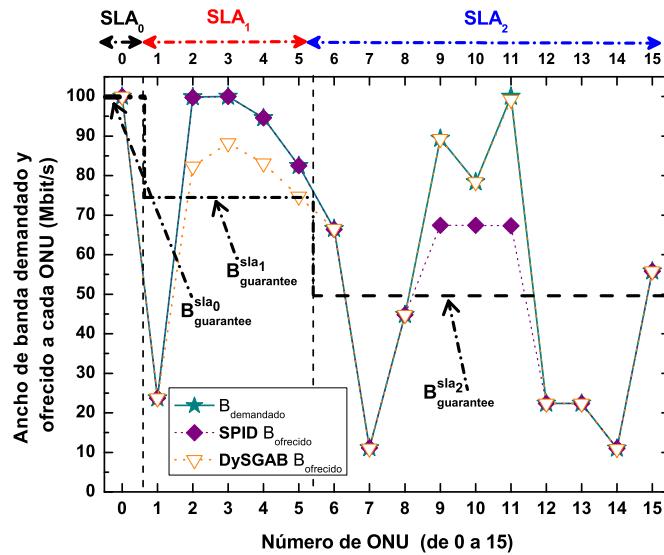


Figura 4.18: Ancho de banda demandado y ofrecido a cada ONU por SPID y DySGAB para las condiciones del Escenario 1 de la Tabla 4.8.

donde el 4% es generado por la ONU del SLA₀, el 30% se genera aleatoriamente entre las cinco ONUs del SLA₁ y el 66% restante entre las diez ONUs del SLA₂. Al igual que para el Escenario 1, en este escenario tanto SPID como DySGAB son capaces de garantizar el ancho de banda mínimo garantizado en aquellas ONUs cuya demanda es mayor que dicho nivel. Además, ambos algoritmos ofrecen el ancho de banda solicitado

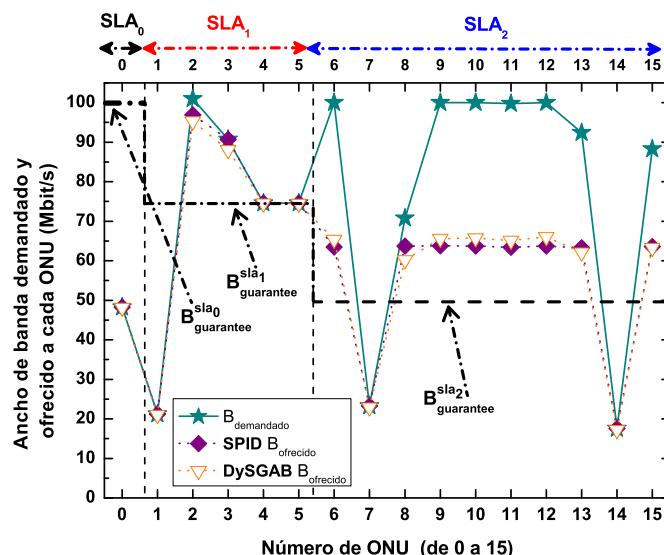


Figura 4.19: Ancho de banda demandado y ofrecido a cada ONU por SPID y DySGAB para las condiciones del Escenario 2 de la Tabla 4.8.

a aquellas ONUs cuya demanda es menor que su nivel garantizado, presentando ambos algoritmos un comportamiento análogo.

Por tanto, se demuestra que SPID, al igual que DySGAB, mantiene su funcionalidad bajo un patrón de tráfico asimétrico, lo cual demuestra su robustez ante un escenario con una distribución de tráfico más real.

4.5.2. Escenarios con un patrón de tráfico dinámico

Para analizar el comportamiento de SPID con tráfico dinámico se ha llevado a cabo el mismo estudio que el presentado en el Apartado 3.4.2 del Capítulo 3 para DySGAB. Por tanto, la variación de la carga a lo largo del tiempo sigue la ecuación (4.5). Además se han escogido dos valores para la frecuencia con la que se cambia el valor de la carga dinámica, para así comprobar el nivel de reacción de SPID ante cambios más rápidos y cambios más lentos de la carga. En concreto, se han escogido los valores de 1800 s y de 900 s al igual que en el Apartado 3.4.2 del Capítulo 3.

$$\rho_{onu}(t) = 0.6 \cdot \left[1 + \frac{1}{2} \sin \left(2\pi \left(\frac{t + 3600}{21600} \right) \right) \right] \quad (4.5)$$

La Figura 4.20 (a), (b) y (c) representa la variación en SPID y en DySGAB del ancho de banda medio asignado a lo largo del tiempo para una ONU del SLA₀, SLA₁ y SLA₂, respectivamente, cuando la frecuencia de los cambios en la carga es de 1800 s. En azul está representada la variación del ancho de banda demandado por cada ONU en Mbit/s, esto es, la variación de la carga de la ONU (referida al eje de la derecha también en azul). En negro aparece la variación del ancho de banda para DySGAB y en verde para SPID (ambos valores referidos al eje de la izquierda). Tal y como puede observarse, el comportamiento de ambos algoritmos es similar y es el esperado. Esto es, para los tres SLAs el ancho de banda medio asignado se corresponde con el ancho de banda demandado, siempre que existan recursos suficientes. En el caso de que la demanda total del ancho de banda sea mayor que la capacidad del canal, ambos algoritmos aseguran que el ancho de banda medio sea igual a su nivel garantizado (100, 75 y 50 Mbit/s para cada SLA, respectivamente).

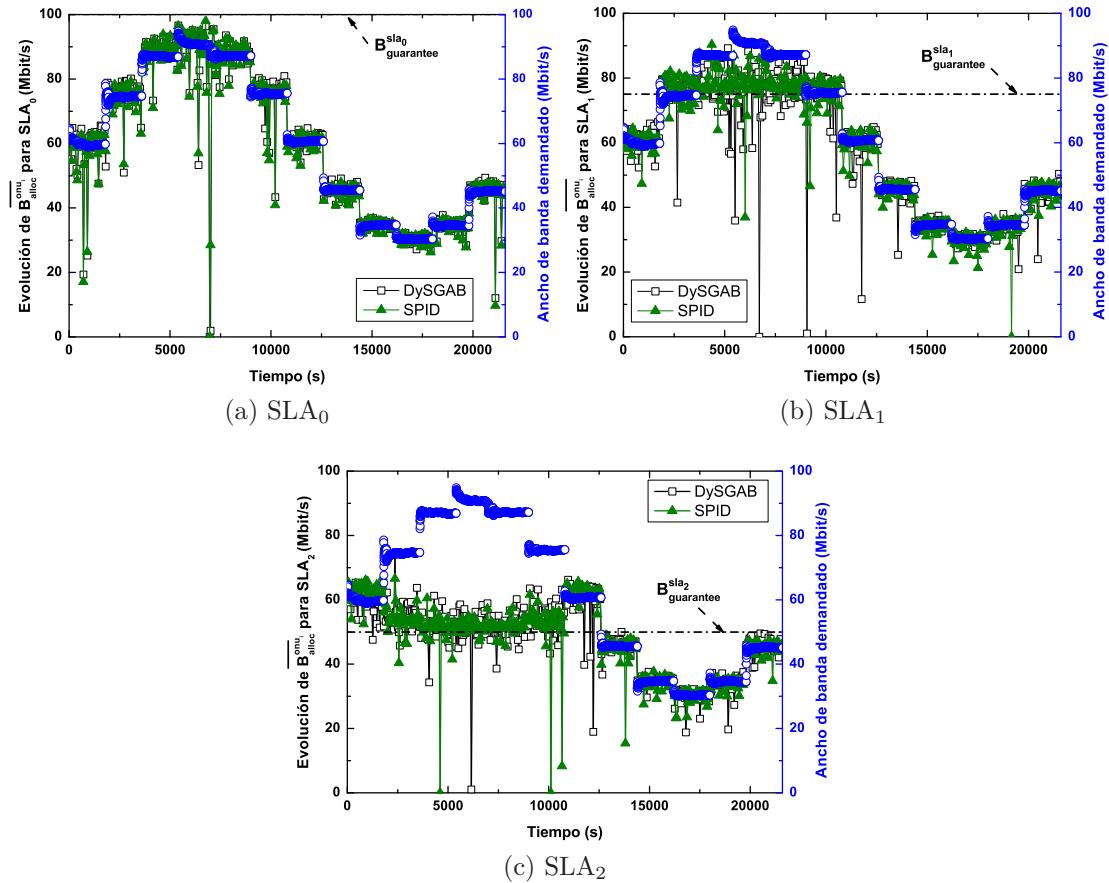


Figura 4.20: Ancho de banda demandado por una ONU (azul) y variación en tiempo real del ancho de banda medio asignado en DySGAB (negro) y SPID (verde) para una frecuencia de cambio de 1800 s. (a) SLA₀ (b) SLA₁ (c) SLA₂.

Por otro lado, la Figura 4.21 (a) (b) y (c) representa la variación en tiempo real del ancho de banda máximo (término B_{max}) en DySGAB y SPID para el SLA₀, SLA₁ y SLA₂, respectivamente, y considerando una frecuencia de cambio de 1800 s. Al igual que en la figura anterior, en azul está representada la variación del ancho de banda demandado por cada ONU en Mbit/s (referida eje de la derecha también en azul). En negro aparece la variación del ancho de banda máximo para DySGAB y en verde para SPID (ambos valores referidos al eje de la izquierda). Tal y como puede observarse, SPID presenta un comportamiento más estable, ya que tal y como se observa, la variación de los anchos de banda máximos es menor que en DySGAB para todas las cargas de red.

Para comparar con más detalle ambos algoritmos, la Figura 4.22 (a), (b) y (c) muestra la desviación estándar del ancho de banda medio asignado con respecto al ancho de banda garantizado para el SLA₀, SLA₁ y SLA₂, respectivamente. Para el cálculo de la

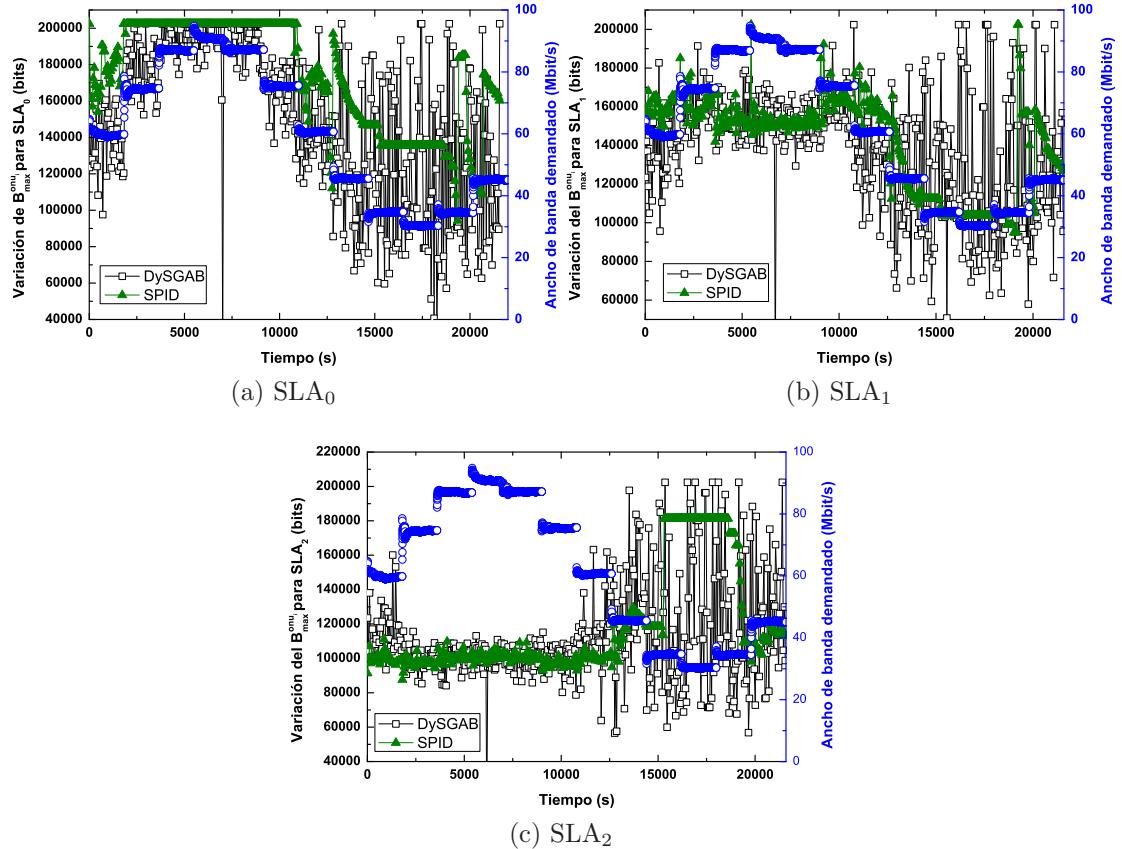


Figura 4.21: Ancho de banda demandado por una ONU (azul) y variación en tiempo real del ancho de banda máximo en DySGAB (negro) y en SPID (verde) para una frecuencia de cambio de 1800 s. (a) SLA₀ (b) SLA₁ (c) SLA₂.

desviación, si el ancho de banda demandado es menor que el garantizado (para cargas de red bajas), se ha calculado la desviación con respecto al ancho de banda demandado. Por el contrario, si la demanda es mayor que el nivel garantizado, la desviación se ha calculado sobre dicho valor. Tal y como puede observarse, la desviación obtenida por SPID es menor que en DySGAB en todos los casos independientemente de la carga y para todos los SLAs, demostrando SPID una mayor robustez y estabilidad que DySGAB.

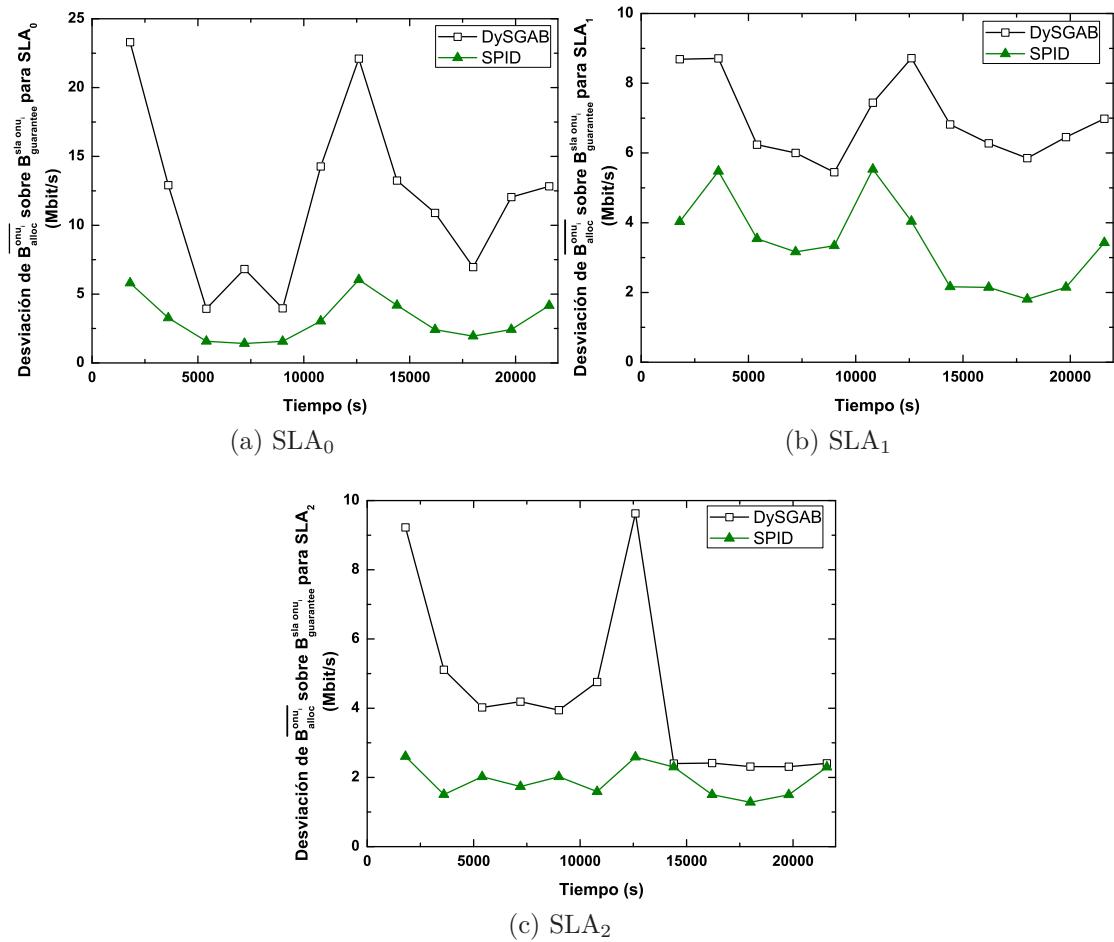


Figura 4.22: Desviación del ancho de banda medio asignado con respecto a nivel garantizado o nivel demandado en DySGAB (negro) y SPID (verde) para una frecuencia de cambio de 1800 s. (a) SLA₀ (b) SLA₁ (c) SLA₂.

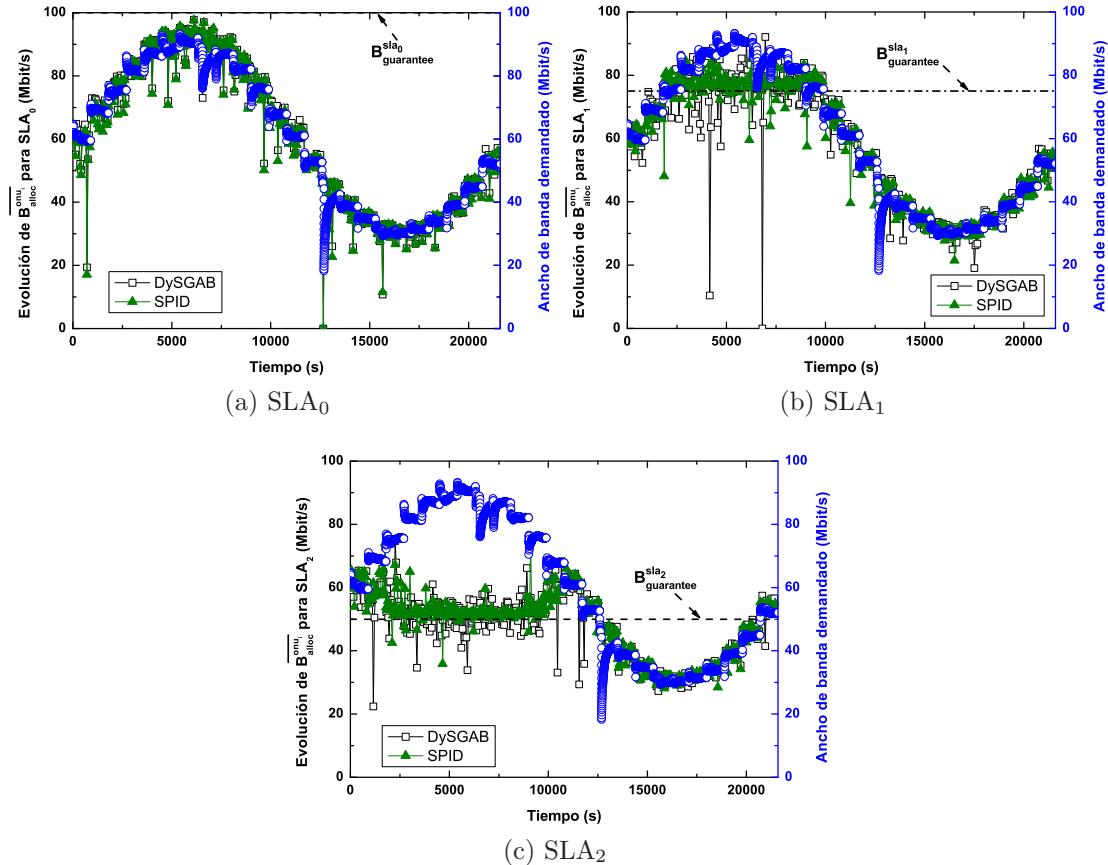


Figura 4.23: Ancho de banda demandado por una ONU (azul) y variación en tiempo real del ancho de banda medio asignado en DySGAB (negro) y SPID (verde) para una frecuencia de cambio de 900 s. (a) SLA₀ (b) SLA₁ (c) SLA₂.

Por otro lado, la Figura 4.23 y la Figura 4.24 representan la variación del ancho de banda medio y del ancho de banda máximo (representadas en negro para DySGAB, en verde para SPID y referidos ambos valores al eje de la izquierda) respectivamente, según la variación de la carga de ONU (representada en azul y referida al eje de la derecha) para una frecuencia de cambio de la carga de 900 s. Al igual que en el caso anterior, ambos algoritmos adaptan el ancho de banda medio asignado al ancho de banda demandado, y cuando la demanda total de todas las ONUs supera la capacidad del canal, se asegura el nivel garantizado establecido por el proveedor de servicios. Por otro lado, tal y como ocurría para la frecuencia de 1800 s, para 900 s SPID presenta una menor variación de los anchos de banda máximos que DySGAB para todas las cargas.

Por último, la Figura 4.25 (a), (b) (c) representa la desviación del ancho de banda medio asignado sobre el nivel garantizado para cada SLA considerado, cuando se realizan

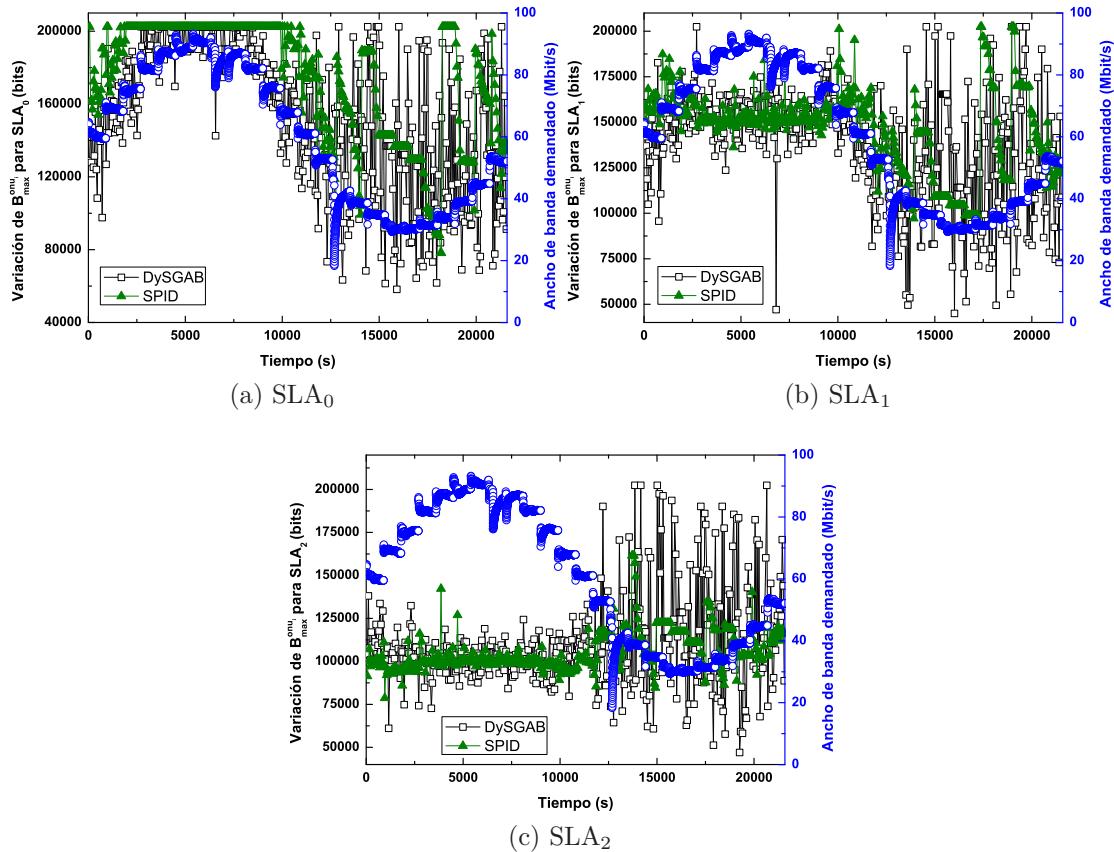


Figura 4.24: Ancho de banda demandado por una ONU (azul) y variación en tiempo real del ancho de banda máximo en DySGAB (negro) y en SPID (verde) para una frecuencia de cambio de 900 s. (a) SLA₀ (b) SLA₁ (c) SLA₂.

cambios cada 900 s. De forma similar que para 1800 s, la desviación de SPID se mantiene por debajo de la de DySGAB en todo momento, por lo que SPID presenta un mejor comportamiento que DySGAB cuando los cambios se producen de forma más rápida.

Por tanto, con este análisis puede concluirse que ambos algoritmos, SPID y DySGAB, son capaces de garantizar el ancho de banda estipulado por el proveedor de servicios bajo escenarios con condiciones de tráfico más reales, como el tráfico dinámico, independientemente de la frecuencia de los cambios en la carga. Sin embargo, el algoritmo SPID presenta más robustez que el algoritmo DySGAB, ya que tanto la variación de los anchos de banda máximos como la desviación con respecto al ancho de banda garantizado son menores en SPID.

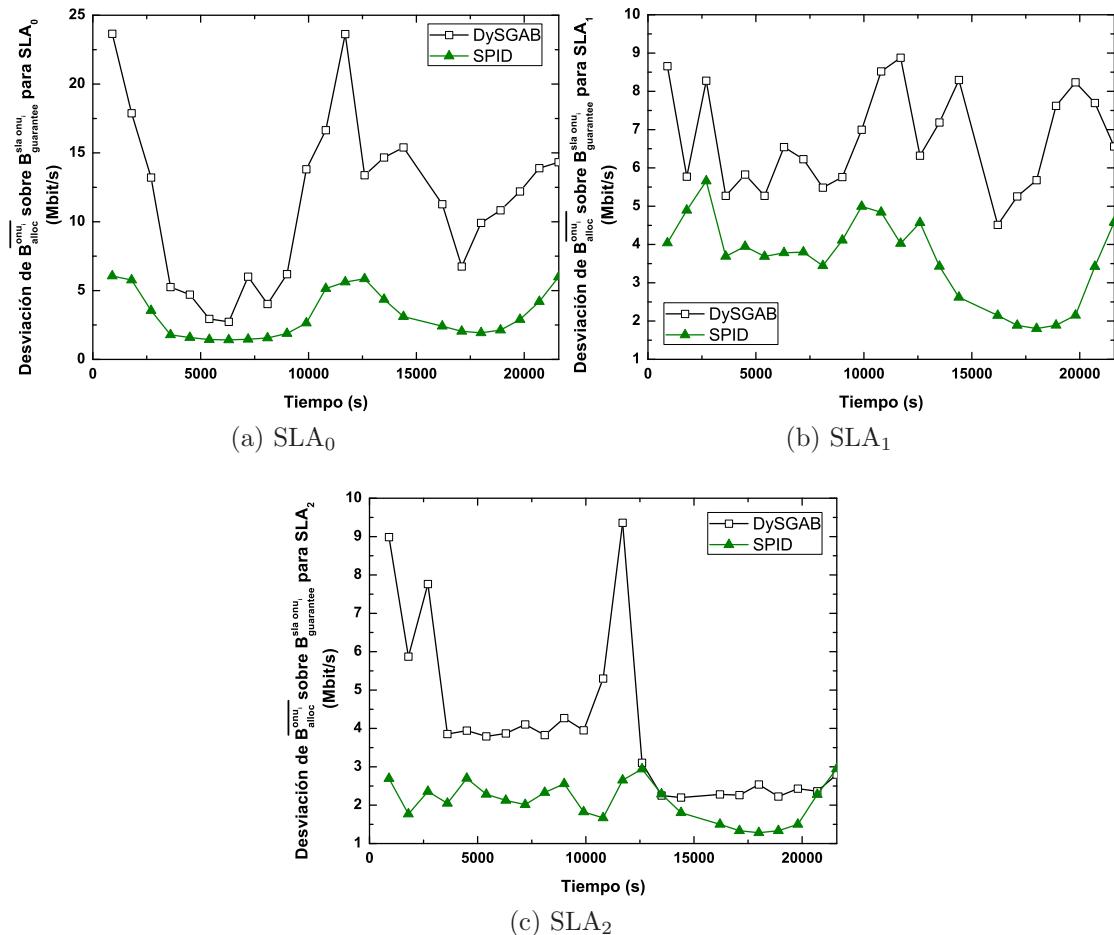


Figura 4.25: Desviación del ancho de banda medio asignado con respecto a nivel garantizado o nivel demandado en DySGAB (negro) y SPID (verde) para una frecuencia de cambio de 900 s.
 (a) SLA₀ (b) SLA₁ (c) SLA₂.

4.6. Controlador PID en el control del retardo. Algoritmo DaSPID (*Delay aware Service level agreement PID*)

En este apartado del capítulo, se va a realizar la descripción del diseño e implementación de un algoritmo que controla el retardo medio del tráfico mediante el uso de controladores PID. En una red EPON y LR-PON, tal y como se comentó en el Capítulo 2, se puede llevar a cabo una diferenciación de servicios o de clientes, pero para ofrecer de forma más completa una mejor calidad de servicio en el canal de acceso compartido, es necesario una combinación adecuada de ambos paradigmas. Sin embargo, tal y como se vio en el análisis del estado del arte del Apartado 2.3.4 del Capítulo 2, solo dos algoritmos aseguran cotas máximas al retardo medio para tráfico con prioridad alta y media y además realizan de forma simultánea diferenciación de usuarios. Esto es, dependiendo de la prioridad del SLA estos algoritmos distinguen entre diferentes niveles de retardo mínimo garantizado para dos clases de servicio prioritario.

El primer algoritmo [78] utiliza en la asignación del ancho de banda dos algoritmos diferentes, uno para la clase de tráfico más prioritaria y otro para la de prioridad media. Esto tiene como desventaja una mayor complejidad en el proceso de asignación de ancho de banda, además de que puede dar lugar a problemas de escalabilidad si se quiere extender el control del retardo a otros tipos de tráfico. Así mismo, se permite el envío de los paquetes de la clase de servicio más restrictiva antes de que llegue a la ONU el mensaje *Gate* con su asignación de ancho de banda para el siguiente ciclo, ya que considera que estos paquetes se generan a una tasa constante. Además, para llevar a cabo la diferenciación de usuarios, a las ONUs que pertenecen a los SLAs más prioritarios se les asigna ancho de banda con una mayor frecuencia. Esto es, mientras que a las prioritarias siempre se les asigna ancho de banda en cada ciclo de transmisión, no todas las ONUs con baja prioridad transmiten en todos los ciclos. De esta forma, el retardo que obtienen estas ONUs prioritarias es menor que el de las ONUs de menor prioridad.

El segundo algoritmo, DySLa (*Dynamic Service Level agreement algorithm*) [79], consigue que el retardo medio de los servicios más sensibles se sitúe por debajo de un límite máximo permitido para cada SLA, distribuyendo el ancho de banda disponible a partir de un conjunto de pesos que se van a ir modificando en tiempo real. Sin embargo, esta evolución de los pesos se realiza restando una cantidad fija, denominada α , del peso asociado a los abonados que cumplen con los requisitos y sumando esa misma cantidad al peso de los abonados que no cumplen con las restricciones. Esta cantidad se mantiene siempre fija, lo cual puede no resultar óptimo para adaptar los retardos en todos los escenarios de red. Por ejemplo, si α tiene un valor pequeño y el error cometido a la hora

de acotar el retardo es muy elevado, DySLa tardará más tiempo en conseguir mantener los retardos por debajo del límite máximo estipulado. Además, una elección errónea del valor de α por parte del proveedor de servicios puede llevar no solo a una mayor lentitud en la consecución de los objetivos, sino incluso a un mal funcionamiento del algoritmo, de forma que no se cumplan las restricciones de retardo máximo establecidas.

Por tanto, se ha diseñado y desarrollado un algoritmo basado en un sistema de control PID que soluciona los inconvenientes que presentan los algoritmos hasta ahora existentes, especialmente DySLa, ya que la modificación del ancho de banda asignado a cada usuario dependiendo de su SLA se hace en función de la diferencia entre el retardo obtenido y el retardo estipulado para los servicios más sensibles, es decir, en función del error cometido. Este algoritmo se ha denominado DaSPID (*Delay aware Service level Agreement PID*) [111].

En los siguientes apartados del capítulo, se va a describir el funcionamiento de este nuevo algoritmo y a analizar sus prestaciones a través de un estudio de simulación. Para ello, en primer lugar, se describe como se realiza la diferenciación de servicios y de usuarios en DaSPID. Después, se explicará el proceso de asignación de ancho de banda, así como la adaptación dinámica que se lleva a cabo para garantizar que el retardo de las clases de servicio prioritarias se sitúe por debajo del límite máximo estipulado según la prioridad del SLA contratado. A continuación, se realiza el proceso de sintonización del PID en DaSPID y el análisis a nivel de simulación para decidir el tipo de controlador que mejor se adapta al cometido de DaSPID. Por último, se establecerán los valores óptimos de los parámetros relacionados con el diseño del algoritmo y se evaluarán sus prestaciones y comportamiento para diferentes condiciones de red.

4.6.1. Diferenciación de servicios y de perfiles de abonado en DaSPID

Para llevar a cabo la diferenciación de servicios en DaSPID se va a utilizar un esquema de prioridad de colas (*Priority Queue*) con compartición de memoria al igual que en DySLa [79] y otros algoritmos de la literatura [17, 112]. En este esquema, las ONUs clasifican y separan el tráfico entrante en diferentes colas según su prioridad. Por tanto, los paquetes, dependiendo de la clase de servicio a la que pertenezcan, son insertados en sus correspondientes colas de prioridad, compartiendo todas las colas de los diferentes servicios la misma memoria virtual. De este modo, mientras la memoria no se encuentre llena, los paquetes son insertados en la cola correspondiente a su clase de servicio. Pero si la memoria total se encuentra llena, un paquete entrante podría ser insertado en su cola, siempre y cuando existan paquetes de menor prioridad que puedan ser borrados. Así pues, si la memoria está llena y llega un nuevo paquete, las colas se comprueban de

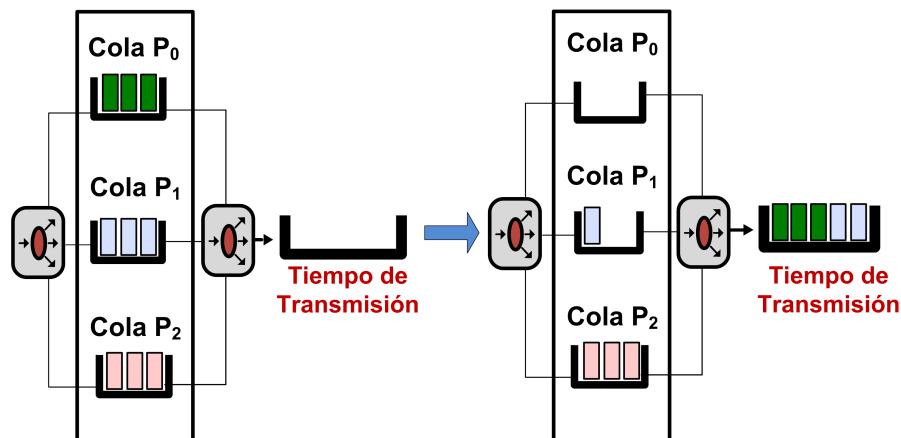


Figura 4.26: Método de extracción de paquetes basado en el esquema de prioridad de colas estrictas (*Strict Priority Queue*).

menor a mayor prioridad hasta la prioridad del paquete recibido, analizando si tienen o no paquetes. En la primera cola de menor prioridad que tenga paquetes almacenados se borra uno de ellos para insertar el entrante de mayor prioridad. De esta forma, se consigue mejorar la probabilidad de pérdida de paquetes del tráfico más sensible.

Aparte del método anterior, que se encarga de organizar el tráfico dentro de las ONUs, es necesario otro esquema para planificar la transmisión de los paquetes. En este caso, se va a utilizar un sistema de prioridad de colas estrictas (*Strict Priority Queue*). Este esquema con prioridad estricta tiene un controlador interno localizado dentro de la ONU que extrae paquetes de cada cola durante el tiempo de transmisión asignado por el OLT en un ciclo, siempre y cuando las colas de prioridad más elevada estén vacías [113]. Este esquema se muestra en la Figura 4.26. En este caso P₀ representa la clase de servicio con mayor prioridad y P₂ la menos prioritaria. Este esquema de transmisión otorga una mayor preferencia al tráfico con la prioridad más alta, lo que resulta beneficioso para servicios muy sensibles al retardo, siendo además un esquema de fácil implementación. Este esquema, además se emplea típicamente en métodos de asignación de ancho de banda limitado [113] como el utilizado en DaSPID, con un número de usuarios elevado.

Por otro lado, la diferenciación de usuarios en DaSPID se va a llevar a cabo igual que los algoritmos SPID y DySGAB y otros algoritmos existentes en la literatura [56, 64, 65]. Esto es, mediante la contratación de SLAs por parte de los usuarios finales. Sin embargo, a diferencia de los algoritmos anteriores, en este caso no se estipula de forma directa un ancho de banda determinado para garantizar. Como lo que se pretende garantizar son retardos máximos para cada tipo de tráfico, se van a establecer diferentes cotas de retardo dependiendo tanto de los requisitos estipulados por los estándares para

cada clase de servicio, como de la prioridad del perfil o SLA contratado, al igual que en [78, 79]. Por tanto, dependiendo no solo de la prioridad del SLA, sino también del tipo de tráfico que soportan, se establecerán unos niveles de retardo adecuados para que en ningún caso se sobrepasen los retardos máximos impuestos por los estándares, para así garantizar un buen nivel de calidad de servicio y calidad de experiencia percibida por los usuarios finales.

4.6.2. Asignación y adaptación del ancho de banda en DaSPID para controlar el retardo

El algoritmo DaSPID sigue el mismo esquema de asignación de ancho de banda que los algoritmos SPID y DySLa. Es por tanto un algoritmo de *polling* o de testeo continuo donde el OLT asigna a cada ONU el ancho de banda de forma individualizada en cada ciclo cuando recibe su mensaje *Report*. También implementa un esquema limitado, siguiendo por tanto la ecuación (4.3) para el cálculo del ancho de banda a asignar a cada ONU. Por otro lado, DaSPID realiza la diferenciación de usuarios dependiendo de la prioridad del SLA contratado, por lo que el ancho de banda máximo permitido para cada ONU ($B_{max}^{onu_i}$ en la ecuación (4.3)) se calcula en función de dicha prioridad.

Así mismo, para determinar los anchos de banda máximos iniciales en DaSPID, se va a utilizar la ecuación (4.4), al igual que en DySLa [79] para realizar una comparación más realista de los dos algoritmos.

Para implementar el proceso de adaptación del máximo ancho de banda permitido a cada ONU ($B_{max}^{onu_i}$), DaSPID, al igual que SPID, va a utilizar un controlador PID, debido al buen funcionamiento que ha demostrado en este último algoritmo. Consecuentemente, la novedad que implementa DaSPID frente a DySLa, es que en DaSPID los valores de $B_{max}^{onu_i}$ se van a ir modificando cada T segundos mediante un controlador PID en función del error cometido a la hora de garantizar los retardos de cada una de las clases de servicio prioritarias. Mientras, en DySLa, estos valores se calculan siguiendo la ecuación (4.4) a partir de los pesos W^{sla_j} , que son los que se modifican una cierta cantidad fija (α) ciclo tras ciclo para garantizar que los retardos se sitúen por debajo de las cotas máximas.

Por tanto, el diagrama de bloques de un proceso controlado por un PID representado en la Figura 4.1 es directamente aplicable al algoritmo DaSPID. En el caso de DaSPID, como controla el nivel de retardo en vez del ancho de banda medio asignado, el punto de referencia y_0 se va corresponder con el retardo máximo estipulado por el proveedor de servicios para las j clases de servicio con restricciones en el retardo, dependiendo del SLA k contratado, esto es $R_{P_j}^{sla_k}$. Así mismo, el término y se corresponde con el retardo medio actualizado en cada instante de tiempo para cada una de esas j

clases de servicio prioritarias de cada uno de los k SLAs ($\overline{r_{P_j}^{sla_k}}[n]$). El valor del error cometido en cada instante n de una ONU que ha contratado un SLA k , viene dado por $e[n] = \sum_j (R_{P_j}^{sla_k} - \overline{r_{P_j}^{sla_k}}[n])$, donde j representa cada clase de servicio con requisitos de retardo. Esto es, para calcular el error, se hace el sumatorio de los errores individuales cometidos a la hora de restringir el retardo medio de cada una de las j clases de servicio a las cotas máximas estipuladas.

Para obtener la señal de control $u[n]$, se sigue la ecuación (4.2). Sin embargo, como el error obtenido es en segundos y la señal $u[n]$ debe estar en bits para que modifique adecuadamente el término $B_{max}^{onu_i}$, la señal $u[n]$ se multiplica por un factor de conversión. A continuación se pasa por el delimitador para ajustar los anchos de banda sugeridos por el PID para que se ajusten al ancho de banda disponible en el tiempo de ciclo máximo de 2 ms. El diagrama de bloques correspondiente al algoritmo DaSPID es el que se muestra en la Figura 4.27. Tal y como puede observarse, para obtener el nuevo ancho de banda máximo de cada ONU, el valor absoluto de la señal de control se suma o resta del máximo anterior dependiendo de si el valor de $u[n]$ es positivo o negativo. De esta forma, si por ejemplo los retardos medios actuales de los paquetes de la ONU i para los distintos servicios prioritarios j ($\overline{r_{P_j}^{sla_k}}[n]$) son mayores que sus correspondientes retardos máximos estipulados ($R_{P_j}^{sla_k}$), el error $e[n]$ es negativo y también lo es la señal de control $u[n]$. Por lo tanto, como lo que se necesita es aumentar el valor del ancho de banda máximo ($B_{max}^{onu_i}$) para permitir que la ONU i reduzca su retardo y pueda cumplir con la cota impuesta, el valor absoluto de $u[n]$ se suma al ancho de banda máximo anterior. Por el contrario, si los retardos medios de los paquetes de esa ONU i para las j clases de servicio ($\overline{r_{P_j}^{sla_k}}[n]$) son menores que sus cotas máximas estipuladas, el error es positivo y la señal de control también. Esto es, su retardo se sitúa por debajo del límite máximo establecido por lo que el valor absoluto de $u[n]$ se resta del ancho de banda máximo anterior para ajustar el retardo a su cota máxima estipulada.

Para calcular el valor medio del retardo de cada una de las clases de servicio para

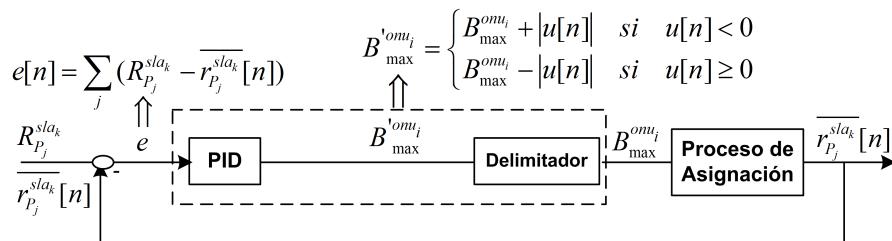


Figura 4.27: Diagrama de bloques del proceso controlado por un PID propuesto para el control del retardo máximo del tráfico prioritario en redes LR-EPON.

cada SLA ($\overline{r_{P_j}^{sla_k}}[n]$), se van a utilizar ventanas temporales deslizantes del mismo tipo que la ventana descrita en el Apartado 3.2.2 del Capítulo 3 (también usadas en SPID), con una duración fija de T_{window} segundos. En concreto existirá una ventana por cada SLA k y cada clase de tráfico con restricción j . En esta ventana se almacenarán las muestras con el retardo medio de los paquetes recibidos durante un ciclo de la clase de servicio j perteneciente al SLA k y el tiempo en el que se obtuvo esa medida de retardo.

Por tanto, el resumen del funcionamiento de DaSPID es el siguiente. Cada tiempo de ciclo, el OLT calcula el retardo medio de los paquetes de cada clase de servicio prioritaria de cada SLA en ese ciclo e introduce dicho valor en la ventana temporal correspondiente. A continuación, realiza la asignación de ancho de banda para el ciclo siguiente siguiendo la ecuación (4.3). Por otro lado, cada T segundos, el OLT realiza la modificación de los máximos permitidos a cada ONU, $B_{max}^{onu_i}$, mediante el uso de un PID y basándose en el retardo medio actual de cada clase de servicio prioritaria y en los requisitos de calidad de servicio estipulados para todos los SLAs.

Por último, cabe destacar que el algoritmo DaSPID se analizará con una estrategia de asignación agresiva. Esto quiere decir que, cuando el ancho de banda disponible no es suficiente para garantizar los retardos estipulados a todas las clases de tráfico restrictivo, los usuarios que hayan contratado el SLA menos prioritario, verán reducido sus niveles de ancho de banda máximos para favorecer que los usuarios de SLAs más prioritarios cumplan con los requisitos de retardo en sus clases prioritarias. No obstante, también se podría aplicar la estrategia conservativa diseñada en SPID, aunque ésta no se ha considerado en el análisis.

4.7. Estudio de simulación del algoritmo DaSPID

Para analizar el comportamiento de DaSPID se va a seguir el mismo procedimiento de análisis que con SPID. Por ello, en primer lugar se explica el entorno de simulación donde se va a realizar el estudio de simulación. Tras ello, se va a realizar la sintonización del controlador y a seleccionar el tipo de controlador más adecuado (P, PI o PID). A continuación, se determinará el valor de los parámetros de funcionamiento del algoritmo diseñado. Por último, se realizará un análisis de las prestaciones de DaSPID comparándolo con el algoritmo DySLa considerando distintos escenarios de red.

4.7.1. Escenario de simulación

La simulaciones de DaSPID, al igual que en los algoritmos anteriormente presentados en el desarrollo de esta Tesis, se han llevado a cabo considerando una red LR-EPON

de 100 km, con 16 ONUs y un usuario conectado a cada ONU. La tasa de transmisión en el canal ascendente entre las ONUs y el OLT es de 1 Gbit/s y entre los usuarios y las ONUs de 100 Mbit/s. Se ha establecido un tiempo de guarda entre transmisiones de $1 \mu s$ y el tiempo máximo de ciclo a 2 ms. Los paquetes tienen un tamaño uniforme entre 64-1538 *bytes* (más 38 *bytes* de cabeceras).

Por otro lado, es necesario definir métodos para la diferenciación del tráfico en distintas clases de servicio. De esta forma, siguiendo la Recomendación ITU-T G.1010 [34] y otros trabajos relacionados [17, 79, 84, 112, 114], se ha considerado que la red soporta tres clases de servicio prioritarias P_0 , P_1 y P_2 , representando al tráfico en tiempo real (*Real-Time*), al tráfico de respuesta rápida (*Responsively*) y al tráfico no crítico (*Best Effort*), respectivamente.

Para soportar la diferenciación de tráfico, las ONUs están provistas de tres colas, una para cada clase de servicio, todas ellas compartiendo la misma memoria de capacidad 10 Mbytes, al igual que en [79]. Además, se considera que P_0 representa el 20 % de la carga de la ONU y P_1 y P_2 el 40 % cada una [17, 79, 112]. De esta forma, se asumen las mismas condiciones de carga que en DySLA, para así poder realizar una mejor comparación.

En cuanto a la diferenciación de clientes, se ha considerado la situación más general, donde todos los usuarios soportan las mismas clases de servicio independientemente del SLA contratado. Sin embargo, cada una de las clases tiene un umbral máximo de retardo, que será distinto dependiendo de la prioridad del SLA contratado. En el escenario de red presentado se consideran tres SLAs, SLA_0 de mayor prioridad, SLA_1 de prioridad intermedia y SLA_2 de prioridad baja. Además 2, 6 y 8 ONUs contratan el SLA_0 , SLA_1 y SLA_2 , respectivamente. Inicialmente, los pesos escogidos para calcular los anchos de banda iniciales de cada uno de los SLAs son $W^{sla_0} = 1$, $W^{sla_1} = 1$ y $W^{sla_2} = 1$. Sin embargo, para probar la flexibilidad del algoritmo DaSPID, estos pesos se modificarán a lo largo de los siguientes apartados para analizar el comportamiento en diferentes escenarios de red.

Para definir las restricciones de retardo para cada una de las clases de servicio prioritario de cada uno de los SLAs, se han utilizado los valores empleados en DySLA [79] donde se aplicaron las recomendaciones ITU-T G.1010 [34] y G.114 [115]. En concreto, la Recomendación G.114, especifica que las aplicaciones en tiempo real deben satisfacer un retardo máximo de 1.5 ms en la red de acceso. Por tanto, como P_0 representa los servicios más sensitivos, se ha impuesto que todos los usuarios deben cumplir con esta restricción independientemente del SLA contratado ($\forall k R_{P_0}^{sla_k} = 1.5$ ms). En cuanto al tráfico de respuesta rápida (P_1), la recomendación G.1010 establece un retardo extremo a extremo máximo de 2 s. Sin embargo, este retardo máximo se refiere a una conexión ex-

tremo a extremo y no al segmento de la red de acceso. Por ello, se han establecido valores muy por debajo de ese nivel, esto es, valores muy restrictivos en el segmento de acceso. En concreto, para los usuarios más prioritarios se ha estipulado un retardo máximo de 5 ms ($R_{P_1}^{sla_0} = 5$ ms), para los de prioridad intermedia de 20 ms ($R_{P_1}^{sla_1} = 20$ ms) y para los de menor prioridad un retardo máximo de 60 ms ($R_{P_1}^{sla_2} = 60$ ms). Sin embargo, se pueden considerar otros valores dependiendo de la decisión tomada por los proveedores de servicio. Por ello, también se ha analizado el comportamiento de DaSPID teniendo en cuenta otras cotas máximas. Finalmente, para el tráfico no crítico (P_2) no es necesario cumplir ningún requisito de retardo máximo.

En la Tabla 4.9 se recogen los requisitos básicos de cada uno de los servicios considerados, tanto los valores extremo a extremo propuestos por la Recomendación G.1010, como los umbrales en la red de acceso propuestos en la Recomendación G.114 para el tráfico en tiempo real. Así mismo, también se recogen las estimaciones realizadas para el resto de servicios y las principales aplicaciones en las que se puede clasificar cada clase de tráfico.

Segmento de red	Recomendación	Clase de Servicio	Requisitos de retardo	Aplicaciones
Extremo a Extremo	ITU-T G.1010	P_0 (Tiempo Real, Interactivo)	150 ms	VoIP, videoconferencia, juegos interactivos, <i>Telnet</i>
Acceso	ITU-T G.114		1.5 ms	
Extremo a Extremo	ITU-T G.1010	P_1 (Respuesta rápida)	2 s	Mensajes de voz, navegador, HTML, correo electrónico, servicios de transacción
Acceso	Estimación		SLA ₀ =5 ms SLA ₁ =20 ms SLA ₂ =60 ms	
Extremo a Extremo	ITU-T G.1010	P_2 (No crítico)	-	Datos sin prioridad
Acceso	Estimación		-	

Tabla 4.9: Retardos máximos considerados en el escenario de simulación en DaSPID para las distintas clases de servicio.

4.7.2. Proceso de sintonización y comparativa de estrategias P-PI-PID

Al igual que en el caso de SPID, en DaSPID el controlador tiene que ser sintonizado de forma adecuada para obtener unos valores óptimos para los términos K_p , T_i y T_d , dependiendo del tipo de controlador escogido, P, PI o PID.

Para sintonizar el controlador también se ha utilizado en este caso el método de Ziegler-Nichols de respuesta en frecuencia. Por tanto, se han llevado a cabo los mismos pasos explicados en el Apartado 4.4.2. En concreto, la Figura 4.28 (a) y (b) muestra la variación en tiempo real del retardo medio de los paquetes ($r_{P_j}^{slak}[n]$) para la clase de servicio P_1 para el SLA₁ y el SLA₂, respectivamente. Aunque no se han representado los resultados del perfil más prioritario, el comportamiento es análogo al de resto de perfiles. Tal y como puede observarse, para valores pequeños de K_u (0.02, 0.05), DaSPID no evoluciona adecuadamente el retardo medio de esta clase de servicio para ambos perfiles de abonado a sus correspondientes niveles estipulados (20 ms y 60 ms para el SLA₁ y SLA₂, respectivamente). Este comportamiento es perceptible especialmente para el SLA₂, donde se observa que los valores se mantienen muy por debajo del nivel estipulado. Por el contrario, para valores elevados (1.8, 3) el retardo presenta una gran oscilación a lo largo del tiempo y además para el SLA₂ dicha oscilación no se produce en torno al valor estipulado. Por tanto, es necesario seleccionar un valor intermedio de este parámetro (0.6, 1) para obtener un comportamiento correcto en todos los perfiles. De este modo, se ha optado por usar un valor de $K_u = 0.6$ que produce un periodo de oscilación igual a $T_u = 62$ s. Finalmente, aplicando las relaciones de Ziegler-Nichols de la Tabla 4.3 se obtienen los valores de sintonización K_p , T_i y T_d mostrados en la Tabla 4.10 para los

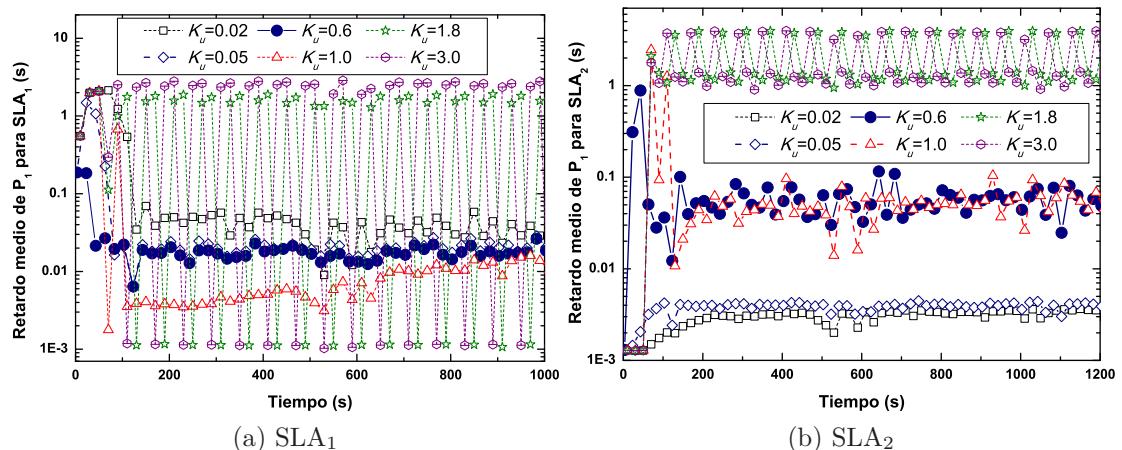


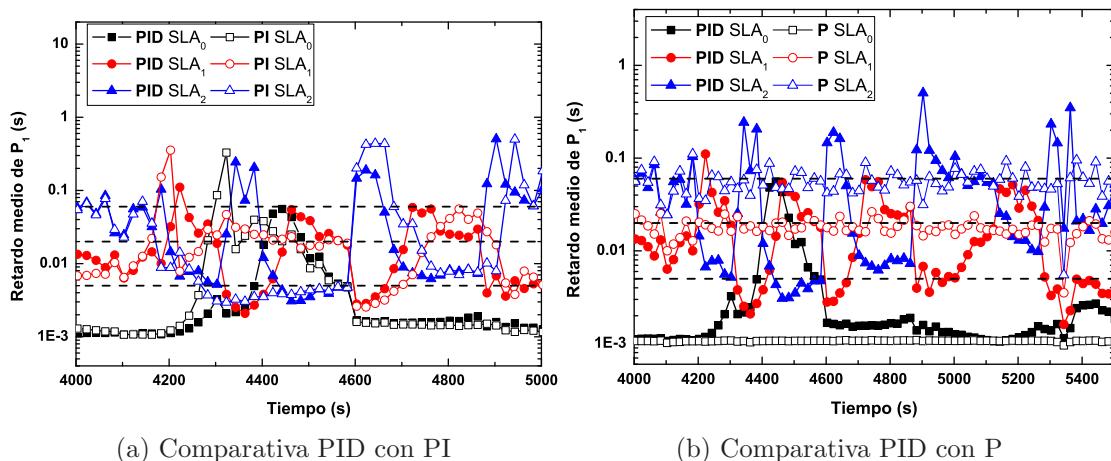
Figura 4.28: Variación en tiempo real del retardo medio asignado para distintos valores de K_u . (a) SLA₁ (b) SLA₂.

Controlador	K_p	T_i	T_d
P	0.3		
PI	0.24	49.6	
PID	0.36	31	7.75

Tabla 4.10: Valores de sintonización de los controladores P, PI y PID en DaSPID.

tres tipos de controladores.

A continuación, se va a analizar el comportamiento de los controladores P, PI y PID para elegir aquel que permita una mejor gestión del retardo medio. En concreto, la Figura 4.29 (a) y (b) muestra el retardo medio a lo largo del tiempo de los paquetes de la clase de servicio P_1 cuando se comparan el controlador PID con el PI (Figura 4.29 (a)) y el PID con el P (Figura 4.29 (b)). Tal y como se puede observar en la Figura 4.29 (a), ambos controladores PI y PID presentan inestabilidades y no son capaces de estabilizar el retardo medio. Sin embargo, tal y como se aprecia, las oscilaciones del controlador PID son ligeramente menores que las del controlador PI. Por este motivo, se ha comparado el comportamiento de este controlador PID con el del controlador P. Tal y como se observa en la Figura 4.29 (b), el controlador P presenta el mejor comportamiento, ya que es mucho más estable que el controlador PID para los tres perfiles de abonado y consigue garantizar los niveles de retardo estipulados. Por tanto, para la gestión del retardo el controlador P presenta un mejor comportamiento que los controladores PI y PID, por lo que ha sido seleccionado para su uso en DaSPID.

Figura 4.29: Variación en tiempo real del retardo medio de P_1 para los tres SLAs considerados cuando se comparan los controladores (a) PID con PI (b) PID con P.

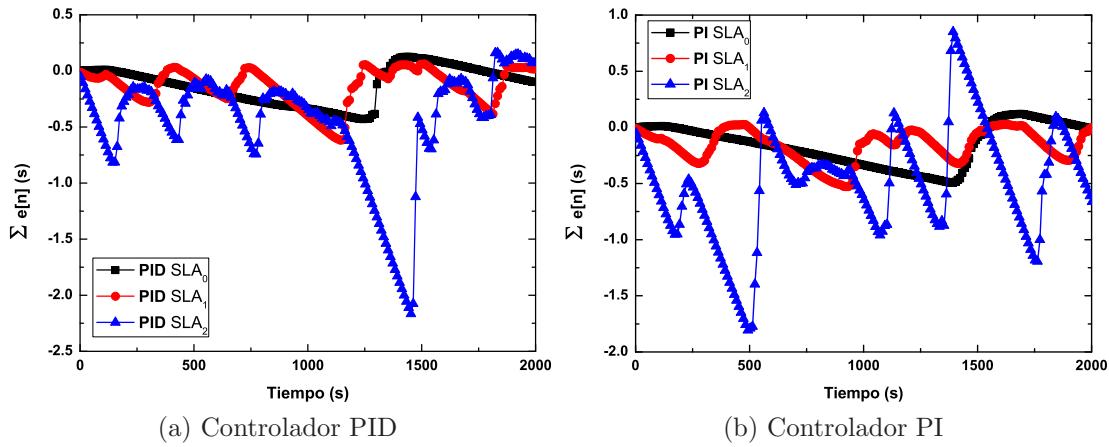


Figura 4.30: Evolución en tiempo real del término integral. (a) Controlador PID (b) Controlador PI.

En este sentido, se ha profundizado más en el análisis del comportamiento de los controladores PI y PID para identificar las causas de su mal funcionamiento en la gestión del retardo. Por tanto, se ha estudiado la contribución de cada término en la señal de control ($u[n]$), que se usa para modificar el ancho de banda máximo permitido. En concreto, la Figura 4.30 (a) y (b) representa la evolución en tiempo real del término integral en el controlador PID y PI, esto es, el sumatorio de los errores cometidos a lo largo del tiempo. Como puede verse en ambas figuras, el término integral es bastante inestable, de forma particular para el perfil SLA₂, ya que la evolución del sumatorio presenta una forma de “diente de sierra”, con bajadas con una cierta pendiente y subidas repentinas. Además, el valor del sumatorio puede llegar a alcanzar valores del orden de segundos, lo que provoca cambios abruptos en la señal de control $u[n]$, ya que la señal de control depende directamente de este sumatorio (ecuación (4.2)). Esto es así porque, al ser los retardos a garantizar del orden de milisegundos, la adición del término integral provoca que el controlador PID o PI se desestabilice, aumentando los errores cometidos.

Para completar el estudio, la Figura 4.31 (a) y (b) muestra el error instantáneo cometido ($e[n]$) a lo largo del tiempo para los tres perfiles de abonado comparando los controladores PID con P, y PI con P respectivamente. Tal y como puede observarse en ambas figuras, para los controladores PI y PID, los cambios abruptos en la señal de control afectan en gran medida al error instantáneo, ya que $e[n]$ alcanza valores altos justo después de los picos en las curvas de la Figura 4.30 (a) y (b). Esto ocurre porque justo después de estos picos, el ancho de banda máximo permitido ($B_{max}^{onu_i}$) sufre un gran cambio provocado por la inestabilidad del término integral. Esto afecta seriamente a las siguientes muestras del error instantáneo de los controladores PID y PI, ya que el

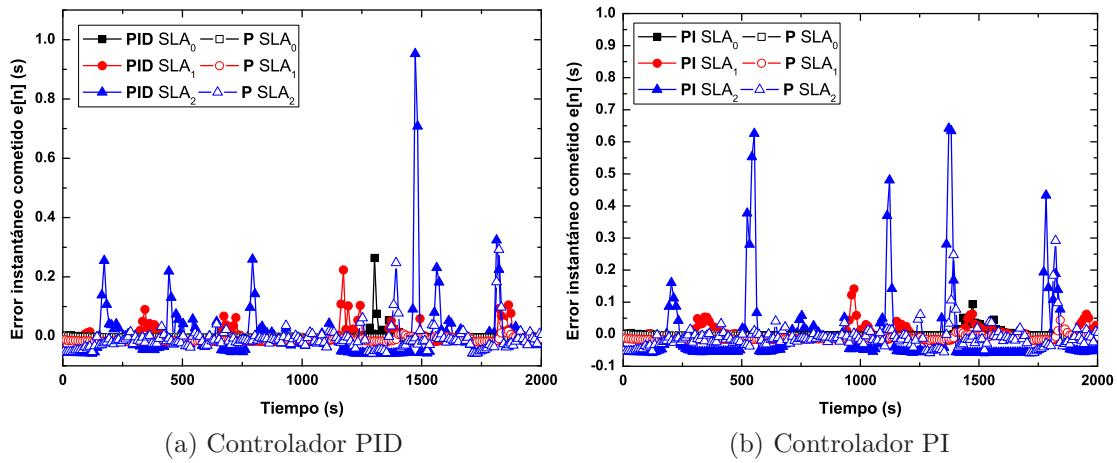


Figura 4.31: Evolución en el tiempo del error instantáneo $e[n]$ comparando un controlador P con un (a) controlador PID (b) controlador PI.

error depende del término $\overline{r_{P_j}^{slak}}$, que a su vez depende del máximo permitido $B_{max}^{onu_i}$. De hecho, si el error es muy elevado o muy inestable, el comportamiento del controlador se ve afectado negativamente. Por lo tanto, este comportamiento del término integral hace que el retardo medio de los paquetes sea más inestable para los controladores PI y PID, tal y como se observa en la Figura 4.29 (a). Por el contrario, como el controlador P no tiene término integral, tanto el error cometido como la señal de control son más estables, dando lugar a un mejor comportamiento de este controlador para esta aplicación específica, tal y como se aprecia en la Figura 4.31 (a) y (b).

Para mitigar la influencia del término integral y comprobar así si se produce una mejora en el comportamiento de los controladores PI y PID, se decidió restringir el término del sumatorio para incluir únicamente los últimos errores cometidos en vez de realizar la suma completa de todos los errores desde el principio de la simulación. Se realizaron diversas pruebas asumiendo distinto número de errores a considerar en el sumatorio. En concreto, se ha considerado el sumatorio de los últimos 5, 15, 50 y 100 errores cometidos. El comportamiento del controlador PID para la clase de servicio P₁ del SLA₁ está representado en la Figura 4.32. Tal y como se aprecia, cuanto menor es el número de errores considerados en el sumatorio, más estable es la respuesta del controlador PID, esto es, a menor peso del término integral, mejor comportamiento.

Por tanto, aunque un controlador PID completo incluye tres términos (P-I-D), existen aplicaciones donde solo se requiere el uso de controladores que constan únicamente de la parte proporcional (P) o de los términos proporcional e integral (PI). Por tanto, como para esta aplicación concreta de gestión del retardo en redes LR-EAPON, el controlador

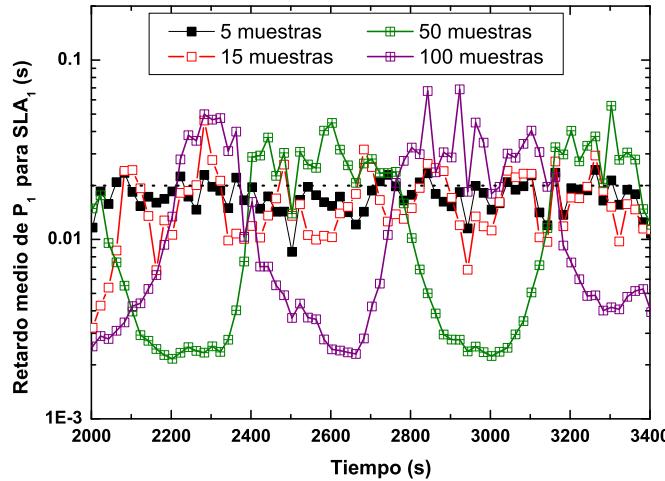


Figura 4.32: Variación del retardo medio de la clase de servicio P_1 del SLA_1 cuando se considera un número diferente de errores en el término integral.

P es el que ha demostrado un mejor comportamiento, se ha optado por usar este tipo de controlador en el diseño de DaSPID. Además, este controlador es más sencillo ya que solo implementa la parte proporcional.

4.7.3. Establecimiento de los parámetros asociados al diseño del controlador en DaSPID

Una vez sintonizado y seleccionado el tipo de controlador que ofrece las mejores prestaciones, es necesario determinar los valores óptimos de los parámetros propios de DaSPID para asegurar un buen funcionamiento del sistema. En este sentido, es necesario establecer dos parámetros, la duración de la ventana donde se almacenan las muestras de los retardos de cada clase de servicio para cada SLA (T_{window}) y el tiempo de muestreo o ejecución (T) que es el tiempo entre adaptaciones consecutivas de los máximos.

El primer parámetro, T_{window} es un parámetro clave ya que en DaSPID la adaptación de los anchos de banda máximos permitidos de cada ONU, $B_{max}^{onu_i}$, se realiza dependiendo de los retardos medios calculados con las ventanas $r_{P_j}^{slak}$. De esta forma, si la duración de la ventana es elevada, el tiempo de respuesta de DaSPID crecerá, e incluso, si existen cambios muy abruptos y rápidos, la adaptación puede no realizarse a tiempo. Por el contrario, si la ventana tiene una duración demasiado pequeña, las muestras recogidas en la ventana pueden no ser significativas y no representar el retardo medio real de forma adecuada. Por tanto, es necesario escoger un valor intermedio para conseguir una buena aproximación del retardo medio sin ralentizar la respuesta del algoritmo. Para determinar dicho valor se han llevado a cabo simulaciones con distintos tiempos de ventana.

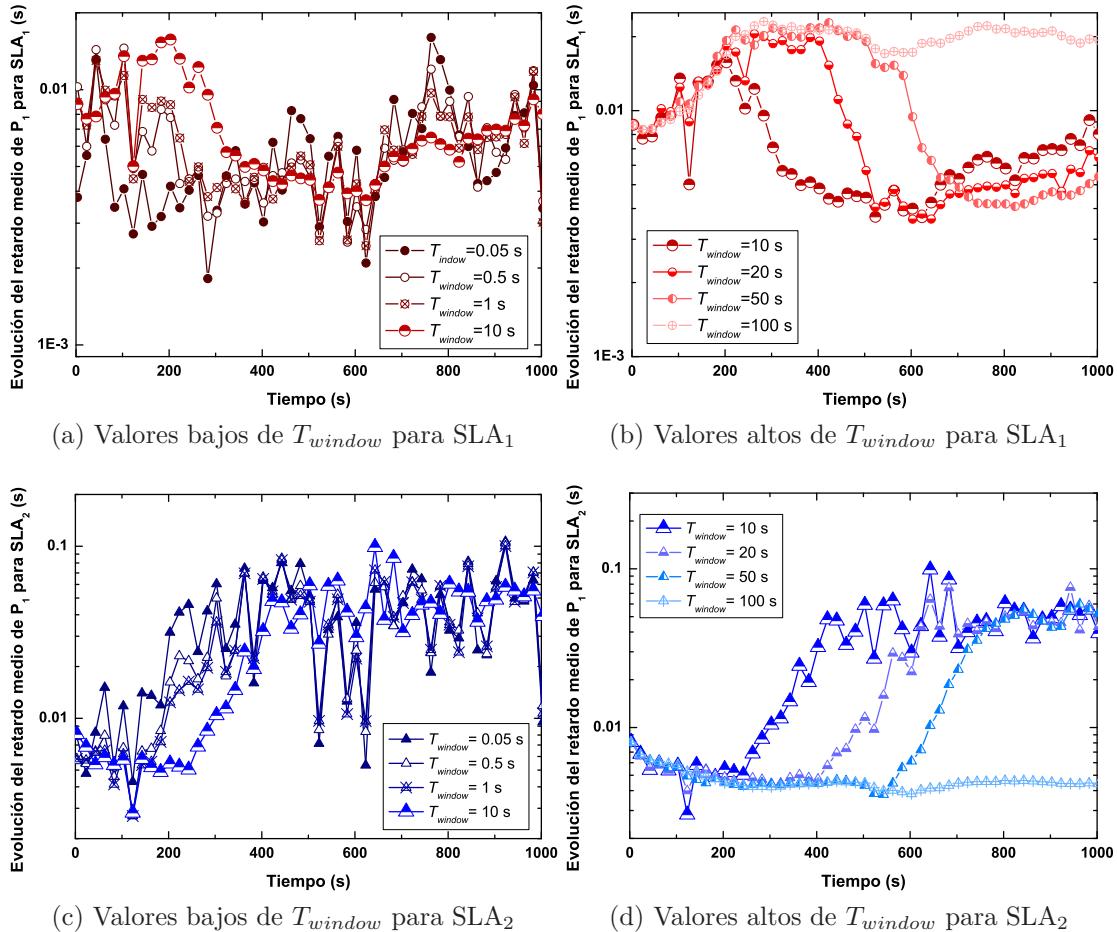


Figura 4.33: Evolución en tiempo real del retardo medio de P_1 para distintos valores de T_{window} (a) SLA₁ valores bajos de T_{window} (b) SLA₁ valores altos de T_{window} (c) SLA₂ valores bajos de T_{window} (d) SLA₂ valores altos de T_{window} .

En concreto, la Figura 4.33 (a), (b), (c) y (d) muestra la evolución a lo largo del tiempo del retardo medio de la clase de servicio P_1 para el SLA₁ cuando se consideran valores pequeños del parámetro T_{window} en la Figura 4.33 (a) y valores elevados en la Figura 4.33 (b), y para el SLA₂ con valores de T_{window} pequeños en la Figura 4.33 (c) y elevados en la Figura 4.33 (d). Solo se han representado estos dos SLAs para reducir el número de gráficas, pero los resultados para el SLA más prioritario son similares. Para ambos perfiles de abonado puede verse que si la duración de la ventana es muy pequeña (Figura 4.33 (a) y (c)), la evolución en tiempo real del retardo medio presenta una mayor inestabilidad (valores de 0.05 s y 0.5 s). Por el contrario, si se consideran valores elevados de tiempo de duración (20 s, 50 s, 100 s), la evolución hacia los límites de retardo máximo estipulados es muy lenta (Figura 4.33 (b) y (d)). Estas dos situaciones no son deseables,

ya que se necesita que el retardo medio evolucione de forma estable pero que además converja de forma rápida. Por tanto, se ha escogido un valor intermedio de 10 s, ya que cumple con estas dos premisas.

El segundo parámetro clave, está relacionado con la frecuencia con la que los anchos de banda máximos permitidos se van a actualizar. Este tiempo se ha denominado tiempo de muestreo, T en la ecuación (4.2). La determinación de un valor óptimo para este parámetro es fundamental, ya que si es demasiado pequeño, los anchos de banda máximos permitidos se modificarán con demasiada rapidez, sin tener apenas en cuenta los efectos de las modificaciones previas. Por el contrario, si el tiempo es elevado, el algoritmo tardará demasiado en adaptarse a los cambios. Por tanto, es necesario elegir un valor de compromiso que no suponga ni la desestabilización del sistema ni una respuesta demasiado lenta.

Para determinar un valor óptimo para este parámetro, se han llevado a cabo simulaciones con distintos tiempos T . En la Figura 4.34 (a), (b), (c), (d) se muestran los resultados obtenidos para la clase de servicio P_1 cuando se consideran tiempos de muestreo pequeños en el SLA_1 y SLA_2 (Figura 4.34 (a) y (c), respectivamente) y cuando se consideran tiempos de muestreo más elevados en dichos SLAs (Figura 4.34 (b) y (d), respectivamente). Para reducir el número de gráficas no se ha representado el SLA_0 , pero los resultados para este perfil son similares. Tal y como puede apreciarse en la Figura 4.34 (a) y (c), para tiempos de muestreo pequeños (0.05 s, 0.5 s y 1 s) se observan fluctuaciones en el retardo medio muy elevadas. En concreto, para 0.05 s el algoritmo no es capaz de evolucionar el retardo al límite establecido para ambos SLAs, fijado en 20 ms y 60 ms para el SLA_1 y el SLA_2 , respectivamente, obteniendo valores de retardo más elevados. Por el contrario, para valores elevados (20 s, 50 s, 100 s) el retardo medio evoluciona hacia los niveles límites muy lentamente. Esto ocurre sobre todo para el SLA de menor prioridad SLA_2 , tal y como se aprecia en la Figura 4.34 (d), ya que su retardo tiene que evolucionar hasta un nivel más elevado (60 ms). Por lo tanto es necesario escoger un valor que permita simultáneamente, una evolución estable y rápida hacia los niveles establecidos. Por ello, se ha escogido un tiempo de muestreo T igual a 10 s, ya que cumple ambas condiciones, tal y como puede verse en la Figura 4.34 (a), (b), (c), (d), para todos los SLAs considerados.

Por otro lado cabe destacar, que para el servicio más prioritario P_0 , los valores de T_{window} y de T no tienen una gran influencia en su comportamiento y que por este motivo no se han utilizado en la selección de los valores óptimos para ambos parámetros. Sin embargo, la tendencia de escoger un valor alto o bajo de estos parámetros es la misma para esta clase de servicio.

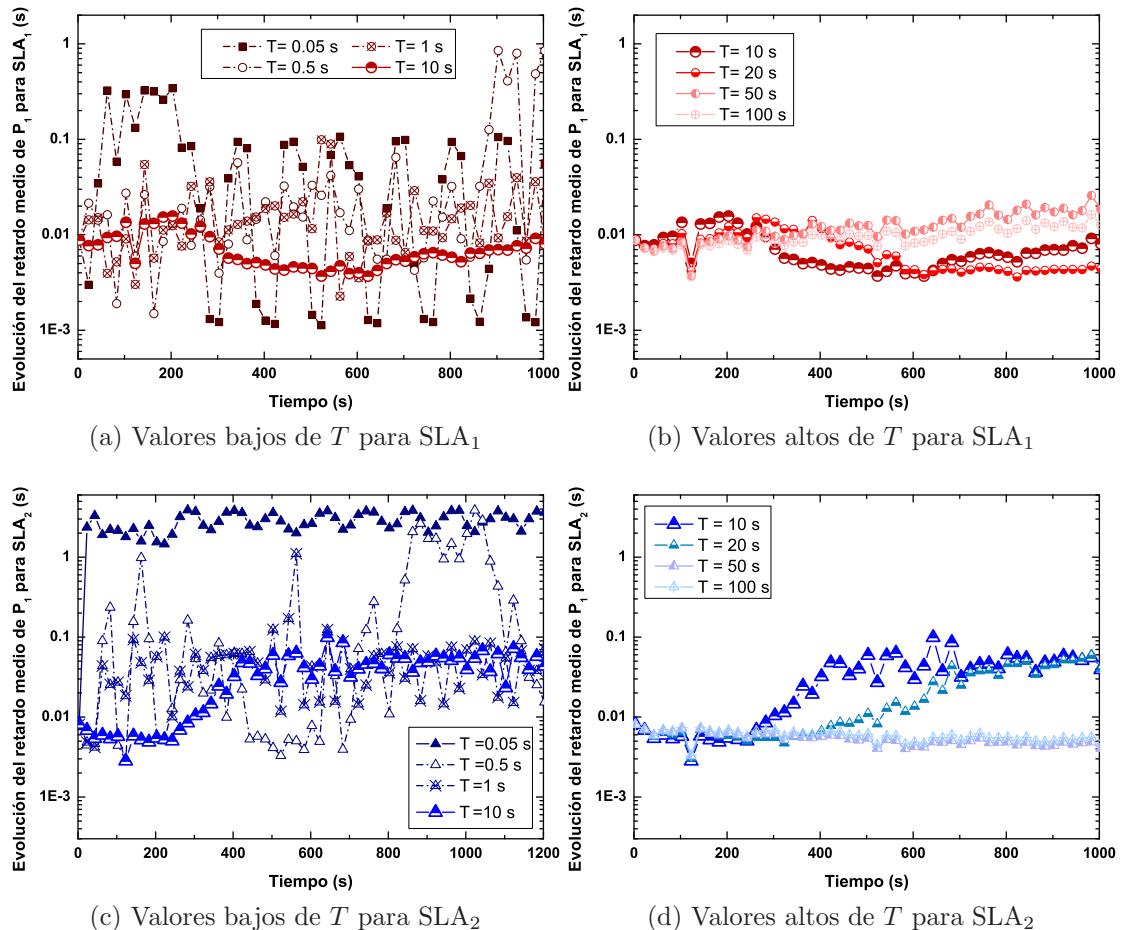


Figura 4.34: Evolución en tiempo real del retardo medio de P_1 para distintos valores de T
 (a) SLA₁ valores bajos de T (b) SLA₁ valores altos de T (c) SLA₂ valores bajos de T (d) SLA₂ valores altos de T .

4.7.4. Análisis de prestaciones de DaSPID

Tal y como se ha comentado anteriormente, la característica más importante de DaSPID es la modificación de los anchos de banda máximos para conseguir que el retardo medio de las diferentes clases de servicio de los distintos SLAs se sitúen por debajo de los máximos estipulados por los proveedores de servicio. Por tanto, para analizar sus prestaciones y ventajas se va a llevar a cabo una comparativa con el algoritmo DySLa. Para ello, se han establecido unos pesos iniciales en ambos algoritmos iguales a $W^{sla_0} = 1.5$, $W^{sla_1} = 1$ y $W^{sla_2} = 0.5$ al igual que en [79] y se ha considerado el escenario presentado en el Apartado 4.7.1. Para comparar ambos algoritmos, se va a analizar en primer lugar la influencia del factor fijo (α) que utiliza DySLa para actualizar los anchos de banda máximos. A continuación, se estudiará la evolución en tiempo real del retardo de ambos algoritmos hacia los límites estipulados y de los anchos de banda máximos permitidos. Seguidamente, se probará la independencia de DaSPID frente a las condiciones iniciales de red y a los límites de retardo estipulados por los proveedores de servicio. Por último, se han llevado a cabo simulaciones en redes EPON y LR-EPON de distinta longitud para probar su funcionamiento en redes de diferente cobertura.

4.7.4.1. Influencia del parámetro α en DySLa

Uno de los principales inconvenientes del algoritmo DySLa, es que la modificación de los máximos se realiza sustrayendo y añadiendo un factor fijo α a los pesos para actualizar los anchos de banda máximos de los distintos SLAs. En [79] se utiliza un valor $\alpha = 0.1$ por ser el que ofrece un comportamiento óptimo en DySLa. Sin embargo, un error en la elección de este valor puede suponer un mal funcionamiento del algoritmo. Esto puede verse en la Figura 4.35, donde se muestra la evolución a lo largo del tiempo del retardo medio de la clase de servicio P_1 considerando diferentes valores del parámetro α para el SLA de prioridad intermedia (SLA₁) en la Figura 4.35 (a) y para el SLA de menor prioridad (SLA₂) en la Figura 4.35 (b), para una carga de red en ambos casos igual a 1.4 (ONUs generando tráfico a una tasa alta de 87.5 Mbit/s). El comportamiento del SLA₀ es similar, pero no se ha representado por reducir el número de gráficos. Como puede observarse en ambas gráficas, la evolución del retardo medio exhibe una gran dependencia del valor α . En concreto, para el SLA₁ (Figura 4.35 (a)), puede observarse que valores altos de α incrementan la inestabilidad del retardo medio. Por el contrario, valores bajos de este factor dan lugar a una convergencia más lenta hacia el límite máximo estipulado para este perfil de abonado (20 ms). Con respecto a los usuarios de menor prioridad, el comportamiento es similar, tal y como se observa en la Figura 4.35 (b). Esto es, para valores altos de α el algoritmo no reacciona ($\alpha = 0.2$), mientras que

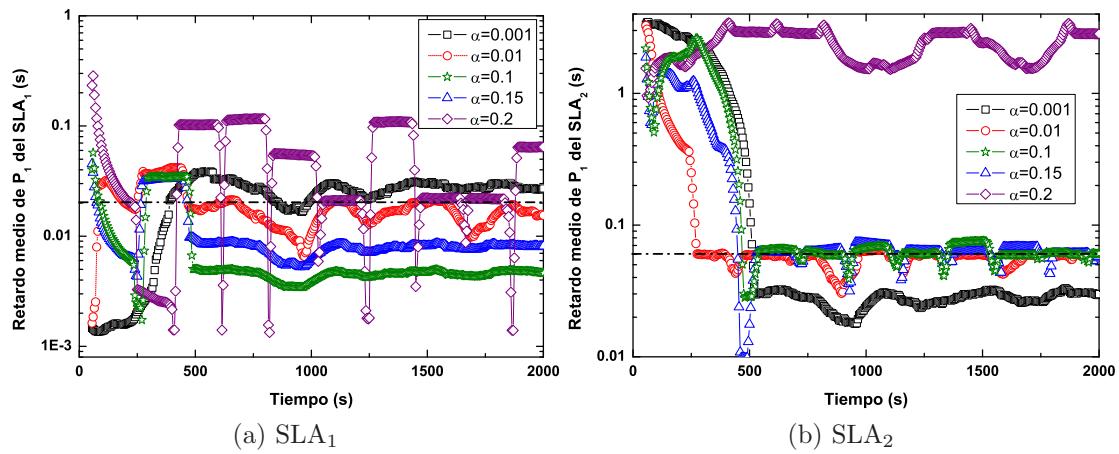


Figura 4.35: Evolución en tiempo real del retardo medio de P_1 de los dos SLAs de menor prioridad considerando diferentes valores de α en DySLA. (a) SLA₁ (b) SLA₂.

para valores pequeños la velocidad de convergencia hacia el límite máximo estipulado para este SLA (60 ms) se reduce.

Para completar el análisis, en la Figura 4.36 (a) y (b) se presenta la evolución en tiempo real del ancho de banda máximo permitido a las ONUs para el SLA₁ y SLA₂, respectivamente. Como puede observarse, los anchos de banda máximos permitidos no evolucionan hacia los mismos valores y además, dependiendo del valor de α , tienen una mayor oscilación. En ambos SLAs las mayores oscilaciones se producen para un valor de α igual a 0.2 y 0.15, porque estos valores hacen que el término $B_{max}^{onu_i}$ cambie de forma más brusca y eso afecta inherentemente a la evolución y estabilidad del retardo medio,

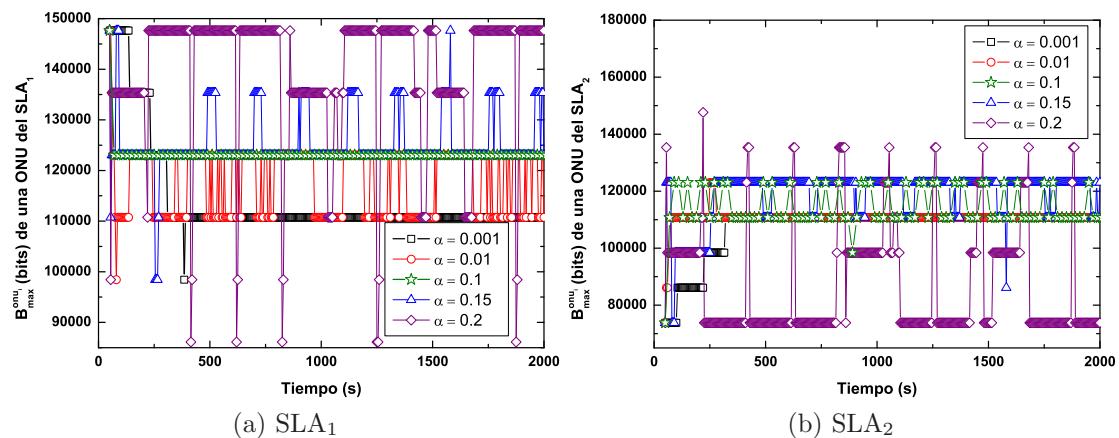


Figura 4.36: Evolución en tiempo real del ancho de banda máximo permitido a cada ONU del SLA₁ y SLA₂ considerando diferentes valores de α en DySLA. (a) SLA₁ (b) SLA₂.

tal como se observa en ambas figuras.

Por el contrario, DaSPID, al modificar los anchos de banda dependiendo del error cometido y no en una cantidad fija, presenta un tiempo de reacción menor que en DySLa independientemente del escenario elegido, así como una evolución en tiempo real de los anchos de banda máximos permitidos mucho más estable. Este comportamiento se demostrará con más detalle en las siguientes secciones.

4.7.4.2. Análisis del retardo medio en DaSPID

En este apartado se va a comparar el funcionamiento de DaSPID y DySLa, ya que ambos algoritmos han sido diseñados con el mismo propósito; esto es, garantizar que los retardos estén por debajo de unos límites máximos establecidos teniendo en cuenta de forma simultánea tanto la prioridad del tráfico como la de los perfiles de abonado contratados. Para llevar a cabo esta comparativa se ha escogido en DySLa un valor de $\alpha = 0.1$ [79], por ser el valor que ofrece el mejor comportamiento en este algoritmo. A continuación, se estudia el retardo obtenido para cada clase de tráfico y SLA.

Retardo medio de paquetes del servicio P_0

En la Figura 4.37 se representa el retardo medio de la clase de servicio P_0 para DaSPID y DySLa, para los perfiles de abonado SLA_0 , SLA_1 y SLA_2 frente a la carga de red. Tal y como se observa en la gráfica, los dos algoritmos tienen un comportamiento

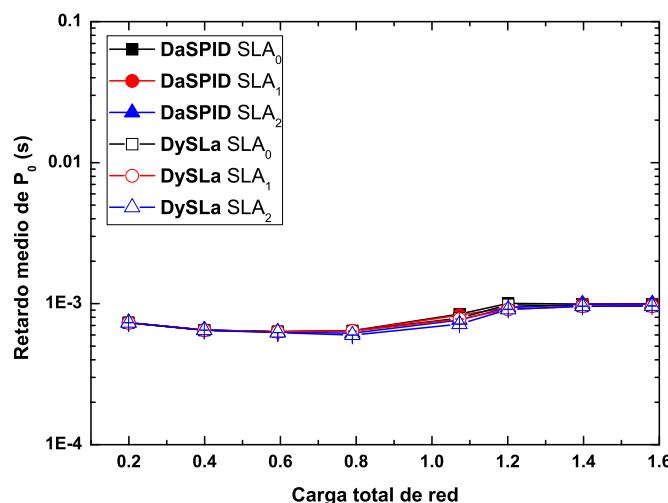


Figura 4.37: Retardo medio de la clase de servicio P_0 frente a la carga de red al comparar DySLa y DaSPID.

similar para esta clase de servicio, cumpliendo ambos con la restricción de 1.5 ms estipulada. Además, al tener la red una distancia de 100 km, el retardo para esta clase de servicio siguiendo el esquema de *Strict Priority Queue* viene limitado por el tiempo de ida y vuelta de la red, que es igual a 1 ms. Esto es así, porque al usar este esquema de colas, los paquetes de este servicio son siempre los primeros en salir, ya que es el tráfico más prioritario. De esta forma, todo el tráfico P_0 acumulado se envía en el mismo ciclo, luego la restricción de retardo para esta clase de servicio viene determinada por la distancia extremo a extremo de la red.

Retardo medio de paquetes del servicio P_1

Respecto a la clase de servicio de prioridad media P_1 , la Figura 4.38 representa el retardo medio de esta clase de servicio frente a la carga total de red cuando se comparan los algoritmos DySLa y DaSPID para los tres SLAs considerados.

Como se observa, ambos algoritmos consiguen realizar la diferenciación de servicios y usuarios de forma correcta, consiguiendo que los retardos se sitúen por debajo de los límites máximos estipulados para cargas de red bajas y medias. Por el contrario, para aquellas cargas de red que hacen que el canal de subida esté saturado (cargas mayores de 1 Gbit/s), el comportamiento de ambos algoritmos es diferente. De hecho, cuando el ancho de banda disponible no es suficiente para que los retardos de los servicios sensibles se encuentren por debajo del límite máximo, DySLa no es capaz de cumplir con los requisitos de retardo para el SLA₂ (cargas de red de 1.2 y 1.4), ya que, tal y como se

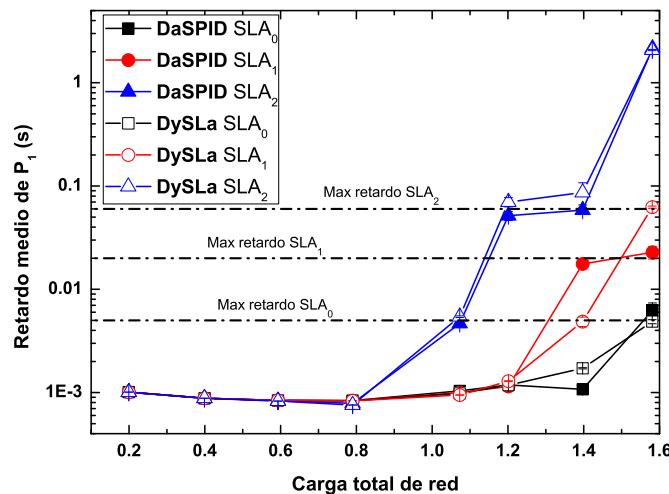


Figura 4.38: Retardo medio de la clase de servicio P_1 frente a la carga de red al comparar DySLa y DaSPID.

aprecia en la Figura 4.38, el retardo medio para estos usuarios es ligeramente superior que el máximo permitido (establecido en 60 ms), mientras que DaSPID sí consigue ajustarlos a su nivel máximo. También se puede observar otra diferencia en el comportamiento de ambos algoritmos. Mientras que DySLa comprueba únicamente que los retardos están por debajo de las cotas estipuladas pero no los evoluciona hacia dichos niveles, DaSPID intenta siempre evolucionar los retardos hacia dichos límites estipulados. De esta forma, los usuarios más prioritarios se van a beneficiar de retardos más bajos en clases de servicio no prioritarias, esto es, van a tener un comportamiento global de retardo mejor en todas las clases de servicio. De esta forma, tal y como se observa en la Figura 4.38, DySLa ofrece un retardo al SLA₁ por debajo de su umbral, mientras que DaSPID lo evoluciona hasta su retardo límite correspondiente (fijado en 20 ms). Por ello, el SLA₀ se beneficiará más, puesto que la clase de servicio P₂ de este perfil presentará retardos menores si se compara con DySLa, tal y como se verá en el siguiente apartado.

Por último, destacar que cuando los recursos disponibles no son suficientes para garantizar que el retardo medio de las clases prioritarias esté por debajo del límite establecido, DaSPID está diseñado para aplicar una política agresiva. En este sentido, DaSPID va a conseguir que el retardo medio esté en el límite estipulado para los usuarios más prioritarios pero en detrimento del SLA de menor prioridad. Esto es, los usuarios del SLA de menor prioridad ceden ancho de banda a los usuarios de mayor prioridad hasta que éstos alcanzan las cotas de retardo máximo para las clases de servicio con restricción en retardo. En concreto, para el escenario de simulación escogido, el ancho de banda disponible no es suficiente para garantizar los retardos para la carga de red de 1.6, esto es, todas las ONUs transmitiendo a su tasa máxima de 100 Mbit/s. Para esta carga máxima de red, todos los usuarios cumplen con las restricciones de P₀ independientemente del SLA tal y como se vio en la Figura 4.37. En cuanto al comportamiento de la clase de servicio P₁, tal y como se aprecia en la en la Figura 4.38, DaSPID mantiene el retardo del SLA menos prioritario (SLA₂) entorno a los 3 s mientras que el retardo de los dos SLAs más prioritarios se mantiene en las cotas estipuladas por el proveedor de servicios. Por otro lado, DySLa sigue una estrategia parecida consiguiendo garantizar el retardo de P₁ del SLA más prioritario. No obstante, el SLA de prioridad intermedia (SLA₁) ya no cumple con el retardo límite, y consecuentemente, tampoco el SLA₂.

Retardo medio de paquetes del servicio P₂

Respecto a la clase de servicio menos prioritaria P₂, la Figura 4.39 representa el retardo medio frente a la carga de red para los tres perfiles de usuario al comparar DySLa y DaSPID. Esta clase de servicio no tiene que cumplir con ninguna restricción respecto al

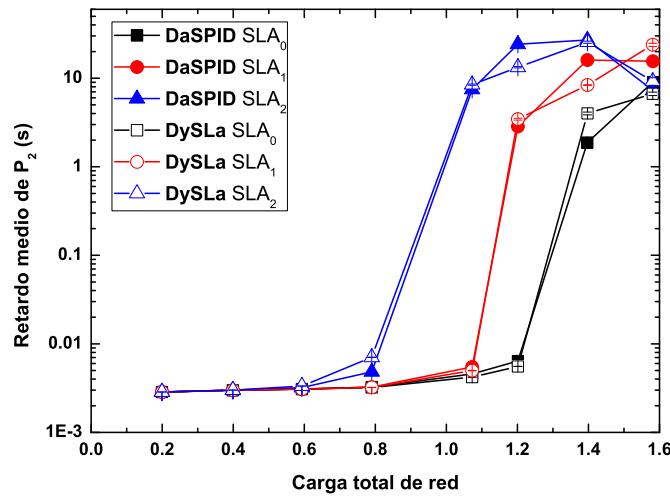


Figura 4.39: Retardo medio de la clase de servicio P_2 frente a la carga de red al comparar DySLa y DaSPID.

retardo. Comparando ambos algoritmos se observa que los resultados son similares. En concreto, las únicas diferencias aparecen para la carga de red de 1.4, en donde DaSPID consigue un retardo menor para el SLA_0 , mientras que DySLa lo hace para el SLA_1 . Esta diferencia radica en el hecho comentado anteriormente de que en DaSPID, una vez que todos los retardos están por debajo del umbral establecido, se intentan evolucionar hacia los límites máximos para beneficiar a las clases no prioritarias del SLA más prioritario, que es lo que ocurre en este caso.

4.7.4.3. Independencia de las condiciones iniciales de ancho de banda

Otro de los puntos fuertes de DaSPID es que ha sido diseñado, al igual que DySLa, para ser independiente de las condiciones iniciales de ancho de banda. Sin embargo, DaSPID, al modificar los anchos de banda máximos dependiendo del error cometido, es capaz de adaptarse de forma más rápida independientemente de los pesos utilizados por ambos algoritmos.

Para demostrar la independencia frente a las condiciones iniciales, se han llevado a cabo diferentes simulaciones comparando DySLa y DaSPID bajo dos escenarios de red distintos. En concreto, se ha estudiado la velocidad de convergencia a las cotas de retardo máximas consideradas cuando se utilizan diferentes anchos de banda máximos iniciales. La Tabla 4.11 recoge los pesos utilizados para calcular los anchos de banda iniciales en los dos escenarios, siguiendo ambos algoritmos la ecuación (4.4).

	W^{sla_0}	W^{sla_1}	W^{sla_2}
Escenario 1	1.5	1	0.5
Escenario 2	0.5	1	1.5

Tabla 4.11: Pesos utilizados en los escenarios considerados para establecer los anchos de banda iniciales en DaSPID y DySLa para los tres perfiles de usuario considerados.

Por tanto, la Figura 4.40 (a), (b) y (c) representa la evolución en tiempo real del retardo medio de los paquetes para la clase de servicio de prioridad intermedia (P_1) para el SLA_0 , SLA_1 y SLA_2 , respectivamente, cuando se considera una carga de red de 1.4 (ONUs transmitiendo a una tasa de 87.5 Mbit/s) en los dos escenarios de la Tabla 4.11. Tal y como puede observarse, DySLa depende en gran medida del escenario. De hecho, para el perfil SLA_0 , Figura 4.40 (a), puede observarse como DySLa presenta un peor comportamiento que DaSPID para el Escenario 2, ya que alcanza el retardo máximo estipulado para este SLA (5 ms) en aproximadamente 500 s. Es importante destacar que este valor es muy elevado, ya que los usuarios del SLA_0 tienen la máxima prioridad en la red. Por el contrario, DaSPID alcanza el valor de retardo estipulado en menos de 40 s. Para el SLA_1 (Figura 4.40 (b)), se puede observar como DySLa consigue que los retardos se sitúen por debajo del nivel máximo estipulado en 300 s y 200 s aproximadamente para el Escenario 1 y Escenario 2, respectivamente, mientras que DaSPID lo consigue en tan solo 50 s. Por último, el mismo comportamiento puede observarse para el SLA de menor prioridad (SLA_2), donde DaSPID consigue que los retardos se sitúen por debajo del límite establecido (60 ms) para todos los escenarios de forma más rápida que DySLa. En concreto, DySLa presenta un peor comportamiento para el Escenario 1, donde consigue reducir el retardo por debajo del límite en 300 s. Por lo tanto, se puede concluir, que el comportamiento de DaSPID no depende en gran medida de las condiciones iniciales del ancho de banda y que además el algoritmo hace que los retardos evolucionen hacia los valores estipulados para todos los perfiles de forma más rápida que DySLa. Por el contrario, el comportamiento de DySLa depende en gran medida del escenario considerado, en concreto, de las condiciones de ancho de banda iniciales. Particularmente, existen escenarios donde DySLa tarda tanto en conseguir los requisitos de retardo, que su aplicación en un contexto de red real puede no ser posible.

Para completar el estudio es importante analizar la evolución en tiempo real del ancho de banda máximo, ya que esta variación es un indicador de la estabilidad de los algoritmos. Esto es, cuánto más pequeñas sean las variaciones de los máximos, existirá una mejor adaptación a los límites máximos establecidos y una mayor estabilidad del retardo medio. La Figura 4.41 (a), (b) y (c) muestra la evolución instantánea del ancho de banda máximo permitido para cada ONU del SLA_0 , SLA_1 y SLA_2 respectivamente,

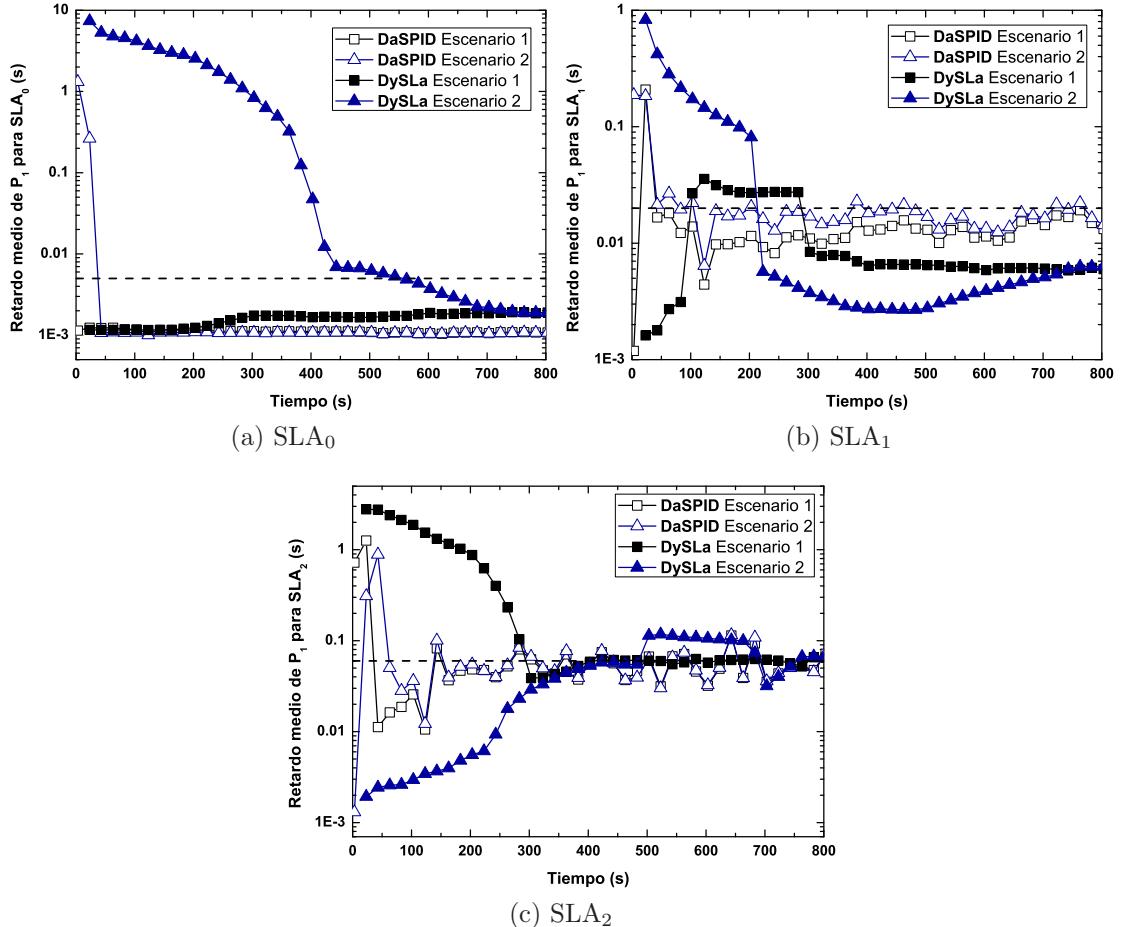


Figura 4.40: Evolución en tiempo real del retardo medio de la clase de servicio P₁ para las condiciones iniciales de ancho de banda de la Tabla 4.11 en DySLa y DaSPID. (a) SLA₀ (b) SLA₁ (c) SLA₂.

al comparar DySLa y DaSPID. Tal y como se puede apreciar, DaSPID muestra un comportamiento más estable independientemente del escenario considerado para todos los perfiles de abonado. Por el contrario, DySLa modifica continuamente el ancho de banda máximo a lo largo del tiempo, obteniendo una mayor inestabilidad, en concreto para el SLA₀ y SLA₂ (Figura 4.41 (a) y (c)). Por tanto, DaSPID presenta una mayor robustez que DySLa, obteniendo una mejor adaptación a los límites máximos estipulados por el proveedor de servicios para el tráfico más sensible al retardo.

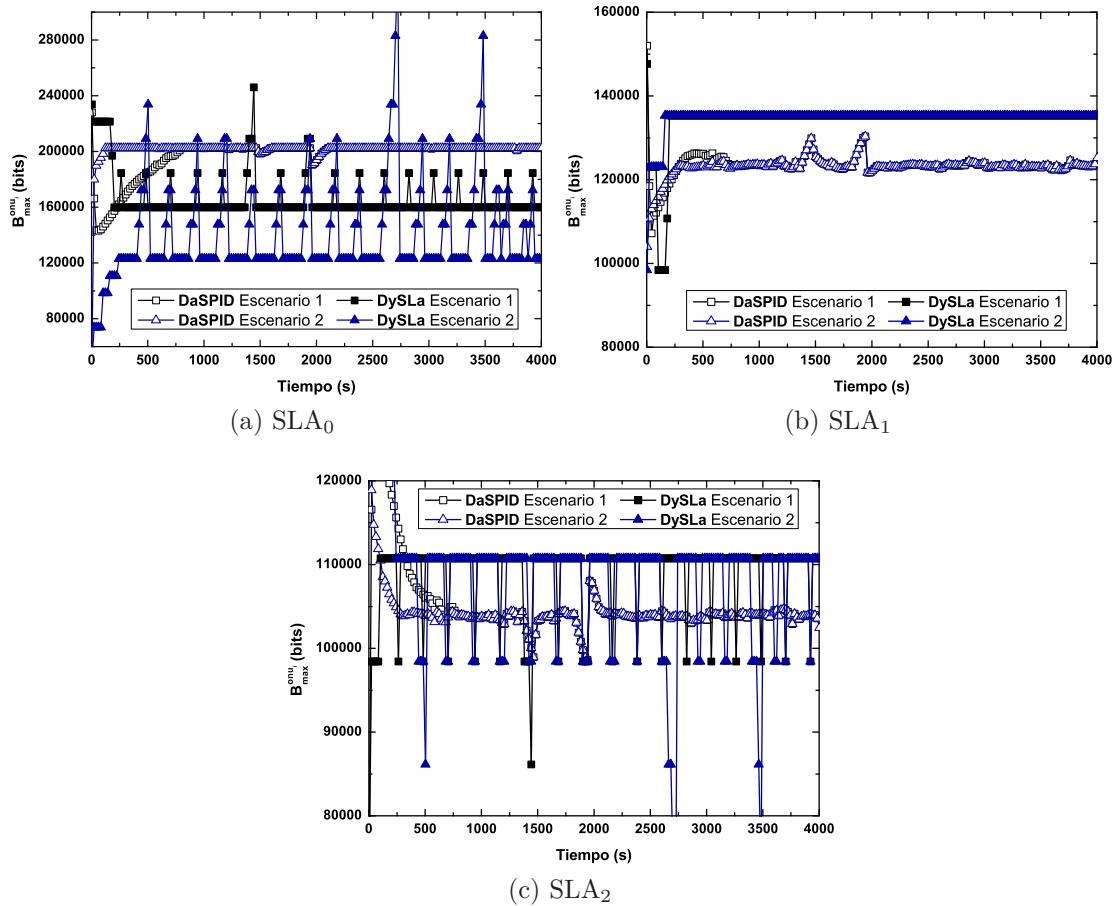


Figura 4.41: Evolución en tiempo real del ancho de banda máximo permitido para los escenarios de la Tabla 4.11 en DySLa y DaSPID. (a) SLA₀ (b) SLA₁ (c) SLA₂.

4.7.4.4. Independencia de las cotas de retardo de las clases de servicio

DaSPID es también capaz de reaccionar con diferentes condiciones iniciales de ancho de banda y con diferentes cotas máximas del retardo según distintos valores establecidos por los proveedores de servicio. Por ello, a continuación se analiza el comportamiento de DaSPID y DySLa cuando se considera otro conjunto de cotas para los servicios prioritarios. Así, en la Tabla 4.12 se muestran los nuevos valores establecidos para el nuevo escenario de simulación. Se han considerado unos pesos iniciales iguales a $W^{sla_0} = 1.5$, $W^{sla_1} = 1$ y $W^{sla_2} = 0.5$ para analizar el comportamiento de ambos algoritmos.

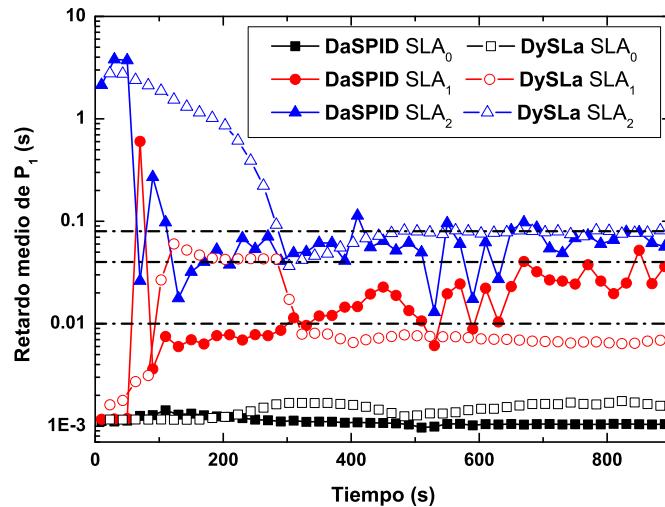
En la Figura 4.42 se muestra la evolución en tiempo real de los retardos de los distintos SLAs para la clase de servicio P₁ cuando las ONUs transmiten a una tasa

Clase de servicio	SLA	Retardo
P_0	SLA_0	
	SLA_1	1.50 ms
	SLA_2	
P_1	SLA_0	10 ms
	SLA_1	40 ms
	SLA_2	80 ms
P_2	SLA_0	
	SLA_1	
	SLA_2	-

Tabla 4.12: Cotas máximas de retardo consideradas en el nuevo escenario de simulación.

alta de 87.5 Mbit/s. Solo se ha representado la adaptación para esta clase de servicio porque para P_0 no se han modificado las cotas de 1.5 ms con respecto a las simulaciones consideradas anteriormente. Como puede apreciarse en la Figura 4.42, ambos algoritmos consiguen adaptarse a los nuevos niveles establecidos. Sin embargo, DaSPID lo hace de una forma más rápida que DySLa. Por tanto, el comportamiento de DaSPID es igual al obtenido con los niveles de retardo considerados en el apartado anterior.

Por consiguiente, se ha demostrado no solo la mejor adaptabilidad de DaSPID frente a distintas condiciones iniciales de ancho de banda, sino también la independencia con respecto a las cotas de retardo estipuladas por el proveedor de servicios para las distintas clases de servicio y perfiles de abonado considerados.

Figura 4.42: Evolución en tiempo real del retardo medio de la clase de servicio P_1 para DySLa y DaSPID cuando las ONUs transmiten a una tasa de 87.5 Mbit/s considerando las cotas de retardo máximo de la Tabla 4.12

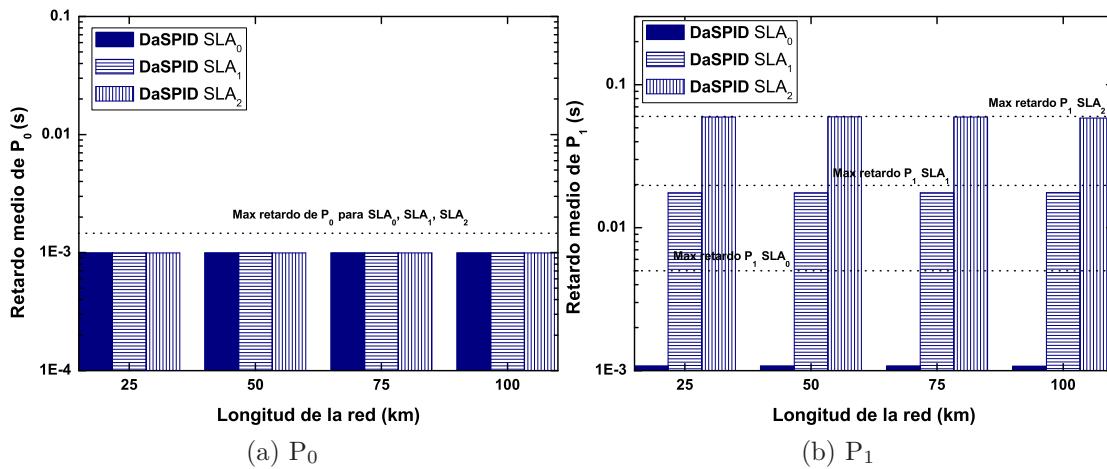


Figura 4.43: Retardo medio obtenido en DaSPID para las clases de servicio prioritarias considerando redes LR-EAPON y EPON de diferente longitud. (a) P₀ (b) P₁.

4.7.4.5. Análisis de DaSPID considerando redes de distinta longitud

Por último, es importante destacar que DaSPID puede utilizarse no solo en redes de largo alcance LR-EAPON, sino también en redes EPON de menor longitud. Para demostrar esto, la Figura 4.43 (a) y (b) representa el retardo medio de las clases de servicio prioritarias P₀ y P₁, respectivamente, para cada SLA cuando se consideran distancias de 25, 50, 75 y 100 km para la carga de red de 1.4 (ONUs transmitiendo a 87.5 Mbit/s). Las cotas de retardo máximo establecidas son las recogidas en la Tabla 4.9, esto es 1.5 ms para la clase de servicio P₀ de los tres SLAs, y 5, 20 y 60 ms para la clase de servicio P₁ del SLA₀, SLA₁ y SLA₂, respectivamente. Tal y como se observa, independientemente de la distancia considerada, DaSPID consigue que el retardo medio para cada clase de servicio y SLA considerado estén por debajo de la cota establecida. Por tanto, el algoritmo DaSPID puede utilizarse en redes LR-EAPON y EPON de diferente cobertura manteniendo un comportamiento óptimo.

4.8. Análisis de prestaciones de DaSPID bajo patrones de tráfico asimétrico y dinámico

En este apartado del capítulo se van a analizar las prestaciones del algoritmo DaSPID bajo condiciones de tráfico más realistas, como el asimétrico y el dinámico. De esta forma se pretende analizar el comportamiento y demostrar la robustez de DaSPID ante diferentes patrones de tráfico.

4.8.1. Escenarios con un patrón de tráfico asimétrico

Para realizar el análisis con tráfico asimétrico, se han llevado a cabo simulaciones fijando las cotas máximas de retardo recogidas en la Tabla 4.9, pero considerando que cada ONU transmite a una tasa aleatoria entre 0 y 100 Mbit/s (máxima tasa de transmisión) y limitando la contribución de cada SLA a la carga total de red, de igual modo que el mostrado en el apartado 4.5.1. En concreto, la contribución de cada SLA a la carga total de red de los dos nuevos escenarios considerados aparece recogida en la Tabla 4.13.

	Carga de red	SLA ₀	SLA ₁	SLA ₂
Escenario 1	1.0	10 %	30 %	60 %
Escenario 2	1.2	4 %	30 %	66 %

Tabla 4.13: Contribución de cada SLA a la carga total de red para los distintos escenarios considerados con tráfico asimétrico en DaSPID.

La Figura 4.44 muestra el retardo medio de las clases de servicio prioritarias (P_0 en negro y P_1 en rojo) para cada ONU en las condiciones del Escenario 1. En este escenario, la carga de total es de 1.0 (el ancho de banda demandado por todas las ONUs es de 1 Gbit/s), donde el 10 % es generado aleatoriamente por las dos ONUs del SLA₀, el 30 % lo generan las seis ONUs del SLA₁ y el 60 % restante entre las ocho ONUs del SLA₂. Por tanto, el ancho de banda demandado por cada ONU (representado en la Figura 4.44 por una estrella azul y referido al eje de la derecha también en azul) es diferente al tener un carácter aleatorio. Como se puede apreciar, el comportamiento de DaSPID es adecuado a las condiciones de red estipuladas. Esto es, el algoritmo es capaz de mantener los retardos

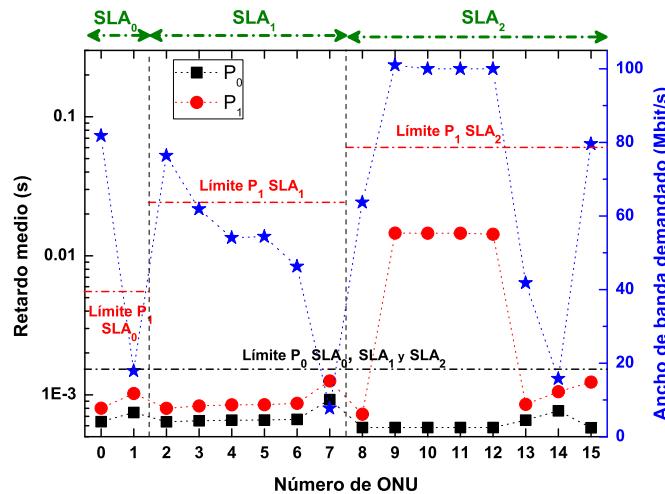


Figura 4.44: Retardo medio de P_0 y P_1 para cada ONU para el Escenario 1 de la Tabla 4.13.

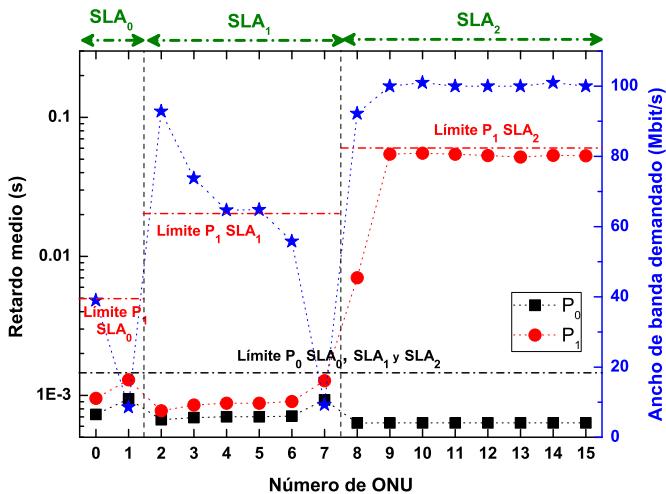


Figura 4.45: Retardo medio de P_0 y P_1 para cada ONU para el Escenario 2 de la Tabla 4.13.

de ambas clases de servicio prioritarias por debajo de la cota máxima establecida para cada SLA.

De forma similar, la Figura 4.45, muestra el retardo medio de cada ONU en las condiciones del Escenario 2. En este caso, la carga de red considerada para este escenario es de 1.2 (la demanda total de las ONUs es de 1.20 Gbit/s) donde el 4% es generado aleatoriamente entre las dos ONUs del SLA₀, el 30% por las seis ONUs del SLA₁ y el 66% restante entre las ocho ONUs del SLA₂. Al igual que ocurría en el Escenario 1, en el Escenario 2 DaSPID también consigue que los retardos medios de las clases de servicio prioritarias P_0 y P_1 estén por debajo de las cotas máximas establecidas en función de la prioridad del SLA.

Por tanto, se demuestra que DaSPID mantiene su funcionalidad bajo un patrón de tráfico asimétrico, lo cual demuestra su robustez ante un escenario con una distribución de tráfico más realista.

4.8.2. Escenarios con un patrón de tráfico dinámico

En esta sección se va a analizar el comportamiento de DaSPID en un escenario con tráfico dinámico. Para ello, se ha asumido que las ONUs varían su carga siguiendo la ecuación (4.5), al igual que en SPID y DySGAB. En este caso, y por reducir el número de gráficos, solo se ha utilizado una única frecuencia de cambios en tiempo real de la carga, la cual se ha fijado a 900 s.

En todas las figuras que se muestran a continuación, aparece representado en azul y referido al eje de la derecha, el ancho de banda demandado por cada ONU a lo largo del

tiempo. Por otro lado, los valores de retardo obtenidos por DaSPID en tiempo real para los distintos anchos de banda demandados se muestran en verde y están referidos al eje de la izquierda.

La Figura 4.46 (a), (b) y (c) ilustra la evolución del retardo de la clase de servicio P_0 para los tres SLAs considerados, SLA_0 , SLA_1 y SLA_2 respectivamente. Tal y como se puede observar, el retardo medio obtenido a lo largo del tiempo para las distintas cargas de la ONU, esto es, para los distintos anchos de banda demandados (representados en azul), está siempre por debajo de la cota máxima establecida para esta clase de servicio (1.5 ms) en todos los SLAs. Por tanto, DaSPID es capaz de garantizar que el retardo del servicio más prioritario de la red se sitúe por debajo del límite establecido para todos los SLAs.

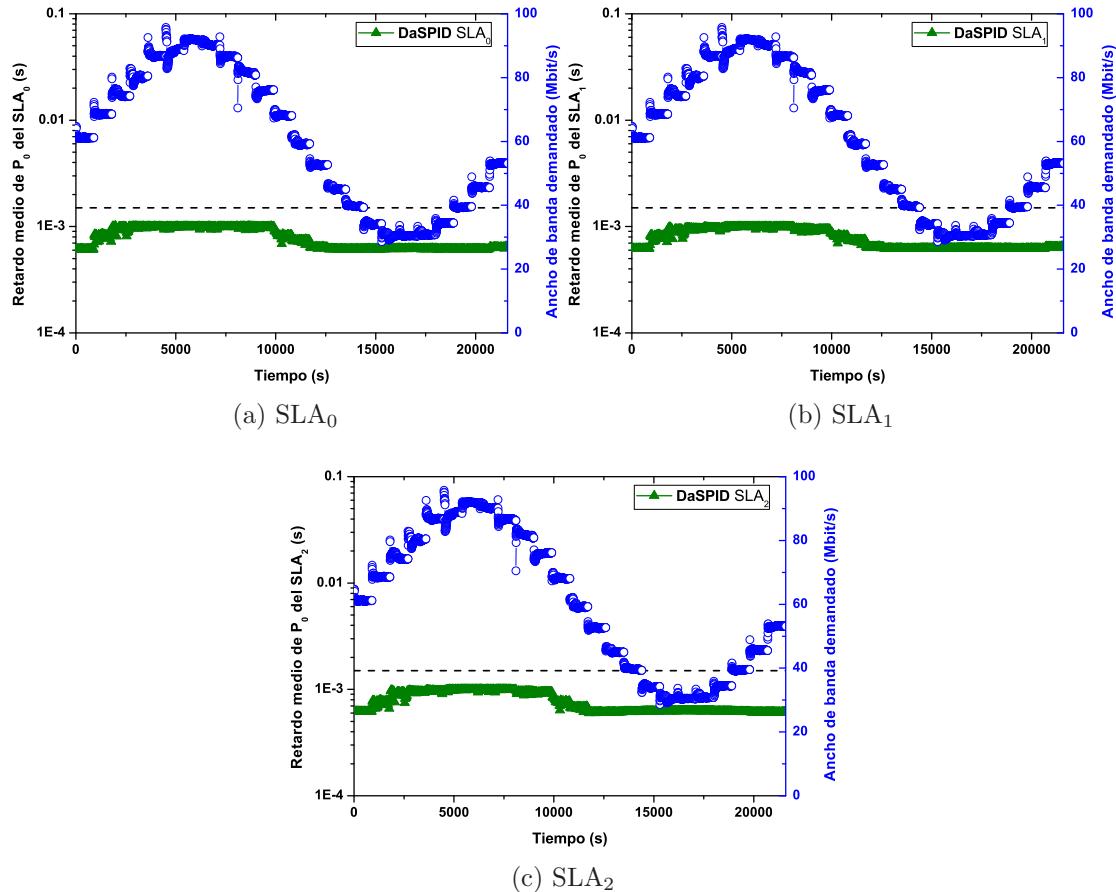


Figura 4.46: Retardo medio obtenido en DaSPID para la clase de servicio P_0 cuando varía el ancho de banda demandado a lo largo del tiempo. (a) SLA_0 (b) SLA_1 (c) SLA_2 .

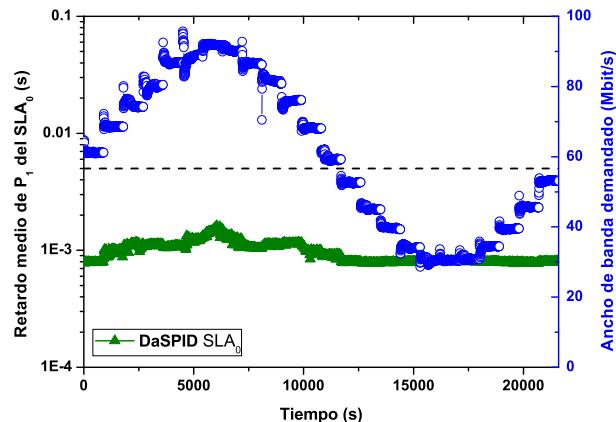


Figura 4.47: Retardo medio obtenido en DaSPID para la clase de servicio P_1 y el SLA_0 cuando varía el ancho de banda demandado a lo largo del tiempo.

Con respecto a la clase de servicio de prioridad intermedia P_1 , la Figura 4.47 muestra el retardo medio obtenido a lo largo del tiempo cuando el ancho de banda demandado varía a lo largo del tiempo para el SLA_0 . Como puede verse, para el perfil de abonado de mayor prioridad, DaSPID consigue que el retardo medio esté siempre por debajo del límite establecido para este SLA (5 ms), independientemente del ancho de banda demandado.

La Figura 4.48 presenta el retardo medio obtenido a lo largo del tiempo para el SLA_1 . Para este perfil de abonado, el retardo medio se encuentra por debajo del límite máximo establecido para este SLA (20 ms). No obstante, en ciertos instantes de tiempo el retardo se dispara bruscamente, en concreto entorno a los 4500 s. Tal y como se puede

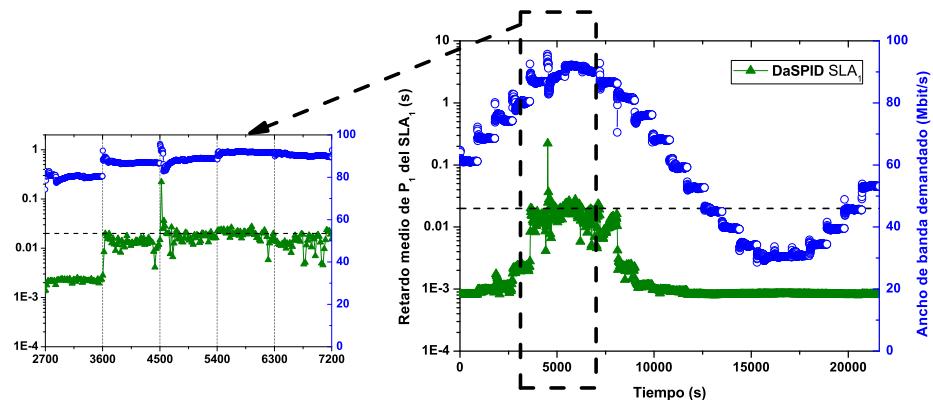


Figura 4.48: Retardo medio obtenido en DaSPID para la clase de servicio P_1 y el SLA_1 cuando el ancho de banda demandado varía a lo largo del tiempo.

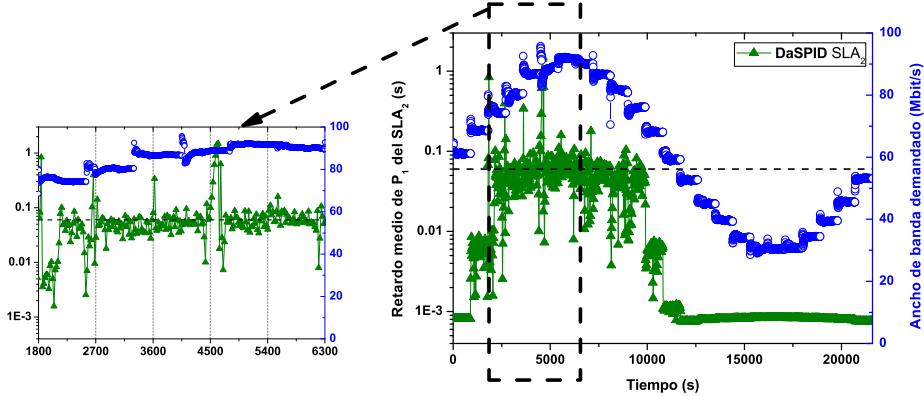


Figura 4.49: Retardo medio obtenido en DaSPID para la clase de servicio P_1 y el SLA_2 cuando el ancho de banda demandado varía a lo largo del tiempo.

observar en el zoom situado a la izquierda de la imagen, se trata de una inestabilidad producida por un cambio repentino en la carga. Este cambio produce que el retardo aumente eventualmente de forma brusca, pero DaSPID lo corrige de forma rápida para que el retardo se sitúe en el límite establecido.

Para el SLA_2 (Figura 4.49), se puede ver como DaSPID consigue situar el retardo medio por debajo del límite máximo establecido para este SLA (60 ms) para todas las cargas de red. Sin embargo, y de forma similar a lo que ocurría con el SLA_1 , en algunos instantes en los que la carga cambia repentinamente, el retardo aumenta de forma brusca debido a dicha fluctuación súbita de la carga de red, pero DaSPID es capaz de hacer que evolucione al límite estipulado de forma rápida.

Por último, la Figura 4.50 muestra la variación en tiempo real del ancho de banda máximo permitido para cada ONU ($B_{max}^{onu_i}$) cuando varía el ancho de banda demandado. Como puede observarse, los máximos de cada SLA permanecen estables. Las únicas variaciones se aprecian para cargas de red altas (ancho de banda demandado por encima de los 65 o 70 Mbit/s), donde el máximo del SLA_0 y del SLA_1 descienden y el del SLA_2 aumenta para que este perfil pueda cumplir con el límite máximo del retardo medio.

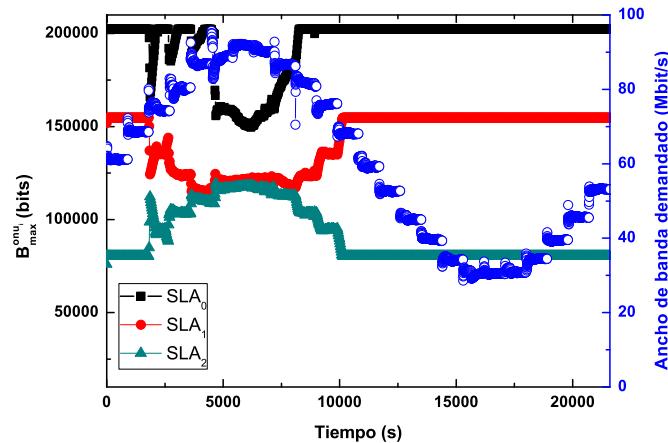


Figura 4.50: Variación del ancho de banda máximo en DaSPID para los tres SLAs considerados cuando el ancho de banda demandado varía a lo largo del tiempo.

4.9. Conclusiones

En este capítulo de la Tesis se han incorporado controladores PID en la gestión de recursos en redes LR-EPON. En este sentido, en la primera parte del capítulo, se ha presentado un innovador algoritmo de asignación de ancho de banda dinámico, denominado SPID (*Service level agreement PID*). Este algoritmo, utiliza un controlador PID en la asignación dinámica de ancho de banda, algo que hasta el momento no se había utilizado nunca en las redes LR-EPON y EPON. El algoritmo SPID ofrece diferenciación de usuarios garantizando de forma óptima los niveles de ancho de banda estipulados por el proveedor de servicios. En concreto, SPID actualiza los anchos de banda máximos permitidos de forma automática y dinámica mediante un controlador PID, para que todas las ONUs cumplan con sus requisitos de ancho de banda en tiempo real. De esta forma, como el PID tiene en cuenta el error cometido en el presente, los errores pasados y una estimación del error futuro a la hora de ajustar el ancho de banda, es más estable y robusto que otros algoritmos con una finalidad similar, tales como DySGAB.

Para demostrar las funcionalidades y prestaciones del algoritmo SPID, en primer lugar se ha realizado el proceso de sintonización del controlador PID y se han fijado los parámetros relacionados con el diseño del controlador para obtener un comportamiento óptimo del algoritmo. Tras ello, se han llevado a cabo una serie de simulaciones bajo diferentes escenarios. Los resultados de las simulaciones han demostrado que no solo SPID auto-adapta de forma eficiente los recursos cuando se producen cambios en algunos parámetros de red o en los propios niveles de ancho de banda estipulados por los proveedores, sino que además lo hace de forma más estable y rápida que DySGAB.

En concreto, se ha visto que DySGAB presenta una gran dependencia con el valor de algunos parámetros, como el ancho de banda involucrado en la adaptación de los máximos, la frecuencia de ejecución o incluso las condiciones iniciales de ancho de banda. Por el contrario, SPID gracias al uso del PID presenta una mayor independencia de tales parámetros, lo que implica una mejor adaptabilidad a diferentes escenarios y condiciones de red. Además, se ha demostrado que SPID se puede emplear en redes LR-EPON y EPON con diferente distancia manteniendo sus funcionalidades, así como su aplicabilidad bajo patrones de tráfico más reales como el asimétrico y el dinámico. Por tanto, el uso de un controlador PID ofrece buenos resultados para garantizar los anchos de banda estipulados.

Debido al buen funcionamiento que ha demostrado SPID, en la segunda parte de este capítulo, se han llevado a cabo las modificaciones necesarias para introducir el uso de controladores PID en la gestión del retardo medio de las clases de tráfico prioritario. En este sentido, se ha diseñado el algoritmo DaSPID (*Delay aware Service level agreement PID*) para controlar el retardo medio de los paquetes de las clases de servicio sensibles al retardo. Este algoritmo, lleva a cabo una doble diferenciación, a nivel de usuarios y a nivel de clases de servicio, para ofrecer una mejor calidad de servicio y calidad de experiencia. Para ello, el algoritmo DaSPID utiliza un controlador P que modifica constantemente el ancho de banda máximo permitido a cada SLA con el fin de que cada perfil de abonado cumpla con las cotas de retardo máximo establecidas para las distintas clases de servicio dependiendo de la prioridad del perfil de abonado.

En primer lugar, se realizó el proceso de sintonización y se analizó el tipo de controlador que mejor se adapta al proceso del control de retardo en una red LR-EPON. En concreto, el controlador P es el que mejor comportamiento ha presentado para esta aplicación concreta, por lo que fue seleccionado. Tras la elección del tipo de controlador, se llevaron a cabo simulaciones para analizar las prestaciones del algoritmo DaSPID, que ha sido comparado con el algoritmo DySLa, un algoritmo existente en la literatura con la misma finalidad. Los resultados de las simulaciones demostraron que DySLa presenta una gran dependencia de un parámetro interno que utiliza para actualizar los anchos de banda máximos con el fin de garantizar los retardos de las clases de servicio sensibles. Además, este algoritmo también es dependiente de las condiciones iniciales de la red a la hora de mantener los retardos por debajo de los límites establecidos. Por el contrario, DaSPID no solo demostró que es independiente de las condiciones iniciales de ancho de banda, sino que además es más estable y rápido que DySLa en la evolución de los retardos a las cotas máximas dependiendo de la prioridad del servicio y de los usuarios. En concreto, se mostró que para ciertos escenarios de red, el uso de DySLa no

es posible ya que evoluciona los retardos por debajo de las cotas en cientos de segundos, haciendo que la transición sea demasiado lenta para un contexto de red real. Consecuentemente, ha quedado patente que el comportamiento de DaSPID supera el de otros algoritmos como DySLa, en el control del retardo medio de los paquetes de los servicios con restricciones en el retardo, dependiendo de las características y prioridad del perfil de abonado. Además, se verificó que DaSPID puede ser utilizado en redes LR-EAPON y EPON de diferente cobertura manteniendo su buen comportamiento y funcionalidades, así como su aplicabilidad bajo patrones de tráfico más reales como el tráfico asimétrico y dinámico.

Por lo tanto, se puede concluir que la aplicación de controladores PID en la gestión de recursos en redes LR-EAPON, tales como garantizar niveles de ancho de banda o cotas máximas en los retardos según la prioridad del tráfico y del usuario, ofrece unos resultados mejores que otros algoritmos existentes en la literatura. Esto además, es algo indispensable por parte del proveedor de servicios para garantizar una calidad de servicio integral y óptima.

Además, la introducción de los controladores PID en la gestión de recursos, puede extenderse de forma fácil para el control de otros parámetros de red, como la probabilidad de bloqueo o el *jitter*.

Capítulo 5

Métodos de sintonización automática de PIDs basados en algoritmos genéticos

5.1. Introducción

El método de sintonización de Ziegler-Nichols, tal y como se ha visto en el Capítulo 4, es un método sencillo que proporciona unos resultados adecuados en la sintonización de controladores PID para la gestión de recursos en redes LR-EAPON. En particular, en los casos de uso analizados en esta Tesis, esto es, en el control del ancho de banda y del retardo medio del tráfico prioritario, donde no existe una función de transferencia o un modelo matemático que represente adecuadamente el comportamiento del sistema. Sin embargo, este método presenta una gran desventaja, que es su carácter manual. Es decir, es necesaria la realización de un análisis experimental manual donde se va cambiando el valor de la constante proporcional y se observa tanto si se producen oscilaciones, como la magnitud de las mismas. Por tanto, es un proceso laborioso que puede consumir mucho tiempo y con el que se pueden llegar a obtener resultados poco precisos.

Por este motivo, en este capítulo se desarrollan técnicas de sintonización automática basadas en algoritmos genéticos. De esta forma, el proceso de sintonización del PID se automatiza, obteniendo unos valores de sintonización mucho más precisos y en la mayoría de los casos este proceso de sintonización se realiza de forma mucho más rápida. Se ha optado por el uso de algoritmos genéticos, ya que han demostrado resolver de forma efectiva distintos problemas de optimización en una gran diversidad de sistemas. En este sentido, existen diferentes ejemplos en la literatura que plantean el uso de al-

goritmos genéticos en la sintonización de controladores PID, y que han conseguido una optimización del proceso de sintonización gracias al uso de esta técnica. Por tanto, en este capítulo, tras la exposición de algunos conceptos básicos de los algoritmos genéticos, se explica cómo estos se han incorporado a la sintonización de los controladores PID en los algoritmos desarrollados en el capítulo anterior, esto es, en SPID y en DaSPID para el control del ancho de banda y del retardo, respectivamente. Para ello se explicará el diseño del algoritmo genético propuesto y a continuación se establecerán mediante simulaciones los parámetros específicos del mismo (tamaño de la población, número de generaciones a considerar, etc.) para después comparar su funcionamiento con el método de sintonización de Ziegler-Nichols.

5.2. Fundamentos básicos de los algoritmos genéticos

Los algoritmos genéticos [116, 117] son métodos adaptativos, generalmente usados en problemas de búsqueda y optimización de parámetros, basados en mecanismos de selección natural y genética. En la selección natural de las especies, los individuos más fuertes o mejor adaptados al entorno suelen sobrevivir, mientras que los más débiles suelen morir. Esto es lo que se denomina un entorno competitivo. Además, esto implica que los genes de los individuos mejor adaptados se propagarán a un número cada vez mayor de individuos de las sucesivas generaciones. En concreto, los algoritmos genéticos se basan en los siguientes conceptos:

- Noción Darwiniana de aptitud (*fitness*) que influye en generaciones futuras. Esto es, los individuos se adaptarán mejor a la solución requerida generación tras generación.
- Apareamiento que produce descendientes en generaciones futuras.
- Operadores genéticos que determinan la configuración genética de los descendientes (tomada de los padres).

La principal característica de los algoritmos genéticos es que representan cada solución del espacio de soluciones como un individuo definido por un conjunto de parámetros. Estos parámetros se denominan genes, por alusión a los genes en los seres vivos que determinan las características de cada individuo. El conjunto de todos los genes de un individuo se codifica en una cadena de valores denominada cromosoma [117]. De forma similar a los genes en la vida real, algunas características son pasadas de padres a hijos produciendo nuevas soluciones (hijos) mediante la combinación de soluciones actuales (padres). Para considerar qué individuo se adapta mejor al entorno hay que definir una

función objetivo que permita evaluar la bondad de cada individuo. A esta bondad se le denomina comúnmente salud (*fitness*) y a la función que permite calcular la salud de cada individuo se la conoce como función de salud.

El proceso de adaptación en los algoritmos genéticos no se realiza cambiando incrementalmente un solo individuo, sino manteniendo una población de individuos sobre los cuales se aplican los operadores genéticos para generar los nuevos individuos. Cada individuo de la población tiene asociado un valor de salud o aptitud. Estos valores de salud se emplean para determinar qué individuos son los más aptos para formar nuevos individuos. Por analogía a la vida real, se supone que aquellos individuos más sanos formarán más individuos que los menos sanos y así las buenas propiedades de cada individuo se irán propagando generación a generación.

Uno de los puntos clave en el diseño de un algoritmo genético es la representación del espacio de soluciones como un conjunto (cromosoma) de parámetros (genes). La codificación suele hacerse mediante valores binarios. Se asigna un determinado número de bits a cada parámetro y se realiza una discretización de la variable representada por cada gen. El número de bits asignados dependerá del grado de ajuste que se desee alcanzar. Además, no todos los parámetros tienen por qué estar codificados con el mismo número de bits. También pueden existir representaciones que codifiquen directamente cada parámetro con un valor entero, real o punto flotante. Aunque a estas representaciones se las acusa de degradar el paralelismo implícito de las representaciones binarias, permiten el desarrollo de operadores genéticos más específicos al campo de aplicación del algoritmo genético [118].

El proceso de los algoritmos genéticos empieza con la creación de una población inicial, con un tamaño denotado por N , que generalmente se crea de forma aleatoria. Sin embargo, los algoritmos genéticos pueden incluir en la población inicial determinados individuos, previamente conocidos y que tienen una función de salud muy alta. De esta forma, se mejoran las características de dicha población inicial y se favorece la evolución hacia la población final óptima de forma más rápida y efectiva.

En cada generación, la población evoluciona por medio de los operadores genéticos para formar la siguiente generación, que se espera sea mejor que la generación de sus padres. Los operadores genéticos más comunes suelen ser el cruce y la mutación. Para determinar qué individuos de la población padre van a ser seleccionados para formar la población hijo se emplea un procedimiento de selección cada vez que se aplica un operador genético. Lo más habitual es que la probabilidad de que un individuo sea seleccionado dependa de la evaluación de su función objetivo, esto es, de su salud. Los individuos más aptos tienen una mayor probabilidad de contribuir en la siguiente ge-

neración. No obstante, no se deben eliminar por completo las opciones de reproducción de los individuos menos aptos, ya que en pocas generaciones la población se volvería demasiado homogénea.

El proceso de selección se puede implementar de varias formas. La forma más común es utilizar la selección por ruleta [117, 118]. En este método, se asigna a cada individuo una parte de la ruleta de forma proporcional a su salud, donde la suma de todos los porcentajes es la unidad. Los mejores individuos recibirán por tanto una porción de la ruleta mayor que la recibida por los peores. De esta forma, las buenas propiedades se transmiten de generación a generación. Otra alternativa es usar la selección por torneo [118]. La idea principal de este método consiste en realizar la selección en base a comparaciones directas entre individuos. Es decir, se escogen aleatoriamente un número de individuos (generalmente 2) que compiten entre sí, seleccionando el mejor. En esta alternativa el mejor individuo siempre es seleccionado.

Una vez seleccionados los individuos, éstos son recombinados para producir la descendencia que se insertará en la siguiente generación. Para ello se utiliza el operador genético de cruce, que permite la generación de dos nuevos individuos tomando características de dos individuos padres. En este operador genético se seleccionan dos individuos mediante un proceso de selección, y se determina una posición de cruce aleatoria entre los genes de los cromosomas. Tras ello se intercambian las cadenas entre la posición inicial y el punto de cruce y el punto de cruce y la posición final. Se puede realizar un solo cruce o varios dentro del cromosoma. En la Figura 5.1 (a) se muestra el ejemplo del operador de cruce entre dos individuos con un único punto de cruce y en la Figura 5.1 (b) con dos

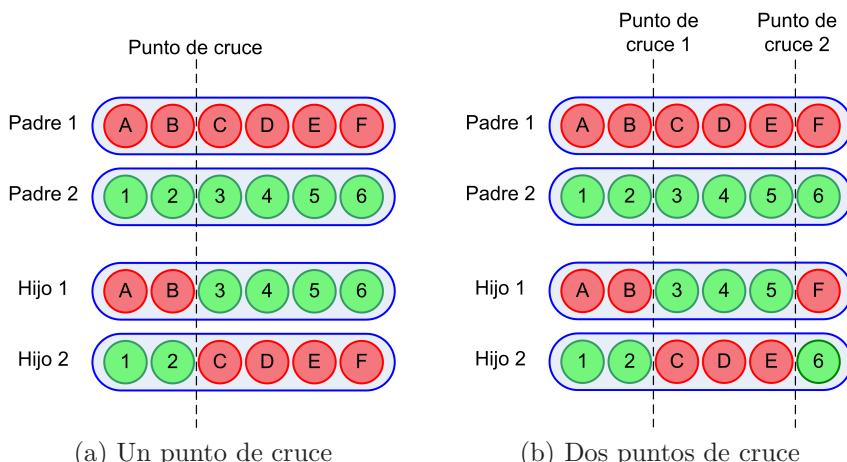


Figura 5.1: Ejemplo del operador de cruce de dos individuos en un algoritmo genético. (a) Con un único punto de cruce (b) Con dos puntos de cruce.

puntos de cruce.

Cuando la generación de descendientes está formada, se puede aplicar a los individuos resultantes el operador de mutación. La mutación de un individuo provoca que alguno de sus genes varíe su valor aleatoriamente. Para llevar a cabo la mutación, se recorre uno a uno los genes de cada descendiente y se decide si el gen va a ser mutado de acuerdo a una probabilidad, denominada $P_{mutacion}$. Si se decide realizar mutación, el gen se intercambia por otro creado de forma aleatoria entre el resto de posibles valores que puede tomar ese gen. En la Figura 5.2 se puede observar el ejemplo de una mutación sobre un gen para dar lugar a un cromosoma mutado. Mediante la mutación se garantiza que ningún punto del espacio de búsqueda tenga una probabilidad nula de ser examinado, evitando con ello que el algoritmo genético caiga en un mínimo local.

En cuanto a qué individuos van a formar la nueva población, existen varias estrategias, que se pueden resumir en la siguiente clasificación:

- La siguiente población estará formada únicamente por los individuos que acaban de ser generados con los operadores genéticos (esquema generacional). Tiene el problema de que se eliminan buenos individuos de la población padre.
- Se mantienen los mejores individuos de una población a la siguiente (esto suele ser la mejor opción). Si se conserva solo el mejor individuo se llama estrategia elitista.
- Existe una opción intermedia en la que se incluye un nuevo operador genético, la reproducción. En la reproducción se crea una copia literal de un individuo de la población padre en la población de descendientes. Este operador se aplica al mismo tiempo que el cruce. Por tanto, cada vez que se selecciona un individuo para formar un descendiente se decide con probabilidad P_{cruce} si dicho individuo se va a reproducir o a cruzar.

Por otro lado, cada vez que se crea un individuo, es necesario evaluar la salud o aptitud de dicho individuo de acuerdo al objetivo de optimización. Si el número de individuos creados en la población de descendientes es mayor que el número de individuos

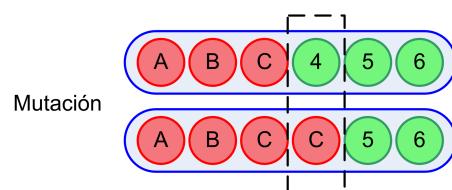


Figura 5.2: Ejemplo del operador mutación en un algoritmo genético cuando solo un gen del cromosoma es mutado.

que se desea que tenga la población, será necesario seleccionar los individuos más sanos para formar la población padre de la nueva generación. El proceso de evolución del algoritmo se repite un número predefinido de veces o hasta que se satisfaga otro criterio de optimización. Como los individuos de cada generación son más y más saludables, la población final puede encontrar la solución óptima o bien una muy cercana a ella. Los criterios de parada más comunes son los siguientes:

- Cuando se alcanza el número de generaciones máximo especificado.
- Cuando los mejores individuos de la población representan soluciones suficientemente buenas para el problema que se desea resolver bajo unos criterios estipulados.
- Cuando la población ha convergido. Se considera que un gen ha convergido cuando un porcentaje elevado de los individuos de la población, típicamente del 95 %, tienen el mismo valor para dicho gen, en caso de trabajar con codificaciones binarias. Si se trabaja con otro tipo de codificaciones, se considera que un gen ha convergido cuando un porcentaje elevado de la población (típicamente el 95 %) tiene un valor para ese gen dentro de un rango especificado. Una vez que todos los genes del cromosoma alcanzan la convergencia, se dice que la población ha convergido. Cuando esto ocurre, la media de bondad de la población se aproxima a la bondad del mejor individuo.

Finalmente, hay que destacar que los algoritmos genéticos son métodos genéricos que deben ser particularizados para cada problema que se deseé solucionar. En concreto, según el problema a tratar se debe diseñar la codificación de las soluciones como un conjunto de genes y adaptar los operadores genéticos.

Los algoritmos genéticos, al resolver problemas de optimización de forma adecuada, han sido utilizados en diversas ocasiones para la sintonización de controladores PID. En concreto, en la literatura es posible encontrar numerosos ejemplos de sintonización mediante algoritmos genéticos de controladores PID utilizados en diversos campos y aplicaciones. Por ejemplo, en [119–122] los autores aplican satisfactoriamente algoritmos genéticos en la sintonización de PIDs en plantas de distinto orden, obteniendo en todos los casos mejores respuestas que con el método de sintonización de Ziegler-Nichols. En [123] se utiliza un algoritmo genético para sintonizar un controlador PID aplicado a un proceso no lineal y se presenta un ejemplo de su utilización en un proceso de neutralización de la acidez en disoluciones o pH. De forma similar, los autores en [124] implementan un algoritmo genético para el control de los niveles de agua en tanques. Por último, en el campo de las telecomunicaciones, también se ha utilizado la sintonización de PIDs mediante algoritmos genéticos en el control de la congestión mediante el manejo activo

de colas (AQM) en redes TCP/IP con el objetivo de reducir la pérdida de paquetes y mejorar la utilización de la red [125].

5.3. Sintonización de SPID mediante algoritmos genéticos. Algoritmo GA-SPID

En este apartado del capítulo se va a proceder a describir el proceso de sintonización diseñado para el controlador PID del algoritmo SPID incorporando un algoritmo genético. Este nuevo algoritmo se ha denominado GA-SPID (*Genetic Algorithm Service level agreement PID*) [126, 127]. El esquema general del proceso pasa a ser el que se observa en la Figura 5.3. La diferencia con respecto al esquema de SPID es que los valores de los parámetros para la sintonización del PID, la constante proporcional K_p , el tiempo integral T_i y el tiempo derivativo T_d , son proporcionados de forma automática por un algoritmo genético.

Para la implementación del algoritmo GA-SPID se han seguido los pasos genéricos de los algoritmos genéticos. Por tanto, se ha definido la codificación que se va a usar, los criterios para generar la población inicial, cómo se va a evaluar la bondad de los individuos y qué operados genéticos se emplearán para obtener las nuevas generaciones de individuos [119–121]. A continuación se explicará de forma más específica cada uno de estos pasos en el algoritmo genético diseñado:

1. **Codificación:** La implementación del proceso de sintonización por medio de un algoritmo genético comienza con la definición de la representación del cromosoma.

Tal y como se ilustra en la Figura 5.4, el cromosoma está formado por tres valores, que se corresponden con los tres parámetros de sintonización que hay que hallar para conseguir un comportamiento óptimo del PID (K_p , T_i y T_d). Los valores de estos tres parámetros están limitados inicialmente a un rango de valores para acelerar la respuesta del algoritmo genético y asegurar que muchos de los individuos generados para la sintonización del PID no produzcan inestabilidades en el sistema

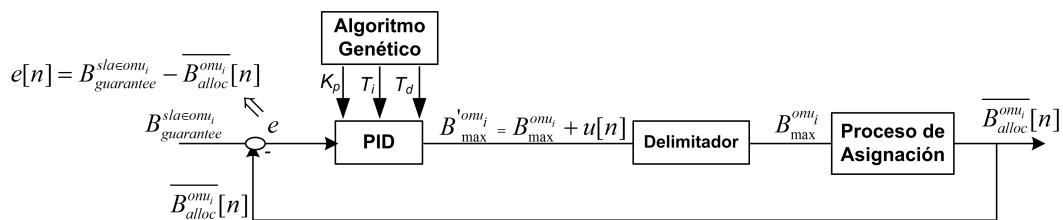


Figura 5.3: Diagrama de bloques del algoritmo GA-SPID incorporando un algoritmo genético para la sintonización del controlador PID.

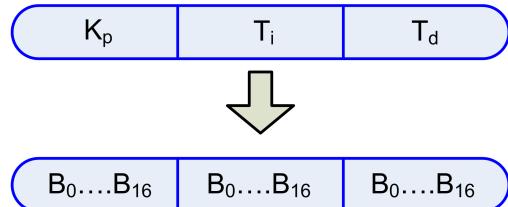


Figura 5.4: Diseño del cromosoma y codificación en el algoritmo genético desarrollado.

[120]. Estos tres valores se codifican formando una cadena binaria. Aunque esta aplicación concreta permite el uso de una codificación real como la usada en [119], se ha optado por una codificación binaria en vez de asignar a cada parámetro un valor real, con el fin de mejorar la eficiencia del algoritmo genético [119–121, 123, 125]. En concreto, se ha utilizado una codificación de 16 bits, para obtener una precisión adecuada. Los valores máximos y mínimos de los tres parámetros se han obtenido mediante simulaciones que se explicarán en la siguiente sección.

2. **Creación de una población inicial:** Se genera una población inicial aleatoria, cuyos valores para cada uno de los genes estén comprendidos entre los valores máximos y mínimos establecidos. El número de individuos que componen la población es un parámetro muy importante a determinar para un buen funcionamiento del algoritmo genético. En este caso particular, tal y como se explica en los siguientes apartados, existe una dependencia muy fuerte entre el tamaño de la población y el tiempo de convergencia del algoritmo. En este sentido, cuanto menor sea la población, más rápida será la evolución del algoritmo, pero al mismo tiempo la exploración de soluciones se reduce. Por lo tanto, se optará por tamaños pequeños de población para acelerar el proceso de sintonización pero que no reduzcan en gran medida el espacio de soluciones.
3. **Evaluación de los individuos y obtención de su salud:** En esta fase, se decodifican los valores de cada uno de los individuos que conforman la población en sus valores decimales y se utilizan como parámetros K_p , T_i y T_d del PID. Para el cálculo de la salud o bondad de cada individuo, se ha definido una función objetivo basada en el error cometido durante la aplicación de cada individuo al PID [121, 122]. En concreto, todos los individuos de la población sintonizan al PID durante un tiempo determinado, medido en el número de iteraciones o ejecuciones del PID. Es decir, durante m iteraciones del PID (por tanto un tiempo en segundos igual a $m \cdot T$, donde T es el tiempo de muestreo del PID) se utilizan los valores de un individuo de la población como los parámetros del controlador PID y se suma

el error cometido por cada una de las ONUs de la red durante ese periodo. Después se divide dicha suma entre el número de ONUs y entre el número de iteraciones del PID. Ésta será la función objetivo que el algoritmo genético tratará de minimizar. De esta forma, lo que se minimiza es el error medio cometido en cada ONU durante cada iteración del PID en la que un individuo concreto es usado. Así, se obtiene una medida del error general cometido por ese individuo, teniendo en cuenta todas las iteraciones y el efecto que ese individuo tiene sobre todas las ONUs.

Por tanto, la función objetivo para un determinado individuo de la población viene definida por la ecuación (5.1), donde m es el número de iteraciones del PID, N_{onus} es el número de ONUs en la red, y el término $e_i[m]$ es el error cometido por la ONU i en la iteración m del PID.

$$F = \frac{1}{m} \cdot \frac{1}{N_{onus}} \sum_m \sum_{i=0}^{N_{onus}} |e_i[m]| \quad (5.1)$$

4. Aplicación de los operadores genéticos para obtener una nueva generación: En cada generación, el algoritmo crea una población completamente nueva de individuos generados a partir de la selección y cruce de individuos de la población anterior, para producir la descendencia que conforma la nueva generación. En concreto, se han utilizado los siguientes operadores genéticos:

- **Selección:** Se ha optado por utilizar el método de la ruleta para seleccionar a los individuos de la población actual que servirán de padres para formar la siguiente población. Se ha elegido este método por ser uno de los métodos de selección más utilizados en general [117, 118] y más concretamente en los algoritmos desarrollados para la sintonización de PIDs [119, 120, 122, 123, 125]. Tal y como se comentó en el Apartado 5.2 de este capítulo, la probabilidad de selección de cada individuo depende de su salud. Por tanto, como en este caso el algoritmo genético trata de minimizar el error cometido, valores pequeños de la función de salud implican mayores probabilidades de reproducción.
- **Elitismo:** Se ha optado por introducir elitismo en las poblaciones al igual que en los métodos presentados en [119, 121, 123]. Esto quiere decir que el mejor individuo de cada generación estará presente sin ser modificado en la siguiente generación. De esta forma, se asegura que la mejor solución encontrada no se pierde, ya que puede que los individuos producidos en la siguiente generación mediante los operadores de cruce y mutación tengan una bondad menor que

la mejor solución encontrada hasta el momento.

- Cruce: Dos cromosomas padre se cruzan para producir un hijo. Sin embargo, no todos los individuos seleccionados se cruzarán. Se ha elegido un valor elevado de probabilidad de cruce debido al uso de valores relativamente pequeños en el tamaño de la población; en concreto $P_{cruce} = 0.9$ tal y como se especifica en [128] para poblaciones de tamaño relativamente pequeño (del orden de 30 individuos). Además se utiliza el cruce en un único punto al igual que en los métodos presentados en [119, 123].
- Mutación: La mutación cambia la estructura de la cadena binaria, cambiando el valor de algún bit elegido aleatoriamente. Habitualmente la probabilidad de mutación es pequeña (normalmente se encuentra entre 0 y 0.1). Una probabilidad de mutación mayor de 0.1 implica una búsqueda aleatoria. Por todo ello, para este caso se ha fijado la probabilidad de mutación $P_{mutacion} = 0.01$, tal y como se indica en [128] para poblaciones de tamaño relativamente pequeño (del orden de 30 individuos).

Una vez realizado este proceso se comprueba si se cumple el criterio de parada. En nuestro caso se ha optado por utilizar un criterio de parada basado en el número de generaciones alcanzado. De esta forma, se puede controlar el tiempo de convergencia del algoritmo estableciendo así una duración fija del proceso de sintonización, ya que como se ha mencionado anteriormente el tiempo de sintonización es crucial en nuestro sistema. Por tanto, cuando se llega a un número determinado de generaciones, se da por concluida la evolución, por lo que se selecciona el mejor individuo de la generación final que será el que se utilice para sintonizar el PID. Por el contrario, si no se ha alcanzado dicho número de generaciones se vuelve al paso 3 y se calculan los nuevos valores de salud para los individuos de la nueva población. Un esquema de los pasos seguidos puede verse en el diagrama de la Figura 5.5.

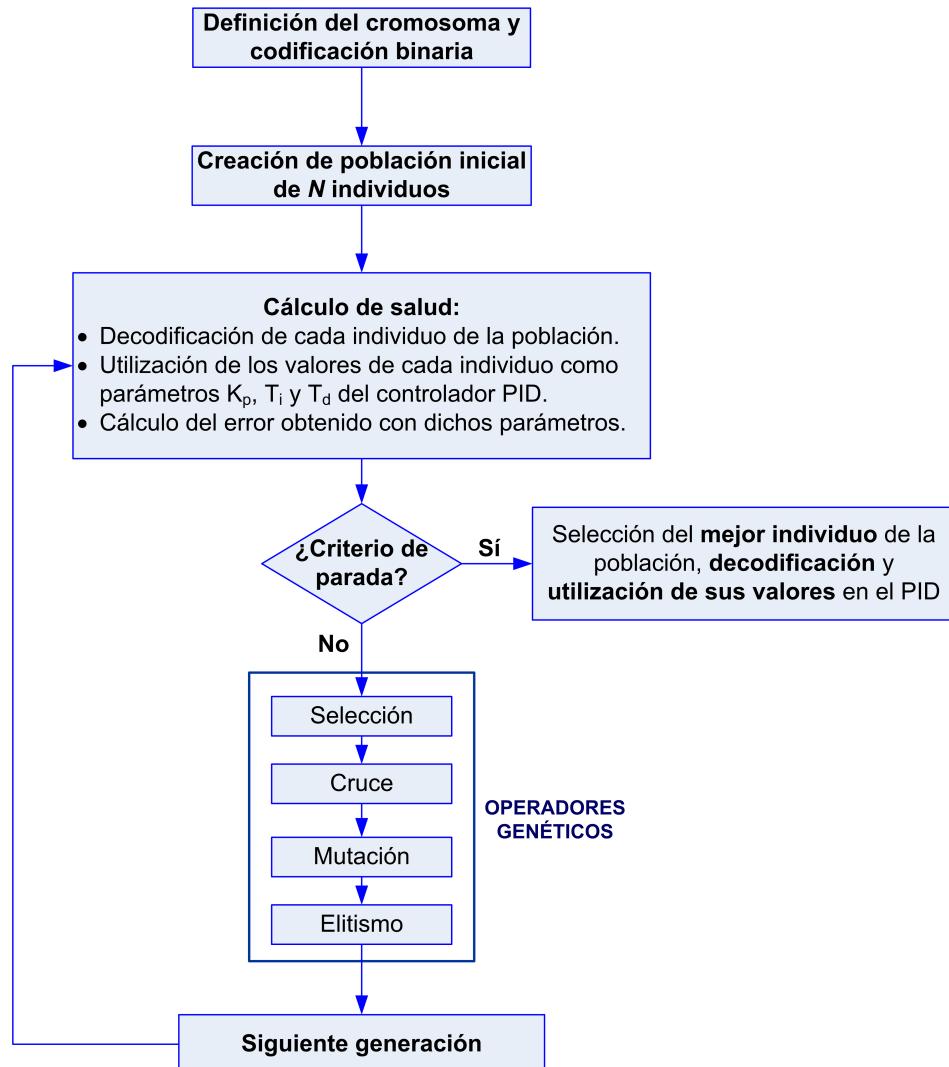


Figura 5.5: Diagrama de flujo del proceso del algoritmo genético para la sintonización de controladores PID en la gestión de recursos en redes LR-EPON.

5.4. Estudio de simulación del algoritmo GA-SPID

En este apartado se va a proceder a realizar el análisis de las prestaciones del algoritmo GA-SPID. Para ello, en primer lugar se determinarán los parámetros más adecuados relacionados con el diseño del algoritmo genético. Después se realizará un análisis comparativo del tiempo de sintonización empleado en GA-SPID comparado con el método de Ziegler-Nichols. Tras este análisis, se estudiará la adaptación en tiempo real hacia los niveles garantizados de ancho de banda y la independencia con respecto a dichos niveles cuando se compara la sintonización llevada a cabo mediante algoritmos genéticos con el método de Ziegler-Nichols.

5.4.1. Escenario de simulación

El escenario de simulación en GA-SPID es el mismo que el considerado en SPID (especificado en el Apartado 4.4.1 del Capítulo 4). La Tabla 5.1 recoge un resumen de los principales parámetros de red y los valores considerados.

Parámetros de simulación	Valores
Número de ONUs	16 ONUs (N_{onus})
Tasa de transmisión canal ascendente	1 Gbit/s ($EPON_Line_rate$)
Tasa de transmisión ONU-abonado	100 Mbit/s ($User_Line_rate$)
RTT	1 ms
Tiempo de guarda	1 μ s
Tiempo de ciclo máximo	2 ms
Distribución de tráfico	Fuente de Pareto H=0.8
Tamaño de paquete	Variable entre 64-1500 bytes (más 38 bytes de cabeceras)
<i>Service Level Agreement (SLA)</i>	SLA ₀ : 1 ONU Ancho de banda garantizado: 100 Mbit/s
	SLA ₁ : 5 ONUs Ancho de banda garantizado: 75 Mbit/s
	SLA ₂ : 10 ONUs Ancho de banda garantizado: 50 Mbit/s

Tabla 5.1: Parámetros de red considerados en el escenario de simulación para GA-SPID.

Para la implementación del algoritmo genético de GA-SPID en C++, se ha utilizado la librería *GAlib* [129]. En concreto, se ha utilizado la clase *GASimpleGA* que se corresponde con el algoritmo genético descrito por Goldberg en [116] y denominado “simple”. Este algoritmo usa poblaciones no solapadas. Esto es, en cada generación, el algoritmo crea una población completa de nuevos individuos. Además permite la opción de incorporar elitismo. Por defecto, esta clase utiliza la ruleta como método de selección y un

factor de escalado lineal para la función de salud, esto es, para calcular la aptitud de cada individuo. Para la codificación se ha utilizado la clase *GABin2DecGenome*, que permite la conversión de cadenas binarias a valores decimales y viceversa. Por defecto, esta clase utiliza un mapa binario a decimal (en base 2), que es el que ha sido utilizado en nuestro algoritmo. Para crear el mapa, se han establecido unas fronteras para los valores decimales que se van a considerar y se ha fijado el número de bits usados para representar dichas fronteras en 16.

5.4.2. Establecimiento de los parámetros asociados al diseño del algoritmo genético en GA-SPID

Para conseguir un funcionamiento óptimo de GA-SPID es necesario determinar los parámetros de los cuáles depende su ejecución. En concreto, en el algoritmo genético es necesario determinar una cota máxima y mínima de los valores de los parámetros para codificar el cromosoma, así como el tamaño de la población y el número de generaciones para establecer el criterio de parada. Los dos últimos parámetros tienen además una gran influencia en el tiempo necesario para llevar a cabo el proceso de sintonización, lo cual será especialmente crítico en nuestro sistema. Además, también será necesario determinar el tiempo que cada individuo será utilizado en el controlador PID para calcular su nivel de salud. Para analizar todos estos parámetros, se han promediado los resultados de 20 simulaciones, por ser este un número suficientemente representativo para realizar un buen análisis estadístico y hallar los valores medios e intervalos de confianza de la salud de cada individuo. Por otro lado, las gráficas de este apartado muestran la media del error cometido por el mejor individuo de cada población según avanzan las generaciones en las 20 simulaciones. De esta forma, se analiza cómo se va incrementando la aptitud de la mejor solución generación tras generación.

5.4.2.1. Análisis de la cota de los parámetros del cromosoma

Tal y como se ha mencionado anteriormente, es conveniente que los valores de los parámetros que componen el cromosoma (K_p , T_i y T_d) tengan unas cotas máximas y mínimas establecidas para favorecer una convergencia más rápida hacia los valores óptimos. Por ello, se han llevado a cabo simulaciones considerando diferentes valores para dichas cotas. Si el intervalo considerado para los parámetros es muy pequeño, el espacio de soluciones posibles se reduce, pudiendo existir soluciones factibles fuera de este intervalo. Por el contrario, un intervalo muy grande implica un mayor tiempo de convergencia hacia los valores óptimos. Por tanto, a la hora de establecer dichos intervalos es bueno conocer un rango de valores de las posibles soluciones. En nuestro caso, utili-

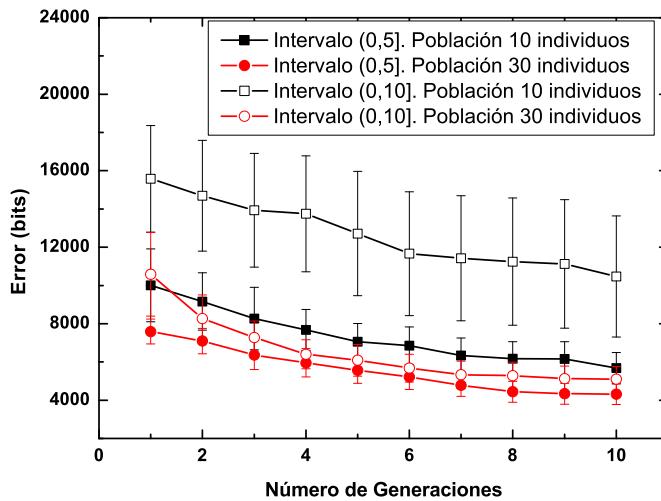


Figura 5.6: Media del error del mejor individuo para los dos intervalos analizados para las cotas de los parámetros y considerando distintos tamaños de población.

zando como guía los valores previos obtenidos con el método de Ziegler-Nichols, se han seleccionado dos posibles intervalos, concretamente de $(0,5]$ y de $(0,10]$. La Figura 5.6 representa la media del error cometido por el mejor individuo de la población cuando se considera que los tres parámetros toman valores entre $(0,5]$ y entre $(0,10]$ según avanza las generaciones y considerando dos tamaños de población distintos, 10 y 30 individuos. Tal y como puede observarse, los errores cometidos son menores (individuos con mejor salud) cuando los valores de los parámetros se encuentran dentro del intervalo $(0-5]$ para cualquier número de generaciones como criterio de parada y para ambos tamaños de población. Por tanto, este es el intervalo de valores que se ha escogido como cota de los parámetros K_p , T_i y T_d . También cabe destacar que a mayor tamaño de población y mayor número de generaciones, el error es decreciente. No obstante, el aumento de estos parámetros lleva asociado una ralentización del proceso de sintonización que puede ser demasiado crítico en nuestro sistema.

5.4.2.2. Análisis del número de iteraciones del PID para calcular la salud de cada individuo

Tal y como se explicó en el apartado anterior, para calcular la salud de cada individuo, se sintoniza el PID con los valores concretos de los parámetros K_p , T_i y T_d de cada individuo y se espera un tiempo para determinar el error cometido por el controlador durante ese periodo. Por tanto, es necesario fijar el periodo de tiempo en el que se utiliza cada individuo para obtener su aptitud. Este tiempo se ha definido en función

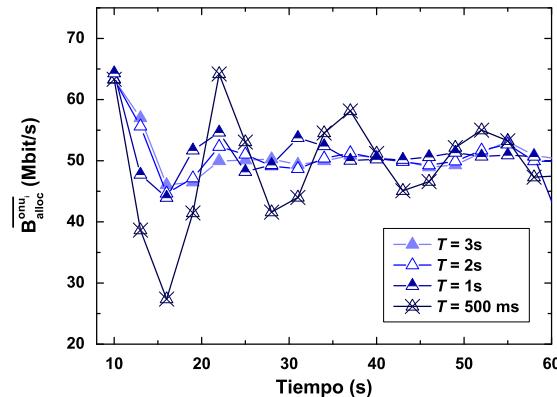


Figura 5.7: Evolución del ancho de banda medio asignado para distintos tiempos de muestreo T del PID en el algoritmo SPID.

del número de iteraciones del PID, es decir, el número de veces que el PID actualiza los anchos de banda máximos para cada ONU. Por tanto este tiempo es igual a $m \cdot T$, donde T es el tiempo de muestreo (ecuación (4.1) del Capítulo 4) y m representa el número de iteraciones del PID.

Por otro lado, el tiempo en el que se utiliza cada individuo en el PID es un parámetro estrechamente relacionado con el tiempo empleado en realizar la sintonización. Si el tiempo es muy pequeño, no se podrá determinar con demasiada exactitud la bondad del individuo en el proceso asignación de ancho de banda. Por el contrario, si el tiempo es muy elevado el proceso de sintonización global será más lento. Por tanto, es necesario un compromiso en el número de iteraciones o ejecuciones del PID para evaluar las prestaciones de cada individuo. Por ello, en primer lugar se ha estudiado el efecto en la reducción del tiempo de muestreo T . En el Capítulo 4 se llevó a cabo este estudio en el Apartado 4.4.3. Tal y como se vio en dicho apartado, si el tiempo de muestreo se reducía demasiado (500 ms), la oscilación de los niveles de ancho de banda era mucho mayor. Sin embargo, tal y como puede verse en la Figura 5.7 para tiempos de muestreo de 1 s y 2 s el comportamiento de la evolución del ancho de banda es muy similar al obtenido con el valor de 3 s escogido en el Capítulo 4. Por tanto, para reducir el tiempo de sintonización total del algoritmo genético, ya que cada individuo se utilizará durante m ejecuciones del PID, se ha optado por escoger un tiempo de muestreo T igual a 1 s, esto es, el mínimo posible.

Una vez reducido el tiempo de muestreo, se va a analizar el número de ejecuciones del PID para determinar el tiempo mínimo necesario para calcular la salud de cada individuo de una forma relativamente precisa. Para seleccionar dicho valor, se han llevado a cabo simulaciones variando el número de iteraciones desde 2 hasta 4. Los resultados se

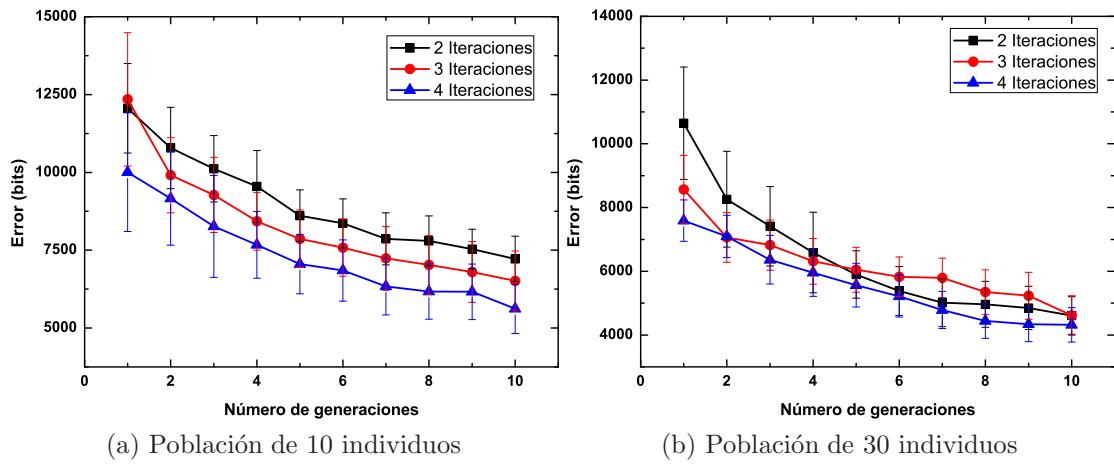


Figura 5.8: Media del error del mejor individuo al variar el número de iteraciones del PID.
 (a) Para una población de 10 individuos (b) Para una población de 30 individuos.

muestran en la Figura 5.8. En concreto, la Figura 5.8 (a) muestra la media del error del mejor individuo considerando distinto número de iteraciones del PID cuando el tamaño de la población es de 10 individuos y en la Figura 5.8 (b) cuando el tamaño de la población es igual a 30. En ambos casos, a medida que el número de generaciones aumenta, se observa cómo para un mayor número de iteraciones durante las cuales el PID usa cada individuo, el error decrece, lo que implica que la aptitud de cada individuo es mejor. Sin embargo, ya que el tiempo de sintonización depende proporcionalmente del número de iteraciones del PID, se ha optado por utilizar un número de iteraciones igual a 2 ó 3 para minimizar el impacto de ese tiempo. El valor definitivo en el número de iteraciones se decidirá en los subapartados siguientes, al analizar cualitativamente cómo afecta este parámetro combinado con el tamaño de la población y con el número de generaciones.

5.4.2.3. Análisis del tamaño de la población

El tamaño de la población es un parámetro muy importante en el funcionamiento del algoritmo genético. En principio, con tamaños de población mayores el algoritmo genético explora un mayor número de posibles soluciones, pudiendo obtener unos parámetros más óptimos. Sin embargo, esto conlleva un mayor tiempo de ejecución del algoritmo genético. Este hecho es especialmente crítico en nuestro sistema, donde para obtener la salud de un individuo es necesario sintonizar el PID con ese individuo y observar el error que comete durante un tiempo. Esto implica que si consideramos tamaños de población muy grandes, se tardará demasiado en sintonizar el PID. Por tanto, también existe un fuerte compromiso entre el tiempo de sintonización y el tamaño de la población.

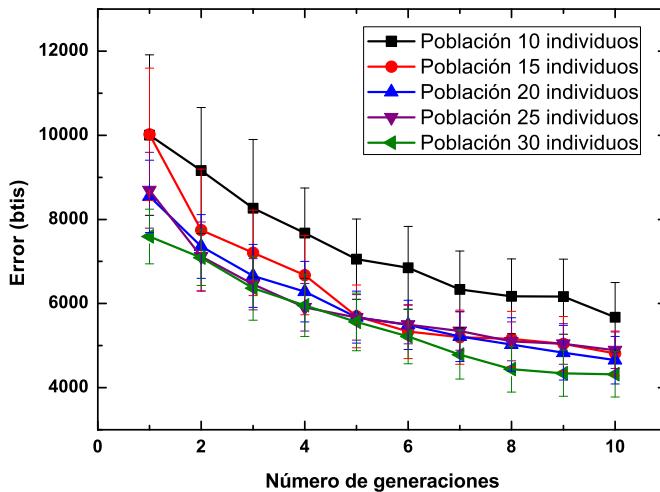


Figura 5.9: Evolución de la media del error del mejor individuo para distintos tamaños de población considerados (10, 15, 20, 25 y 30) a lo largo de sucesivas generaciones.

Para determinar el tamaño de la población más idóneo se han llevado a cabo simulaciones con tamaños que varían entre 10 y 30 individuos. En concreto, la Figura 5.9 representa la media del error del mejor individuo de la población cuando se consideran distintos tamaños de población. Como puede observarse, los individuos con una peor salud se obtienen para tamaños de población de 10 individuos, lo cual es lógico ya que se reduce el espacio de soluciones. Para el resto de tamaños considerados se observa cómo en las primeras generaciones obtienen mejores resultados aquellas poblaciones con un mayor número de individuos. Sin embargo, a medida que aumenta el número de generaciones los resultados de las distintas poblaciones se parecen cada vez más. Por tanto, para reducir el tiempo de sintonización se va a trabajar con una población de 15 ó 20 individuos, ya que presentan un comportamiento muy similar y no tienen un tamaño demasiado grande. El valor definitivo de este parámetro se decidirá en el siguiente subapartado, para así analizar la influencia de este parámetro en el número de generaciones necesario para fijar el criterio de parada.

5.4.2.4. Análisis del número de generaciones para fijar el criterio de parada

Al igual que el tamaño de la población, este parámetro está estrechamente relacionado con el tiempo de sintonización del algoritmo, ya que al evolucionar los individuos un mayor número de generaciones, será necesario un mayor tiempo para el proceso de sintonización. Además, se ha optado por utilizar un criterio de parada basado en el número de generaciones para tener limitado el tiempo empleado en la sintonización a un tiempo determinado y fijo.

Para establecer el número de generaciones necesario para finalizar el algoritmo genético se han llevado a cabo simulaciones para dos tamaños de población (15 y 20 individuos) y dos números diferentes de ejecuciones del algoritmo PID (2 y 3 iteraciones), escogidos como valores relativamente óptimos en apartados anteriores. Estas poblaciones se han dejado evolucionar hasta 20 generaciones y se ha analizado la evolución de la salud (error medio cometido) del mejor individuo de cada población. Como puede observarse en la Figura 5.10, la mejora de la aptitud de los individuos es apreciable para todos los casos hasta las 10 generaciones. A partir de esta generación, las mejoras son notablemente menores, por lo que la similitud en su comportamiento es muy alta. Por tanto, se ha optado por establecer el número de generaciones para parar el algoritmo genético igual a 10, esto es, el mínimo posible para reducir así el tiempo de sintonización final.

Además, tal y como puede apreciarse, los resultados para la población de 15 individuos con 3 iteraciones y de 20 individuos con dos y tres iteraciones son bastante similares. Por tanto, los parámetros finales escogidos son una población de 20 individuos, con 2 iteraciones del PID y un número de generaciones de 10. Así pues, esta combinación escogida requiere un menor tiempo de sintonización que utilizando una población de 15 individuos con un número de iteraciones del PID igual a 3 y considerando 10 generaciones para el tiempo de parada.

Finalmente, en la Tabla 5.2 se muestra un resumen de los parámetros finales seleccionados para el diseño del algoritmo genético que sintonizará el PID del algoritmo SPID.

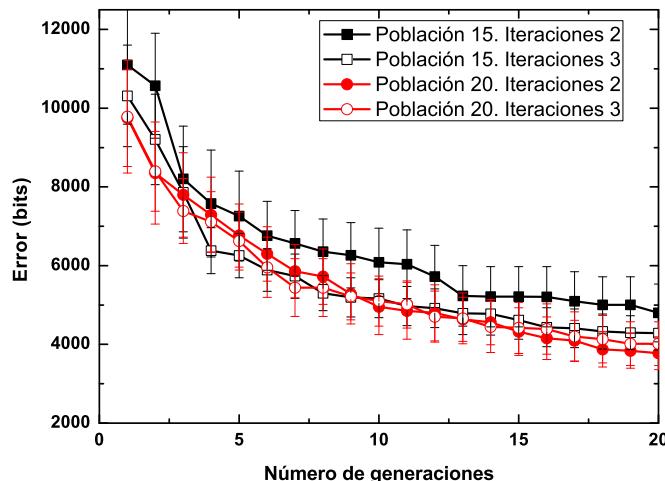


Figura 5.10: Evolución de la media del error cometido por el mejor individuo de la población de cada generación para distintos tamaños de población (15, 20) e iteraciones del PID (2, 3).

Parámetros del algoritmo genético	Valores
Selección	Método de la Ruleta
Intervalo de los parámetros	(0,5]
Probabilidad de cruce (P_{cruce})	0.9
Tipo cruce	1 punto
Tamaño población	20 individuos
Elitismo	Sí
Probabilidad de mutación ($P_{mutacion}$)	0.01
Criterio de parada	Número de generaciones igual a 10
Iteraciones del PID para obtener la salud	2

Tabla 5.2: Parámetros finales del algoritmo genético diseñado para GA-SPID.

5.4.3. Análisis de prestaciones de GA-SPID

En el siguiente apartado de la Tesis se van a analizar las bondades de la sintonización automática basada en algoritmos genéticos, comparando sus prestaciones con el método de sintonización de Ziegler-Nichols. En primer lugar, se va a estudiar el tiempo empleado por ambos métodos para llevar a cabo el proceso de sintonización del PID, lo cual resulta especialmente crítico en nuestro sistema. Después se comprobará la eficiencia de las dos técnicas de sintonización en la adaptación de los niveles de ancho de banda, llevada a cabo en SPID, hacia los niveles de ancho de banda garantizados para asegurar los requisitos de QoS.

5.4.3.1. Comparación del tiempo de sintonización del método de Ziegler-Nichols respecto al algoritmo genético

Tal y como se ha comentado anteriormente, una de las principales ventajas de GA-SPID frente a SPID es la automatización del proceso de sintonización. En este sentido, mientras que el método de Ziegler-Nichols, utilizado en SPID, es completamente manual y basado en la observación de las oscilaciones producidas para distintos valores del parámetro K_p , la sintonización llevada a cabo en GA-SPID mediante algoritmos genéticos es automática. Por tanto, al ser el método de Ziegler-Nichols un método totalmente manual y experimental, no se puede comparar directamente el tiempo de sintonización necesario por esta técnica con el empleado en los algoritmos genéticos.

Por este motivo, y para poder realizar una comparación justa entre ambas técnicas, se ha desarrollado un proceso y metodología que automatiza en cierto grado el método de sintonización del Ziegler-Nichols. En concreto, se va a automatizar la búsqueda del término K_u . Para ello, en primer lugar se considera únicamente la parte proporcional del controlador y se selecciona un intervalo en el que se modificará el término K_p para

obtener un valor de K_u óptimo. Por analogía al algoritmo GA-SPID, este intervalo se ha definido como $(0,5]$. Por tanto, se va a generar un número aleatorio en este intervalo, denominado δ , que será igualado al término K_p del controlador. A continuación, se deja que el ancho de banda medio evolucione durante un tiempo utilizando ese valor de K_p y se comprueban las fluctuaciones obtenidas. El tiempo escogido ha sido de 300 s para dejar que el controlador PID se estabilice y que las fluctuaciones se deban únicamente al término K_p utilizado. Si se escogen tiempos menores, se reduciría el tiempo total de sintonización, pero se podrían descartar soluciones válidas porque todavía no se hubiesen estabilizado los niveles medios del ancho de banda. Por el contrario, para tiempos mayores, el proceso de sintonización se ralentizaría considerablemente. De este modo, si las fluctuaciones obtenidas al cabo de esos 300 s superan un umbral superior e inferior del ancho de banda a garantizar, se considera que las fluctuaciones son elevadas y que es necesario probar con otro valor de K_p . El umbral inferior y superior fijado está relacionado con un cierto porcentaje del nivel de ancho de banda garantizado. En concreto, se han escogido dos porcentajes, 20 % y 10 %. Para porcentajes mayores, las oscilaciones permitidas son mayores, por lo que se pierde precisión. Por tanto, si las oscilaciones obtenidas se encuentran por encima o por debajo del porcentaje fijado sobre el nivel garantizado, se incrementa el valor de K_p una cierta cantidad, siempre y cuando el nuevo valor de K_p esté contenido dentro del intervalo y no se haya alcanzado en el proceso el valor máximo del intervalo. En este caso, el valor escogido para este incremento es de 0.1. Si se utilizan valores mayores, el tiempo del proceso de sintonización sería menor, pero se perdería precisión a la hora de obtener el parámetro óptimo. Por otro lado, para incrementos menores, se ganaría precisión pero se ralentizaría el proceso. Si el valor de K_p alcanza el máximo del intervalo, se vuelve a igualar el valor de K_p al número aleatorio inicial, y a partir de ese momento, para obtener nuevos valores de K_p , se resta la misma cantidad (0.1). Por ejemplo y para los valores concretos seleccionados, si el número aleatorio inicial es 4, se irá sumando a este valor 0.1 y se comprobarán cada 300 s las oscilaciones obtenidas. Si el valor de K_p llega a 5 porque los valores anteriormente utilizados han generado oscilaciones superiores al margen fijado, el siguiente valor a probar será de 3.9 (valor K_p inicial menos 0.1) y a partir de este momento se restará 0.1 para obtener los nuevos valores de K_p hasta encontrar un valor apropiado, esto es, se avanza hacia atrás en el intervalo.

De forma esquemática este proceso se refleja en el diagrama de flujo de la Figura 5.11. Una vez determinado un valor de K_p que produzca fluctuaciones del ancho de banda medio dentro del intervalo considerado, este valor se iguala a K_u . A partir de aquí el proceso vuelve a ser manual, ya que con ese valor de K_u se tiene que examinar el periodo

de oscilación para obtener el término T_u , y poder así aplicar las relaciones del método de Ziegler-Nichols de la Tabla 4.3 del Capítulo 4.

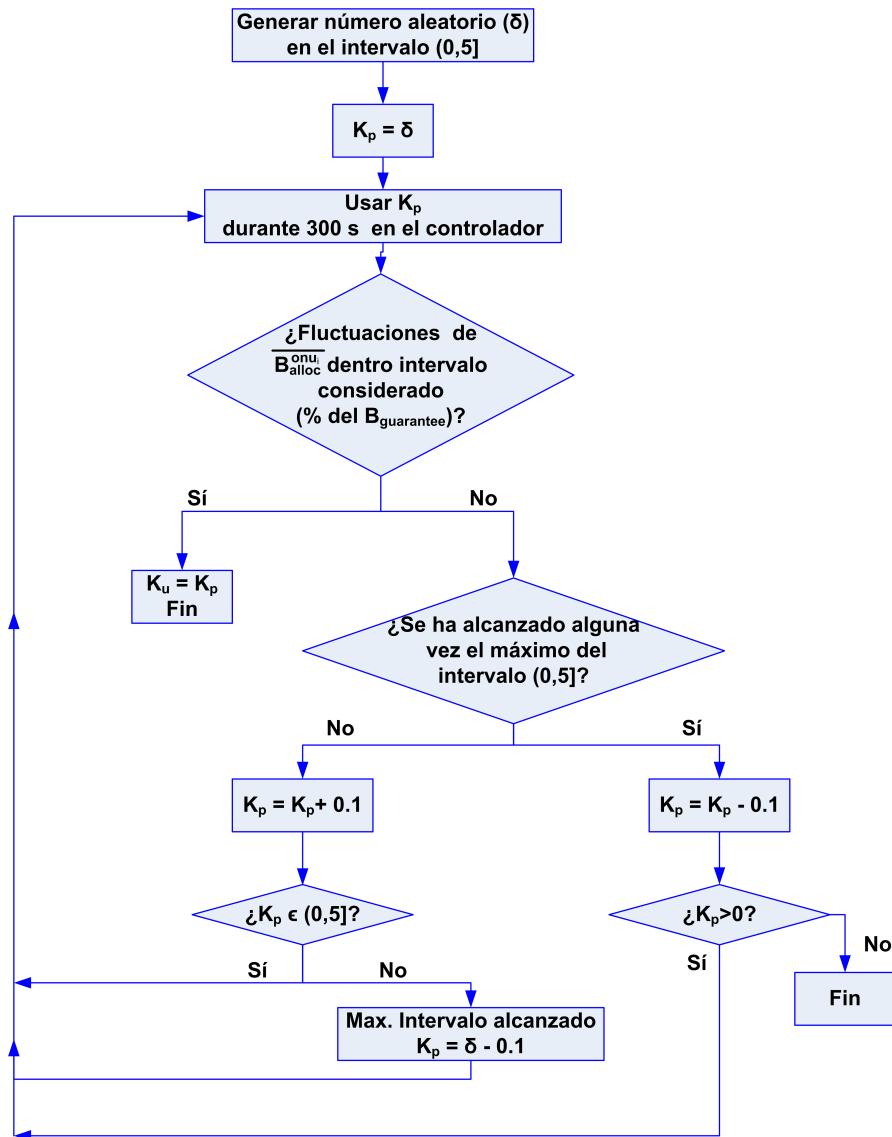


Figura 5.11: Esquema para la sintonización automática siguiendo el método de Ziegler-Nichols en SPID.

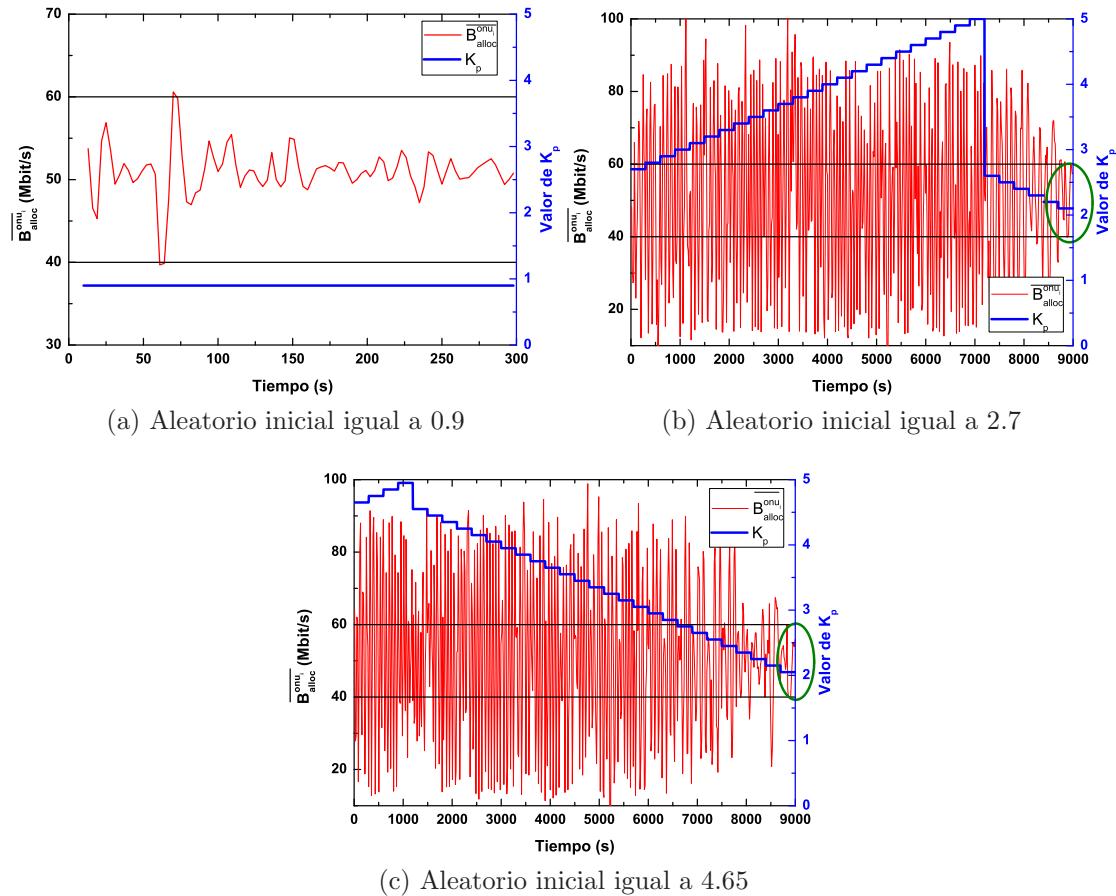


Figura 5.12: Evolución en tiempo real del ancho de banda medio asignado para el algoritmo de sintonización automática basado en el método de Ziegler-Nichols considerando un intervalo de fluctuaciones de ancho de banda del 20 % por encima y por debajo del nivel garantizado del SLA₂ (50 Mbit/s). (a) K_p inicial igual a 0.9 (b) K_p inicial igual a 2.7 (c) K_p inicial igual a 4.65.

La Figura 5.12 (a), (b) y (c) muestra los resultados del proceso de sintonización automática explicado anteriormente. En concreto, se representa la evolución del ancho de banda medio para el perfil SLA₂ cuando se utilizan tres números aleatorios diferentes como valor inicial de K_p , esto es, $K_p = 0.9$ en la Figura 5.12 (a), $K_p = 2.7$ en la Figura 5.12 (b) y $K_p = 4.65$ en la Figura 5.12 (c). Además, se ha considerado un margen de fluctuaciones del ancho de banda asignado del 20 % por encima y por debajo del nivel garantizado para este SLA (50 Mbit/s). Esto es, se considerará que las fluctuaciones generadas por un cierto K_p son válidas para este perfil de abonado si están entre 40 Mbit/s y 60 Mbit/s. En azul, y referido al eje de la derecha está representada la variación en tiempo real del valor del parámetro K_p , mientras que en rojo y referido al eje de la izquierda se muestra la variación del ancho de banda medio. Tal y como

puede observarse, en la Figura 5.12 (a) el tiempo de sintonización es pequeño, ya que el número aleatorio inicial permite que el ancho de banda medio fluctúe dentro del intervalo considerado en un corto espacio de tiempo, por lo que el valor óptimo de K_u se obtiene en 300 s. Sin embargo, para los otros dos valores iniciales de K_p se observa que el tiempo de sintonización es mucho mayor. En concreto, para un valor inicial de K_p igual a 2.7, la determinación del valor de K_u se consigue tras 9000 s de simulación, ya que desde este valor inicial, el término K_p tiene que aumentar hasta un valor de 5 y después disminuir desde 2.6 hasta 2.1, valor que obtiene unas fluctuaciones dentro del intervalo considerado. De forma similar, para un aleatorio inicial igual a 4.65 el tiempo empleado en realizar la sintonización es de 9000 s. Este comportamiento analizado demuestra que, el tiempo de sintonización siguiendo esta técnica depende en gran medida del número aleatorio inicial, y puede llegar a ser del orden de miles de segundos si no se escoge adecuadamente, por lo que puede no ser un método adecuado en un escenario de red real.

Por otro lado, si el margen del intervalo de fluctuación del ancho de banda se reduce, los tiempos aumentarán ya que los requisitos para elegir el valor de K_p son más restrictivos. En este sentido, la Figura 5.13 (a), (b) y (c) ilustra la evolución del ancho de banda medio del SLA₂ para los tres valores K_p iniciales considerados anteriormente pero cuando el intervalo de fluctuaciones permitidas es tan solo del 10 % por encima y por debajo del nivel garantizado del SLA₂, esto es, de 45 Mbit/s a 55 Mbit/s. Tal y como se puede ver, exceptuando el aleatorio inicial igual a 0.9, para los otros dos casos el tiempo de sintonización se incrementa por encima de los 10000 s, ya que los valores K_u obtenidos se han reducido a 1.7 y 1.65 respectivamente.

Por tanto, con este sistema diseñado para realizar una sintonización automática basada en Ziegler-Nichols, se demuestra que los tiempos de sintonización del controlador PID pueden llegar a ser extremadamente elevados si el valor de K_p inicial se aleja del intervalo de soluciones óptimas. Además, en este caso, se han considerado incrementos de K_p de 0.1. Si se escogen valores más pequeños para obtener una mayor precisión, el tiempo de sintonización se verá incrementado todavía en mayor medida. Si por el contrario, los valores son mayores, el tiempo de sintonización puede reducirse pero la precisión de los valores obtenidos será menor, de forma que puede dar lugar a sintonizaciones no óptimas.

De esta forma, el tiempo de sintonización del método de Ziegler-Nichols presenta una gran dependencia según el valor de K_p inicial escogido para comprobar las fluctuaciones, del valor del incremento/decremento del término K_p , del margen de fluctuaciones del ancho de banda permitido y del tiempo con el que se prueba cada valor del término K_p .

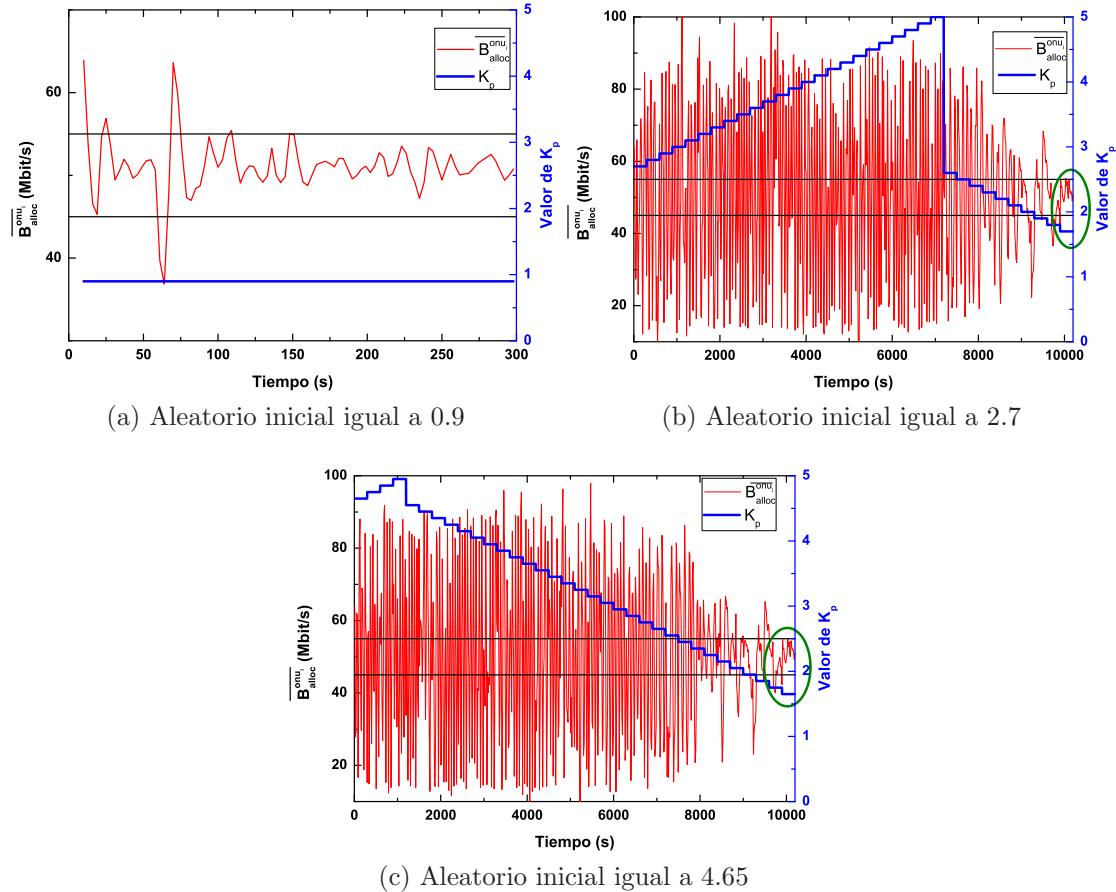


Figura 5.13: Evolución en el tiempo del ancho de banda medio asignado para el algoritmo de sintonización automática basado en el método de Ziegler-Nichols con un intervalo de fluctuaciones de ancho de banda del 10 % por encima y por debajo del nivel garantizado del SLA₂ (50 Mbit/s). (a) K_p inicial igual a 0.9 (b) K_p inicial igual a 2.7 (c) K_p inicial igual a 4.65.

Esto se debe al carácter manual y experimental de este método.

Por el contrario, utilizando un método basado en algoritmos genéticos, aunque la población inicial también es generada aleatoriamente, el tiempo de sintonización será por lo general mucho menor. En este sentido, tal y como se ha explicado a lo largo de este capítulo, este tiempo va a depender del tamaño de la población, del número de iteraciones del PID para calcular la bondad y del número de generaciones consideradas como criterio de parada del algoritmo. En concreto, el tiempo de sintonización viene dado por la ecuación (5.2), donde $TamPob$ es el tamaño de la población seleccionado, m es el número de iteraciones, T es el tiempo de muestreo y $NumGen$ es el número de generaciones del criterio de parada.

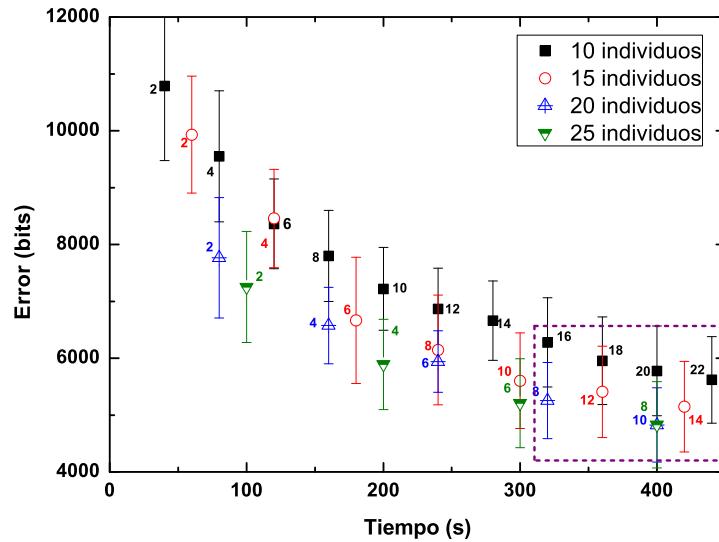


Figura 5.14: Media del error del mejor individuo para distinto tamaño de población y número de generaciones en función del tiempo de sintonización en GA-SPID.

$$T_{sintonizacion} = TamPob \cdot m \cdot T \cdot NumGen \quad (5.2)$$

De este modo, la Figura 5.14 representa el error cometido por el algoritmo genético frente al tiempo de sintonización empleado considerando diferentes combinaciones de tamaño de población y número de generaciones en el criterio de parada cuando se fija un número de iteraciones m igual a 2 (tal y como se escogió anteriormente). El número que aparece al lado de cada punto se corresponde con la generación en la que ese error ha sido obtenido. Tal y como se puede observar, a medida que se incrementa el tiempo, los errores obtenidos son menores, bien porque se ha aumentado el tamaño de la población o bien el número de generaciones. Sin embargo, para tiempos de sintonización a partir de 300 s, los errores ya se minimizan significativamente por lo que dan lugar a individuos que, tal y como se comprobará en el siguiente apartado, ofrecen un comportamiento adecuado en la sintonización del controlador PID. Como ejemplo, para una población de 20 individuos con un número de generaciones igual a 10, que son los parámetros seleccionados anteriormente, se obtienen errores pequeños en un tiempo de sintonización de 400 s.

Por tanto se puede concluir que, el método de Ziegler-Nichols, al ser un método manual y experimental, además de resultar extremadamente laborioso y en ocasiones poco preciso, puede llegar a consumir una gran cantidad de tiempo. Al automatizar en cierto grado este método, se puede evaluar a nivel cualitativo dicho tiempo de sintonización

y se ha demostrado que el tiempo empleado depende de varios factores, como el valor inicial del término K_p , el valor del incremento/decremento de dicho factor, el tiempo que se utiliza cada valor para comprobar las fluctuaciones y el margen permitido de dichas fluctuaciones. No obstante, tal y como se ha demostrado, este tiempo puede incrementarse en gran medida dando lugar a una gran ralentización del proceso de sintonización, lo cual puede resultar extremadamente ineficiente en nuestro sistema. Por el contrario, gracias al uso de algoritmos genéticos, el tiempo de sintonización empleado se reduce considerablemente, dando lugar a sintonizaciones más precisas, rápidas y automáticas.

5.4.3.2. Evolución del ancho de banda a los niveles garantizados por el proveedor de servicios

Una de las principales características de GA-SPID, al igual que SPID, es el control dinámico del ancho de banda para garantizar los niveles estipulados por el proveedor de servicios. Por ello, en las siguientes gráficas se presenta la adaptación del ancho de banda con dos individuos óptimos que se han obtenido como resultado del algoritmo genético. Esto es, de dos simulaciones distintas realizadas previamente, se ha escogido el mejor individuo de la población final y sus valores de K_p , T_i y T_d se han utilizado para sintonizar el PID bajo las condiciones del escenario presentado en la Tabla 5.1. En concreto, se va a comparar la evolución del ancho de banda que se obtiene con el PID sintonizado mediante el método de Ziegler-Nichols (SPID), con el PID sintonizado con dos individuos óptimos obtenidos mediante el algoritmo genético y con dos configuraciones en las que el PID no está sintonizado de forma óptima. Los valores de cada uno de los parámetros K_p , T_i y T_d para las distintas sintonizaciones están recogidos en la Tabla 5.3.

	Valores
A. Genético Solución 1	$K_p = 0.248$, $T_i = 4.3$, $T_d = 3.8$
A. Genético Solución 2	$K_p = 0.3$, $T_i = 3.22$, $T_d = 3.06$
Ziegler-Nichols	$K_p = 0.66$, $T_i = 11$, $T_d = 2.75$
Sintonización no optimizada 1	$K_p = 2.1$, $T_i = 4.86$, $T_d = 0.3$
Sintonización no optimizada 2	$K_p = 4.3$, $T_i = 4.9$, $T_d = 0.15$

Tabla 5.3: Valores de los parámetros K_p , T_i y T_d seleccionados para cada tipo de sintonización.

La Figura 5.15 (a), (b) y (c) ilustra la variación en tiempo real del ancho de banda medio asignado a una ONU del SLA₀, SLA₁ y SLA₂ respectivamente, para las sintonizaciones recogidas en la Tabla 5.3, cuando los niveles del ancho de banda garantizado son 100/75/50 Mbit/s y la configuración de pesos iniciales es $W^{sla_0} = 1$, $W^{sla_1} = 1$ y $W^{sla_2} = 1$. Además, la carga de red es la más alta 1.6, esto es, las ONUs transmiten a 100 Mbit/s. Tal y como se puede apreciar, excepto en el caso del SLA más prioritario,

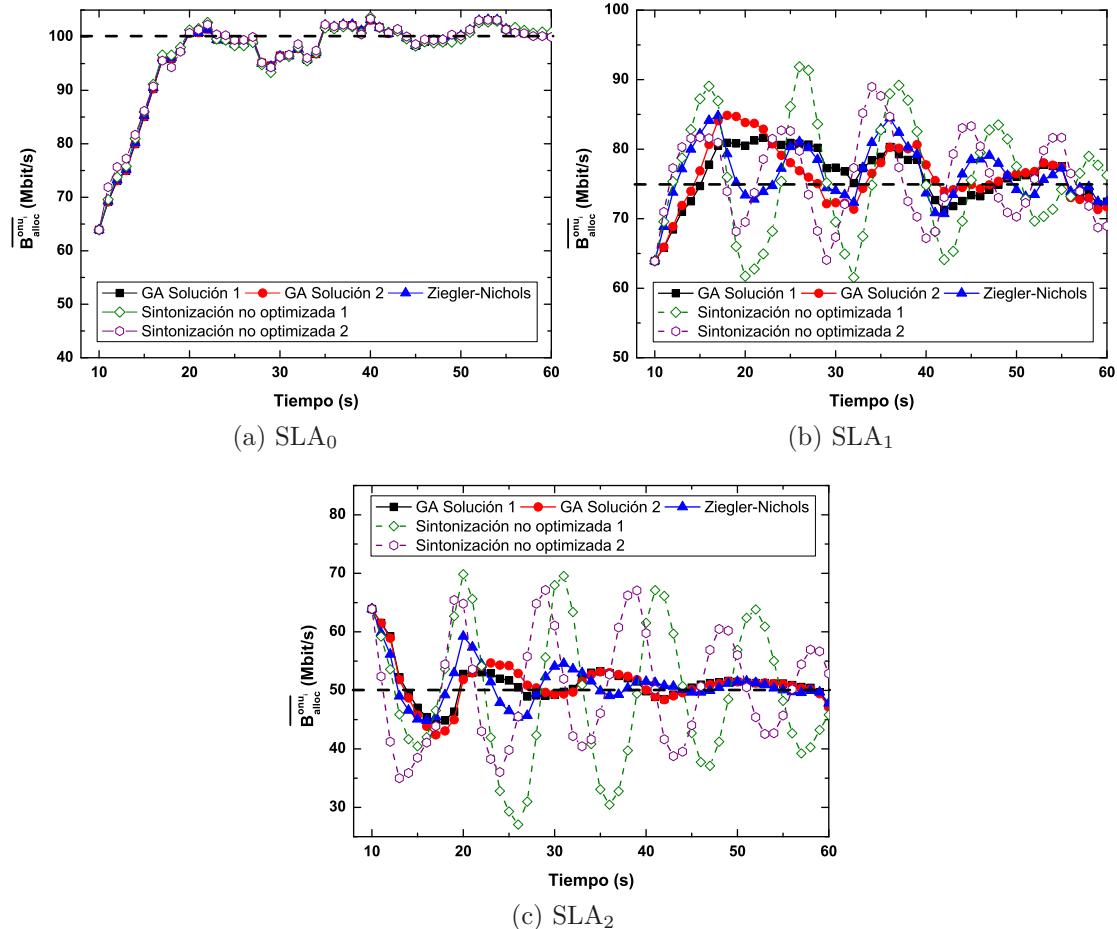


Figura 5.15: Evolución en tiempo real del ancho de banda medio asignado a una ONU de cada SLA para las sintonizaciones de la Tabla 5.3. (a) SLA₀ (b) SLA₁ (c) SLA₂.

donde todas las configuraciones obtienen un ancho de banda muy parecido (Figura 5.15 (a)), el nivel de fluctuación del ancho de banda medio asignado en los primeros segundos es muy diferente dependiendo de las sintonizaciones. En concreto, se observa como el ancho de banda medio asignado a los dos SLAs menos prioritarios (SLA₁ y SLA₂) oscila considerablemente para las dos sintonizaciones no óptimas (Figura 5.15 (b) y (c)). Para el caso de la sintonización mediante el método de Ziegler-Nichols, se aprecia cómo las fluctuaciones de dicho parámetro son también ligeramente mayores que para el caso de las dos sintonizaciones realizadas mediante el algoritmo genético.

Estas variaciones instantáneas en el ancho de banda medio asignado a cada ONU se deben principalmente a que los anchos de banda máximos actualizados por el PID en cada una de las sintonizaciones sufren fluctuaciones. Para comprobarlo, la

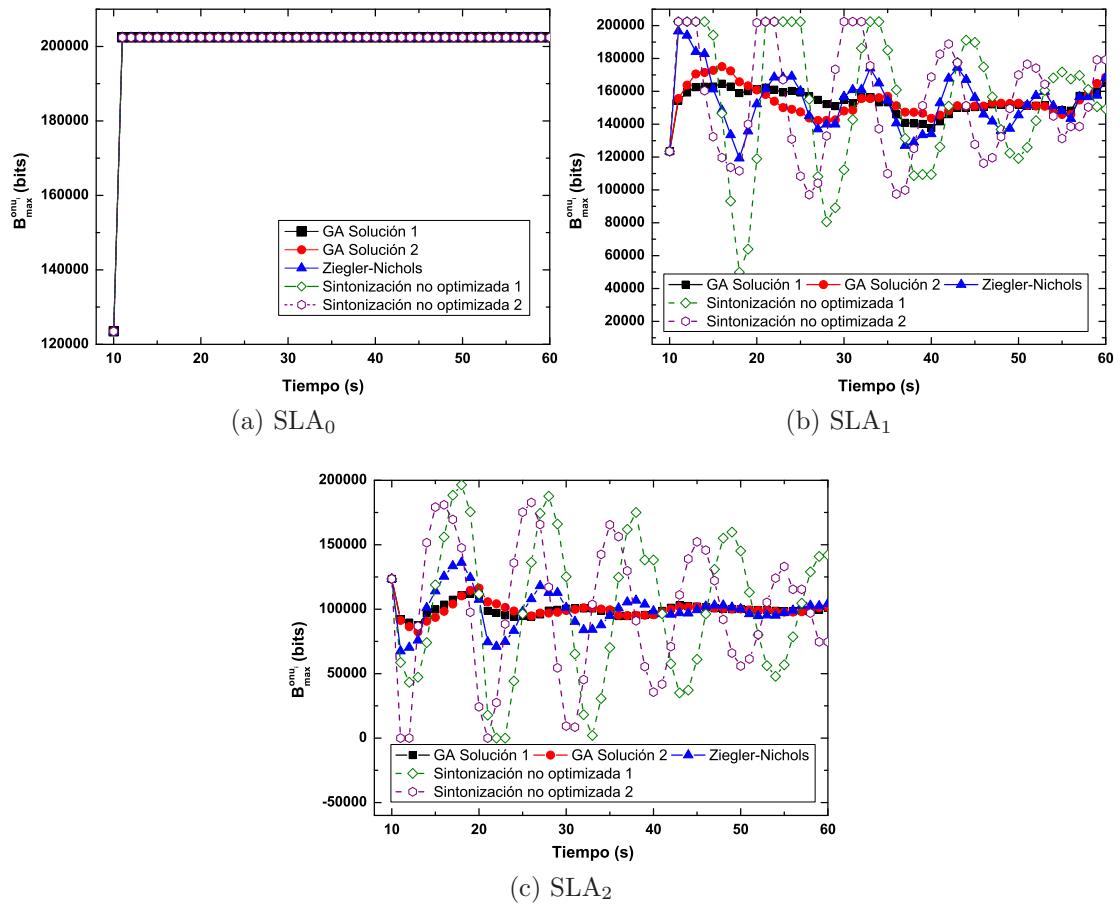


Figura 5.16: Evolución en tiempo real del ancho de banda máximo permitido a una ONU de cada SLA para las sintonizaciones de la Tabla 5.3. (a) SLA₀ (b) SLA₁ (c) SLA₂.

Figura 5.16 (a), (b) y (c) muestra la variación en tiempo real del ancho de banda máximo permitido de una ONU de cada SLA para los primeros segundos de simulación. Tal y como puede observarse, excepto para el SLA de mayor prioridad, existe una gran oscilación del ancho de banda máximo para las sintonizaciones no optimizadas. Para el caso de Ziegler-Nichols, al igual que ocurría con la evolución del ancho de banda de la ventana, se observa cómo existe una oscilación ligeramente mayor a la obtenida con las sintonizaciones del algoritmo genético.

Para finalizar el análisis, la Figura 5.17 (a), (b) y (c) muestra la evolución instantánea de la desviación estándar del ancho de banda medio asignado sobre el nivel garantizado cuando se consideran los diferentes tipos de sintonizaciones. Este valor proporciona una medida aproximada de la exactitud con la que cada sintonización alcanza el nivel de ancho de banda garantizado. Como se observa, en la Figura 5.17 (a) la desviación para

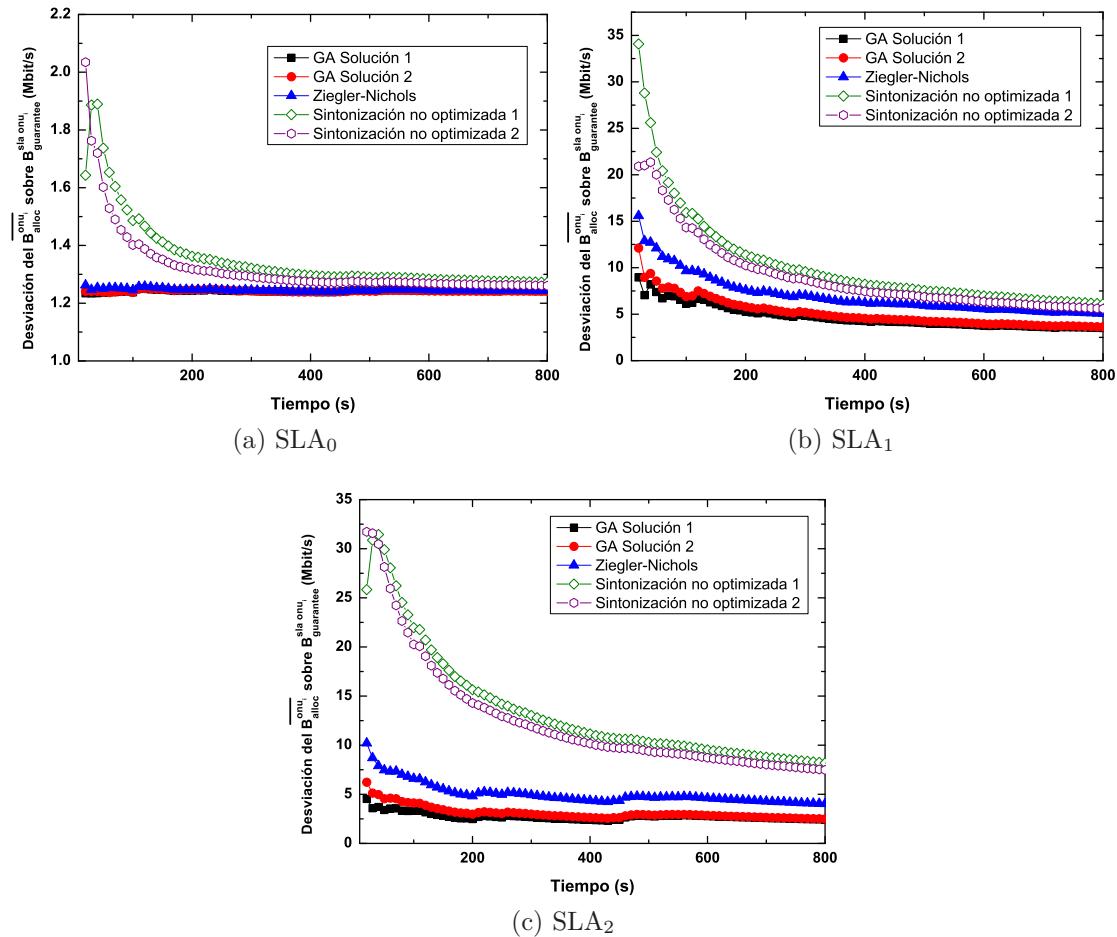


Figura 5.17: Evolución en tiempo real de la desviación estándar del ancho de banda medio asignado sobre el ancho de banda garantizado para las sintonizaciones de la Tabla 5.3. (a) SLA₀ (b) SLA₁ (c) SLA₂.

el SLA de mayor prioridad (SLA₀) es muy pequeña para todas las sintonizaciones consideradas. Esto se debe a que para todas las sintonizaciones el ancho de banda medio de este perfil evoluciona de forma rápida al nivel estipulado (100 Mbit/s). Por el contrario, para el SLA de prioridad intermedia (SLA₁), es importante destacar que la sintonización llevada a cabo por el algoritmo genético en GA-SPID es más estable que el resto, ya que desde un principio la desviación con respecto a su nivel garantizado es menor que en el resto de sintonizaciones. Además se observa cómo para las sintonizaciones no óptimas, la desviación al comienzo de las simulaciones es bastante elevada. Con el paso de tiempo, el PID ajusta el ancho de banda al garantizado con mayor exactitud y las diferencias entre las sintonizaciones son más pequeñas. De forma similar, en el SLA₂ las desviaciones obtenidas por las sintonizaciones con algoritmos genéticos son menores que las obtenidas

con la sintonización de Ziegler-Nichols y con las sintonizaciones no optimizadas.

Por tanto, de los gráficos anteriores se puede extraer la conclusión de que con un PID correctamente sintonizado el ancho de banda medio asignado evoluciona hacia los niveles de ancho de banda garantizados de forma más rápida y con menor fluctuación, esto es, de un modo robusto y estable. Además, se observa que con la sintonización llevada a cabo por el algoritmo genético se alcanzan los valores deseados de forma más rápida, precisa y estable que con la sintonización de Ziegler-Nichols, tal y como se demostró en el apartado anterior. A pesar de eso, pasado un tiempo, el PID es capaz de adaptar el ancho de banda asignado a cada usuario a su nivel garantizado independientemente de la sintonización escogida, aunque esto no resultaría eficiente si se producen cambios en las condiciones del escenario.

5.4.3.3. Independencia de los niveles de ancho de banda garantizado

Para mostrar la bondad del algoritmo genético en el proceso de sintonización del PID, se han realizado simulaciones modificando los niveles de ancho de banda estipulados por el proveedor de servicios para cada SLA. En este sentido, se han considerado cuatro nuevos escenarios, cuyos valores de ancho de banda garantizado para cada SLA aparecen recogidos en la Tabla 5.4.

	Nivel de ancho de banda garantizado a cada SLA		
	SLA ₀	SLA ₁	SLA ₂
Escenario 1	100 Mbit/s	75 Mbit/s	50 Mbit/s
Escenario 2	70 Mbit/s	100 Mbit/s	40 Mbit/s
Escenario 3	90 Mbit/s	70 Mbit/s	55 Mbit/s
Escenario 4	60 Mbit/s	90 Mbit/s	45 Mbit/s

Tabla 5.4: Niveles de ancho de banda garantizado estipulados para cada SLA en los escenarios considerados.

En primer lugar, se va a analizar la evolución en tiempo real de los anchos de banda máximos asignados a una ONU de cada SLA para los distintos perfiles realizado por SPID y GA-SPID. En la Figura 5.18 (a), (b), (c) y (d) se muestra dicha evolución para los tres perfiles de abonado en cada uno de los escenarios considerados. La principal conclusión que se puede obtener de los cuatro gráficos, es que la sintonización en GA-SPID realizada con el algoritmo genético es mejor que la obtenida en SPID (realizada mediante Ziegler-Nichols). En todos los escenarios se observa cómo las fluctuaciones en tiempo real del ancho de banda máximo de GA-SPID son menores que en SPID.

Por tanto, una buena sintonización asegura una evolución hacia los niveles de ancho de banda garantizados más rápida y estable. Esto también puede verse en la Figura 5.19,

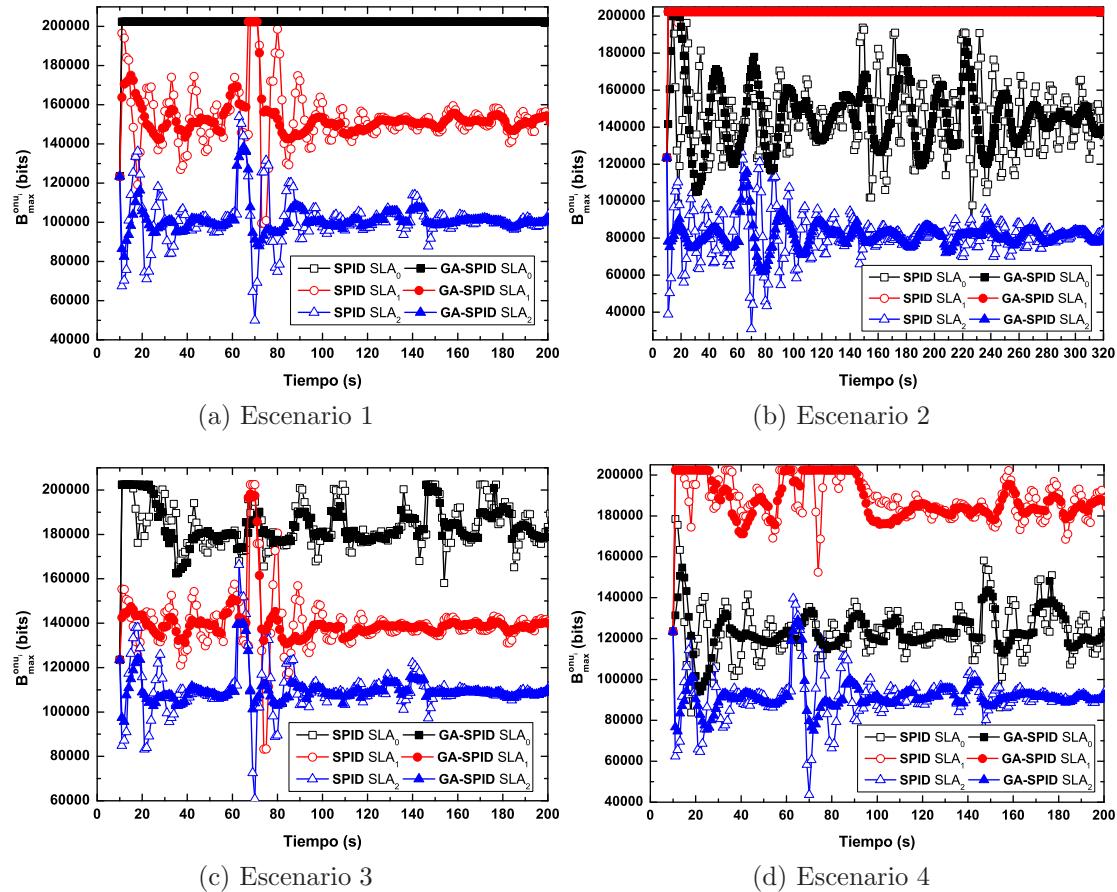


Figura 5.18: Variación en tiempo real del ancho de banda máximo para los distintos escenarios de la Tabla 5.4. (a) Escenario 1 (b) Escenario 2 (c) Escenario 3 (d) Escenario 4.

donde aparece la desviación instantánea del ancho de banda medio asignado a una ONU de cada SLA sobre su nivel de ancho de banda garantizado en cada escenario considerado. En general, puede verse cómo GA-SPID ajusta mejor que SPID el ancho de banda asignado al nivel estipulado para los diferentes SLAs. En concreto, GA-SPID consigue menores desviaciones para el SLA₂ en todos los escenarios, para el SLA₁ en el Escenario 1 y Escenario 3 y para el SLA₀ en el Escenario 3 y Escenario 4.

En resumen, se puede concluir una vez más que mientras el método de sintonización de Ziegler-Nichols es un proceso manual que puede consumir una gran cantidad de tiempo y puede resultar poco preciso, el uso de algoritmos genéticos en la sintonización del PID no solo automatiza el proceso, sino que obtiene sintonizaciones más rápidas y precisas, consiguiendo un funcionamiento más estable y robusto del controlador PID.

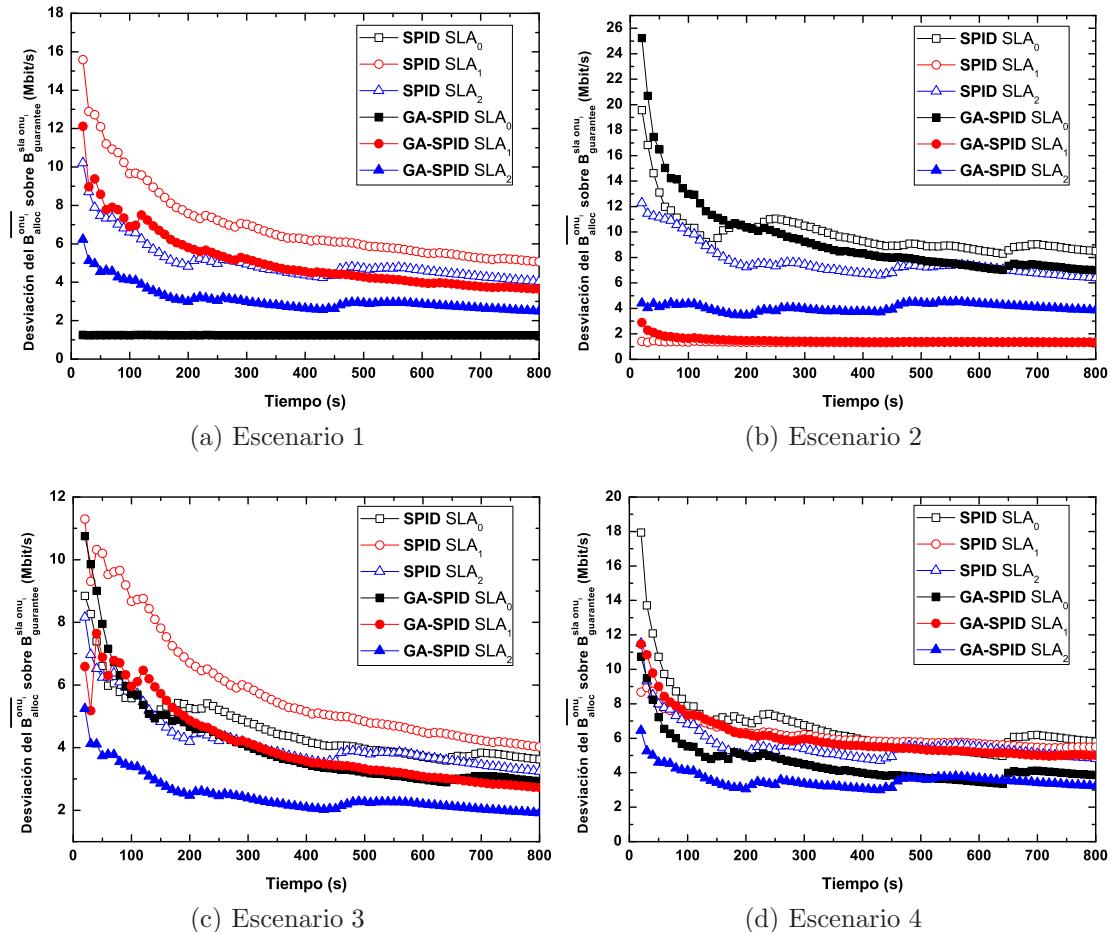


Figura 5.19: Variación en tiempo real de la desviación del ancho de banda medio asignado sobre el ancho de banda garantizado para los distintos escenarios de la Tabla 5.4. (a) Escenario 1 (b) Escenario 2 (c) Escenario 3 (d) Escenario 4.

5.5. Sintonización de DaSPID mediante algoritmos genéticos. Algoritmo GA-DaSPID

En este apartado de la Tesis se presenta el diseño de un algoritmo genético para realizar la sintonización automática del algoritmo DaSPID. La mecánica de este nuevo algoritmo, denominado GA-DaSPID (*Genetic Algorithm Delay aware SPID*) [127], es básicamente la misma que la de GA-SPID. Esto es, a partir de un algoritmo genético se va a sintonizar el controlador P que utiliza DaSPID, tal y como puede verse en el diagrama de bloques representado en la Figura 5.20.

La principal diferencia entre GA-SPID y GA-DaSPID se encuentra en el cromosoma utilizado. En este caso, DaSPID utiliza un controlador P, por lo que el único parámetro a determinar en la sintonización es la constante proporcional K_p . Por tanto, el cromosoma va a estar formado solo por este parámetro, que se traduce en una cadena binaria codificada con 16 bits.

Por otro lado, el intervalo establecido para este valor de K_p es el (0,5], como en GA-SPID, y la población inicial va a estar compuesta por individuos generados aleatoriamente dentro de este intervalo. En cuanto a los operadores genéticos, se va a utilizar la selección por ruleta, el elitismo, el cruce en un único punto y la mutación, estos últimos operadores con las mismas probabilidades que en GA-SPID, esto es, $P_{cruce} = 0.9$ y $P_{mutacion} = 0.01$. Así mismo, para determinar la aptitud de cada individuo en GA-DaSPID se va a utilizar la misma función objetivo que en GA-SPID y que viene dada por la ecuación (5.1). No obstante, para GA-DaSPID el error viene determinado por el sumatorio de la diferencia del retardo de cada una de las clases de servicio con restricciones en el retardo respecto a su cota máxima estipulada.

Por tanto, el algoritmo genético diseñado para la sintonización en SPID es directamente aplicable a DaSPID sin apenas modificaciones, lo que presupone el diseño de un único algoritmo genético integrado que gestiona el proceso de sintonización con diferentes PIDs que controlan distintos parámetros de red únicamente cambiando internamente

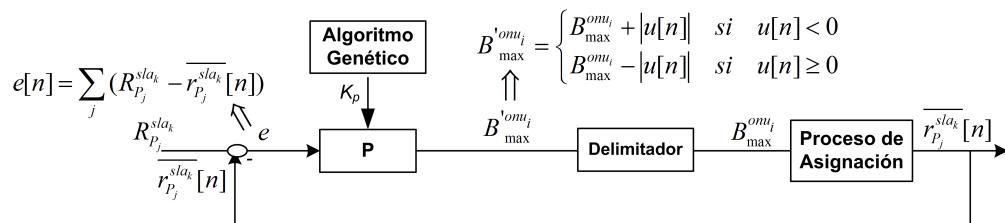


Figura 5.20: Diagrama de bloques del algoritmo GA-DaSPID con sintonización del controlador P mediante un algoritmo genético.

algunos parámetros de diseño según la función objetivo que se desea alcanzar. Esto es, bajo el mismo diseño de algoritmo integrado, se pueden sintonizar automáticamente diferentes controladores PIDs, en concreto, el desarrollado para la gestión de ancho de banda y para la gestión de retardo de los servicios prioritarios.

5.6. Estudio de simulación del algoritmo GA-DaSPID

En este apartado se va a proceder a realizar el análisis de las prestaciones del algoritmo GA-DaSPID. En primer lugar, se determinarán los valores óptimos de parámetros relacionados con el diseño del algoritmo genético para el control óptimo del retardo. A continuación, se estudiará y comparará el tiempo empleado en realizar la sintonización con GA-DaSPID y con Ziegler-Nichols. Después se realizará un estudio de simulación para analizar la adaptación del retardo hacia las cotas establecidas para cada perfil de abonado y clase de servicio, así como la independencia frente a dichas cotas, comparando el proceso de sintonización mediante algoritmos genéticos con el método de Ziegler-Nichols.

5.6.1. Escenario de simulación

El escenario de simulación en GA-DaSPID es el mismo que el considerado en DaSPID (especificado en el Apartado 4.7.1 del Capítulo 4). La Tabla 5.5 recoge un resumen de los principales parámetros de red y los valores considerados para cada uno de ellos.

Parámetros de simulación	Valores		
Número de ONUs	16 ONUs (N_{onus})		
Tasa de transmisión canal ascendente	1 Gbit/s ($EPON_Line_rate$)		
Tasa de transmisión ONU-abonado	100 Mbit/s ($User_Line_rate$)		
Diferenciación de Servicios	$P_0 = 20\%$		
	$P_1 = 40\%$		
	$P_2 = 40\%$		
<i>Service Level Agreement (SLA)</i>	SLA ₀ : 2 ONUs	Retardo P_0 : 1.5 ms Retardo P_1 : 5 ms Retardo P_2 : -	
	SLA ₁ : 6 ONUs	Retardo P_0 : 1.5 ms Retardo P_1 : 20 ms Retardo P_2 : -	
	SLA ₂ : 8 ONUs	Retardo P_0 : 1.5 ms Retardo P_1 : 60 ms Retardo P_2 : -	

Tabla 5.5: Parámetros de red considerados en el escenario de simulación para GA-DaSPID.

5.6.2. Establecimiento de los parámetros asociados al diseño del algoritmo genético en GA-DaSPID

Al igual que ocurría en GA-SPID, es necesario establecer los valores adecuados de los parámetros propios del algoritmo genético para asegurar un buen funcionamiento del mismo a la hora del sintonizar el controlador P. Para ello, se va a determinar el número de iteraciones mínimo que debe ejecutarse el controlador P para realizar un cálculo adecuado de la salud de cada individuo, el tamaño de población y el número de generaciones para establecer el criterio de parada. Al igual que ocurría en GA-SPID, estos tres parámetros tienen una gran influencia en el tiempo de sintonización final con el algoritmo genético.

5.6.2.1. Análisis del número de iteraciones del controlador P para calcular la salud de cada individuo

Tal y como se ha comentado anteriormente, es necesario fijar el número de veces que el individuo proporcionado por el algoritmo genético va a ser utilizado en la sintonización del controlador P para poder calcular su aptitud de forma relativamente precisa u óptima. Por tanto, el tiempo que se va a emplear para calcular la aptitud de cada individuo es $m \cdot T$, donde m es el número de iteraciones del controlador y T es el tiempo de muestreo (que para el caso de DaSPID, tiene un valor de 10 s fijado en el Capítulo 4). En este sentido, un número de iteraciones muy pequeño, puede dar lugar a un cálculo erróneo de la aptitud ya que el individuo se ha utilizado durante un tiempo demasiado corto. Por el contrario, si el número de iteraciones es muy elevado, el tiempo de sintonización se incrementará mucho. Por tanto existe un compromiso entre el número de iteraciones necesarias para evaluar de manera óptima la aptitud del individuo y el tiempo de sintonización.

Para determinar un valor relativamente óptimo, se han llevado a cabo simulaciones considerando distinto número de iteraciones del controlador P (3, 5 y 10). La Figura 5.21 muestra la media del error del mejor individuo frente al número de generaciones cuando se considera un tamaño de población de 15 individuos. Tal y como puede observarse, para el menor número de iteraciones considerado, esto es 3, el error cometido es mayor que para el resto de iteraciones. Además, se aprecia cómo a medida que el número de iteraciones aumenta, el error se reduce. Sin embargo, esto lleva consigo un gran aumento en el tiempo de sintonización, ya que cada incremento en el número de iteraciones añade 10 s al tiempo de evaluación de cada individuo y si la población o el número de generaciones es relativamente alto, este tiempo puede ser inviable en un escenario de red real. Por tanto, se ha optado por un número de iteraciones igual a 5, que obtiene unos errores

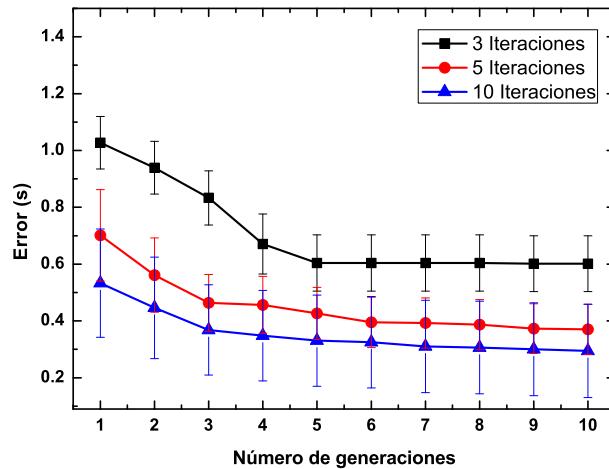


Figura 5.21: Media del error del mejor individuo considerando distinto número de iteraciones del controlador P.

ligeramente superiores a los de 10 iteraciones, pero supone una reducción del 50 % en el tiempo de sintonización comparado con 10 iteraciones.

5.6.2.2. Análisis del tamaño de la población

Al igual que en GA-SPID, el tamaño de la población es un parámetro muy importante a tener en cuenta. Para tamaños de población grandes, el número de soluciones exploradas es mayor. Sin embargo, en nuestro caso, un tamaño de población elevado supone un gran incremento en el tiempo de sintonización, ya que todos los individuos son evaluados durante un tiempo en el controlador P para poder obtener su salud.

La Figura 5.22 muestra la media del error del mejor individuo de cada población, para un número de iteraciones del controlador P igual a 5 (fijado en el apartado anterior) a medida que el número de generaciones aumenta. Los tamaños de población considerados son iguales a 10, 15 y 20 individuos. Tal y como se puede apreciar, la población formada por 10 individuos es la que peores soluciones ofrece, ya que el error que presenta es mayor. Por el contrario, para una población de 20 individuos el error es relativamente menor. No obstante, debido al compromiso existente entre el tiempo de sintonización y el tamaño de la población se ha optado por escoger el tamaño intermedio de 15 individuos, que supone una reducción del 25 % del tiempo de sintonización comparado con un tamaño de población de 20 individuos y además, la diferencia en los errores no es muy significativa a medida que el número de generaciones evoluciona.

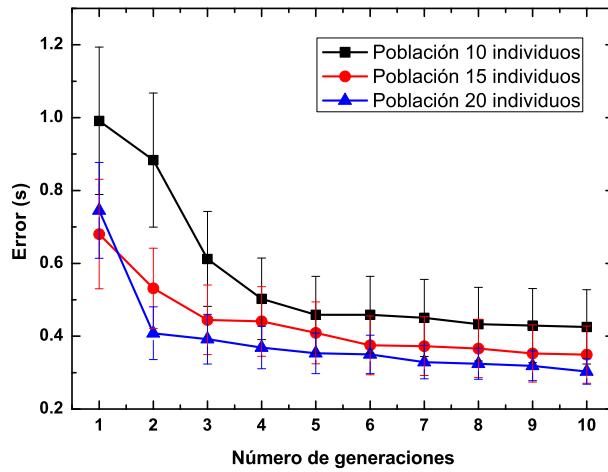


Figura 5.22: Media del error del mejor individuo considerando distinto tamaño de población.

5.6.2.3. Análisis del número de generaciones para el criterio de parada

El último parámetro por establecer es el número de generaciones que es necesario evolucionar el algoritmo. Al igual que los otros dos parámetros, el número de generaciones está también estrechamente relacionado con el tiempo de sintonización.

Para determinar su valor más óptimo se han llevado a cabo simulaciones considerando un número de iteraciones igual a 5 y una población de 15 individuos, dejando evolucionar el algoritmo genético hasta 20 generaciones. Tal y como se observa en la Figura 5.23, donde se muestra la media del error del mejor individuo en cada generación, a partir de la décima generación, los errores no descienden significativamente. Por tanto, se ha optado por escoger un número de generaciones para el criterio de parada igual a 10, esto

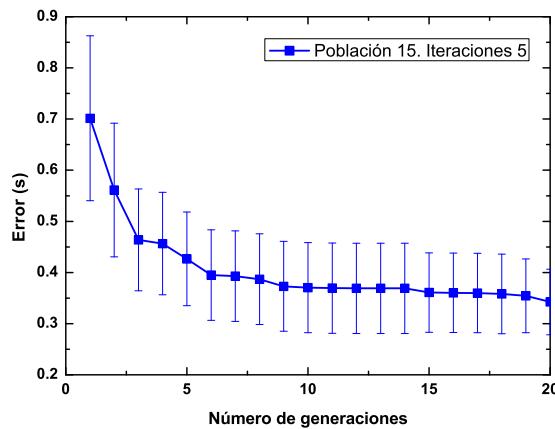


Figura 5.23: Media del error del mejor individuo a lo largo de 20 generaciones.

es, lo mínimo posible para obtener un tiempo de sintonización lo más pequeño posible.

A modo de resumen, en la Tabla 5.6 aparecen los valores finales seleccionados para los distintos parámetros del algoritmo genético diseñado para sintonizar el controlador P en GA-DaSPID.

Parámetros del algoritmo genético	Valores
Selección	Método de la Ruleta
Intervalo de los parámetros	(0,5]
Probabilidad de cruce (P_{cruce})	0.9
Tipo cruce	1 punto
Tamaño población	15 individuos
Elitismo	Sí
Probabilidad de mutación ($P_{mutacion}$)	0.01
Criterio de parada	Número de generaciones igual a 10
Iteraciones del controlador P para obtener la salud	5

Tabla 5.6: Parámetros finales del algoritmo genético diseñado para sintonizar el controlador P en GA-DaSPID.

5.6.3. Análisis de prestaciones de GA-DaSPID

Una vez establecidos los parámetros de los que depende la ejecución y diseño del algoritmo genético, se va a proceder a realizar un estudio de las prestaciones de la sintonización automática que ofrece el algoritmo GA-DaSPID para el control del retardo. Así mismo, se va a comparar con el método de sintonización manual de Ziegler-Nichols en términos de tiempo de sintonización y evolución en tiempo real de los retardos medios a las cotas máximas estipuladas para las clases de servicio prioritarias.

5.6.3.1. Comparación del tiempo de sintonización del método de Ziegler-Nichols respecto al algoritmo genético

Al igual que ocurría para el algoritmo GA-SPID, no se puede llevar a cabo una comparación directa en términos de tiempo de sintonización entre el método basado en el algoritmo genético y el método manual de Ziegler-Nichols. En este sentido, se ha aplicado la adaptación automática del método de Ziegler-Nichols realizada en el Apartado 5.4.3.1 para el control del retardo. Por tanto, el método explicado en dicha sección es aplicable aquí, con la excepción de que es el retardo medio de las clases de servicio prioritarias lo que se va a controlar para observar si sus oscilaciones se encuentran dentro de un cierto intervalo, medido en porcentaje por encima y por debajo de la cota máxima estipulada

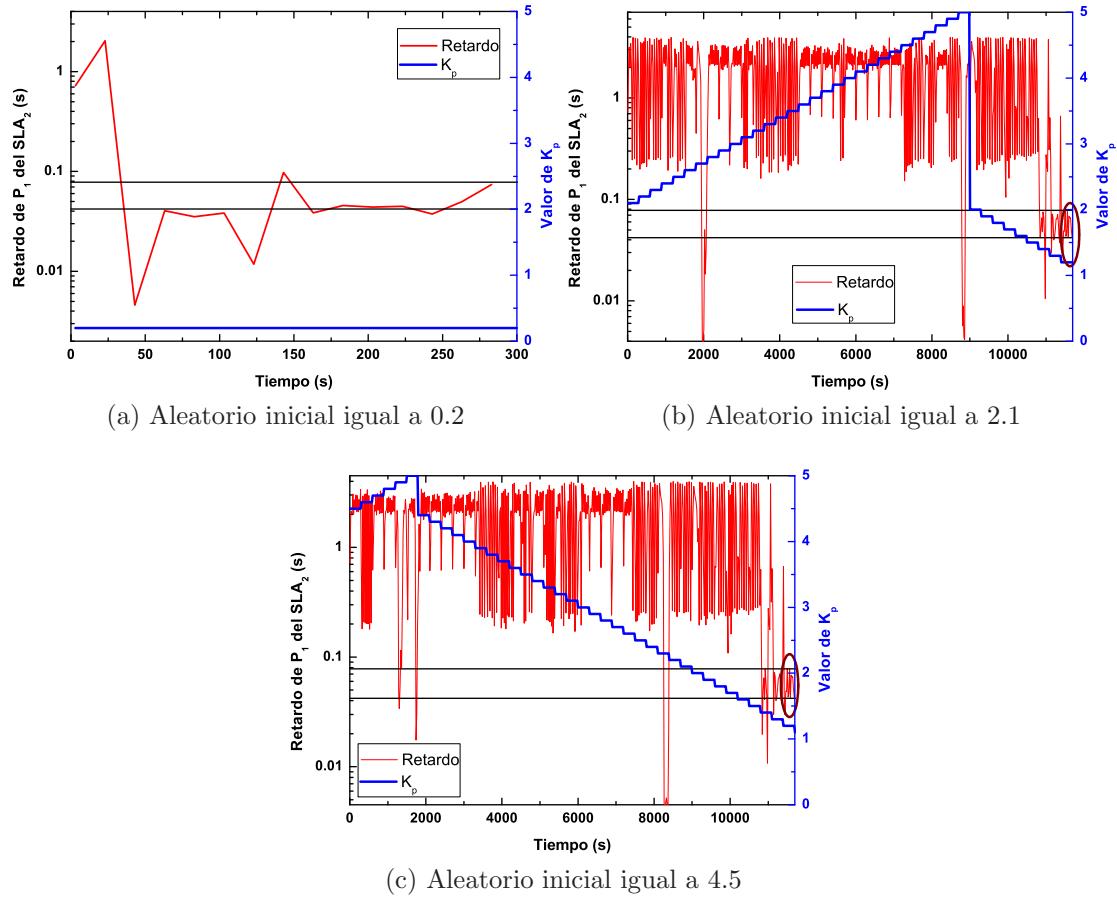


Figura 5.24: Evolución en tiempo real del retardo medio de P_1 del SLA_2 para el algoritmo de sintonización automática basado en el método de Ziegler-Nichols considerando un intervalo de oscilaciones del retardo del 30 % por encima y por debajo de la cota máxima de retardo para este perfil y clase de servicio. (a) K_p inicial igual a 0.2 (b) K_p inicial igual a 2.1 (c) K_p inicial igual a 4.5.

para una determinada clase de servicio y perfil de abonado. Para reducir el número de gráficos, solamente se representan las oscilaciones del retardo medio de la clase de servicio P_1 del SLA_2 . Por otro lado, en este caso concreto, la sintonización automática de Ziegler-Nichols considerará que ha encontrado el valor de K_u cuando el retardo medio de esta clase de servicio se encuentre en el intervalo [0.042 s, 0.078 s], esto es, un 30 % por encima y por debajo de la cota establecida para este SLA y clase de servicio (60 ms).

La Figura 5.24 (a), (b) y (c) muestra la evolución en tiempo real del retardo obtenido para la clase P_1 del SLA_2 para cada uno de los valores de K_p aleatorios iniciales utilizados en la sintonización del controlador P. En concreto, los valores aleatorios para el valor inicial de K_p se han establecido en 0.2 para la Figura 5.24 (a), 2.1 para la Figura 5.24

(b) y 4.5 para la Figura 5.24 (c). En azul y referido al eje de la derecha está representada la variación del valor del parámetro K_p , mientras que en rojo y referido al eje de la izquierda se muestra la variación del retardo medio de la clase P₁ del perfil SLA₂. Tal y como puede observarse, en la Figura 5.24 (a), el tiempo de sintonización es pequeño, ya que el número aleatorio inicial consigue fluctuaciones dentro del intervalo considerado, por lo que el valor óptimo de K_u se obtiene en 300 s. No obstante, para los otros dos valores iniciales de K_p el tiempo de sintonización es mucho mayor. En concreto, para el caso de 2.1 la determinación del valor de K_u se consigue tras 11700 s de simulación, ya que desde este valor inicial, la K_p aumenta hasta un valor de 5 y después disminuye desde 2.1 hasta 1.2, valor donde se obtienen oscilaciones dentro del intervalo considerado. De forma similar, para un aleatorio inicial de 4.5 el tiempo empleado en realizar la sintonización es de 11700 s. Esto demuestra, una vez más, que el tiempo de sintonización siguiendo esta técnica de sintonización manual depende en gran medida del número aleatorio inicial, pero puede ser de hasta decenas de miles de segundos, tiempos que resultarían inviables en un escenario real.

Por el contrario, en el algoritmo GA-DaSPID el tiempo de sintonización viene determinado por la ecuación (5.2) al igual que en el caso de GA-SPID. La Figura 5.25 representa el error cometido frente al tiempo de sintonización empleado para diferentes combinaciones de tamaño de población y número de generaciones en el criterio de parada cuando se considera un número de iteraciones m igual a 5 (fijado en el apartado anterior). El número que aparece al lado de cada punto se corresponde con la generación en

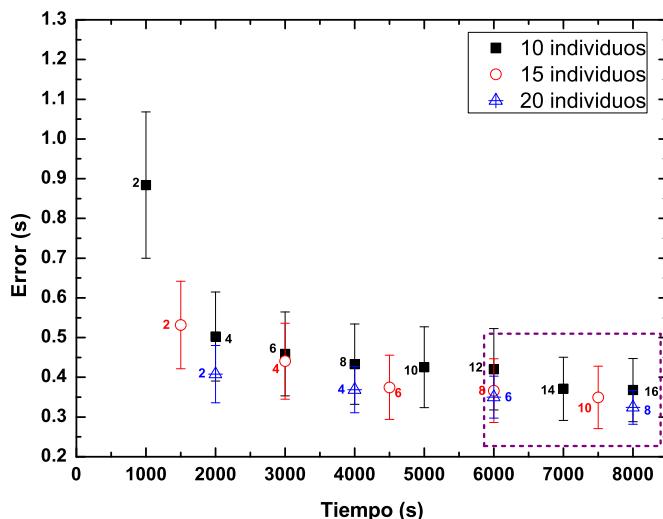


Figura 5.25: Media del error del mejor individuo considerando distinto tamaño de población y número de generaciones en función del tiempo de sintonización en GA-DaSPID.

la que ese error ha sido obtenido. Tal y como se puede observar, a medida que avanza el tiempo, los errores obtenidos van disminuyendo, bien porque se ha aumentado el tamaño de la población o bien el número de generaciones. Sin embargo, a partir de tiempos de sintonización en torno a 7000 s los errores no disminuyen tan significativamente y el algoritmo genético proporciona individuos que, tal y como se comprobará en el siguiente apartado, ofrecen una buena sintonización del controlador P.

Por tanto, se demuestra de nuevo que la sintonización con el método de Ziegler-Nichols, al ser un método manual y experimental, puede resultar laborioso y consumir una gran cantidad de tiempo, obteniendo resultados que pueden no ser suficientemente precisos. No obstante, con los algoritmos genéticos, al igual que ocurría en GA-SPID el tiempo de sintonización se reduce, y se obtienen sintonizaciones más precisas y robustas.

5.6.3.2. Evolución del retardo de las clases de servicio prioritarias a las cotas establecidas por el proveedor de servicios

GA-DaSPID lleva a cabo la sintonización automática de un controlador P destinado a gestionar de forma óptima y dinámica el retardo de las clases de servicio prioritarias con el fin de mantener su valor por debajo de las cotas máximas estipuladas por el proveedor de servicios. Por este motivo, las gráficas que se presentan a continuación muestran los resultados obtenidos en la adaptación del retardo cuando en la sintonización del controlador P se utilizan dos individuos óptimos obtenidos de la ejecución del algoritmo genético. Esto es, se ha escogido el mejor individuo de la última generación en dos simulaciones distintas y se ha utilizado el valor de K_p de cada uno de esos individuos en el controlador P. Además, estos resultados se van a comparar con el algoritmo que utiliza una sintonización basada en el método de Ziegler-Nichols (DaSPID) y con dos sintonizaciones no optimizadas. Los valores del parámetro K_p para las distintas sintonizaciones consideradas aparecen recogidos en la Tabla 5.7.

	Valores
A. Genético Solución 1	$K_p = 0.35$
A. Genético Solución 2	$K_p = 0.462$
Ziegler-Nichols	$K_p = 0.3$
Sintonización no optimizada 1	$K_p = 4.22$
Sintonización no optimizada 2	$K_p = 2.29$

Tabla 5.7: Valores del parámetro K_p seleccionado para cada tipo de sintonización.

La Figura 5.26 (a), (b) y (c) muestra la variación en tiempo real del retardo medio de la clase de servicio P_0 para el SLA_0 , SLA_1 y SLA_2 respectivamente, y para las

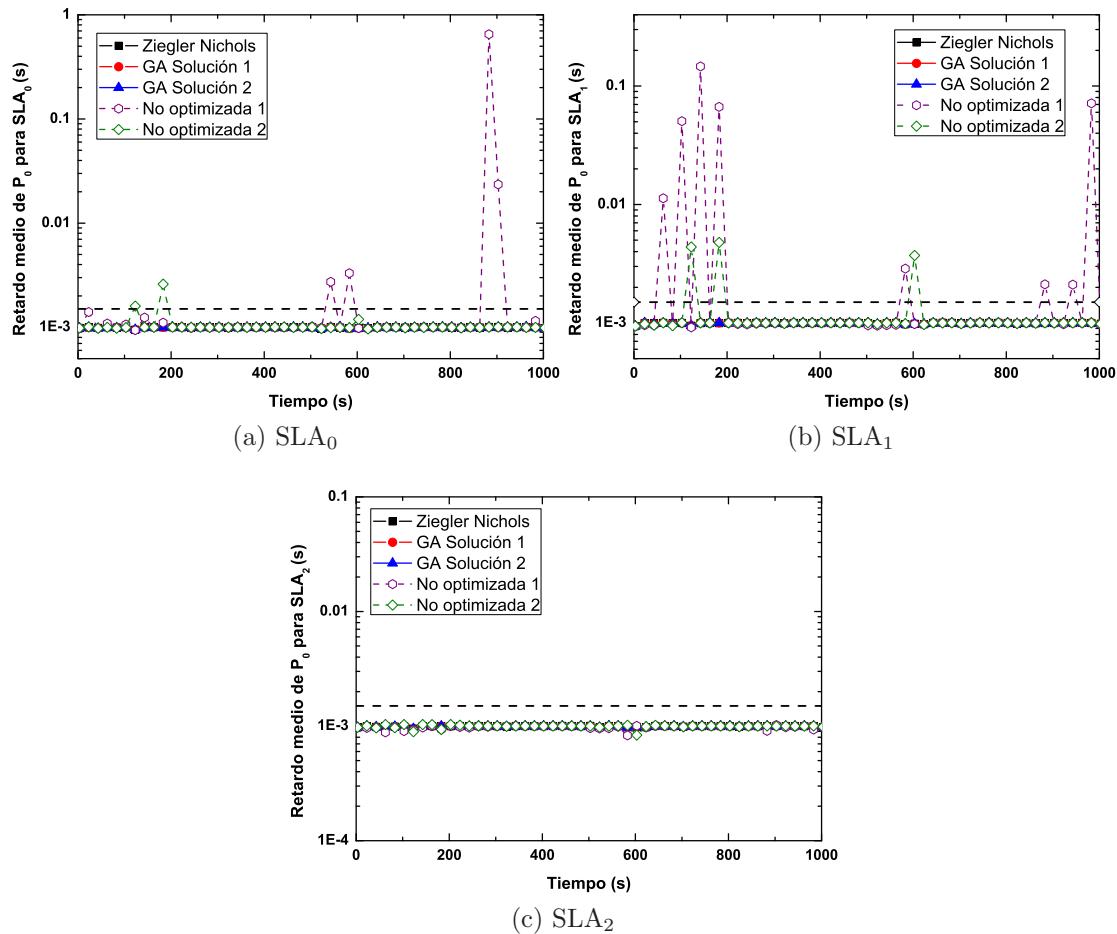


Figura 5.26: Evolución en tiempo real del retardo medio de la clase de servicio P_0 de cada SLA para las sintonizaciones de la Tabla 5.7. (a) SLA₀ (b) SLA₁ (c) SLA₂.

sintonizaciones recogidas en la Tabla 5.7, considerando una configuración de pesos para calcular el ancho de banda máximo inicial igual a $W^{sla_0} = 1.5$, $W^{sla_1} = 1$ y $W^{sla_2} = 0.5$. Además, se ha supuesto una carga de red de 1.4, esto es, todas las ONUs transmitiendo a 87.5 Mbit/s. Tal y como se puede apreciar, para los tres SLAs, tanto la sintonización llevada a cabo mediante Ziegler-Nichols o algoritmos genéticos hacen que el controlador P mantenga el retardo de la clase de servicio más prioritaria P_0 por debajo del límite máximo establecido (1.5 ms). Sin embargo, para las dos soluciones no optimizadas se producen aumentos bruscos en dicho retardo. Por tanto una sintonización del controlador P no adaptada da lugar a un mal funcionamiento del algoritmo que gestiona el retardo.

Para la clase de servicio de prioridad intermedia, P_1 , representada en la Figura 5.27

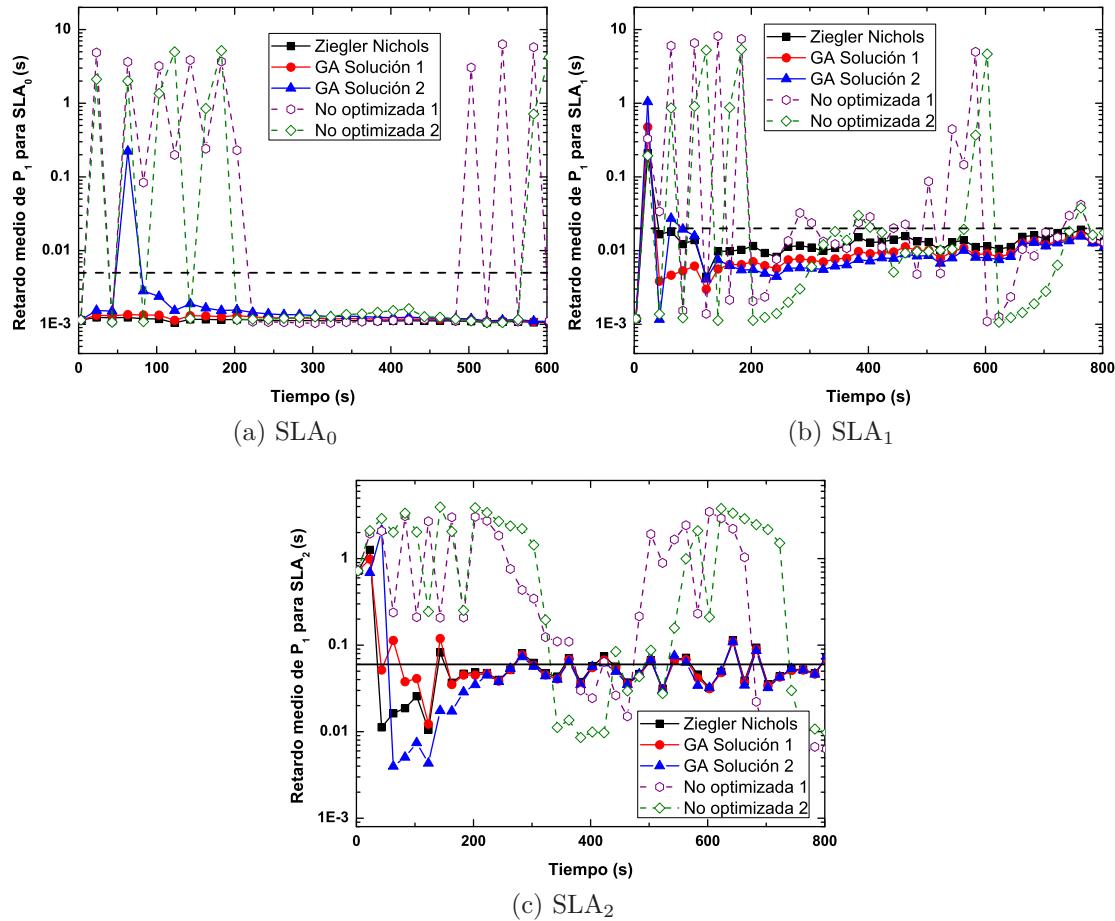


Figura 5.27: Evolución en tiempo real del retardo medio de la clase de servicio P_1 de cada SLA para las sintonizaciones de la Tabla 5.7. (a) SLA₀ (b) SLA₁ (c) SLA₂.

(a), (b) y (c) para los tres SLAs considerados, se observa una situación similar. Para todos los SLAs puede apreciarse cómo la sintonización llevada a cabo con el método de Ziegler-Nichols o con los algoritmos genéticos mantiene el retardo por debajo de las cotas máximas de cada uno de los SLAs. Por el contrario, las sintonizaciones no optimizadas producen fluctuaciones bruscas en el retardo, de forma que incluso a veces no se consigue mantener el retardo por debajo de la cota máxima estipulada para cada SLA. Por otro lado, cabe mencionar que las diferencias obtenidas entre la sintonización con el método de Ziegler-Nichols y GA-DaSPID son muy pequeñas. Esto se debe al hecho de que las soluciones propuestas por los algoritmos genéticos son muy parecidas a la ofrecida por el método de Ziegler-Nichols, tal y como se observa en la Tabla 5.7. Por tanto, se ha llegado a una solución parecida en ambos métodos pero con la ventaja añadida de que en GA-DaSPID el proceso de sintonización es automático y los tiempos de sintonización

son en la mayoría de los casos considerablemente inferiores.

5.6.3.3. Independencia frente a las cotas de retardo de las clases de servicio

Para demostrar la capacidad que ofrece GA-DaSPID para mantener el retardo por debajo de diferentes cotas máximas estipuladas, se han realizado simulaciones considerando los niveles máximos de la Tabla 5.8. Para ello, se ha ejecutado el algoritmo GA-DaSPID para obtener una sintonización previa óptima de acuerdo a los nuevos límites de retardo máximo establecidos para los servicios prioritarios P_0 y P_1 .

Clase de servicio	SLA	Retardo
P_0	SLA_0	1.50 ms
	SLA_1	
	SLA_2	
P_1	SLA_0	10 ms
	SLA_1	40 ms
	SLA_2	80 ms
P_2	SLA_0	-
	SLA_1	
	SLA_2	

Tabla 5.8: Cotaz máximas de retardo consideradas para las diferentes clases de servicio en el nuevo escenario de simulación.

En concreto, la Figura 5.28 muestra la evolución en tiempo real del retardo de la clase de servicio P_1 para los tres SLAs cuando se compara la sintonización con el método

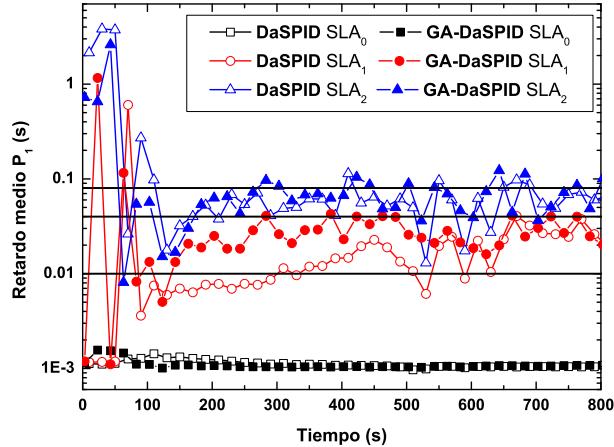


Figura 5.28: Evolución en tiempo real del retardo medio para la clase de servicio P_1 de los tres SLAs considerados comparando la sintonización de DaSPID y GA-DaSPID para las cotas máximas de retardo recogidas en la Tabla 5.8.

de Ziegler-Nichols (algoritmo DaSPID) y la realizada por GA-DaSPID. Tal y como se observa, las diferencias entre ambos métodos son pequeñas. Sin embargo, GA-DaSPID evoluciona los retardos ligeramente más rápido que la sintonización de Ziegler-Nichols, ofreciendo por tanto una mayor rapidez de adaptación a los niveles máximos estipulados.

Por tanto, la sintonización del controlador P llevada a cabo mediante algoritmos genéticos proporciona no solo una automatización del proceso de sintonización frente al método manual de Ziegler-Nichols, sino una sintonización ligeramente más rápida y precisa, que da lugar a un funcionamiento más estable del controlador P que gestiona el control del retardo de los servicios prioritarios.

5.7. Conclusiones

En este capítulo se ha diseñado e implementado un sistema de sintonización óptimo basado en algoritmos genéticos para sintonizar de forma automática controladores PID que gestionan diferentes parámetros de calidad en redes LR-EPON. En concreto, la técnica desarrollada es directamente aplicable a los algoritmos SPID y DaSPID para el control del ancho de banda y del retardo medio, respectivamente. El nuevo sistema, que nunca antes había sido implementado en la sintonización de controladores PID en redes PON y LR-PON, está basado en el uso de algoritmos genéticos para explorar un rango de soluciones y obtener valores que ofrezcan una sintonización óptima. Los nuevos algoritmos, denominados GA-SPID y GA-DaSPID, generan en primer lugar, una población de individuos aleatoria. Después, se obtiene la bondad de cada individuo y se aplican los operadores genéticos, con el fin de obtener en cada generación unos parámetros que sintonicen de forma más adecuada y estable el controlador.

Además, estos algoritmos proporcionan una sintonización automática, a diferencia del método de Ziegler-Nichols, donde es necesario un operador manual que realice el proceso de sintonización, lo que puede resultar tedioso y consumir una gran cantidad de tiempo. Por el contrario, la sintonización basada en algoritmos genéticos ha demostrado reducir dicho tiempo de sintonización de manera considerable, consiguiendo además sintonizaciones más precisas que conllevan un comportamiento más estable de los controladores.

Para el desarrollo de esta técnica, se han establecido en primer lugar, los parámetros de funcionamiento relacionados con el diseño del algoritmo genético, tales como, las fronteras de los parámetros K_p , T_i y T_d en el caso de GA-SPID y K_p en GA-DaSPID. Así mismo, se han establecido los valores adecuados para el tamaño de la población, el número de ejecuciones del controlador para calcular la salud del individuo y el número

de generaciones para determinar el criterio de parada del algoritmo. Tras este proceso de selección, se ha llevado a cabo un estudio de simulación exhaustivo que ha demostrado que la sintonización realizada con el algoritmo genético ofrece mejores resultados y prestaciones que el método manual de Ziegler-Nichols. En concreto, las fluctuaciones obtenidas con GA-SPID para alcanzar los niveles de ancho de banda garantizados y en GA-DaSPID para mantener los niveles de retardo por debajo de las cotas estipuladas, son menores y más estables. Esto implica una adaptación más rápida y precisa a los niveles estipulados de ancho de banda y a las cotas de retardo, respectivamente. Sin embargo, las principales ventajas de GA-SPID y GA-DaSPID, vienen dadas por la automatización del proceso de sintonización y por una reducción considerable del tiempo empleado en llevarla a cabo, tal y como ha quedado demostrado en el estudio de simulación desarrollado a lo largo del capítulo.

Capítulo 6

Métodos de sintonización adaptativa de PIDs basados en redes neuronales

6.1. Introducción

En este capítulo de la Tesis se presenta el diseño e implementación de un método de sintonización adaptativo basado en redes neuronales para ser utilizado en la gestión de recursos en redes LR-EAPON. A diferencia de los métodos presentados anteriormente (Ziegler-Nichols y algoritmos genéticos), este método no realiza una sintonización previa al funcionamiento del controlador, sino que los parámetros de sintonización del controlador PID se van modificando en tiempo real para adaptarse a las diferentes condiciones de red, consiguiendo una sintonización óptima del controlador PID en cada momento.

Se ha optado por utilizar redes neuronales para llevar a cabo esta sintonización adaptativa ya que han demostrado un buen funcionamiento en diferentes campos. Además, en la literatura existen estudios donde las redes de neuronales se utilizan en el proceso de sintonización de controladores PID en diferentes aplicaciones. Por tanto, en la primera parte del capítulo se describirán los conceptos básicos de las redes neuronales, así como ejemplos de controladores PID utilizados en otros ámbitos y sintonizados mediante esta técnica adaptativa. Después, se procederá a describir el proceso de sintonización desarrollado en el sistema que nos atañe. Para ello, se presentará el diseño de la red neuronal y de la arquitectura elegida para llevar a cabo la sintonización de PIDs para la gestión de recursos en redes LR-EAPON.

Por último, se incorporará la red neuronal al controlador PID y se presentarán los

resultados obtenidos de las simulaciones realizadas en distintos escenarios. En concreto, y a modo de ejemplo se realizará el diseño de la red neuronal y el estudio de simulación sobre el algoritmo SPID, encargado de garantizar diferentes niveles de ancho de banda estipulados por los proveedores de servicio a diferentes perfiles de abonado.

6.2. Fundamentos básicos de las redes neuronales artificiales

6.2.1. Introducción

Las redes neuronales artificiales son modelos matemáticos simplificados que intentan reproducir el funcionamiento del sistema nervioso. Tratan de extraer las capacidades del cerebro para resolver ciertos problemas complejos. El objetivo de las redes de neuronas artificiales es llegar a diseñar máquinas con elementos neuronales de procesamiento en paralelo, de modo que el comportamiento global de esa red emule, de la forma más fiel posible, los sistemas neuronales de los animales [130, 131].

En los sistemas neuronales biológicos, existen billones de neuronas que se comunican a través de conexiones llamadas sinapsis. Cada neurona está compuesta por el cuerpo, del que sale una rama principal denominada axón y varias ramificaciones más cortas, las dendritas. De forma general, la neurona recibe los estímulos, y a través de las dendritas llegan al núcleo del cuerpo, donde se genera una respuesta o impulso que se propaga por el axón. Este impulso se ramifica y llega a las dendritas de otras células a través de su sinapsis. La sinapsis, que es la unión entre axón y dendritas, es un líquido con ciertas concentraciones iónicas que potencia o inhibe la señal procedente de los axones. Por tanto, una red neuronal es una enorme malla que propaga señales de una neurona a otras y que va modificando sucesivamente la concentración de iones de la sinapsis para ir activando o inhibiendo las siguientes neuronas.

Las redes de neuronas artificiales están basadas en esta idea pero son mucho más simples. La neurona artificial, al igual que en las redes neuronales biológicas, es la unidad de proceso de información fundamental. Posee un estado interno, llamado nivel de activación, y recibe señales que le permiten, en su caso, cambiar de estado. Para ello, las neuronas artificiales poseen una función, llamada función de activación, que les permite cambiar de nivel de activación a partir de las señales que reciben. La entrada total a la neurona se calcula como la suma de todas las entradas ponderadas por ciertos valores. En la Figura 6.1 se muestra un modelo de neurona artificial, donde tal y como se puede observar, las entradas se corresponden con las señales de la sinapsis de una neurona biológica. Cada señal se multiplica por un peso asociado antes de ser aplicado el suma-

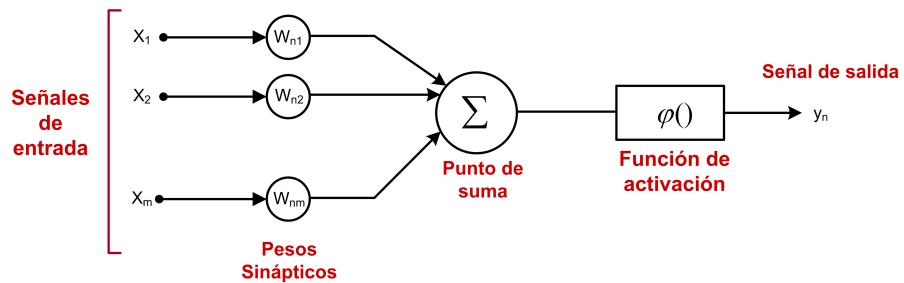


Figura 6.1: Modelo de neurona artificial.

torio. El peso sináptico define la fuerza de una conexión sináptica entre dos neuronas. Los pesos sinápticos pueden tomar valores positivos, negativos o cero. En caso de una entrada positiva, un peso positivo actúa como excitador, mientras que un peso negativo actúa como inhibidor. En caso de que el peso sea cero, no existe comunicación entre el par de neuronas. Mediante el ajuste de los pesos sinápticos, la red es capaz de adaptarse a cualquier entorno y realizar una determinada tarea. Después, se suman todas las entradas ponderadas algebraicamente y posteriormente la señal es procesada por una función llamada función de activación, produciendo la salida de la neurona. Una vez calculada la salida de una neurona, ésta se propaga a las neuronas destino con las que esta neurona está conectada [130, 131].

6.2.2. Topología de las redes neuronales artificiales

La organización y disposición de las neuronas dentro de una red neuronal se denomina topología, y viene dada por el número de capas, la cantidad de neuronas por capa, el grado de conectividad y el tipo de conexión entre neuronas. Las neuronas suelen agruparse en unidades funcionales denominadas capas. Una red neuronal artificial está compuesta por una o más capas, las cuales se encuentran interconectadas entre sí. Una red puede estar formada por una única capa de neuronas (redes monocapa), en las que las neuronas que conforman dicha capa cumplen la función de neuronas de entrada y salida simultáneamente. Cuando la red está compuesta por dos o más capas, ésta se denomina multicapa. La estructura básica de interconexión entre neuronas de la red multicapa está representada en la Figura 6.2. Como se puede apreciar en dicha figura, el primer nivel lo constituyen las neuronas de la capa de entrada, las cuales reciben la información procedente desde el exterior. A continuación hay una serie de capas intermedias, llamadas capas ocultas. Puede haber una o varias capas ocultas. El último nivel es la capa de salida. La salida de estas unidades sirve como salida de toda la red. La red neuronal puede estar formada por tantas capas de tantas neuronas como sean necesarias.

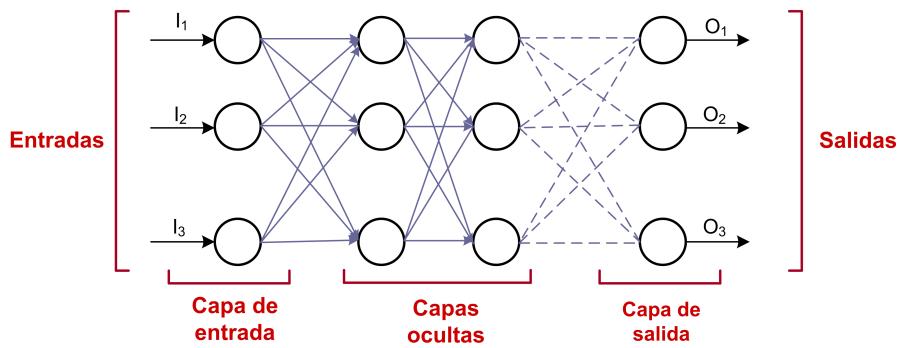


Figura 6.2: Esquema de una red neuronal multicapa totalmente interconectada.

Tal y como se ha comentado anteriormente, entre un par de neuronas de la red neuronal artificial pueden existir conexiones. Estas conexiones son las sinapsis, que tienen asociadas un peso sináptico y son direccionales. Se habla de redes neuronales con conexión hacia delante (redes *feedforward*) cuando las conexiones entre las distintas neuronas de la red siguen un único sentido, desde la entrada de la red hacia la salida de la misma. Cuando las conexiones pueden ser tanto hacia delante como hacia atrás las redes se denominan recurrentes (redes *feedback*).

Cada interconexión entre neuronas actúa como ruta de comunicación. A través de estas interconexiones viajan valores numéricos de una neurona a otra. Estos valores son evaluados por los pesos de las conexiones. Dichos pesos se ajustan durante el proceso de aprendizaje para producir la red neuronal artificial final.

El funcionamiento de una red neuronal se basa en que cada vector de entrada es introducido en la red copiando cada valor de dicho vector en la neurona de entrada correspondiente. Cada neurona de la red, una vez recibida la totalidad de sus entradas, las procesa y genera una salida que es propagada a través de las conexiones, llegando como entrada a la neurona destino. Una vez que la entrada ha sido completamente propagada por toda la red, se producirá un vector de salida, cuyos componentes son cada uno de los valores de salida de las neuronas de salida.

6.2.3. Funciones de activación en las redes neuronales artificiales

Otro parámetro a determinar en el proceso de diseño de una red neuronal artificial es la función de activación. Las funciones de activación definen la salida de la neurona en función del potencial de activación. Las funciones más utilizadas son la función sigmoidal y la función tangente hiperbólica. Dichas funciones poseen como imagen un rango continuo de valores dentro de los intervalos $[0,1]$ y $[-1,1]$, respectivamente. Ambas funciones son crecientes con dos niveles de saturación, el máximo, que proporciona salida 1, y el

mínimo, que proporciona salida 0 para la función sigmoidal y salida -1 para la tangente hiperbólica [130, 131]. La funciones vienen dadas por las expresiones mostradas en la ecuación (6.1) para la función sigmoidal y la ecuación (6.2) para la tangente hiperbólica.

$$f(x) = \frac{1}{1 + e^{-x}} \quad (6.1)$$

$$f(x) = \frac{1 - e^{-x}}{1 + e^{-x}} \quad (6.2)$$

6.2.4. Algoritmos de aprendizaje en las redes neuronales artificiales

En una red neuronal se pueden distinguir claramente dos fases o modos de operación: la fase de aprendizaje o entrenamiento, y la fase de operación o ejecución. Durante la primera fase, la fase de aprendizaje, la red es entrenada para realizar un determinado tipo de procesamiento. Una vez alcanzado un nivel de entrenamiento adecuado, se pasa a la fase de operación, donde la red es utilizada para llevar a cabo la tarea para la cual fue entrenada.

El aprendizaje es la parte más importante de una red de neuronas artificiales. Consiste en la determinación de los valores precisos de los pesos para todas sus conexiones, de tal forma que la red neuronal sea capaz de resolver un problema de manera eficiente. Existen dos tipos básicos aprendizaje, el aprendizaje supervisado y el no supervisado. La diferencia fundamental entre ambos tipos radica en la existencia o no de un agente externo o supervisor que controla el proceso de aprendizaje de la red [132]. A continuación se presentan las características más importantes de ambos tipos de aprendizaje.

- Aprendizaje supervisado. Se caracteriza por la presencia de un agente externo que controla el proceso de entrenamiento, comprobando si la salida de la red es la deseada. Si existen diferencias entre ambos valores, se ajustan iterativamente los pesos hasta que la salida tienda a ser la deseada, utilizando para ello información del error cometido en cada caso. Dentro de este modelo de aprendizaje se distinguen tres tipos:

- *Aprendizaje por corrección de error.* Es el tipo de aprendizaje supervisado más utilizado. Su funcionamiento se basa en el ajuste de los pesos de las conexiones de la red a partir de la diferencia entre los valores deseados y los obtenidos, esto es, en función del error cometido en la salida. Si dicha diferencia es grande y por tanto los datos son muy diferentes, la modificación de los pesos será mayor que en el caso de diferencias pequeñas. Por tanto, para poder utilizar este esquema de aprendizaje se necesitan los datos de

aprendizaje propiamente dichos y cierta información relativa a la solución del problema (las salidas que se desean obtener).

- *Aprendizaje por refuerzo.* En este caso no se dispone de información concreta del error cometido por la red para cada ejemplo de aprendizaje, sino que simplemente se determina si la salida producida para dicho patrón es o no correcta. Por tanto el supervisor indica mediante una señal de refuerzo si la salida se ajusta o no a la deseada y en función de ello se procede al ajuste de los pesos utilizando un mecanismo basado en probabilidades. Este tipo de aprendizaje resulta más lento que el anterior, ya que no se indica de forma expresa la salida ante una determinada entrada.
- *Aprendizaje estocástico.* Este tipo de aprendizaje introduce cambios aleatorios en los valores de los pesos de la red neuronal, y se evalúa su efecto a partir de la salida deseada y de una distribución de probabilidad. Por tanto, si el efecto del cambio del valor es positivo, entonces se acepta la modificación. En caso contrario, la aceptación del cambio depende de una distribución de probabilidad preestablecida.
- *Aprendizaje no supervisado.* Este aprendizaje no requiere información externa para ajustar los pesos de las conexiones neuronales. En este aprendizaje se presenta a la red un conjunto de patrones sin adjuntar la respuesta deseada, por lo que no hay nada que permita guiar el proceso de aprendizaje. En este caso, la red modificará los valores de los pesos en base a la correlación existente entre los datos de entrada, tratando de determinar características de los datos del conjunto de entrenamiento (rasgos significativos, regularidades o redundancias).

De los distintos esquemas de aprendizaje explicados anteriormente, los más comunes son los basados en ejemplos. En concreto, el que mejores características presenta es el aprendizaje supervisado por corrección de error, ya que tiene en cuenta la magnitud del error para modificar en mayor o menor medida los pesos, logrando un aprendizaje más rápido. En este caso, la capacidad de una red para resolver un problema estará ligada al tipo de ejemplos de que dispone en el proceso de aprendizaje. El proceso de aprendizaje debe ser significativo, es decir, debe haber un número suficiente de ejemplos para que la red sea capaz de adaptar sus pesos de forma eficaz. Además, este número de ejemplos debe ser representativo, es decir, los componentes del conjunto de aprendizaje deben ser diversos para que sea general y no se especialice en un subconjunto concreto de datos [130]. Partiendo de un conjunto de pesos sinápticos aleatorio, el proceso de aprendizaje busca un conjunto de pesos que permitan a la red desarrollar correctamente una de-

terminada tarea. Durante el proceso de aprendizaje se va refinando iterativamente la solución hasta alcanzar un nivel de operación suficientemente bueno. El proceso general de aprendizaje consiste en ir introduciendo paulatinamente todos los ejemplos del conjunto de aprendizaje, y modificar los pesos de las conexiones siguiendo un determinado esquema de aprendizaje. Una vez introducidos todos los ejemplos se comprueba si se ha cumplido cierto criterio de convergencia y de no ser así se repite el proceso. El criterio de convergencia depende del tipo de red utilizado o del tipo de problema a resolver. La finalización de la fase de aprendizaje se puede determinar usando los siguientes criterios:

- Fijando el número de ciclos o iteraciones. Se decide a priori cuántas veces será introducido todo el conjunto, y una vez superado dicho número se detiene el proceso y se da por aceptada la red resultante.
- Cuando el error descienda por debajo de una cantidad preestablecida. En este caso hay que definir una función de error. Se decide a priori un valor aceptable para dicho error, y solo se para el proceso de aprendizaje cuando la red obtenga un error por debajo del nivel prefijado.
- Cuando la modificación de los pesos sea irrelevante. En el proceso de aprendizaje, llegará un momento en el que ya no se produzcan variaciones en los valores de los pesos de ninguna conexión. En ese momento se dice que la red ha convergido y se detiene el proceso de aprendizaje.

Una vez finalizada la fase de aprendizaje, la red puede ser utilizada para realizar la tarea para la que fue entrenada. Una de las principales ventajas es que la red aprende la relación existente entre los datos, adquiriendo la capacidad de generalizar conceptos. Se trata de que, tras un periodo de entrenamiento determinado, la red esté lista para funcionar correctamente y calcular la salida para cualquier caso que se presente. De esta manera, una red neuronal puede tratar con información que no le fue presentada durante la fase de entrenamiento, es decir, la red será capaz de calcular la salida correspondiente a cualquier entrada que se presente y no solo a los datos pertenecientes al conjunto de entrenamiento.

6.2.5. Modelo del perceptrón multicapa

De los distintos modelos de redes neuronales existentes, en esta Tesis nos vamos a centrar en el estudio del perceptrón multicapa, por ser el modelo más utilizado en la práctica tanto para la resolución de problemas de clasificación como de regresión, ya que ha demostrado su condición de aproximador universal de funciones [132].

Este modelo de red surgió en los años 80 como extensión del perceptrón simple (primer modelo en poseer un mecanismo de entrenamiento que permitía determinar automáticamente los pesos sinápticos) para superar las limitaciones de dicha arquitectura. En concreto, para subsanar la imposibilidad del perceptrón para aprender clases de funciones no linealmente separables.

El perceptrón multicapa es un modelo neuronal con propagación hacia adelante, que se caracteriza por su organización en capas de celdas disjuntas, de forma que ninguna salida neuronal constituye una entrada para las neuronas de la misma capa o de capas previas, evitándose así las conexiones hacia atrás o autorecurrentes. Su topología está definida por un conjunto de capas ocultas, una capa de entrada y una de salida. Por tanto, la topología representada en la Figura 6.2, es realmente la topología de un perceptrón multicapa. Además, aunque la red puede contener un número indeterminado de capas ocultas, se ha demostrado que la inclusión de una única capa oculta resulta suficiente para que la red actúe como un aproximador universal de funciones [132].

6.2.5.1. Métodos de entrenamiento en el perceptrón multicapa

La mayoría de los métodos de entrenamiento utilizados en este tipo de redes neuronales con conexión hacia delante proponen una función de error que mide el rendimiento actual de la red en función de los pesos sinápticos. Por tanto, se define una función $E(W)$ que proporciona el error que comete la red en función del conjunto de pesos sinápticos (W). El objetivo del aprendizaje será encontrar la configuración de pesos que corresponda al mínimo global de la función de error, aunque en muchos casos es suficiente encontrar un mínimo local lo suficientemente bueno. La función de coste a minimizar puede definirse de distintas formas, si bien en la práctica suele considerarse el error cuadrático medio [132]. Por tanto, el error $E(W)$ se define según la ecuación (6.3), donde el término $Y(n) = (y_1(n), \dots, y_{n_c}(n))$ se corresponde con el vector de salida de las n_c neuronas de la capa de salida para el patrón n y $S(n) = (s_1(n), \dots, s_{n_c}(n))$ es el vector de salidas deseadas para ese mismo patrón. Además, N es el número de muestras o patrones considerados en el aprendizaje.

$$E(W) = \frac{1}{2} \sum_{n=1}^N \sum_{i=1}^{n_c} (s_i(n) - y_i(n))^2 \quad (6.3)$$

Existen múltiples algoritmos de aprendizaje en el perceptrón multicapa, que pueden clasificarse según distintos criterios. Uno de los criterios más útiles en la práctica es el basado en el orden de la derivada de la función de error, que permite distinguir cuatro grandes grupos de algoritmos de aprendizaje [132]:

- Métodos sin derivadas. Emplean únicamente información local sobre los valores de la función $E(W)$, bien ante la imposibilidad de utilizar derivadas parciales (como en el caso de las funciones no diferenciables) o bien ante el excesivo coste de su cálculo. Entre ellos, destacan los algoritmos *inertial shaker* y *affine shaker* que llevan a cabo la búsqueda del vector óptimo de pesos mediante la definición de una región de búsqueda que se va modificando según el desempeño de la red.
- Métodos que emplean la primera derivada o métodos de primer orden. Estos algoritmos de aprendizaje consideran la primera derivada de la función de error respecto al vector de pesos para la actualización de los pesos neuronales $\left(\frac{\partial E(W)}{\partial W}\right)$. Entre los métodos de primer orden se incluyen los algoritmos basados en el gradiente descendente, como el algoritmo de retropropagación (*backpropagation*).
- Métodos que emplean la segunda derivada o métodos de segundo orden. Los métodos de segundo orden consideran la segunda derivada para la actualización de los parámetros de la red $\left(\frac{\partial^2 E(W)}{\partial^2 W}\right)$. Si bien los resultados obtenidos suelen ser muy satisfactorios, requieren un elevado tiempo de computación y muchos recursos de memoria, lo que limita su aplicación a problemas con un número limitado de pesos (en ocasiones inferiores a 100). En este grupo de métodos se incluyen los algoritmos basados en el gradiente conjugado, el método de Newton y los métodos Quasi-Newton.
- Métodos heurísticos. Esta categoría incluye algoritmos que llevan a cabo el ajuste del vector de parámetros de la red mediante métodos heurísticos basados en prueba y error. Dentro de este grupo pueden incluirse los algoritmos de aprendizaje basados en algoritmos genéticos y búsqueda evolutiva.

A continuación se explican brevemente algunos de los métodos de entrenamiento más comunes en redes de neuronas artificiales.

6.2.5.1.1. Método de Retropropagación

El algoritmo de retropropagación o *backpropagation* es el método de entrenamiento más utilizado en redes multicapas con conexión hacia delante. Está basado en el método del gradiente descendente, que es uno de los métodos de optimización de funciones multivariantes más conocidos. Esta técnica realiza una adaptación de los parámetros siguiendo una cierta dirección de búsqueda, que es la dirección negativa del gradiente de la función E , pues conforme al cálculo de varias variables, ésta es la dirección en la que la función decrece. Por tanto, aplicando el método de descenso del gradiente, cada conjunto

de pesos W se modifica de acuerdo con la ley de aprendizaje de la ecuación (6.4), donde $\nabla E(W)$ es el gradiente de la función de error que se quiere minimizar.

$$W(n+1) = W(n) - \alpha \nabla E(W) \quad (6.4)$$

En la ecuación (6.4), α es el parámetro de velocidad de aprendizaje. Este parámetro determina la magnitud de desplazamiento de los pesos, influyendo en la velocidad de convergencia del algoritmo. Si se toma un valor pequeño, el proceso de entrenamiento puede resultar muy lento, mientras que si el valor es muy grande pueden producirse oscilaciones en torno al punto mínimo. Este valor, puede permanecer constante durante todo el proceso o puede ser recalculado en cada iteración. Por ejemplo, es posible incorporar un mecanismo que permita usar valores grandes de α al comienzo y que a medida que se avance en el proceso de optimización haga que este valor decrezca [131, 133].

De forma simplificada, el algoritmo de retropropagación consiste en el aprendizaje de un conjunto predefinido de patrones de entrada-salida, empleando un ciclo de propagación-adaptación con dos fases diferenciadas. En la primera fase o fase de aprendizaje hacia delante, se aplica un patrón de entrada que se propaga por las distintas capas que componen la red hasta producir la salida de la misma. Esta salida se compara con la salida deseada y se calcula el error cometido por cada neurona de salida. En la segunda fase o fase de aprendizaje hacia atrás, estos errores se transmiten hacia atrás, desde la capa de salida hacia las neuronas de la capa intermedia. El proceso se repite capa por capa hasta que todas las neuronas de la red hayan recibido un error que describa su aportación relativa al error de la salida final. A partir del valor de error recibido, se procede al reajuste de los pesos de cada neurona para minimizar el error cometido por la red [130–133].

6.2.5.1.2. Método del Gradiente Conjugado

En este método la dirección de búsqueda se realiza a lo largo de direcciones conjugadas, lo que produce una convergencia más rápida que el método del gradiente descendente. Estas direcciones de entrenamiento están conjugadas con respecto a la matriz Hessiana. De esta forma, si se usa una función de error cuadrática y se usan direcciones de búsqueda no interfiriéntes y linealmente independientes, la función de error se minimiza respecto a la totalidad de subespacio, y el aprendizaje debería terminar en un máximo de N iteraciones (con N el total de pesos de sistema) [132].

6.2.5.1.3. Método de Newton y Quasi-Newton

El método de Newton, al calcular las derivadas de segundo orden de la función

objetivo, obtiene un mejor conocimiento de la topología de la función, lo que da lugar a una elección más eficiente de la dirección descendente a partir del cálculo de la inversa de la matriz Hessiana. Obtener una solución de este tipo en cada iteración puede ser una tarea computacionalmente prohibitiva ya que se requieren muchas operaciones para evaluar la matriz Hessiana y computar su inversa. Por ello, el método Quasi-Newton evita esta dificultad, aproximando directamente la inversa de la matriz Hessiana a partir de la derivada de primer orden. Este método es uno de los más populares en las técnicas de optimización sin restricciones, obteniendo muy buenos resultados en el entrenamiento de redes neuronales [132–135].

6.2.5.1.4. Método de Levenberg-Marquardt

Este algoritmo es uno de los más eficientes para entrenar redes neuronales con un número moderado de parámetros de red. Está basado en el método de los multiplicadores de Lagrange, para la optimización con restricciones de igualdad. Por otro lado, la desventaja del algoritmo radica en el cálculo, almacenamiento e inversión de la matriz Hessiana [131].

De todos estos métodos, el gradiente descendente en el que se basa el algoritmo de retropropagación es uno de los métodos más usados. No obstante, aunque es robusto, su buen funcionamiento está sujeto a un tiempo de entrenamiento que puede llegar a ser muy elevado, lo que puede no ser adecuado en muchas aplicaciones. Por otro lado, el método Quasi-Newton se considera más potente y su aplicación en el entrenamiento de perceptrones multicapa ha tenido mucho éxito [134, 136].

6.2.6. Estado del arte de redes neuronales para la sintonización de PIDs

Tal y como se ha comentado anteriormente, los métodos de sintonización de controladores PID presentados hasta ahora en esta Tesis (método de Ziegler-Nichols y algoritmos genéticos), son métodos de sintonización que hay que ejecutar de forma previa al funcionamiento de la red para hallar los valores óptimos de los parámetros del controlador. Además, una vez fijados dichos valores éstos permanecen fijos. En nuestro caso, esto quiere decir que, si los parámetros de la red cambian a lo largo del tiempo, la sintonización puede no llegar a ser del todo óptima para esas nuevas circunstancias de la red, de forma que, aunque el PID seguiría presentando un buen funcionamiento, éste podría perfeccionarse. Para mejorar este aspecto, en este capítulo se presenta el diseño e implementación de una red neuronal que, dependiendo del estado actual de la red, proporciona

unas constantes de sintonización diferentes y adecuadas a las nuevas circunstancias de la misma.

En este sentido, en la literatura existen diversos ejemplos de controladores PID sintonizados mediante redes neuronales en diferentes campos de aplicación. Por ejemplo, en [137] se presenta un controlador PID para ajustar la altura de corte de una máquina, cuya sintonización se realiza mediante una red neuronal basada en un algoritmo de auto-aprendizaje adaptativo. Por otro lado, en [138] se utiliza un controlador PID con sintonización adaptativa basada en una red neuronal de retropropagación, pero en este caso en el campo de los vehículos eléctricos. Los autores de [139] presentan un PID con sintonización adaptativa mediante una red neuronal para desacoplar sistemas no lineales de múltiples variables. Por otro lado, en los artículos [140, 141] se propone el diseño de un controlador PID adaptativo basado en una red neuronal de retropropagación aplicado al control de la presión de un sistema hidráulico. En todos ellos, el PID con una sintonización adaptativa basada en redes neuronales ofrece un mejor comportamiento que un PID tradicional, ya sea porque los sistemas que controlan sean no lineales o variantes en el tiempo, o simplemente porque ofrezcan una mayor robustez a cambios en el entorno. Por lo tanto, la incorporación de redes neuronales en el proceso de sintonización puede resultar interesante en la gestión de recursos en redes LR-EAPON y EPON, algo que, por otro lado, nunca se había realizado.

6.3. Sintonización adaptativa de SPID mediante redes neuronales. Algoritmo NN-SPID

En este apartado se va a proceder a describir el proceso de diseño de la red neuronal para su incorporación en la sintonización de controladores PID en la gestión de recursos en redes LR-EAPON. Como ejemplo, la red neuronal diseñada se va a integrar en el controlador PID del algoritmo SPID para llevar a cabo una gestión eficiente del ancho de banda. Este nuevo algoritmo se ha denominado NN-SPID (*Neural Network Service level agreement PID*).

Para el diseño de la red neuronal que va a sintonizar al PID, se va a utilizar una representación alternativa de la ecuación del PID en tiempo discreto presentada en la ecuación (4.2) del Capítulo 4. En concreto, la nueva expresión viene dada por la ecuación (6.5), donde K_p es la constante del término proporcional, K_i es la constante del término integral y K_d es la constante del término derivativo.

$$u[k] = u[k - 1] + K_p (e[k] - e[k - 1]) + K_i e[k] + K_d (e[k] - 2e[k - 1] + e[k - 2]) \quad (6.5)$$

Por tanto, los parámetros de sintonización del PID utilizando redes neuronales son los términos K_p , K_i y K_d . La relación entre las constantes K_i y K_d de la ecuación (6.5) con los términos T_i y T_d de la ecuación (4.2) viene dada por la ecuación (6.6), donde T representa el tiempo de muestreo.

$$K_i = K_p \frac{T}{T_i} \quad K_d = K_p \frac{T_d}{T} \quad (6.6)$$

Por tanto, el esquema general del proceso controlado por un PID con una sintonización adaptativa mediante redes neuronales es el representado en la Figura 6.3. Tal y como se observa, se ha incorporado una red neuronal para proporcionar el valor de las constantes de sintonización (K_p , K_i y K_d). Además, al ser una sintonización adaptativa, la red neuronal necesita también conocer el error en tiempo real que se está produciendo para adecuar la sintonización a las circunstancias actuales.

Por otro lado, para llevar a cabo el diseño de la arquitectura de una red neuronal es necesario determinar el valor de una serie de parámetros, tales como, el número de capas en la red, el número de neuronas de cada capa, la función de activación a emplear y el método de entrenamiento. El primer paso consiste en decidir el número de capas que va a formar la red neuronal. Para la sintonización del PID es indispensable tener una capa de entrada y una de salida, además de tantas capas ocultas como sean necesarias. El estudio bibliográfico de la utilización de redes neuronales en controladores PID reveló que en casi todos los artículos se utilizaba una red neuronal con una sola capa oculta [137–143]. Por tanto, se ha optado por una red neuronal de tres capas, una de entrada, una oculta y otra de salida, lo cual resta complejidad al diseño final de la red.

A continuación, es necesario determinar el número de neuronas por capa. Como el número de parámetros a sintonizar en el PID son tres (K_p , K_i y K_d), la red neuronal debe proporcionar tres coeficientes de salida y por tanto, tiene que tener tres neuronas en la capa de salida. Por otro lado, se ha escogido un número de entradas igual a tres. En

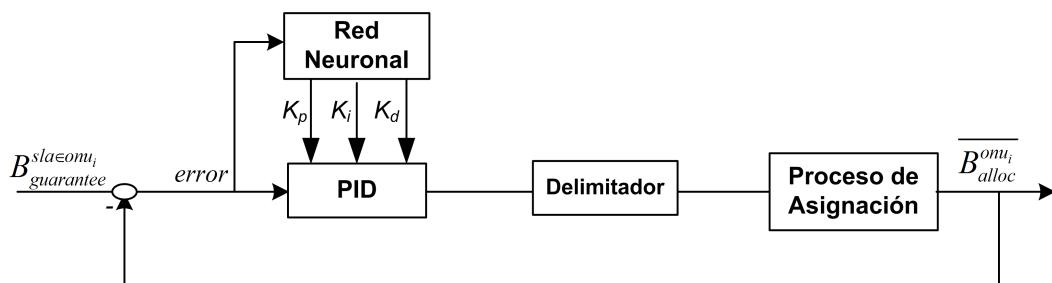


Figura 6.3: Diagrama de bloques del controlador PID para la gestión de ancho de banda en SPID con sintonización adaptativa mediante redes neuronales.

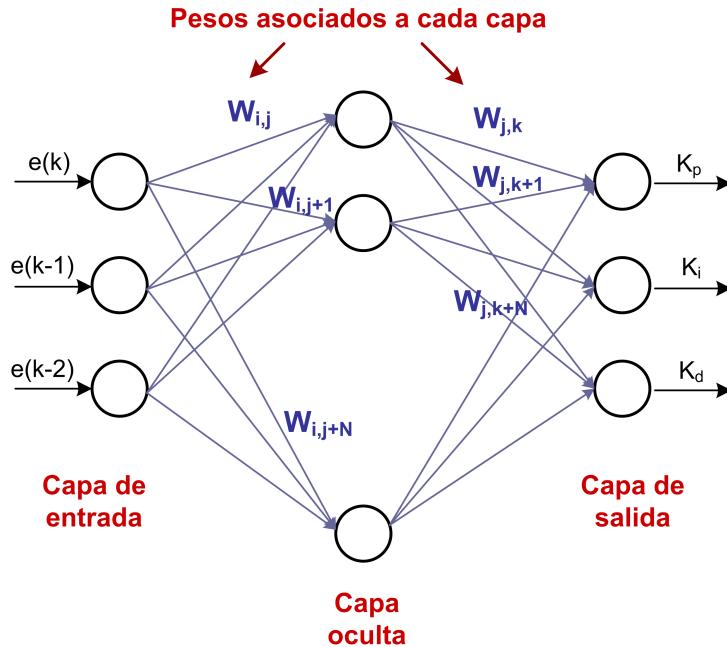


Figura 6.4: Diseño final de la red neuronal para la sintonización adaptativa del controlador PID.

concreto, las tres entradas de la red se corresponden con los errores cometidos en las dos últimas iteraciones del PID y el error de la iteración actual, esto es, $e[k - 2]$, $e[k - 1]$ y $e[k]$ al igual que en [137]. Respecto al número de neuronas que debe tener la capa oculta, dependerá de la exactitud que se quiera obtener (hasta llegar a un límite máximo, a más neuronas más precisión). Sin embargo, se puede concluir, a partir de la bibliografía analizada, que una capa oculta de entre cuatro y ocho neuronas podría ser suficiente para que la red neuronal funcione con la exactitud necesaria. En concreto, en nuestro modelo se han utilizado cinco neuronas en la capa oculta al igual que en [138, 140–142], aunque también se ha llevado a cabo un análisis de simulación para estudiar el impacto del número de neuronas en la capa oculta. Por lo tanto, el diagrama genérico final de la red neuronal que sintonizará al controlador PID en SPID se presenta esquematizado en la Figura 6.4.

La elección de la función de activación se suele hacer basándose en el recorrido deseado. Puesto que K_p , K_i y K_d no pueden ser negativos, se utiliza como función de activación de la capa de salida la función sigmoidal no negativa, mientras que para la capa oculta se utiliza la función tangente hiperbólica, al igual que en [137, 138, 141, 143].

Por último, cabe destacar que para la implementación de una primera aproximación de red neuronal, se ha utilizado un aprendizaje supervisado, presentando a la red una serie de ejemplos para conseguir que la red realice un buen aprendizaje y pueda funcionar

de forma óptima a partir de éstos. Para ello, los ejemplos utilizados se corresponden con buenos individuos proporcionados previamente por el algoritmo genético. De esta forma se pretende que la red neuronal consiga deducir y extraer las características que hacen que esos individuos sean buenos, para que después se proporcionen sintonizaciones adaptadas a las diferentes circunstancias de la red. Además, se ha optado por utilizar el método de Quasi-Newton en la etapa de aprendizaje, ya que presenta tiempos de convergencia menores que el algoritmo de retropropagación [134, 135, 144].

Por tanto, el funcionamiento del sistema de red neuronal diseñado es el siguiente. En primer lugar se procede al entrenamiento de la red neuronal. Para ello es necesario fijar un número de iteraciones en el aprendizaje. Si el número de iteraciones es demasiado pequeño la red puede no alcanzar un buen aprendizaje, mientras que si el número es elevado, puede existir un sobreajuste de los datos de entrenamiento y la capacidad de generalización se puede ver reducida. Para el entrenamiento de la red neuronal se disponen de una serie de patrones que combinan los valores de los valores K_p , K_i y K_d de buenos individuos obtenidos con el algoritmo genético en distintos instantes y con condiciones de red diferentes (distinta carga de red, anchos de banda garantizados, etc.) junto con los errores cometidos con dichos valores en tres iteraciones consecutivas del PID. Estos individuos se han obtenido partiendo del algoritmo genético desarrollado para SPID en el Capítulo 5. De esta forma, cada patrón de entrenamiento incluye los tres valores de entrada de la red neuronal ($e[k]$, $e[k - 1]$ y $e[k - 2]$) y las salidas deseadas (K_p , K_i y K_d). Todos estos valores son escalados antes del entrenamiento en un rango entre $[-1, 1]$, para así poder calcular correctamente las salidas en función de las entradas de la red neuronal. El entrenamiento se realiza una única vez al comienzo de todo el proceso y tiene una duración de unos pocos segundos. Como resultado se obtiene un perceptrón multicapa con los pesos ajustados para ofrecer una sintonización del PID en función de los errores cometidos. A partir de aquí, comienza el funcionamiento normal del controlador PID. Para ello, inicialmente se fijan unos valores iniciales para K_p , K_i y K_d . El PID va a ajustar de forma dinámica los anchos de banda tal y como se describió en el Apartado 4.3 del Capítulo 4. La única diferencia, es que periódicamente (cada m iteraciones del PID), se va a ejecutar la red neuronal para proporcionar unos valores de sintonización más adecuados a las circunstancias de la red de modo que los parámetros de simulación se actualizan. Por ello, cada cierto tiempo (equivalente a ejecutar m iteraciones del PID) se recogen los tres últimos errores cometidos y la red neuronal, a partir de estas entradas, ofrece unos nuevos parámetros de sintonización (salidas de la red neuronal). Un diagrama del proceso completo puede verse en la Figura 6.5.

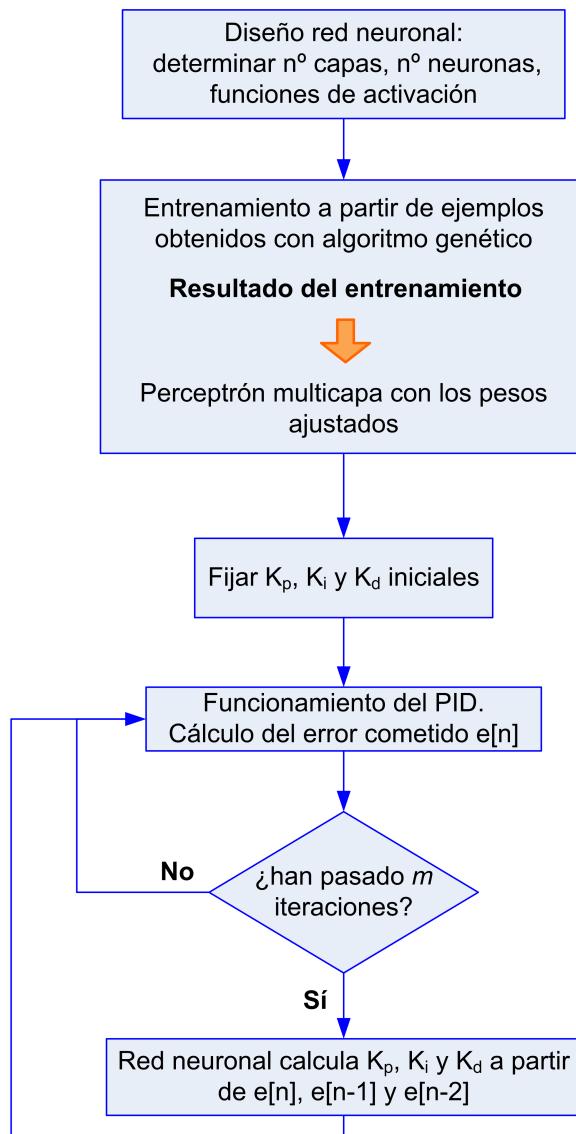


Figura 6.5: Diagrama de flujo del proceso de funcionamiento del algoritmo NN-SPID.

6.4. Estudio de simulación del algoritmo NN-SPID

En este apartado de la Tesis se va a mostrar el análisis de las prestaciones del algoritmo NN-SPID. Para ello, en primer lugar se determinarán los valores óptimos de los parámetros de diseño de la red neuronal y después se comprobará su comportamiento bajo distintos escenarios de red con condiciones iniciales y niveles de ancho de banda garantizado diferentes, comparando esta técnica de sintonización con las presentadas en capítulos anteriores.

6.4.1. Entorno de simulación

El entorno de simulación genérico en NN-SPID es el mismo que el considerado en SPID (Apartado 4.4.1 del Capítulo 4). La Tabla 6.1 recoge un resumen de los principales parámetros de red y los valores elegidos.

Parámetros de simulación	Valores
Número de ONUs	16 ONUs (N_{onus})
Tasa de transmisión canal ascendente	1 Gbit/s ($EPON_Line_rate$)
Tasa de transmisión ONU-abonado	100 Mbit/s ($User_Line_rate$)
RTT	1 ms
Tiempo de guarda	1 μ s
Tiempo de ciclo máximo	2 ms
Distribución de tráfico	Fuente de Pareto H=0.8
Tamaño de paquete	Variable entre 64-1500 bytes (más 38 bytes de cabeceras)
<i>Service Level Agreement (SLA)</i>	SLA ₀ : 1 ONU Ancho de banda garantizado: 100 Mbit/s
	SLA ₁ : 5 ONUs Ancho de banda garantizado: 75 Mbit/s
	SLA ₂ : 10 ONUs Ancho de banda garantizado: 50 Mbit/s
Tiempo muestreo del PID (T)	3 s
Configuración pesos iniciales	$W^{sla_0} = W^{sla_1} = W^{sla_2} = 1$

Tabla 6.1: Parámetros de red considerados en el entorno de simulación de NN-SPID.

Para la implementación de la red neuronal de NN-SPID en C++, se ha utilizado la librería *OpenNN* disponible en [133, 145]. Se trata de una librería genérica de redes neuronales donde se implementan diferentes métodos de entrenamiento (gradiente descendente, gradiente conjugado, método Quasi-Newton, etc.), así como diferentes funciones objetivo para el cálculo del error (error cuadrático medio, error cuadrático normalizado, error Minkowski, etc.). En concreto, en nuestro diseño se han utilizado las clases corres-

pondientes al perceptrón multicapa, junto con los métodos del error cuadrático medio y el método de entrenamiento Quasi-Newton.

6.4.2. Establecimiento de los parámetros asociados al diseño de la red neuronal en NN-SPID

En primer lugar se van a determinar los valores óptimos de los parámetros que rigen el comportamiento de la red neuronal. Algunos ya se han especificado en secciones anteriores, tales como el número de capas, el número de neuronas de entrada y salida, las funciones de activación y el método de entrenamiento. Sin embargo, para establecer otros parámetros propios del diseño de la red neuronal se ha optado por analizar su comportamiento a nivel de simulación en el sistema de red concreto en el que va a ser aplicada. En concreto, los parámetros a determinar de forma experimental serán el tiempo entre actualizaciones de los valores de sintonización, el número de iteraciones en la etapa de entrenamiento y el número de neuronas de la capa oculta. Para ello, en las siguientes secciones se presenta un estudio de simulación en el que se determinan los valores óptimos para dichos parámetros de la red neuronal.

6.4.2.1. Análisis del número de iteraciones del PID para actualizar los parámetros de sintonización

Tal y como se explicó en el apartado anterior, las variables de sintonización del PID (K_p , K_i y K_d), se modifican periódicamente cada cierto tiempo. Este tiempo se ha definido en función del número de iteraciones del PID, es decir, el número de veces que el PID actualiza los anchos de banda máximos para cada ONU en función de los errores cometidos. Por tanto este tiempo es igual a $m \cdot T$ donde T es el tiempo de muestreo que marca la ejecución del PID y m representa el número de iteraciones del PID. Para determinar el valor óptimo se han llevado a cabo simulaciones considerando un número de iteraciones igual a 3, 6, 9, 12, 15 y 20. Además, se ha fijado el número de neuronas de la capa oculta a cinco (apartado anterior) y el número de iteraciones en el entrenamiento a 75. Este valor se ha escogido como ejemplo, por representar un valor medio ni muy elevado ni muy pequeño. No obstante, su valor más adecuado se analizará en el siguiente apartado.

Para analizar el valor óptimo del número de iteraciones del PID, en la Tabla 6.2 se recoge la desviación estándar del ancho de banda medio asignado a cada ONU ($\overline{B_{alloc}^{onu_i}}$) sobre el nivel garantizado de su SLA ($B_{guarantee}^{sla \in onu_i}$) tras 10000 segundos de simulación. Se ha optado por analizar la desviación porque proporciona una idea de la exactitud del algoritmo a la hora de asegurar el nivel de ancho de banda estipulado, de forma que,

It.PID	Neur.	It.Entr.	Desviación de $\overline{B}_{alloc}^{onu_i}$ sobre $B_{guarantee}^{sla \in onu_i}$ (Mbit/s)			
			SLA ₀	SLA ₁	SLA ₂	Total
3	5	75	1.31	2.25	1.12	4.68
6	5	75	1.39	2.18	0.97	4.54
9	5	75	1.38	2.18	0.89	4.45
12	5	75	1.29	2.16	0.81	4.26
15	5	75	1.29	2.24	0.77	4.3
20	5	75	1.45	2.13	0.83	4.41

Tabla 6.2: Comparativa de la desviación estándar del término $\overline{B}_{alloc}^{onu_i}$ sobre el término $B_{guarantee}^{sla \in onu_i}$ considerando diferente número de iteraciones del PID.

a mayor desviación obtenida menor precisión del algoritmo. Como puede observarse, la menor desviación se produce para un número de iteraciones igual a 12, y por tanto este será el número de iteraciones seleccionado. No obstante, las diferencias entre los diferentes resultados no son muy grandes. Sin embargo, es necesario escoger un valor intermedio, ya que para un número de iteraciones muy bajo casi no se dejarían actuar las nuevas constantes de sintonización, mientras que para un número alto de iteraciones, el tiempo de resintonización sería demasiado elevado y el controlador PID podría no reaccionar adecuadamente si los cambios en la red se producen con mucha frecuencia.

6.4.2.2. Análisis del número de iteraciones en la etapa de entrenamiento

Para determinar el número de iteraciones óptimo en la etapa de entrenamiento, se han realizado simulaciones fijando este parámetro a 25, 50, 75, 100, 500 y 1000 iteraciones. Al igual que en el caso anterior, se va a comparar la desviación del término $\overline{B}_{alloc}^{onu_i}$ sobre el término $B_{guarantee}^{sla \in onu_i}$ en los tres SLAs considerados. En concreto, en la Tabla 6.3 se muestran los resultados obtenidos para cada perfil de abonado, cuando se considera un número de neuronas en la capa oculta igual a cinco y un número de iteraciones del

It.PID	Neur.	It.Entr.	Desviación de $\overline{B}_{alloc}^{onu_i}$ sobre $B_{guarantee}^{sla \in onu_i}$ (Mbit/s)			
			SLA ₀	SLA ₁	SLA ₂	Total
12	5	25	1.57	2.38	1.11	5.06
12	5	50	1.79	2.34	0.97	5.1
12	5	75	1.29	2.16	0.81	4.26
12	5	100	1.29	2.19	0.8	4.28
12	5	500	1.29	2.33	1.27	4.89
12	5	1000	1.33	2.75	1.27	5.35

Tabla 6.3: Comparativa de la desviación estándar del término $\overline{B}_{alloc}^{onu_i}$ sobre el término $B_{guarantee}^{sla \in onu_i}$ considerando diferente número de iteraciones en la etapa de entrenamiento.

PID igual a 12, por ser este el mejor valor obtenido en el apartado anterior. Como puede observarse, las menores desviaciones se obtienen para un entrenamiento con 75 iteraciones, por lo que éste será el número de iteraciones de la etapa de entrenamiento seleccionado. Sin embargo, aunque las diferencias entre las distintas iteraciones no son muy grandes, si el número de iteraciones es elevado el proceso se ralentiza demasiado. Por el contrario, si dicho número es muy pequeño no se va a lograr un correcto aprendizaje de la red neuronal. Por lo tanto, se fija un número intermedio de 75 iteraciones llegando así a una solución de compromiso adecuada.

6.4.2.3. Análisis del número de neuronas en la capa oculta

Una vez fijados los dos parámetros anteriores, se procederá a variar el número de neuronas de la capa oculta para ver la influencia de este parámetro en el comportamiento de la red neuronal. Para ello, se han llevado a cabo simulaciones considerando que en la capa oculta existían 3, 5, 6 y 8 neuronas, con un número de iteraciones del PID igual a 12 y considerando 75 iteraciones en la etapa de entrenamiento, ya que son los valores óptimos para ambos parámetros obtenidos en los apartados anteriores. En la Tabla 6.4 se presenta la comparativa de la desviación estándar del ancho de banda medio asignado sobre el garantizado para las distintas configuraciones. Como puede observarse, el valor óptimo de neuronas en la capa oculta es cinco. Además, se aprecia que para este sistema, la influencia del número de neuronas de la capa oculta en el comportamiento de la red es muy pequeña. No obstante, cuanto mayor sea el número de neuronas, mayor complejidad de la red neuronal.

			Desviación de $B_{alloc}^{onu_i}$ sobre $B_{guarantee}^{sla \in onu_i}$ (Mbit/s)			
It.PID	Neur.	It.Entr.	SLA ₀	SLA ₁	SLA ₂	Total
12	3	75	1.29	2.25	0.90	4.44
12	5	75	1.29	2.16	0.81	4.26
12	6	75	1.34	2.23	0.87	4.44
12	8	75	2.19	2.47	1.03	5.69

Tabla 6.4: Comparativa de la desviación estándar del término $\overline{B_{alloc}^{onu_i}}$ sobre el término $B_{guarantee}^{sla \in onu_i}$ considerando diferente número de neuronas en la capa oculta.

Finalmente, en la Tabla 6.5 aparece un resumen de los parámetros finales seleccionados en el diseño de la red neuronal incorporada al algoritmo NN-SPID.

Parámetros de diseño de la red neuronal en NN-SPID	
Número de capas	3 (entrada, capa oculta y salida)
Entradas	$e[n], e[n - 1], e[n - 2]$ (3 neuronas)
Salidas	K_p, K_i, K_d (3 neuronas)
Número de neuronas de la capa oculta	5
Funciones de activación	Capa de salida: función sigmoidal Capa oculta: función tangente hiperbólica
Método de entrenamiento	Quasi-Newton
Número de iteraciones en la etapa de entrenamiento	75
Función objetivo	Error cuadrático medio
Número de iteraciones del PID antes calcular nuevos valores de K_p, K_i y K_d	12

Tabla 6.5: Parámetros finales escogidos para el diseño de la red neuronal en NN-SPID.

6.4.3. Análisis de prestaciones de NN-SPID

En este apartado del capítulo se va a proceder a analizar las bondades de la sintonización adaptativa llevada a cabo por el algoritmo NN-SPID comparando su comportamiento con los métodos de sintonización presentados en capítulos anteriores, esto es, el método manual de Ziegler-Nichols y el método automático basado en algoritmos genéticos. En primer lugar, se va a comprobar la eficiencia de todas las técnicas de sintonización aplicadas al PID diseñado para garantizar diferentes niveles de ancho de banda, considerando el escenario genérico así como otros niveles garantizados. Finalmente, se analizará la influencia de los ejemplos utilizados en la etapa de entrenamiento o aprendizaje de la red neuronal en su comportamiento.

6.4.3.1. Evolución del ancho de banda a los niveles garantizados por el proveedor de servicios

En este primer apartado se va a estudiar la adaptación del ancho de banda máximo que realiza NN-SPID para garantizar el ancho de banda estipulado por el proveedor de servicios a cada uno de los perfiles de abonado soportados. En este sentido, las siguientes gráficas comparan la sintonización adaptativa llevada a cabo por la red neuronal respecto a la sintonización manual de Ziegler-Nichols, con una sintonización óptima realizada por el algoritmo genético para el escenario de simulación recogido en la Tabla 6.1 y con dos sintonizaciones no optimizadas utilizadas también en el capítulo anterior. En concreto, los valores de cada uno de los parámetros K_p, T_i y T_d para las distintas sintonizaciones están recogidos en la Tabla 6.6. Para facilitar una mejor comparación entre las distintas

Método de sintonización	Valores
Ziegler-Nichols	$K_p = 0.66, T_i = 11, T_d = 2.75$
Algoritmo Genético	$K_p = 0.33, T_i = 4.9, T_d = 4.5$
Sintonización no optimizada 1	$K_p = 2.1, T_i = 4.86, T_d = 0.3$
Sintonización no optimizada 2	$K_p = 4.3, T_i = 4.9, T_d = 0.15$
Sintonización redes neuronales	Valores variables dinámicamente

Tabla 6.6: Valores de los parámetros K_p , T_i y T_d seleccionados para cada tipo de sintonización comparada.

técnicas, a lo largo de este análisis se utilizarán los parámetros de sintonización K_p , T_i y T_d . Por tanto, para convertir los valores de sintonización de las redes neuronales K_i y K_d a los valores T_i y T_d se ha utilizado la ecuación (6.6) con un valor de T igual a 3 s, por ser el tiempo de muestreo escogido en SPID en el Capítulo 4.

La Figura 6.6 (a), (b) y (c) muestra la variación en tiempo real del ancho de banda medio asignado a una ONU del SLA₀, SLA₁ y SLA₂, respectivamente, para las sintonizaciones de la Tabla 6.6, bajo las condiciones del escenario genérico (Tabla 6.1) y para una carga de red de 1.6. Tal y como se puede apreciar, excepto para el SLA₀ (Figura 6.6 (a)), se observa un comportamiento diferente dependiendo del tipo de sintonización. En concreto, para el SLA₁ y el SLA₂ se puede ver que tanto Ziegler-Nichols como el algoritmo genético y las redes neuronales obtienen unos resultados parecidos respecto a la velocidad de convergencia y nivel de fluctuaciones del ancho de banda medio asignado sobre su nivel garantizado. Por el contrario, para las dos sintonizaciones no optimizadas se observan grandes oscilaciones en dicho parámetro para ambos SLAs, de forma que no se garantizan adecuadamente los niveles estipulados de ancho de banda para estos dos perfiles de abonado.

Respecto a las variaciones del ancho de banda máximo permitido para cada SLA, el comportamiento de todas las sintonizaciones se muestra en la Figura 6.7 (a), (b) y (c) para el SLA₀, SLA₁ y SLA₂, respectivamente. Tal y como se puede apreciar, para el SLA₁ y SLA₂, las sintonizaciones no optimizadas producen unas fluctuaciones muy grandes, mientras que en las otras tres sintonizaciones éstas son menores y con escasas diferencias entre ellas. Para el SLA más prioritario (SLA₀), las diferencias entre las distintas sintonizaciones son inexistentes.

Por último, la Figura 6.8 (a), (b) y (c) muestra la desviación estándar del ancho de banda medio asignado sobre el correspondiente ancho de banda garantizado para cada SLA. Tal y como se puede observar, los resultados corroboran los datos obtenidos anteriormente. Esto es, excepto para el SLA₀, para el resto de SLAs la desviación obtenida es mayor con las sintonizaciones no optimizadas, ya que éstas aproximan peor la asignación

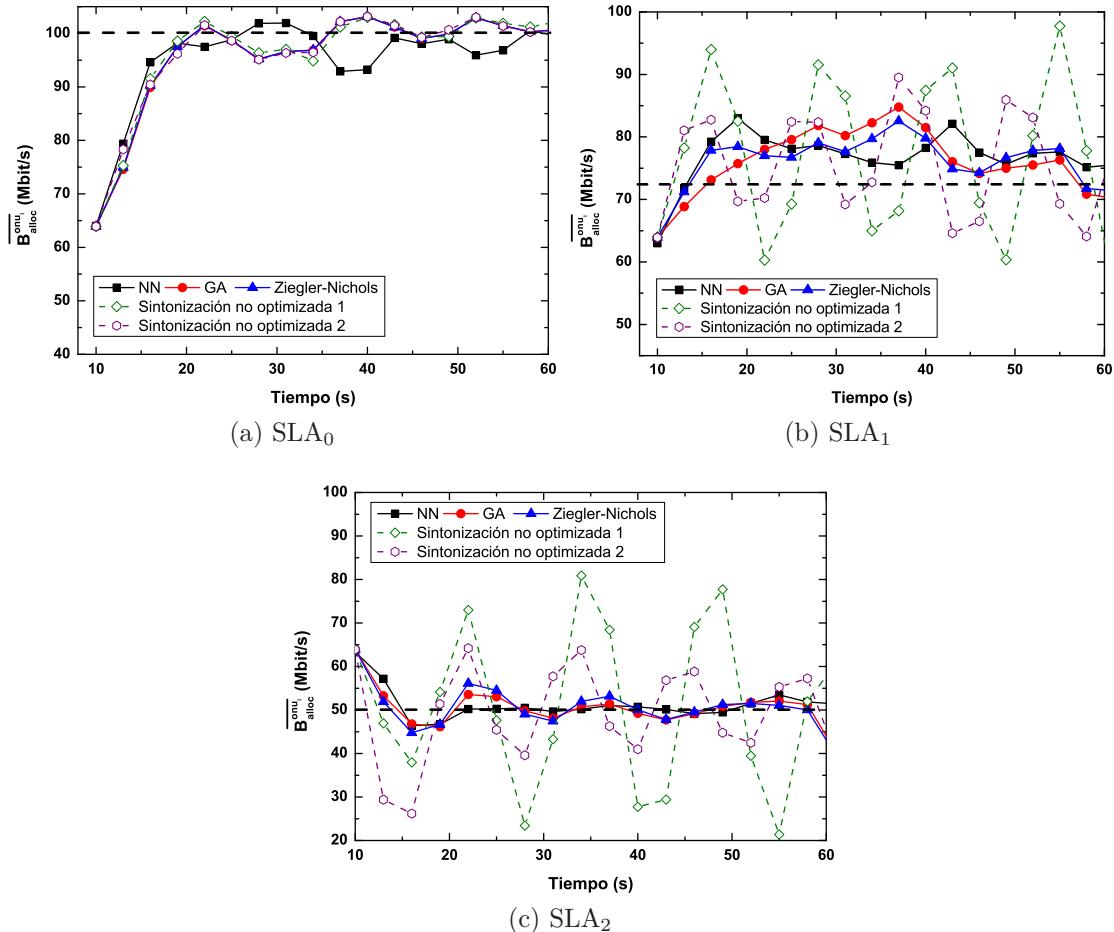


Figura 6.6: Evolución en tiempo real del ancho de banda medio asignado a una ONU de cada SLA para las sintonizaciones de la Tabla 6.6. (a) SLA₀ (b) SLA₁ (c) SLA₂.

del ancho de banda al nivel que deben garantizar.

Por tanto, se ha visto cómo la sintonización llevada a cabo por NN-SPID consigue evolucionar los anchos de banda máximos de forma que consigue garantizar el nivel de ancho de banda estipulado por el proveedor de servicios al igual que los métodos de Ziegler-Nichols y algoritmos genéticos. No obstante, cabe destacar que la sintonización llevada a cabo por las redes neuronales es adaptativa, mientras que en los otros dos métodos la sintonización se realiza de forma previa y los parámetros del controlador PID se mantienen fijos independientemente de posibles cambios en tiempo real que se produzcan en la red.

En el apartado siguiente se va a estudiar la evolución hacia distintos niveles garantizados comparando únicamente las sintonizaciones de Ziegler-Nichols, algoritmo genético

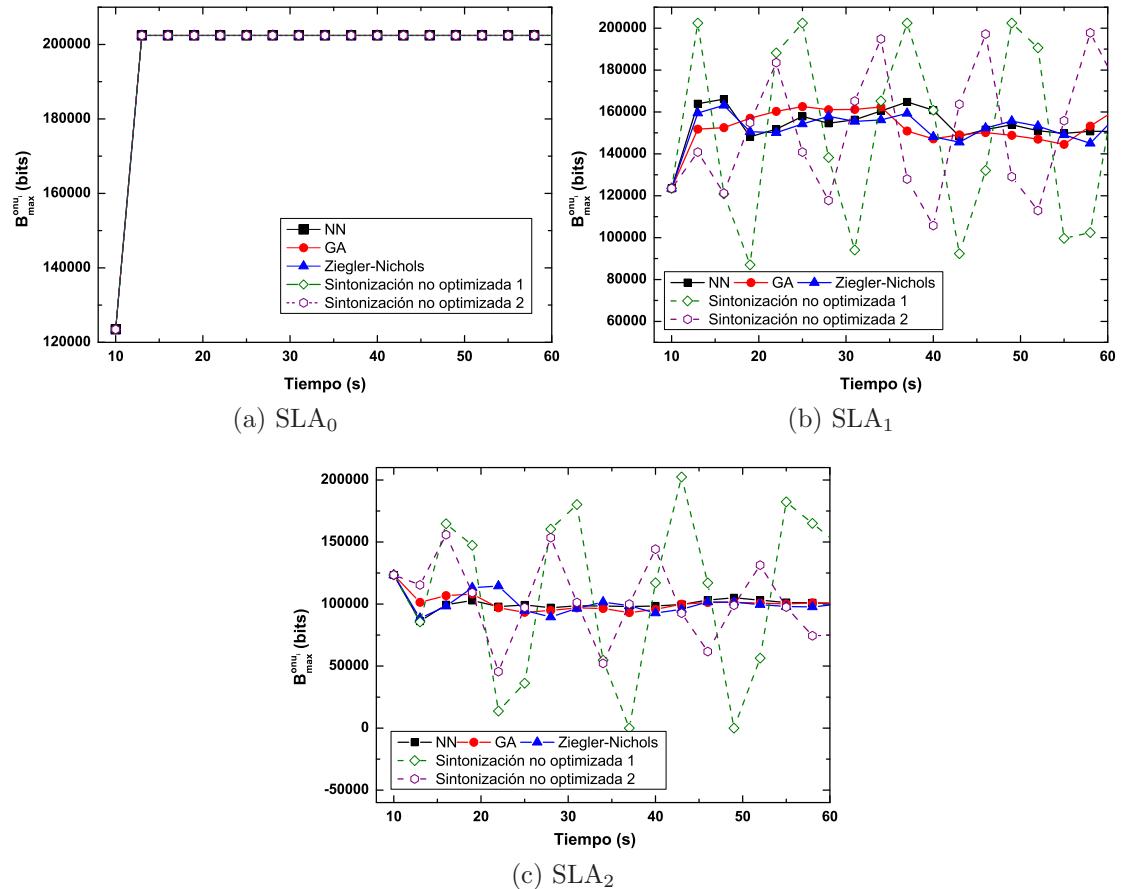


Figura 6.7: Evolución en tiempo real del ancho de banda máximo permitido a una ONU de cada SLA para las sintonizaciones de la Tabla 6.6. (a) SLA₀ (b) SLA₁ (c) SLA₂.

y redes neuronales para comprobar mejor las diferencias entre ellas. El análisis de las sintonizaciones no optimizadas se elimina, ya que ha quedado demostrado que una mala sintonización da lugar a un mal funcionamiento del controlador PID.

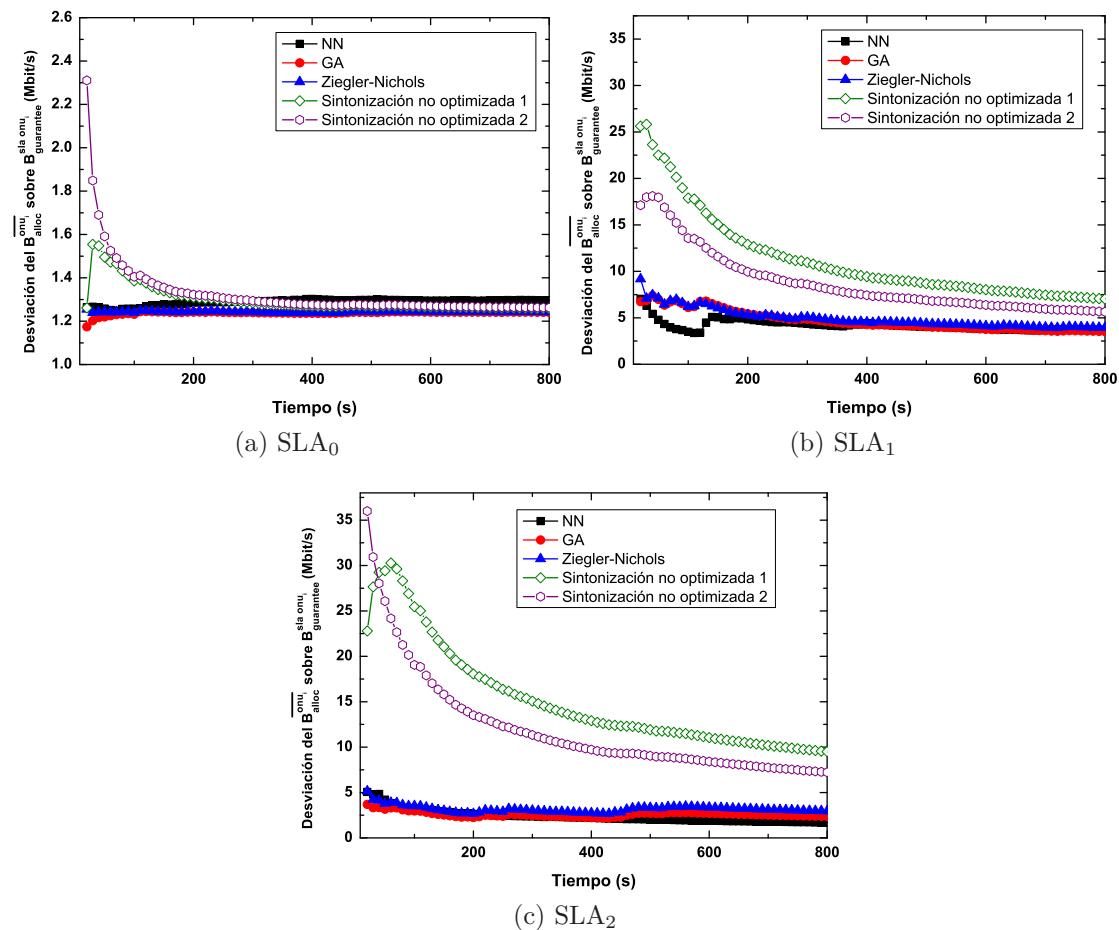


Figura 6.8: Evolución en tiempo real de la desviación estándar del ancho de banda medio asignado sobre el ancho de banda garantizado para las sintonizaciones de la Tabla 6.6. (a) SLA₀ (b) SLA₁ (c) SLA₂.

6.4.3.2. Independencia frente a los niveles de ancho de banda garantizado para cada SLA

NN-SPID tiene que ser capaz de controlar dinámicamente el ancho de banda asignado para garantizar los niveles de ancho de banda estipulados por el proveedor de servicios con independencia de los valores seleccionados. Por ello, se han llevado a cabo simulaciones en cuatro escenarios distintos cuyos valores de ancho de banda garantizado son diferentes. Los valores considerados aparecen recogidos en la Tabla 6.7. Además, para observar diferencias entre la sintonización realizada mediante redes neuronales (NN-SPID), algoritmos genéticos (GA-SPID) y el método de Ziegler-Nichols (SPID), se incluyen en las gráficas los resultados obtenidos por estos tres algoritmos. Para el caso concreto del algoritmo GA-SPID se han calculado individuos óptimos para cada uno de los escenarios de la Tabla 6.7.

Nivel de ancho de banda garantizado a cada SLA			
	SLA ₀	SLA ₁	SLA ₂
Escenario 1	100 Mbit/s	75 Mbit/s	50 Mbit/s
Escenario 2	70 Mbit/s	100 Mbit/s	40 Mbit/s
Escenario 3	90 Mbit/s	70 Mbit/s	55 Mbit/s
Escenario 4	60 Mbit/s	90 Mbit/s	45 Mbit/s

Tabla 6.7: Niveles de ancho de banda garantizado estipulados para cada SLA en los cuatro escenarios considerados.

En la Figura 6.9 se muestra la variación en tiempo real del ancho de banda máximo permitido para una ONU de cada SLA en los primeros 120 segundos de simulación cuando se compara NN-SPID con GA-SPID y SPID. Como puede observarse, para todos los escenarios, NN-SPID presenta una menor oscilación, aunque las diferencias con los otros dos algoritmos son pequeñas. Sin embargo, se observa cómo NN-SPID minimiza la mayoría de las fluctuaciones más abruptas que aparecen en los otros dos algoritmos. Por ejemplo, en el Escenario 1 los cambios bruscos que se producen alrededor de los 70 s en SPID y en GA-SPID tanto para el SLA₁ y SLA₂, no existen en NN-SPID. En el resto de escenarios también se pueden apreciar comportamientos similares. Este comportamiento más estable se consigue gracias a resintonización continua del PID mediante la red neuronal, de forma que se incrementa la robustez ante posibles cambios en la red.

En concreto, la Figura 6.10 muestra la variación en tiempo real de los parámetros de sintonización para los distintos escenarios de la Tabla 6.7 comparando los algoritmos NN-SPID y GA-SPID. Aunque los valores que devuelve la red neuronal son K_p , K_i y K_d , se van a representar los valores de K_p , T_i y T_d , para realizar una comparación directa con los parámetros obtenidos mediante los algoritmos genéticos. Para ello, se

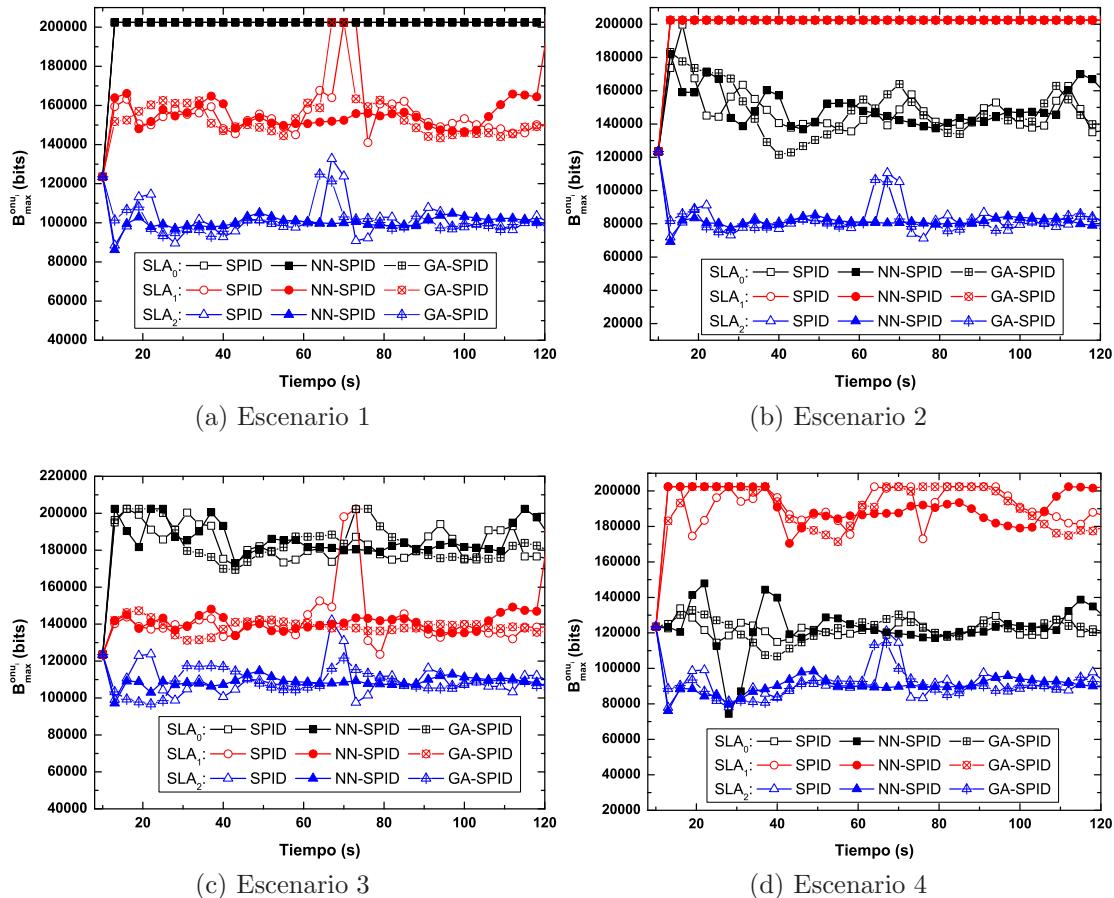


Figura 6.9: Variación en tiempo real del ancho de banda máximo para los escenarios de la Tabla 6.7 comparando SPID, GA-SPID y NN-SPID. (a) Escenario 1 (b) Escenario 2 (c) Escenario 3 (d) Escenario 4.

han utilizado las relaciones de la ecuación (6.6) con un valor de tiempo de muestreo (T) igual a 3 segundos. Por otro lado, los parámetros de sintonización de SPID no se han incluido para simplificar la visualización de las gráficas. Como se aprecia en todas las gráficas, en NN-SPID los valores de los parámetros varían a lo largo del tiempo para todos los escenarios, mientras que en el caso de GA-SPID éstos permanecen fijos. Este comportamiento permite una mayor adaptabilidad a cambios que se puedan producir en tiempo real dentro de la red. Además, tal y como se observa, en general los valores entre los que fluctúan los parámetros calculados por la red neuronal son bastante similares a los obtenidos por los individuos del algoritmo genético, ya que para este escenario no se han considerado cambios en tiempo real en los escenarios.

Por último, se va a proceder a analizar la desviación estándar del ancho de banda

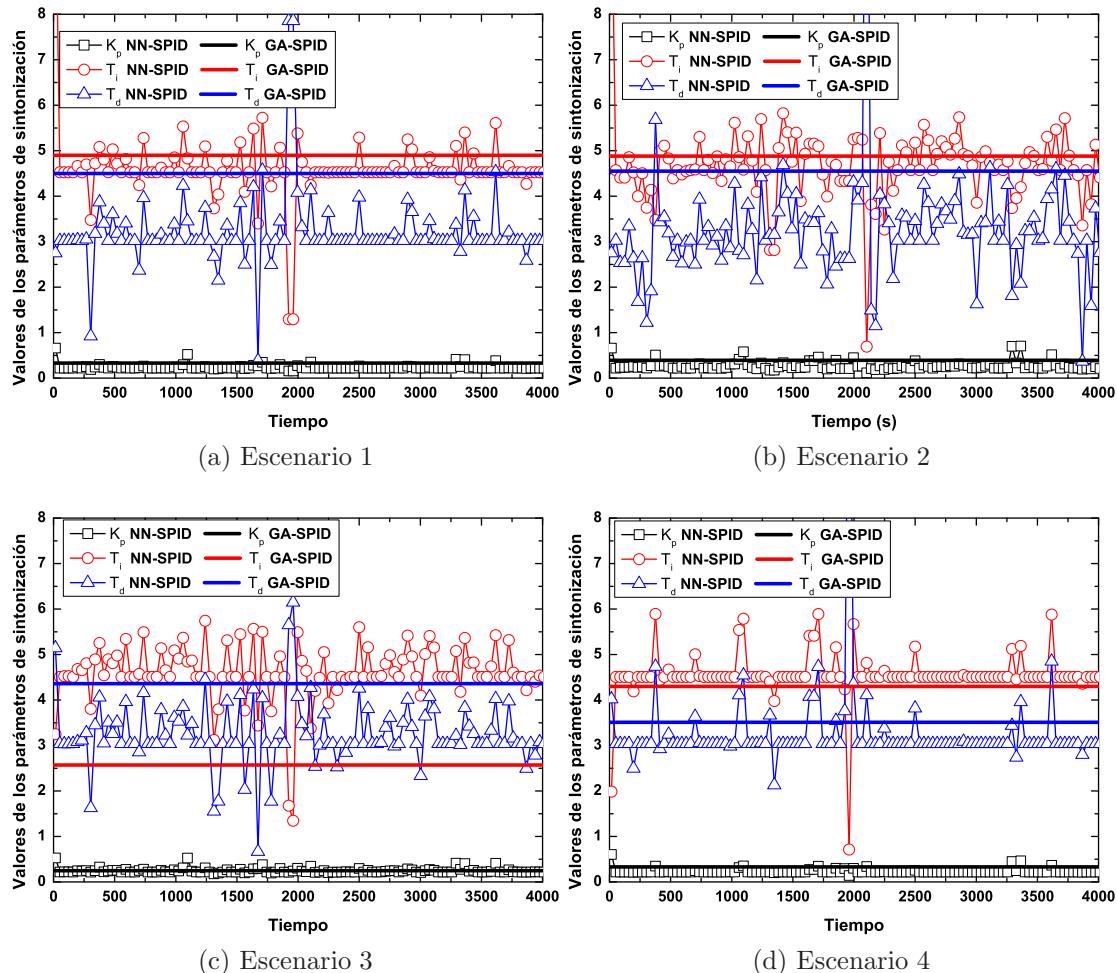


Figura 6.10: Evolución en tiempo real de los parámetros de sintonización K_p , T_i y T_d para los cuatro escenarios considerados en la Tabla 6.7 comparando NN-SPID y GA-SPID. (a) Escenario 1 (b) Escenario 2 (c) Escenario 3 (d) Escenario 4.

medio asignado a las ONUs sobre su nivel garantizado, para, de esta forma medir cuánto de bueno es el ajuste que realiza el algoritmo para alcanzar los niveles garantizados en cada uno de los cuatro escenarios considerados. En este sentido, en la Figura 6.11 (a), (b), (c) y (d) está representada dicha desviación para los tres SLAs considerados y para los cuatro escenarios de la Tabla 6.7, respectivamente, al comparar SPID, GA-SPID y NN-SPID. Tal y como se puede observar en la Figura 6.11 (a), (b), (c) y (d), en general la red neuronal consigue reducir el nivel de la desviación estándar de SPID y GA-SPID. Esta reducción es más apreciable para el SLA₂ en todos los escenarios de red considerados.

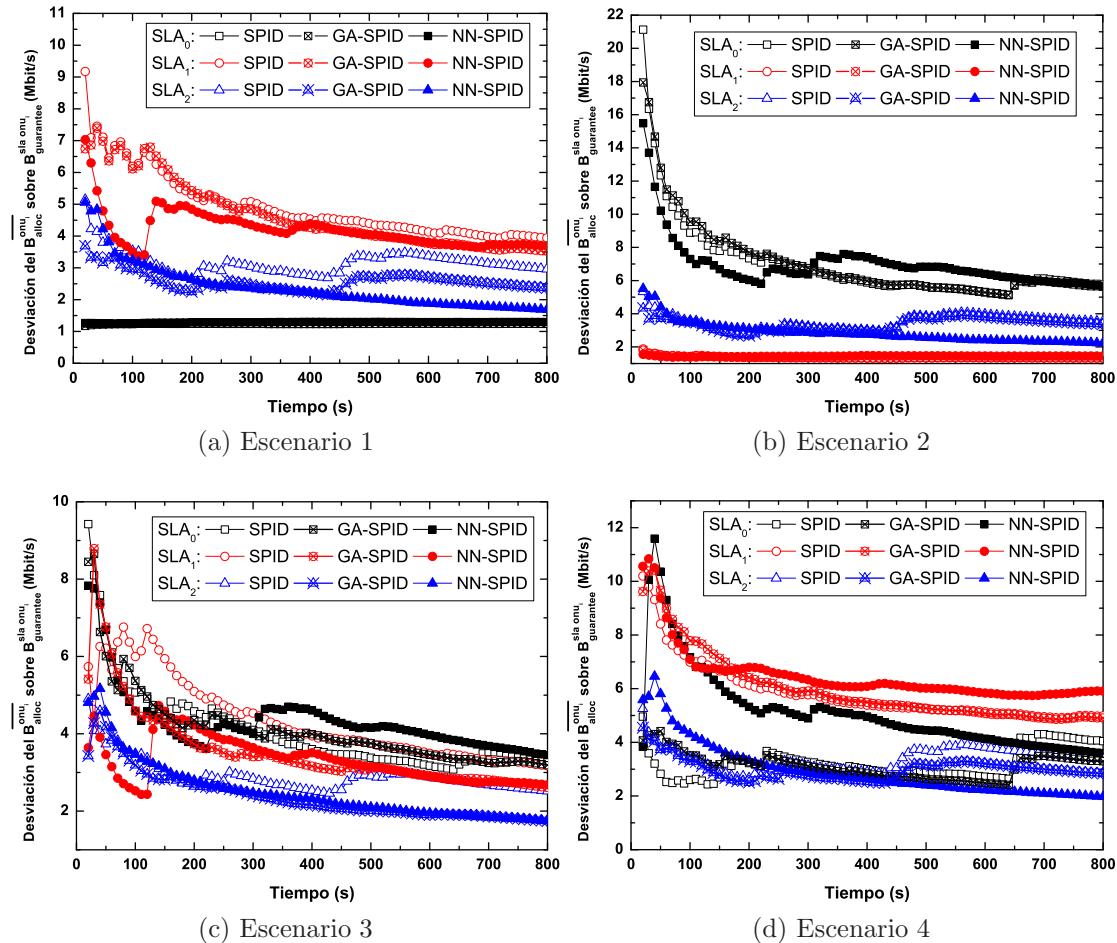


Figura 6.11: Evolución de la desviación estándar en tiempo real del ancho de banda medio asignado sobre el ancho de banda garantizado para los cuatro escenarios considerados en la Tabla 6.7 comparando SPID, GA-SPID y NN-SPID. (a) Escenario 1 (b) Escenario 2 (c) Escenario 3 (d) Escenario 4.

6.4.3.3. Dependencia del comportamiento de NN-SPID con los patrones utilizados en el entrenamiento de la red neuronal

En este apartado del capítulo se estudia el comportamiento de la red neuronal en función del tipo de ejemplos empleados en la etapa de entrenamiento. Obviamente, si los ejemplos de los que aprende la red neuronal son incorrectos o no son suficientes, ésta extraerá unas características erróneas que harán que la red no sintonice de forma correcta el controlador PID.

En primer lugar, se va a estudiar la influencia del número de ejemplos utilizados en la etapa de entrenamiento en el comportamiento de la red neuronal. Para ello, se ha reducido el número de ejemplos un 50 % y un 25 % del total de casos utilizados en el entrenamiento a lo largo de este capítulo. Esto es, de los 16 ejemplos utilizados en el entrenamiento, se van a realizar simulaciones considerando únicamente 8 y 12 ejemplos. Los resultados de la desviación estándar del ancho de banda medio asignado a cada ONU frente al nivel de ancho de banda garantizado, tras simular 10000 s bajo las condiciones del escenario de simulación genérico de la Tabla 6.1 y los parámetros de diseño de la red neuronal recogidos en la Tabla 6.5, se muestran en la Tabla 6.8.

Num. ejemplos	Desviación de $B_{alloc}^{onu_i}$ sobre $B_{guarantee}^{sla \in onu_i}$ (Mbit/s)			
	SLA ₀	SLA ₁	SLA ₂	Total
8	4.95	1.46	2.11	8.52
12	1.79	2.47	1.305	5.56
16	1.29	2.16	0.81	4.26

Tabla 6.8: Comparativa de la desviación estándar del término $\overline{B_{alloc}^{onu_i}}$ sobre el término $B_{guarantee}^{sla \in onu_i}$ considerando diferente número de ejemplos en el entrenamiento de la red neuronal.

Tal y como puede verse en la Tabla 6.8 si los ejemplos considerados no son suficientes (por ejemplo para 8 patrones), la precisión de la red neuronal empeora. Por tanto, una desventaja de este diseño de red neuronal basado en un aprendizaje supervisado, implica que es necesario conocer un número suficiente de ejemplos en la etapa de entrenamiento. Además, estos ejemplos tienen que ser correctos para que la red neuronal aprenda adecuadamente. En este sentido, en este capítulo de la Tesis se ha partido de ejemplos de sintonizaciones que el algoritmo genético ha propuesto como buenas soluciones para distintos escenarios y situaciones de red.

No obstante, para comprobar el funcionamiento de la red cuando se ha realizado el aprendizaje con una serie de ejemplos erróneos, se ha llevado a cabo una simulación donde se utilizan sintonizaciones que el algoritmo genético consideró incorrectas. En este sentido, la Figura 6.12 (a) y (b) muestra la evolución en tiempo real del ancho

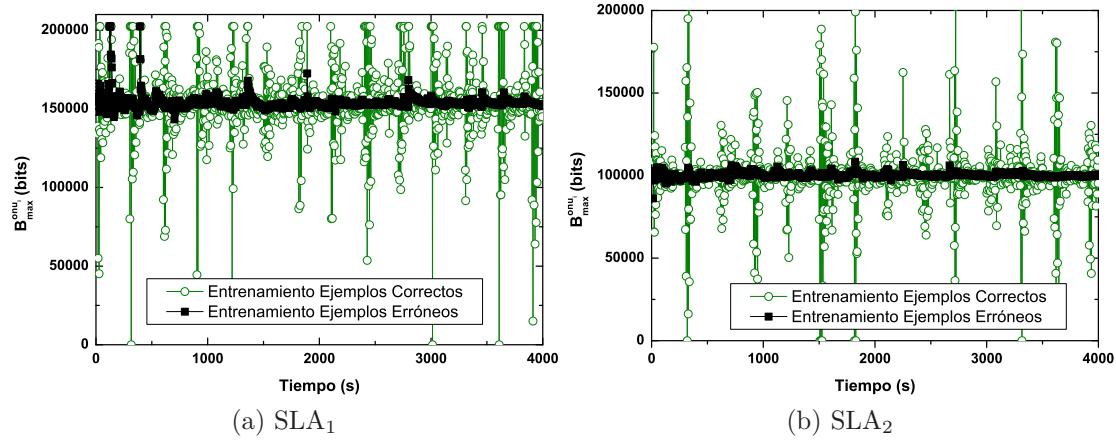


Figura 6.12: Evolución en tiempo real del ancho de banda máximo cuando el entrenamiento de la red neuronal se realiza con ejemplos correctos y erróneos. (a) SLA₁ (b) SLA₂.

de banda máximo permitido para el SLA₁ y SLA₂, respectivamente, cuando se realiza un entrenamiento con ejemplos correctos y erróneos bajo las condiciones del escenario genérico de la Tabla 6.1. Tal y como se puede apreciar, al utilizar un entrenamiento con ejemplos erróneos, el ancho de banda máximo nunca llega a estabilizarse, existiendo para ambos SLAs grandes fluctuaciones. Por el contrario, cuando el aprendizaje se realiza con ejemplos correctos el ancho de banda máximo se estabiliza, tal y como se ha visto a lo largo de este capítulo.

Así mismo, la Figura 6.13 muestra la evolución dinámica de los parámetros K_p , T_i y T_d a lo largo del tiempo obtenidos utilizando un entrenamiento con ejemplos no adecuados al objetivo deseado. En este sentido, puede verse cómo los parámetros de sintonización utilizados difieren completamente de los obtenidos entrenando con sintonizaciones correctas. En concreto, puede verse cómo el valor de K_p y T_d se ha incrementado, mientras que el valor de T_i se ha reducido.

Por tanto, se puede concluir que tanto el número de ejemplos en la fase de entrenamiento, como la adecuación de los mismos a los objetivos deseados, tienen una gran influencia. En nuestro caso se ha partido de individuos buenos seleccionados con el algoritmo genético. Por tanto, esto implica un procesamiento previo, donde se obtengan ejemplos correctos de sintonizaciones para poder llevar a cabo el entrenamiento de la red neuronal y realizar la sintonización adaptativa.

No obstante, una buena solución para mitigar esta influencia de los ejemplos de la fase de entrenamiento en el comportamiento de la red neuronal, sería incluir un algoritmo de auto-aprendizaje continuo. Así pues, partiendo del entrenamiento previo realizado

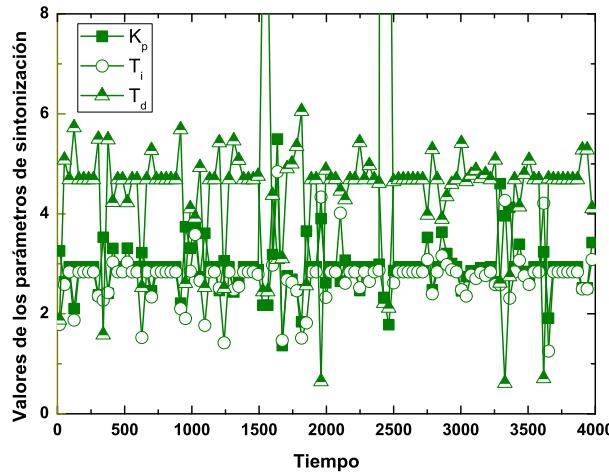


Figura 6.13: Evolución en tiempo real de los parámetros de sintonización K_p , T_i y T_d cuando la etapa de entrenamiento se realiza con ejemplos erróneos.

en este capítulo para obtener unos pesos iniciales para las relaciones entre las neuronas adecuados, éstos se modificarían a lo largo del tiempo con el fin de que los parámetros de sintonización K_p , K_i y K_d proporcionados como salida de la red neuronal minimicen el error cometido por el PID. Esto es, minimizando la diferencia entre el valor de referencia de salida deseado y el valor real obtenido, que en el caso concreto de SPID sería minimizar la diferencia entre el ancho de banda garantizado y el ancho de banda medio asignado. De esta forma, si se producen cambios de red muy diferentes a los contemplados en la fase de entrenamiento, los parámetros de sintonización se modificarían de forma óptima para que el error cometido se minimizase. Además, con este algoritmo de auto-aprendizaje se podría prescindir de la fase de entrenamiento previa. De este modo, en el caso de no disponer de patrones para poder entrenar a la red, se podría aplicar este método de auto-aprendizaje con unos pesos iniciales aleatorios para que evolucionasen hacia los pesos óptimos. Sin embargo, en este caso la red neuronal tardaría un tiempo en proporcionar unos valores de sintonización adecuados.

Para la adaptación en tiempo real de los pesos, se ejecutaría el algoritmo de retro-propagación periódicamente y se definiría una función objetivo relacionada con el error cometido por el PID al ajustar la señal de salida al valor de referencia. De esta forma, los pesos se modificarían siguiendo la dirección negativa del gradiente para conseguir minimizar dicho error a lo largo del tiempo y acorde a los cambios producidos en el escenario de red.

No obstante, en esta Tesis no se aborda la implementación de este algoritmo de auto-aprendizaje, considerándolo una línea futura de investigación.

6.5. Análisis de prestaciones de NN-SPID bajo un patrón de tráfico dinámico

Debido a la naturaleza adaptativa del algoritmo NN-SPID que modifica los parámetros de sintonización a lo largo del tiempo, se va a analizar su comportamiento bajo un patrón de tráfico dinámico. En este sentido, se ha considerado una variación de la carga de cada ONU a lo largo del tiempo igual a la utilizada en el Capítulo 3 y en el Capítulo 4, dada por la ecuación (3.9).

En concreto, la Figura 6.14 (a), (b) y (c) representa la variación del ancho de banda medio asignado a lo largo del tiempo para una ONU del SLA₀, SLA₁ y SLA₂, respectivamente, cuando la frecuencia de los cambios en la carga es de 1800 s. En azul está representado la variación en la carga de ONU en Mbit/s referida al eje de la de-

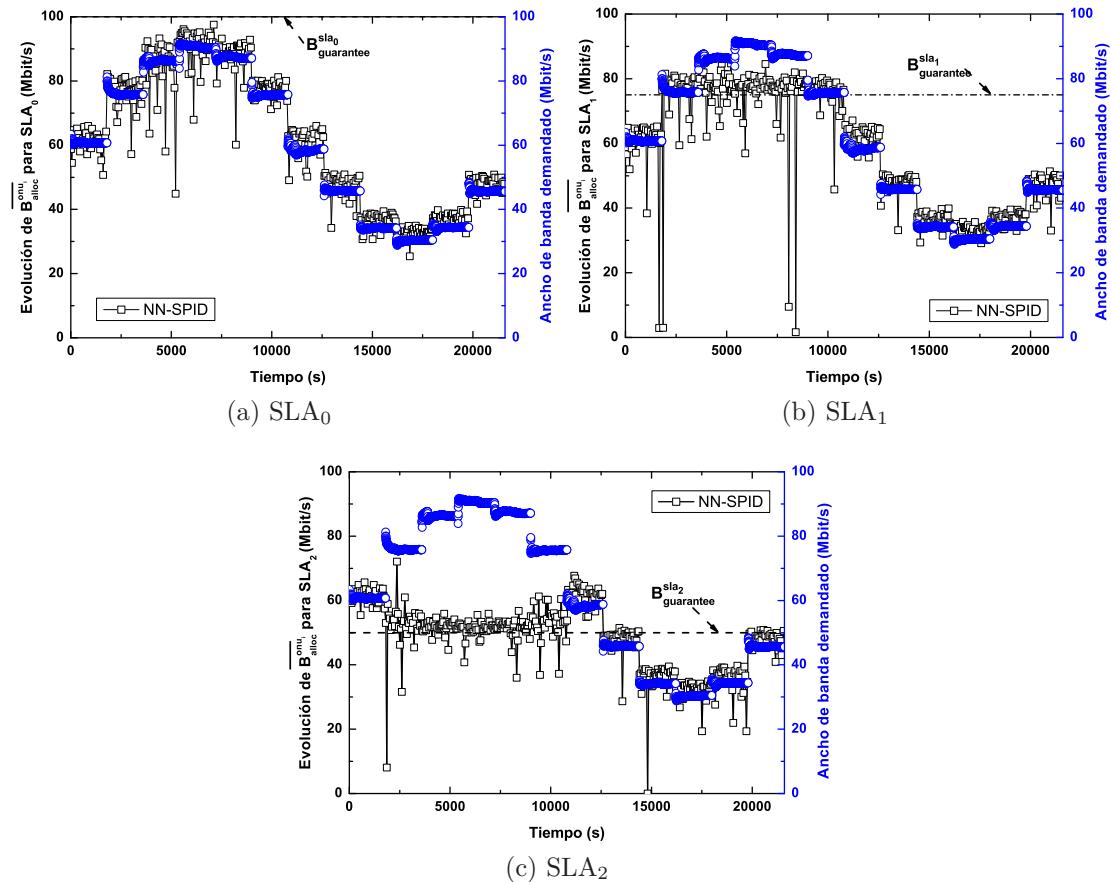


Figura 6.14: Evolución en tiempo real del ancho de banda demandado para una ONU de cada SLA (azul) y variación en tiempo real del ancho de banda medio asignado (negro) considerando una frecuencia de cambio de 1800 s. (a) SLA₀ (b) SLA₁ (c) SLA₂.

recha también en azul (de 0 a 100 Mbit/s). En negro aparece la variación en tiempo real del ancho de banda asignado en NN-SPID (eje de la izquierda). Tal y como puede observarse, para los tres SLAs el ancho de banda medio asignado se corresponde con el ancho de banda demandado, siempre que existan recursos para satisfacer la demanda total. En el caso de que la demanda total del ancho de banda sea mayor que la capacidad del canal, NN-SPID asigna de forma adecuada el nivel mínimo garantizado a cada SLA (100, 75 y 50 Mbit/s respectivamente).

En la Figura 6.15 se representa la variación de los parámetros de sintonización K_p , T_i y T_d a lo largo del tiempo cuando se consideran cambios de la carga cada 1800 s. Así mismo, referido al eje de la derecha en azul, se muestra la variación en tiempo real del ancho de banda demandado por una ONU. Para facilitar la visualización de los datos, solo se han representado 6000 s de simulación. Tal y como se aprecia, los valores de los distintos parámetros varían a lo largo del tiempo para ofrecer una sintonización óptima en cada momento según la carga de red actual.

Por otro lado, para analizar la adaptación de NN-SPID ante cambios de carga más rápidos, la Figura 6.16 (a), (b) y (c) representa la variación del ancho de banda medio asignado (negro) según la variación del ancho de banda demandado por una ONU (azul) cuando se consideran cambios en la carga cada 900 s para el SLA₀, SLA₁ y SLA₂, respectivamente. Tal y como se aprecia, NN-SPID adapta de forma óptima el ancho de banda medio asignado al nivel garantizado por el proveedor de servicios para cada SLA cuando no se puede ofrecer todo el ancho de banda demandado.

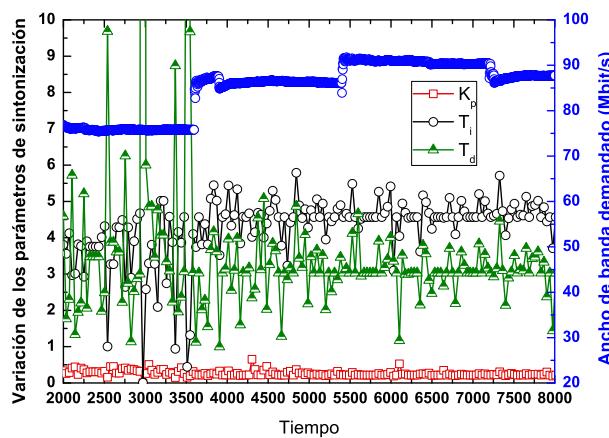


Figura 6.15: Evolución en tiempo real del ancho de banda demandado por cada ONU (azul) y variación en tiempo real de los parámetros de sintonización K_p , T_i y T_d considerando una frecuencia de cambio de 1800 s.

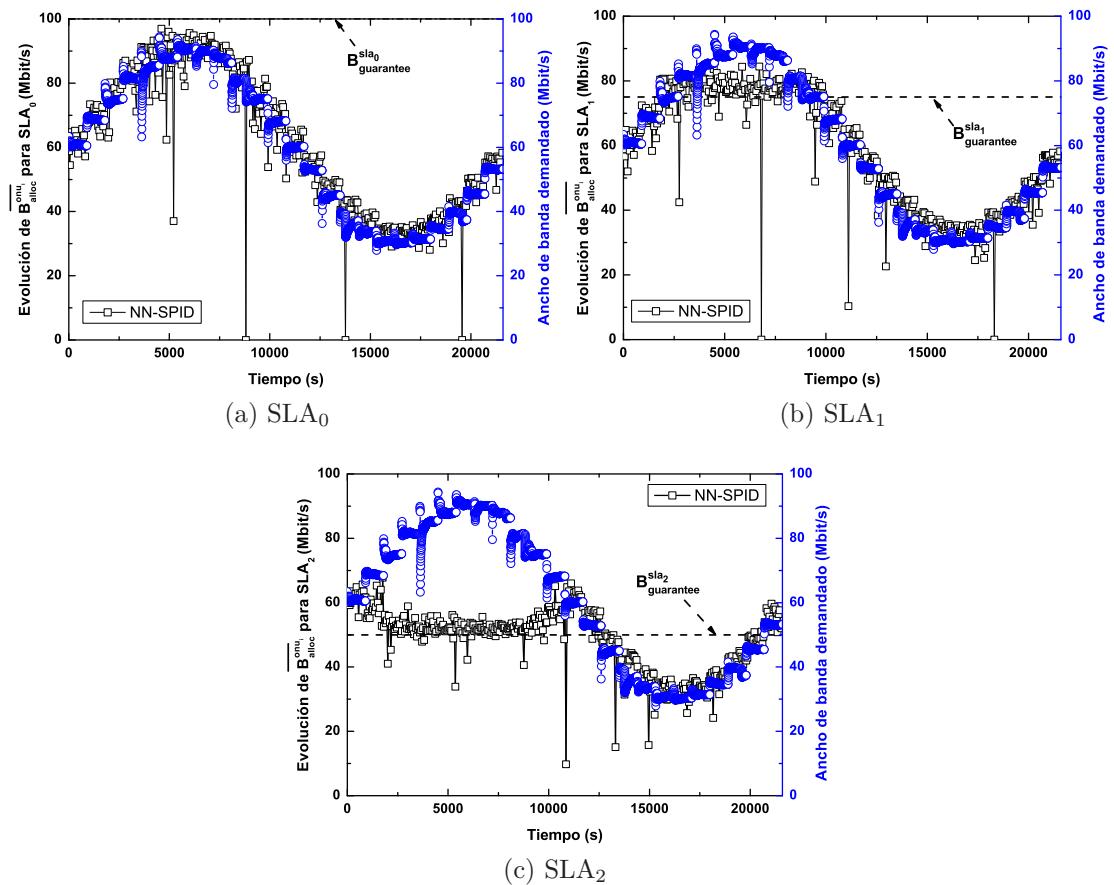


Figura 6.16: Evolución en tiempo real del ancho de banda demandado para una ONU de cada SLA (azul) y variación en tiempo real del ancho de banda medio asignado (negro) considerando una frecuencia de cambio de 900 s. (a) SLA₀ (b) SLA₁ (c) SLA₂.

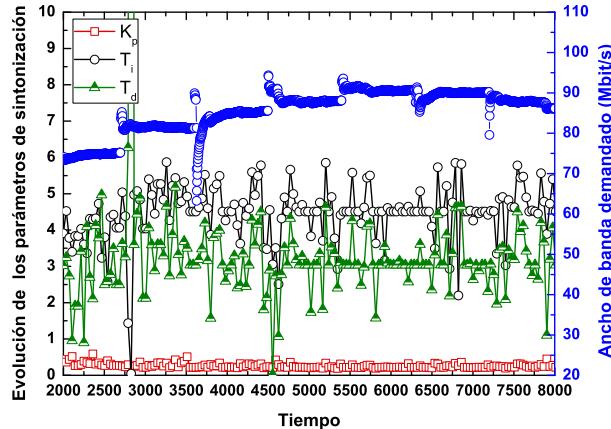


Figura 6.17: Ancho de banda demandado por cada ONU (azul) y variación en tiempo real de los parámetros de sintonización K_p , T_i y T_d considerando una frecuencia de cambio de 900 s.

Así mismo, en la Figura 6.17 donde está representada la variación de los parámetros de sintonización con el ancho de banda demandado, se puede apreciar una modificación en tiempo real de las tres variables similar a la obtenida para una frecuencia de cambio de 1800 s.

6.6. Conclusiones

En este capítulo de la Tesis se ha presentado el diseño e implementación de un método para realizar el proceso de sintonización adaptativa de controladores PID basados en el uso de redes neuronales. En concreto, este nuevo método, que nunca antes se había utilizado en la gestión de recursos en redes de acceso EPON y LR-EPON, se ha aplicado a la sintonización adaptativa del algoritmo SPID que controla el nivel de ancho de banda. No obstante, su integración en controladores PID que gestionen otros parámetros de red, como el caso del retardo medio llevado a cabo en DaSPID, se puede realizar de forma directa. El nuevo algoritmo, denominado NN-SPID, consta de una red neuronal que, a partir de los errores cometidos en la asignación de ancho de banda a lo largo del tiempo, ofrece unas variables de sintonización óptimas para cada situación de la red en tiempo real.

Para ello, en el primer apartado de este capítulo se presentaron los conceptos básicos sobre redes de neuronas artificiales así como los principales métodos de aprendizaje y modelos neuronales, haciendo especial hincapié en aquellos aplicables al perceptrón multicapa, por ser uno de los modelos más utilizados. A continuación, se diseñó la red neuronal para nuestro sistema, fijando la topología de red, el algoritmo de aprendizaje,

la función de error y las funciones de activación. En concreto, se optó por una topología con tres capas, con tres neuronas en la capa de entrada, tres en la capa de salida y cinco en la capa oculta. Posteriormente, se validó el sistema llevando a cabo un análisis de simulación profundo. En concreto, se estudió el comportamiento del sistema modificando distintos parámetros de los que dependía el diseño de la red neuronal del sistema con el fin de obtener los valores óptimos. En particular, se parametrizó el número de iteraciones de la etapa de entrenamiento, el número de iteraciones del PID tras las cuales se ejecuta la red neuronal para realizar una nueva sintonización y el número de neuronas de la capa oculta.

Una vez determinados dichos parámetros se llevó a cabo un estudio de simulación para analizar las prestaciones del algoritmo NN-SPID en diferentes entornos de red. En concreto, se estudió la gestión del ancho de banda realizada para diferentes niveles de ancho de banda garantizado. Los resultados demostraron que NN-SPID, al igual que ocurría con SPID y GA-SPID, es capaz de evolucionar el ancho de banda máximo permitido a cada ONU para garantizar el ancho de banda independientemente de los valores especificados por el proveedor de servicios. Sin embargo, se observó que NN-SPID conseguía una respuesta más estable que SPID y GA-SPID, lo que implica una adaptación más precisa a los niveles garantizados. Esto se consigue gracias a la modificación adaptativa de los parámetros de sintonización del controlador PID, ya que NN-SPID calcula dinámicamente una sintonización óptima en función de las condiciones actuales de red, dando lugar a un controlador más estable y robusto ante posibles cambios en tiempo real.

Por otro lado, se estudió la dependencia del funcionamiento de la red neuronal con el tipo de ejemplos utilizados en la fase de aprendizaje. Como era de esperar, si los ejemplos de los que aprende son erróneos o no son suficientes, el funcionamiento puede no llegar a ser el adecuado. Por tanto, un inconveniente de esta técnica es la necesidad de disponer ejemplos buenos para la fase de entrenamiento. En este sentido, una cuestión a analizar como línea futura sería diseñar una red que autoaprendiese en función de los errores cometidos a lo largo del tiempo.

Finalmente, se analizó la adaptación de NN-SPID bajo escenarios con patrones de tráfico dinámico más realista. En este sentido, se comprobó que el algoritmo diseñado es capaz de adaptar los parámetros de sintonización para conseguir una asignación de ancho de banda óptima y acorde a la carga de la red.

Capítulo 7

Técnicas cognitivas para optimizar la gestión de recursos en redes ópticas troncales y de acceso

7.1. Introducción

Las redes de telecomunicaciones actuales, tal y como se ha comentado en capítulos anteriores, tienen que transportar tráfico de distintos servicios con requisitos diversos y cada vez más restrictivos. Por ello es necesario realizar una gestión eficiente de los recursos disponibles en todos los tramos de la red para poder garantizar la calidad de servicio estipulada para el usuario final.

Por otro lado, está surgiendo un gran interés en la introducción de cognición en las redes de telecomunicaciones. Aunque en un principio la cognición se aplicó solo a las redes inalámbricas (radio cognitiva), en la actualidad existen diversos estudios en la literatura que incorporan cognición en las redes de comunicaciones ópticas [146–148]. La introducción de cognición en una red permite que ésta sea capaz de percibir las condiciones de red en cada momento y planificar y actuar dependiendo de dichas condiciones [149]. Además, la red puede aprender de experiencias pasadas y utilizarlas en la toma de decisiones futuras. De esta forma se dota a la red de cierta inteligencia que se va a ir incrementando para una mejor adaptación y toma de decisiones. Así, la cognición puede ser usada para mejorar la gestión de los recursos, conseguir una calidad de servicio adecuada u otros objetivos relacionados bajo condiciones de red variables.

Por tanto, en este capítulo se presentará el diseño y la implementación de técnicas cognitivas que optimizan la gestión de los recursos en redes ópticas, tanto en las redes troncales como en las de acceso, con el fin de llevar a cabo una mejor gestión de los recursos disponibles para garantizar de manera eficiente la calidad de servicio requerida. La técnica cognitiva que se ha utilizado en esta Tesis es el Razonamiento Basado en Casos (CBR, *Case-Based Reasoning*). Esta técnica de inteligencia artificial se basa en la utilización del conocimiento adquirido en experiencias previas para tomar decisiones sobre un caso particular. En concreto, este mecanismo se ha utilizado en dos ámbitos diferentes de las redes ópticas. En la primera parte del capítulo, se describe cómo se aplicará esta técnica a la predicción de la calidad de transmisión en las redes ópticas troncales, mientras que en la segunda parte del capítulo se describirá cómo se utiliza el CBR en la sintonización adaptativa de un controlador PID en redes de acceso LR-EPON.

La organización de este capítulo es la siguiente. En primer lugar se explica la técnica de razonamiento basado en casos utilizada en el diseño de las técnicas cognitivas desarrolladas en la Tesis. Tras ello, se presenta una breve introducción a las redes troncales ópticas actuales, para pasar después a la descripción del estimador de calidad cognitivo diseñado e implementado en estas redes. A continuación, se analizan las prestaciones del estimador desarrollado. En la segunda parte del capítulo, se aplica la técnica de razonamiento basado en casos para llevar a cabo la sintonización adaptativa de controladores PIDs en redes LR-EPON, tomando como ejemplo de implementación el algoritmo SPID. Tras la descripción del algoritmo se evalúan sus prestaciones y comportamiento.

7.2. Introducción al Razonamiento Basado en Casos (CBR)

El razonamiento basado en casos es un paradigma de resolución de problemas que, a diferencia de otras técnicas de inteligencia artificial, no solo confía en el conocimiento general del dominio del problema sino que también utiliza el conocimiento específico de experiencias previas, es decir, situaciones de un problema concreto (caso) almacenado en una base de conocimiento (KB, *Knowledge Base*). Además, esta base de conocimiento puede ser actualizada para incorporar nuevas experiencias y por tanto conseguir un aprendizaje continuado [150–152].

En el razonamiento basado en casos, un nuevo problema se resuelve siguiendo las siguientes cuatro fases (Figura 7.1):

- Recuperando el caso o casos pasados más similares al nuevo problema. Esto es, retomar de la base de conocimiento la experiencia de un problema que se ha tratado en el pasado y que es similar al que se quiere resolver. A esta fase se la denomina

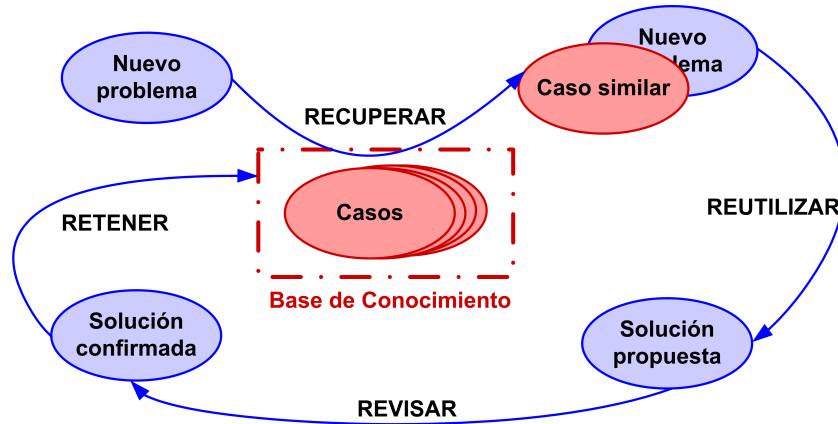


Figura 7.1: Ciclo para la resolución de problemas del razonamiento basado en casos.

Recuperar (*Retrieve*).

- Reutilizando la información y conocimiento de esos casos para resolver el nuevo problema. Esta información se puede utilizar de forma directa (aplicando la misma solución) o realizando algún tipo de modificación para que se adapte mejor al problema actual. A esta fase se la denomina Reutilizar (*Reuse*).
- Revisando la solución propuesta. Esto es, comprobar si la solución aplicada fue satisfactoria o no (*Revising*).
- Guardando en la base de conocimiento el nuevo problema y la solución aplicada una vez que ésta ha sido confirmada y validada, haciéndola disponible para resolver problemas futuros (*Retener* o *Retain*).

Por tanto, un nuevo problema se resuelve recuperando uno o más casos previos (ya experimentados), reutilizando esos casos de una manera u otra, revisando la solución propuesta y finalmente guardando la nueva experiencia, mediante su incorporación a la base de conocimiento existente. Es importante destacar que no todas las fases son necesarias. Por ejemplo, existen aplicaciones donde el conocimiento no se incrementa (no incluyen la fase Retener) o aplicaciones donde la fase Reutilizar y Revisar se fusionan en una sola.

Por otro lado, la técnica de razonamiento basado en casos es particularmente aplicable a problemas donde hay disponibles casos pasados, incluso cuando el dominio del problema no se entiende lo suficiente para obtener un modelo más profundo.

Cada uno de los casos, tanto el nuevo que se quiere resolver como los existentes en la base de conocimiento, se componen de una serie de atributos que definen las características de cada caso. El número de atributos depende del sistema concreto en el que

se aplique el CBR. En la primera fase (Recuperar), se calcula la similitud entre el nuevo caso (x) y cada uno de los casos contenidos en la base de conocimiento (y). En concreto, en las aplicaciones desarrolladas en esta Tesis, se normaliza el valor de los atributos y se calcula la distancia Euclídea ponderada [153, 154] siguiendo la ecuación (7.1) para calcular la similitud. En esta ecuación, el término a representa cada atributo de los casos x e y , el término W_a es el peso asociado a esa característica, y n es el número de atributos. Cada característica viene ponderada por un peso para dar la opción de que ciertas características prevalezcan sobre otras para determinar el caso más parecido. Siguiendo la ecuación (7.1), los valores más próximos a 0 implican una mayor similitud entre los casos.

$$\text{Similitud}(x, y) = -\sqrt{\sum_{a=1}^n W_a^2 \cdot (x_a - y_a)^2} \quad (7.1)$$

Una vez calculada la similitud y recuperado el caso o los casos más similares, éstos se van a reutilizar. La reutilización realizada depende también del sistema concreto donde se aplique el CBR.

Por tanto, gracias al conocimiento adquirido con experiencias previas, el CBR va a resolver nuevas situaciones y a tomar decisiones a partir de la similitud de estos nuevos casos con las experiencias pasadas.

En el Apartado 7.3 se va a incorporar esta técnica en la estimación de la calidad de transmisión en redes troncales. Más adelante, en el Apartado 7.5 la técnica de CBR se aplicará en la red de acceso. Por tanto, en ambos casos se introduce cognición para realizar una gestión eficiente de los recursos disponibles.

7.3. Estimador cognitivo de la calidad de transmisión en redes ópticas troncales

7.3.1. Introducción a las redes ópticas troncales actuales

Las redes troncales o de transporte, tal y como se vio en el Capítulo 2, son las redes que conectan las diferentes redes metropolitanas (Figura 2.1). En estas redes es necesario un gran ancho de banda, puesto que por ellas se transporta el tráfico agregado de varias redes metropolitanas. Por tanto, el gran ancho de banda que la fibra óptica es capaz de proporcionar, unido a la fiabilidad, flexibilidad y transparencia de las redes ópticas, hacen que la fibra sea el medio de transmisión más común en la red troncal. Para explotar de forma óptima y económica toda la capacidad que la fibra es capaz de

ofrecer, se utilizan técnicas de multiplexación por longitud de onda (WDM). De esta forma, diferentes canales de datos pueden ser transmitidos simultáneamente a través de una única fibra utilizando diferentes longitudes de onda (o lo que es lo mismo, diferentes frecuencias). En las redes ópticas de primera generación, solo se utilizaba la multiplexación de longitudes de onda para incrementar la capacidad de transmisión entre nodos adyacentes de la red. De esta forma, cuando se transmitía información entre nodos que no eran adyacentes, era necesario realizar una conversión del dominio óptico al dominio eléctrico para ser procesados y de nuevo convertidos al dominio óptico para ser transmitidos por la fibra hacia el siguiente nodo. Esta conversión Óptica-Eléctrica-Óptica (OEO) en cada uno de los nodos intermedios provocaba cuellos de botella debido al procesamiento eléctrico de todos los flujos de datos que pasaban por los nodos intermedios. Por ello, las redes evolucionaron hacia las redes de segunda generación o “todo ópticas”. En estas redes, la longitud de onda no se utiliza únicamente para incrementar la capacidad de transmisión, sino que también tiene funciones de encaminamiento. De esta forma, las funciones de enrutamiento se realizan en el dominio óptico sin ser necesaria la conversión al dominio eléctrico. En este tipo de redes, denominadas redes ópticas con enrutamiento por longitud de onda (WRON, *Wavelength-Routed Optical Networks*), las conexiones ópticas se establecen entre dos nodos de la red, no necesariamente adyacentes en la topología física (en el conjunto de fibras y nodos que constituyen la red). Estas conexiones, denominadas *lightpaths* o caminos ópticos, se caracterizan por la ruta que siguen y la longitud de onda utilizada en cada enlace. Por tanto, un punto clave en el diseño de redes WRON es el de enrutamiento y asignación de longitud de onda (RWA, *Route and Wavelength Assignment*). Esto es, encontrar una ruta y una longitud de onda (o conjunto de longitudes de onda si se usan conversores) para cada uno de los *lightpaths* que se van a establecer en la red [155]. En la Figura 7.2 está representado un ejemplo sencillo de red con encaminamiento por longitud de onda, donde se han establecido tres *lightpaths* bidireccionales. En este caso, por ejemplo, el tráfico entre los nodos A y C puede transmitirse directamente a través del *lightpath* establecido entre esos dos nodos, sin experimentar la conversión OEO ni procesamiento electrónico (y por tanto sin sufrir retardos) en el nodo intermedio B.

Sin embargo, a medida que las señales ópticas atraviesan las fibras y los nodos y se propagan a través de componentes ópticos activos y pasivos hacia su destino, sufren una serie de perturbaciones que degradan la calidad de la señal. Estas perturbaciones afectan a cada canal óptico de forma individual, pero también causan interferencias entre los canales copropagativos. De esta forma, como no existe conversión en el dominio eléctrico y por tanto no es posible realizar regeneración en los nodos intermedios, la calidad de

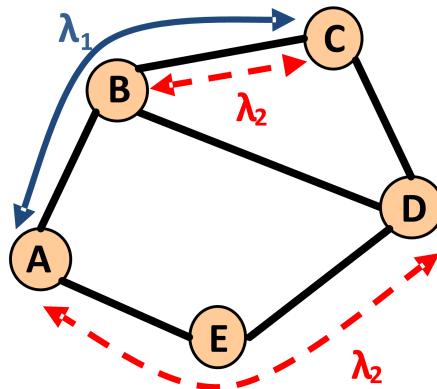


Figura 7.2: Ejemplo de una red con encaminamiento por longitud de onda donde se han establecido tres *lightpaths*.

transmisión (QoT, *Quality of Transmission*) puede verse afectada y puede que no cumpla con los requisitos de calidad previamente estipulados. Por este motivo, en los últimos años ha surgido un gran interés en el diseño de redes ópticas que tengan en cuenta estas perturbaciones físicas. En este sentido, existen ya una serie de propuestas que resuelven el problema RWA a la vez que se asegura que los caminos ópticos o *lightpath* cumplen con los requisitos de QoT [156, 157]. Para llevar a cabo este propósito se necesitan métodos efectivos y eficientes para predecir la calidad de los *lightpaths* antes de que éstos sean establecidos. En este sentido, la herramienta de predicción se puede usar no solo para descartar aquellos caminos ópticos que no cumplen con los requisitos estipulados de QoT, sino también para verificar que el establecimiento de un nuevo camino óptico no tenga un impacto significativo en los ya existentes, evitando de esta forma situaciones problemáticas. En concreto, en [158, 159] se presenta una herramienta capaz de estimar en tiempo real la calidad de transmisión de los *lightpaths* (denominada *Q-Tool*). Esta herramienta combina una serie de modelos analíticos previamente propuestos y verificados en la literatura y, a diferencia de otras aproximaciones, también utiliza el método de paso dividido de Fourier (SSFM, *Split-Step Fourier Method*) para mejorar la precisión de las estimaciones de la calidad de transmisión [158]. A partir de un conjunto de *lightpaths* y de la topología de red (con sus características físicas), la *Q-Tool* calcula el factor Q de cada *lightpath*. El factor Q es un indicador de la calidad de transmisión, que está estrechamente relacionado con la tasa de error de bit de la señal (BER, *Bit Error Rate*), de forma que valores altos del factor Q se corresponden con un valor bajo de BER [160]. De este modo, la *Q-Tool* proporciona estimaciones relativamente precisas del factor Q teniendo en cuenta varios modelos de las perturbaciones lineales y no lineales de la capa física, siendo por tanto un elemento muy útil para el diseño y control de la red óptica.

Sin embargo, esta herramienta presenta una serie de limitaciones. La primera de ellas, es que solo es válida para redes de 10 Gb/s con una modulación OOK (*On-Off Keying*). En segundo lugar, debido a los complicados cálculos requeridos, el tiempo de computación es muy elevado, desde 1 hasta 1000 segundos, dependiendo del escenario, utilizando la implementación *software* descrita en [161]. Por tanto, el uso de esta herramienta puede ser prohibitivo cuando existan restricciones de tiempo, por ejemplo, para el control en tiempo real o para algunas técnicas de planificación como las basadas en algoritmos genéticos [162], ya que evalúan numerosas configuraciones potenciales. Para solventar estos inconvenientes, en [163] se ha propuesto un nuevo modelo capaz de estimar de forma rápida y con exactitud la relación señal a ruido óptica (OSNR, *Optical Signal to Noise Ratio*) de los canales ópticos en sistemas de transmisión coherentes sin compensación, de forma que a partir de este valor se puede obtener la calidad de transmisión. Aunque este trabajo pionero abre las puertas para futuros desarrollos y mejoras, todavía no aborda escenarios de red donde canales provenientes de diferentes ubicaciones sean multiplexados en la fibra óptica, y no es válido tampoco en sistemas donde la dispersión se compensa. Por lo tanto, este modelo, aunque es más rápido, no puede aplicarse todavía a un escenario de red troncal como el que se estudia en esta Tesis.

Por tanto, se propone una estrategia alternativa basada en la cognición para predecir la calidad de transmisión de los *lightpaths* antes de que éstos sean establecidos. De esta forma, explotando el conocimiento adquirido en experiencias previas, se pueden realizar estimaciones de forma rápida y correcta de si un *lightpath* cumplirá con los requisitos de calidad de servicio o no, sin tener que depender de métodos o cálculos complejos. En particular, y tal y como se ha comentado anteriormente, se ha utilizado la técnica cognitiva de razonamiento basado en casos (CBR). Cabe destacar que, puesto que existen en la literatura trabajos en la estimación de la QoS para redes OOK de 10 Gb/s, el estudio realizado se ha centrado en esos escenarios con el fin de tener un punto de referencia en las comparaciones. Teniendo esto en cuenta, se ha seleccionado la *Q-Tool* como herramienta de comparación para evaluar las capacidades del sistema cognitivo bajo condiciones de red lo más realistas posibles. Sin embargo, los fundamentos del estimador cognitivo son lo suficientemente genéricos como para poder ser aplicados en otras redes con mayores tasas de transmisión y diferentes modulaciones [164].

7.3.2. Descripción del estimador de calidad cognitivo diseñado

Se ha desarrollado un estimador cognitivo de la calidad de transmisión capaz de clasificar los *lightpaths* antes de su establecimiento en dos categorías, *lightpaths* con una QoS alta y con una QoS baja [165–168]. Para determinar estas dos categorías se utiliza

un umbral del factor Q definido por el proveedor de servicios ($Q_{threshold}$). Por tanto, si el factor Q de un camino óptico es mayor o igual que este umbral, se considera que pertenece a la categoría de *lightpaths* con QoT alta, y se asume que va a cumplir con los requisitos de calidad establecidos. En caso contrario, el camino pertenece a la categoría de QoT baja y se asume que no cumplirá con los requisitos de calidad. Esta clasificación de la calidad se puede realizar a partir de otros parámetros como el vector de magnitud de error (EVM, *Error Vector Magnitude*) [160, 164]. En el caso de utilizar el parámetro EVM, los valores por debajo del umbral establecido se asociarán con *lightpaths* con QoT alta. Sin embargo, en el estimador desarrollado se ha utilizado el factor Q como parámetro para determinar la categoría de cada *lightpath* para así poder comparar de forma directa con la herramienta *Q-Tool*, que proporciona dicho parámetro Q.

Para determinar la categoría del *lightpath*, el estimador cognitivo diseñado utiliza un mecanismo híbrido. En primer lugar, tiene en cuenta la longitud total del *lightpath*, y después, si es necesario utiliza el razonamiento basado en casos. La motivación de utilizar la longitud como primer elemento para su clasificación se debe al gran impacto que tiene este parámetro en el factor Q. Para mostrar esta relación, se han realizado simulaciones con la implementación en MATLAB de la *Q-Tool* para evaluar el factor Q de los *lightpaths* en la red GÉANT2 [169] configurada como una red con encaminamiento por longitud de onda dinámico, equipada con 32 longitudes de onda por enlace y transceptores OOK de 10 Gb/s. Además se han considerado diferentes cargas de red y un umbral ($Q_{threshold}$) igual a 16.9 dB (que se corresponde a una BER de 10^{-12}). La Figura 7.3 muestra el factor Q de los diferentes *lightpaths* en función de su longitud total. Como puede observarse en la gráfica, los *lightpaths* cuya longitud es menor que un cierto

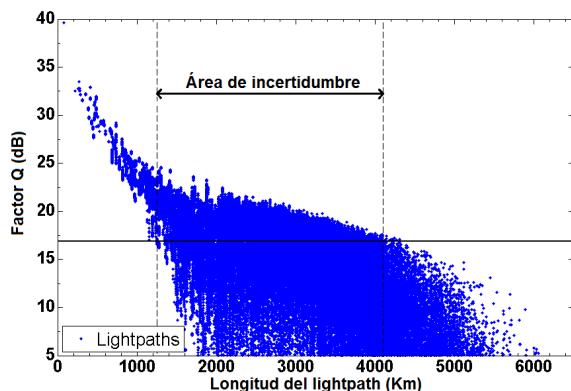


Figura 7.3: Factor Q de los *lightpaths* en función de su longitud para la red GÉANT2 con 32 longitudes de onda considerando distintas cargas de red.

valor (1250 km aproximadamente) generalmente pertenecen a la clase con QoT alta, mientras que aquellos con una longitud muy grande (mayor de 4100 km aproximadamente) pertenecen típicamente a la clase con baja QoT. Sin embargo, existe un área de incertidumbre (un rango de longitudes intermedias) donde el resto de características del *lightpath* adquiere una cierta importancia en la determinación de su factor Q y por tanto en su clasificación. Por tanto, para clasificar los *lightpaths* en esta zona de incertidumbre se utilizará la técnica de razonamiento basado en casos (CBR) cuyo diseño se pasará a explicar en el siguiente apartado.

7.3.3. Diseño del CBR del estimador cognitivo de QoT

En el sistema CBR utilizado en el estimador cognitivo de QoT desarrollado, la base de conocimiento inicial se compone de un número de casos, que consisten en la descripción del *lightpath* (un conjunto de atributos) y su factor Q asociado. La descripción de los *lightpaths* contiene su ruta, esto es, el conjunto de enlaces que atraviesa (representado por el porcentaje de la contribución individual de ese enlace a la longitud total del camino óptico), la longitud de onda seleccionada, la longitud total del camino, la suma de los *lightpaths* copropagativos por enlace y la desviación estándar del número de caminos ópticos copropagativos. Se han tomado estos parámetros por ser los que tienen una mayor influencia a la hora de determinar el factor Q de un camino óptico. Además, el factor Q asociado al *lightpath* que se almacena en la base de conocimiento, es una estimación del factor Q que se ha obtenido utilizando la *Q-Tool*. Para obtener estos casos, se han llevado a cabo una serie de simulaciones previas. Por tanto, los casos almacenados en la base de conocimiento son aquellos *lightpaths* que se han establecido en distintos momentos de las simulaciones y cuyos factores Q se han estimado utilizando la *Q-Tool*. La razón de utilizar la herramienta *Q-Tool* como base para la estimación de la calidad, es porque se trata de una herramienta que combina una serie de modelos y computaciones numéricas que ofrecen una gran exactitud. Sin embargo, es importante destacar que, aunque se haya usado esta herramienta para poblar la base de conocimiento, se pueden introducir datos del factor Q de otras herramientas o incluso datos reales provenientes de monitores durante el funcionamiento de la red [170].

Así pues, en el funcionamiento real de la red, donde se necesita estimar de forma rápida la calidad de un *lightpath* o camino óptico, el estimador de QoT cognitivo funciona de la siguiente manera. En primer lugar, cuando llega una petición de *lightpath*, se resuelve el problema RWA y se calcula la longitud total del camino óptico. Si dicha longitud es menor que el umbral inferior de la zona de incertidumbre, entonces se asume que el *lightpath* cumple con los requisitos de calidad y se establece. Por otro lado, si

la longitud es mayor que el umbral superior de la zona de incertidumbre, entonces se asume que el *lightpath* no cumple con los requisitos y por tanto no se va a establecer. Sin embargo, si la longitud del *lightpath* está dentro de la zona de incertidumbre, se aplica el CBR y se obtiene de la base de conocimiento el *lightpath* más similar al que se quiere estimar la calidad. Para determinar la similitud entre el nuevo *lightpath* (x) y cada uno de los existentes en la base de conocimiento (y), los atributos se normalizan y se calcula la distancia Euclídea ponderada siguiendo la ecuación (7.1) (explicada en el Apartado 7.2). En este caso, los pesos utilizados para cada uno de los atributos (W_a) son los coeficientes de un modelo lineal de regresión por mínimos cuadrados para la base de conocimiento, considerando el factor Q como una variable dependiente. Además, se asume que el factor Q del nuevo *lightpath* es el mismo que el del caso más similar y ese valor se usa para decidir si el nuevo camino óptico cumple con los requisitos de QoT comparándolo con el valor umbral del factor Q ($Q_{threshold}$).

Este proceso constituye la primera versión del estimador cognitivo de QoT, donde la base de conocimiento es completamente estática y por tanto no se actualiza con nuevos casos ni se optimiza. A esta primera versión se le ha denominado R-CBR (*Regular-CBR*)[165, 166]. Sin embargo, en los siguientes apartados se introduce una nueva versión donde se realiza un proceso de optimización de la base de conocimiento.

7.3.4. Optimización de la base del conocimiento

Tal y como se mencionó en el Apartado 7.2 de este capítulo, la base de conocimiento de un sistema de razonamiento basado en casos puede ser actualizada para incluir nuevas experiencias, incorporando la descripción y solución de nuevos problemas resueltos por el CBR. De esta forma, el sistema es capaz de aprender y adaptarse a los cambios que se puedan producir. El aprendizaje tiende a incrementar la efectividad del sistema, ya que la base de conocimiento puede incluir nuevos casos que estimen mejor la calidad de un nuevo *lightpath*. Sin embargo, un aprendizaje excesivo tiene un gran impacto en el tiempo de búsqueda del caso más similar, ya que éste depende directamente del tamaño de la base de conocimiento [171]. Esto se conoce como problema de utilidad [171, 172], que ocurre cuando el coste de mantener y buscar en una base de conocimiento muy grande sobrepasa los beneficios del almacenamiento de ese conocimiento. Por tanto, para evitar el problema de utilidad, es necesario no solo incluir aprendizaje, sino también olvido. Consecuentemente, hay que implementar técnicas que controlen la retención de nuevos casos y la eliminación de aquellos que no mejoran el funcionamiento del sistema. Por tanto, en esta sección se presenta un proceso de optimización de la base de conocimiento que se va a realizar de forma previa a la ejecución real del estimador en la red y que

mejorará el funcionamiento del estimador. Este proceso de optimización consta de una serie de etapas de mantenimiento donde se aprenden y olvidan casos.

Por tanto, en el proceso de optimización, que tiene lugar antes del funcionamiento real del estimador en la red, se ejecuta el estimador cognitivo para estimar la calidad de una serie de *lightpaths*. Si el *lightpath* pertenece al área de incertidumbre, esto es, su clasificación se ha realizado con el CBR, se va a realizar una doble comprobación. En primer lugar, se comprueba si el *lightpath* se ha clasificado en la categoría correcta, esto es, si se ha acertado a la hora de decidir si el factor Q del *lightpath* está por encima o por debajo del límite establecido ($Q_{threshold}$) . En segundo lugar, se calcula el error entre el factor Q estimado por el CBR y su valor real y se comprueba si la diferencia está por debajo de una cierta cantidad, esto es, el error permitido ($\varepsilon_{permitted}$). Si se ha acertado en la clasificación del *lightpath* y además el error obtenido es menor que el permitido, entonces, el CBR es capaz de estimar correctamente el factor Q de dicho *lightpath* y por tanto no es necesario incluirlo en la base de conocimiento. Por el contrario, si no se cumple cualquiera de estas dos condiciones, esto es, si el *lightpath* no se ha clasificado correctamente o si el error es mayor que $\varepsilon_{permitted}$, entonces se almacena el caso en una base de datos auxiliar como *lightpath* candidato a ser incorporado a la base de conocimiento (candidato para ser aprendido). Cuando el CBR ha realizado un cierto número de clasificaciones (correctas o incorrectas), se ejecuta la fase de mantenimiento. Durante esta fase, en primer lugar, se añaden a la base del conocimiento los casos almacenados como candidatos a ser aprendidos y se eliminan de la base de datos auxiliar. Después, se aplica una técnica para eliminar los casos redundantes de la base de conocimiento.

Concretamente, la técnica seleccionada para llevar a cabo la eliminación de los casos redundantes está basada en el método de Reducción de Redundancia Conservativo (CRR, *Conservative Redundancy Reduction*) [173]. El objetivo de este algoritmo es eliminar casos redundantes que no están localizados cerca de los bordes de cada clase. Para realizar este proceso, se calcula el conjunto de cobertura (CS, *Coverage Set*) de cada caso. El conjunto de cobertura de un caso concreto (c) es el conjunto de todos los casos que c es capaz de clasificar de forma correcta [173, 174]. Por tanto, los casos que tienen un conjunto de cobertura grande están situados, probablemente, cerca de un grupo de casos que pertenecen a la misma clase. Por otro lado, si un caso tiene un conjunto de cobertura pequeño, esto indica que tiene pocos vecinos y por tanto, que puede que esté situado cerca del borde de la clase [173]. El pseudocódigo para calcular los conjuntos de cobertura de cada uno de los casos de la base del conocimiento adaptado a las características del estimador de QoT cognitivo es el que se muestra a continuación:

```

para todo caso  $c \in Base\ Conocimiento$  hacer
     $CS(c) \leftarrow \emptyset;$ 
fin para
para todo caso  $q \in Base\ Conocimiento$  hacer
    1:  $c \leftarrow$  siguiente vecino más cercano al caso  $q$ 
    si  $c$  predice a  $q$  correctamente entonces
         $CS(c) \leftarrow CS(c) \cup q;$ 
        Ir a 1.
    fin si
    sino
        Seleccionar siguiente  $q$ 
    fin sino
fin para

```

Nota: c predice q correctamente si c clasifica al caso q en la misma categoría (si ambos pertenecen a la misma clase) y si el error absoluto cometido en la predicción del factor Q (la diferencia entre los factores Q reales de los casos q y c en valor absoluto) es menor que $\varepsilon_{permitted}$.

Una vez que se ha calculado el CS, el algoritmo CRR ordena todos los casos existentes en la base de conocimiento en orden ascendente según el tamaño de su conjunto de cobertura. Después, los casos en la base de conocimiento se analizan comenzando por aquellos con CS pequeños, y los casos contenidos en su conjunto de cobertura se eliminan de la base de conocimiento [173]. Además, si un caso ya ha sido eliminado durante este proceso de la base de conocimiento, éste no se analizará después para borrar su propio conjunto de cobertura. Así mismo, si después de realizar este proceso el tamaño de la base de conocimiento es mayor que su tamaño original (el que tenía antes de comenzar con cualquier proceso de optimización), entonces se eliminan casos de la base de conocimiento empezando por aquellos que tienen un conjunto de cobertura mayor, hasta que el tamaño resultante de la base de conocimiento sea igual al tamaño original. Finalmente, es importante destacar que para llevar a cabo el procedimiento de optimización que se acaba de describir, es necesario comparar la estimación del factor Q de cada *lightpath* con su valor real. Por tanto, para actualizaciones de la base de conocimiento, el estimador cognitivo de QoT tiene que trabajar en colaboración con un sistema de monitorización de red que mida los factores Q de los *lightpaths* establecidos. Sin embargo, el proceso de optimización aplicado aquí es un proceso previo a la ejecución del estimador. De este modo, una vez que la base de conocimiento se genera mediante simulaciones previas, es optimizada aplicando el procedimiento descrito anteriormente.

Para ello, dichas simulaciones previas utilizan la herramienta *Q-Tool* para proporcionar los valores Q “reales”.

Por consiguiente, se van a proponer dos métodos para implementar el CBR del estimador cognitivo. El primero de ellos (R-CBR, *Regular CBR*), explicado en el apartado anterior (Apartado 7.3.3), es un estimador cognitivo de la calidad de transmisión donde no se optimiza la base de conocimiento antes de ponerlo en marcha. El segundo método, denominado FixE-CBR (*Fixed Error CBR*)[\[165, 167\]](#), es un estimador cognitivo que aplica técnicas de aprendizaje y olvido para realizar una optimización de la base de conocimiento previa a la puesta en marcha del estimador. No obstante, cabe destacar que la base de conocimiento asociada al método FixE-CBR no se optimiza más durante la ejecución en tiempo real del estimador.

7.4. Estudio de simulación del estimador cognitivo de QoT

En este apartado se van a analizar las prestaciones de los dos estimadores cognitivos desarrollados. Para ello, en primer lugar se presenta el escenario de simulación utilizado y después se estudiará el funcionamiento y la eficiencia de ambos estimadores en diferentes contextos de redes ópticas troncales.

7.4.1. Escenario de simulación

Para evaluar el funcionamiento del estimador cognitivo de QoT cuando se implementan las dos versiones desarrolladas (R-CBR y FiXE-CBR), se han llevado a cabo simulaciones en dos redes diferentes para mostrar posibles problemas de escalabilidad. Las dos redes escogidas son una red de larga distancia, la red *Deutsche Telekom* (DT) de 14 nodos [\[158\]](#) y la red de ultra larga distancia GÉANT2, con 34 nodos [\[169\]](#). Ambas redes se han configurado como redes de encaminamiento por longitud de onda dinámicas y están equipadas con transceptores OOK de 10 Gb/s. Cada enlace está formado por un conjunto de fibras estándar monomodo (SMF, *Single Mode Fibre*), seguidas por fibras compensadoras de la dispersión (DCF, *Dispersion Compensating Fibre*). Se ha considerado un número igual a 32 y 64 longitudes de onda por enlace. Los resultados se han obtenido analizando los escenarios bajo diferentes cargas de tráfico. Las rutas y longitudes de onda para cada una de las conexiones han sido obtenidas por medio de un algoritmo RWA adaptativo, más concretamente, el algoritmo *Aur-Exhaustive* [\[175\]](#), ya que ofrece una mayor flexibilidad y por tanto una menor probabilidad de bloqueo en escenarios dinámicos que otros métodos basados en la utilización de rutas fijas precalculadas [\[175, 176\]](#).

En el marco del proyecto Europeo DICONET se desarrollaron dos implementaciones de la *Q-Tool* [158, 159], una implementación en MATLAB y una implementación basada en FPGAs (*Field Programmable Gate Array*) que acelera el proceso de estimación de la QoS [161]. Con el propósito de poder realizar una comparación justa en términos de tiempo de computación, el estimador cognitivo se ha implementado en MATLAB y por tanto se comparará con la versión en MATLAB de la *Q-Tool*.

Además, el umbral del factor Q para la clasificación de los *lightpaths* en la categoría de QoS alta y QoS baja ($Q_{threshold}$) se ha fijado a 16.9 dB (BER de 10^{-12}). Los límites inferior y superior de la zona de incertidumbre se han establecido de forma que el 99.9 % de los *lightpaths* fuera de esta zona se clasifiquen correctamente. Esto es, el 99.9 % de los *lightpaths* por debajo del límite inferior tienen una QoS mayor que el valor de $Q_{threshold}$ (QoS alta) y el 99.9 % de los *lightpaths* con una longitud mayor que el límite superior tienen una QoS baja (valor menor que $Q_{threshold}$). Los umbrales de la zona de incertidumbre obtenidos en las dos redes son los que se muestran en la Tabla 7.1.

Red	Número de longitudes de onda	Límite inferior (km)	Límite superior (km)
DT	32	975	1875
	64	975	2050
GÉANT2	32	1250	4125
	64	1175	4225

Tabla 7.1: Límite superior e inferior (en km) de la zona de incertidumbre para las redes DT y GÉANT2.

La base de conocimiento inicial del CBR se ha poblado con diferentes números de casos, desde 500 hasta 5000 en el caso de la DT y desde 5000 hasta 50000 para la red GÉANT2. Para asegurar una comparación justa entre ambas redes, la base de conocimiento de la red GÉANT2 se ha incrementado, ya que tiene un mayor número de nodos (el número de pares origen-destino se incrementa 6 veces con respecto a la DT). Los casos en la base de conocimiento para ambas redes, se han escogido de forma aleatoria de aquellos generados en una simulación previa. Cada base de conocimiento cubre los casos de la zona de incertidumbre para todas las cargas de red, de forma que la misma base de conocimiento se puede usar independientemente de la carga de red.

En el caso de realizar la optimización de la base de conocimiento (FixE-CBR), se han clasificado previamente 6000 nuevos *lightpaths* para la DT y 36000 para GÉANT2 pertenecientes a la zona de incertidumbre, utilizando el método descrito en el Apartado 7.3.4 cada 500 clasificaciones. El error permitido ($\varepsilon_{permitted}$) se ha fijado a 3 dB. Este es un valor de compromiso, ya que valores más bajos implicarían un aumento de los casos a

aprender y valores altos supondrían un mayor error cometido en la estimación del factor Q. Una vez ha finalizado el proceso de optimización, se analiza el funcionamiento del estimador cognitivo. Para ello, se han evaluado otros 6000 *lightpaths* para la DT y 36000 para la GÉANT2 (pertenecientes y no pertenecientes a la zona de incertidumbre). Sin embargo, la base de conocimiento no se ha actualizado durante esta evaluación, es decir, no ha existido ni aprendizaje ni olvido.

Por último, los resultados que se muestran en las siguientes figuras se han obtenido después de repetir este proceso 100 veces con diferentes bases de conocimiento aleatorias. La media de los resultados se ha representado junto con los intervalos de confianza del 95 %, aunque en la mayor parte de los casos, el tamaño de los intervalos de confianza es más pequeño que el tamaño de los símbolos.

7.4.2. Análisis de prestaciones de los estimadores R-CBR y FixE-CBR

A continuación se muestra la evaluación de las prestaciones de R-CBR y FixE-CBR para llevar a cabo las clasificaciones en la zona de incertidumbre en las redes DT y GÉANT2. En primer lugar, se analiza el grado de exactitud del estimador al realizar las clasificaciones de los *lightpaths*. Después se compara el tiempo de computación de los estimadores cognitivos desarrollados con la herramienta *Q-Tool*. Por último, se describen métodos prácticos para poblar la base de conocimiento en un escenario real para la puesta en marcha de la red.

7.4.2.1. Análisis del porcentaje de acierto en las clasificaciones

Para comprobar el funcionamiento del estimador cognitivo bajo el escenario de red presentado anteriormente, se van a realizar primero simulaciones para estudiar la exactitud del estimador a la hora de decidir la categoría del *lightpath* a clasificar. La Figura 7.4 representa el porcentaje de aciertos obtenidos por el estimador cognitivo al clasificar los *lightpaths* en las categorías de baja QoT y alta QoT para la red DT con 32 y 64 longitudes de onda cuando se utilizan los estimadores R-CBR y FixE-CBR. Esto es, se comparan los aciertos obtenidos cuando la base de conocimiento ha sido poblada con casos seleccionados aleatoriamente de simulaciones realizadas previamente (R-CBR) y cuando dicha base se ha optimizado antes del funcionamiento del estimador (FixE-CBR). Los números que aparecen escritos al lado de los puntos asociados al método FixE-CBR en la figura, representan el tamaño inicial de la base de conocimiento (esto es, el tamaño antes de la optimización). Por ejemplo, para FixE-CBR con 32 longitudes de onda se ha partido de una base de conocimiento inicial de 5000 casos, pero tras aplicar el proceso de optimización la base de datos con la que se hace la evaluación del porcentaje de aciertos

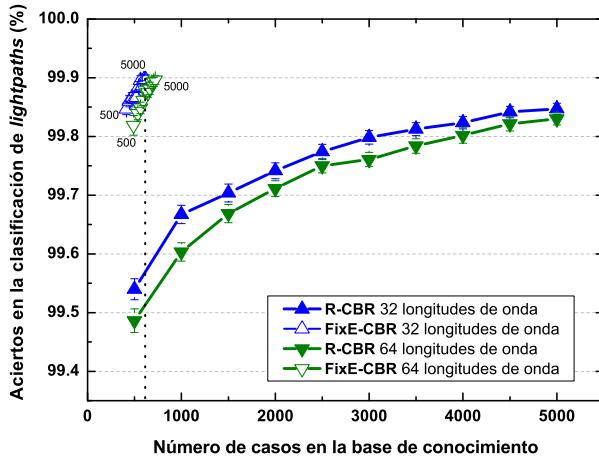


Figura 7.4: Porcentaje de aciertos en la clasificación de la QoT de los *lightpaths* para los métodos R-CBR y FixE-CBR para la red DT. Los números que aparecen al lado de los puntos de FixE-CBR se refieren a los valores iniciales de la base de conocimiento antes de ejecutar el proceso de optimización.

consta solo de 618 casos.

Tal y como se puede observar en la Figura 7.4, el porcentaje de aciertos es muy alto incluso cuando la base de conocimiento no está optimizada (R-CBR). Para el tamaño de base de conocimiento más pequeño (500 casos), el estimador cognitivo alcanza más de un 99.45 % de clasificaciones correctas y ese porcentaje se eleva hasta los 99.8 % en el caso de considerar la base de conocimiento más grande (5000 casos) para ambas longitudes de onda.

Cuando se comparan R-CBR y FixE-CBR, se puede observar cómo no solo es ligeramente mayor el porcentaje de aciertos con la base de datos optimizada (FixE-CBR) que sin optimización (R-CBR), sino que también el número final de casos en la base de conocimiento es mucho menor. Por ejemplo, para 32 longitudes de onda y con un tamaño en la base de conocimiento de 500 casos, R-CBR consigue un 99.53 % de aciertos, mientras que FixE-CBR partiendo de una base de conocimiento con 500 casos, reduce estos casos a 412 y eleva el porcentaje de aciertos hasta 99.84 %. Por otro lado, para un tamaño inicial de la base de conocimiento igual a 5000 casos, para 32 longitudes de onda, FixE-CBR incrementa ligeramente el porcentaje de aciertos desde 99.84 % (R-CBR) a 99.89 %, pero lo más importante es que consigue una reducción muy significativa en el tamaño de la base de conocimiento, ya que el tamaño final es de 618 casos frente a los 5000 casos iniciales (87.64 % de reducción). Tal y como se demostrará después, esta reducción tiene un impacto muy significativo en términos de la reducción del tiempo de computación.

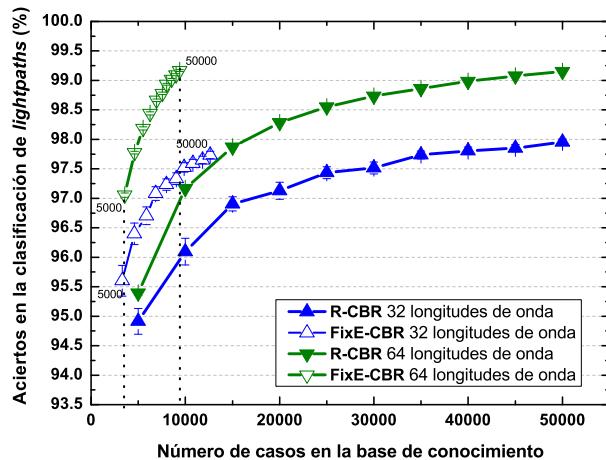


Figura 7.5: Porcentaje de aciertos en la clasificación de la QoT de los *lightpaths* para los métodos R-CBR y FixE-CBR para la red GÉANT2. Los números que aparecen al lado de los puntos de FixE-CBR se refieren a los valores iniciales de la base de conocimiento antes de ejecutar el proceso de optimización.

Para la red GÉANT2 el comportamiento es muy similar. La Figura 7.5 compara la evolución del porcentaje de aciertos en las clasificaciones cuando el tamaño de la base de conocimiento se incrementa para R-CBR y FixE-CBR en esta red. Como puede apreciarse, para el escenario con 32 longitudes de onda, el porcentaje más alto de aciertos que alcanza R-CBR es para 50000 casos y tiene un valor de 98 %, mientras que para el escenario con 64 longitudes de onda este valor es de 99.15 % para ese mismo tamaño de base de conocimiento. Por otro lado, FixE-CBR consigue mejorar el porcentaje de aciertos para las bases de conocimiento con el menor tamaño considerado reduciendo ligeramente el tamaño de la base de conocimiento final. Por ejemplo, para 64 longitudes de onda y un tamaño inicial de base de conocimiento de 5000 casos, los aciertos se elevan desde el 95.4 % (R-CBR) al 97 % (FixE-CBR) mientras que la base de conocimiento se reduce de 5000 a 3560 casos. Sin embargo, las mejoras más significativas que se obtienen con FixE-CBR, se centran de nuevo en una gran reducción del tamaño de base de conocimiento para las bases de conocimiento de mayor tamaño. En este sentido, para 64 longitudes de onda, FixE-CBR reduce el tamaño de 50000 casos iniciales a 9404 casos (81.19 %) manteniendo una tasa de aciertos igual o ligeramente superior.

Estos resultados parecen indicar que existe un problema de escalabilidad, ya que el porcentaje de aciertos para la red GÉANT2 es ligeramente peor que para la red DT. Por ello, se ha analizado este comportamiento en más detalle. El estimador cognitivo, se basa en un sistema híbrido que primero tiene en cuenta el umbral de longitud y después, si es necesario, clasifica el *lightpath* utilizando el sistema CBR. Como se ha

descrito anteriormente, los umbrales de la zona de incertidumbre se han establecido de forma que las clasificaciones acordes a estos umbrales llevan a una tasa de aciertos del 99.99 % para ambas redes. Para la red DT casi todos los casos son resueltos por esta primera aproximación realizada por la longitud del *lightpath* y solo el 8 % de los casos se resuelven con el CBR (tanto para 32 como para 64 longitudes de onda). Por el contrario, para la red GÉANT2, el porcentaje de *lightpaths* resueltos con el CBR se incrementa hasta el 53.8 % y 56 % para 32 y 64 longitudes de onda, respectivamente. Por tanto, la red GÉANT2 plantea una mayor dificultad ya que la longitud de los *lightpaths* no es tan determinante como ocurre con la red DT.

Para analizar más en profundidad este comportamiento, la Figura 7.6 muestra el porcentaje de aciertos al comparar las dos redes cuando la clasificación se realiza mediante el CBR, es decir, cuando los *lightpaths* pertenecen al área de incertidumbre. Es importante destacar que se ha establecido un tamaño de la base de conocimiento igual para ambas redes para facilitar la comparación directa. Como se ilustra en dicha Figura 7.6, para una base de conocimiento de 5000 casos, el porcentaje de aciertos se eleva hasta 96.5 % en la red DT y 90.56 % para la red GÉANT2. Esta diferencia de porcentajes es el resultado de la dependencia de la distribución de los casos en dos categorías. Por tanto, si simplemente se decidiese que un *lightpath* pertenece a la categoría más probable (clase mayoritaria), se tendría en la red DT que todos los *lightpaths* de la zona de incertidumbre pertenecerían a la clase de QoT alta y se acertaría el 89.5 % de las veces. Por otro lado, para la red GÉANT2, si se sigue esta misma filosofía, todos los *lightpaths* de la zona de incertidumbre tendrían una baja QoT y se acertaría únicamente el 57.59 %. Por tanto,

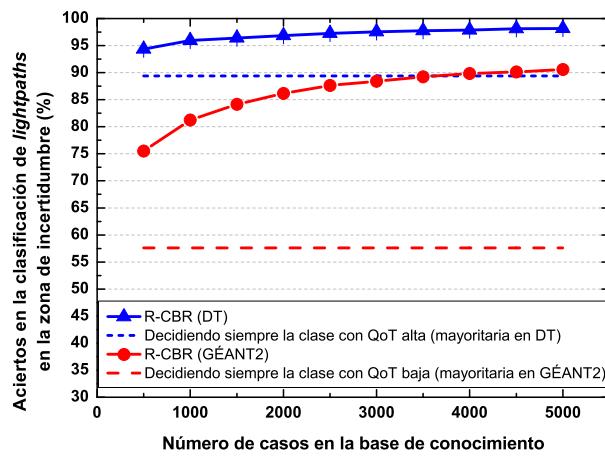


Figura 7.6: Porcentaje de aciertos del CBR en la clasificación de la QoT en la zona de incertidumbre para las redes DT y GÉANT2 con 32 longitudes de onda y aciertos apostando siempre a la clase más probable.

mientras que el mecanismo de CBR mejora los resultados solo en un 8% para la red DT cuando se compara con un estimador que siempre clasificase el *lightpath* en la clase mayoritaria, para la red GÉANT2 la mejoría sería del 32.97%. Como conclusión, estos peores resultados del estimador cognitivo de QoT para la red GÉANT2 comparados con los obtenidos para la red DT, no se deben al incremento del tamaño de la red sino a una estructura más compleja de los datos, esto es, que las características de los casos presentan una mayor dispersión entre sí, lo que hace más complicado una selección correcta del caso más similar.

Por otro lado, la principal razón para considerar un estimador híbrido, que primero tiene en cuenta la longitud del camino óptico para tomar una primera decisión sobre su calidad, frente a un estimador basado únicamente en el CBR, es la considerable reducción del tiempo de computación. Por ejemplo, para una base de conocimiento de 5000 casos, el tiempo medio para clasificar un *lightpath* utilizando únicamente la técnica de CBR es de 6.6 ms para la red DT, mientras que con el sistema híbrido y para el mismo tamaño en la base de conocimiento el tiempo es de 0.5 ms. Por tanto, esta primera fase donde las decisiones se toman considerando únicamente la longitud del *lightpath* aceleran el proceso sin degradar el funcionamiento en término de tasa de aciertos. No obstante, se podría pensar que si la longitud del camino óptico tiene una influencia tan grande, pudiera ser solo necesario utilizar este parámetro para su clasificación en dos categorías sin usar el CBR. Esto sería extremadamente rápido pero no daría buenos resultados. Para demostrar esto, la Tabla 7.2 recoge el porcentaje de aciertos en las clasificaciones para la peor pareja origen-destino, esto es, para la pareja que tiene el menor porcentaje de aciertos. Los resultados mostrados son los obtenidos con el estimador cognitivo sin optimización (R-CBR) comparándolos con los obtenidos con un estimador que solo tiene en cuenta la longitud para clasificar el *lightpath*. Para el primero de ellos, se ha considerado una base de conocimiento de 5000 casos para la red DT y de 50000 para la red GÉANT2. Para el estimador basado únicamente en la longitud de *lightpath*, el umbral de longitud que distingue entre *lightpaths* con alta QoT y baja QoT se ha escogido mediante simulaciones y es igual al umbral que da lugar a los mejores resultados en términos de aciertos globales. Como puede apreciarse, para la red DT, el estimador basado únicamente en la longitud obtiene una tasa de aciertos muy pobre (64% y 59.7% para 32 y 64 longitudes de onda, respectivamente) para la peor pareja origen-destino. Por el contrario, el estimador cognitivo mejora estos valores, ya que los aciertos para la peor pareja origen-destino para ambas longitudes de onda son del 96.1%. La misma tendencia se observa para la red GÉANT2. Es evidente que para esta red los resultados obtenidos con el estimador que tiene en cuenta únicamente la longitud de *lightpath* son

incluso peores. De hecho, hay pares origen-destino cuyos *lightpaths* siempre se clasifican de forma incorrecta. Por el contrario, R-CBR obtiene unos porcentajes de acierto del 77 % y 87 % para 32 y 64 longitudes de onda, respectivamente.

Red	Número de longitudes de onda	Aciertos R-CBR (peor caso)	Estimador basado en distancia	
			Umbral de distancia (km)	Tasa de aciertos (peor caso)
DT	32	96.13 %	1375	64.23 %
	64	96.15 %	1425	59.77 %
GÉANT2	32	77.05 %	1725	0 %
	64	87.15 %	1725	0 %

Tabla 7.2: Porcentaje de acierto en la clasificación de los *lightpaths* para la pareja origen-destino con peor porcentaje de aciertos en la clasificación.

Por tanto, aunque un estimador basado únicamente en la longitud del camino óptico tomaría decisiones de forma instantánea, su porcentaje de aciertos es menor, existiendo *lightpaths* con pares origen-destino con una alta probabilidad de fallo. Por el contrario, la unión de la toma de decisiones mediante longitud con la técnica CBR en la zona de incertidumbre, mejora los resultados y supone un menor tiempo de computación que decidiendo únicamente con CBR.

7.4.2.2. Análisis del tiempo de computación del estimador de QoT cognitivo

Tal y como se ha mencionado anteriormente, el uso de técnicas de aprendizaje y olvido tiende a una reducción del tamaño de la base de conocimiento, lo que se traduce en un tiempo de computación menor, donde el tiempo de computación es el tiempo empleado para estimar la QoT de cada *lightpath*. Esto puede observarse en la Figura 7.7 y Figura 7.8, donde está representado el tiempo de computación (por *lightpath*) de la *Q-Tool* y de los dos estimadores cognitivos (R-CBR y FixE-CBR) frente al tamaño de la base de conocimiento para la red DT y GÉANT2, respectivamente, y para un escenario con 32 longitudes de onda. Las simulaciones de ambas herramientas se han llevado a cabo en una máquina Debian GNU/Linux 6.0 con un procesador *AMD Opteron 6128*.

Es importante mencionar, que en la ejecución dinámica de una red, no solo es necesario determinar la QoT de un nuevo *lightpath* que se va a establecer, sino que también hay que determinar la de los *lightpaths* copropagativos al nuevo *lightpath*, para así verificar que no se van a ver afectados por el nuevo establecimiento. Por lo tanto, se requiere un tiempo de computación bajo, especialmente en redes muy dinámicas. Como puede observarse en la Figura 7.7, para la red DT el estimador cognitivo básico (R-CBR) es alrededor de tres órdenes de magnitud más rápido que la *Q-Tool* cuando la base de co-

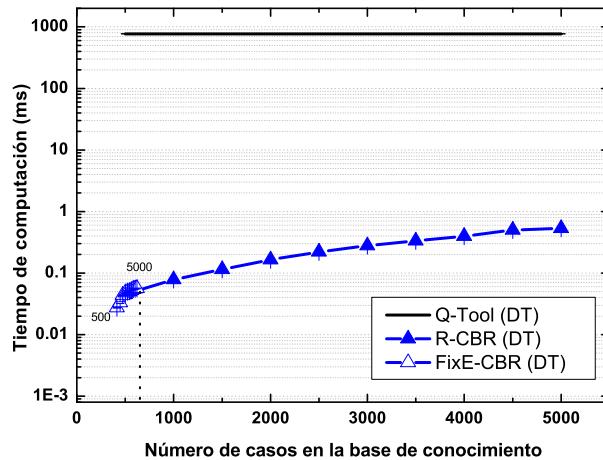


Figura 7.7: Tiempo de computación para la clasificación de un *lightpath* para la red DT con 32 longitudes de onda utilizando la *Q-Tool*, R-CBR y FixE-CBR. Los números que aparecen al lado de los puntos de FixE-CBR se refieren a los tamaños iniciales de la base de conocimiento antes de ejecutar el proceso de optimización.

nocimiento contiene 5000 casos. Sin embargo, el tiempo de computación cuando la base de conocimiento se ha optimizado (FixE-CBR) es incluso más bajo, ya que el tamaño se ha reducido considerablemente (de 5000 a 618 casos). En concreto, el tiempo de computación de FixE-CBR es cercano a cuatro órdenes de magnitud menor que el obtenido con la *Q-Tool*.

Para la red GÉANT2, Figura 7.8, el tiempo de computación es mayor, tanto para la *Q-Tool*, que requiere alrededor de 3.6 s, como para los estimadores cognitivos. En concreto, para los estimadores cognitivos este incremento se debe al hecho de que la red GÉANT2 es una red más compleja, con un mayor número de enlaces. Por tanto, se consideran un mayor número de atributos para calcular la similitud entre los casos y consecuentemente el tiempo para encontrar el caso más similar se ve incrementado. Por otro lado, también se han analizado bases de conocimiento más grandes que para la red DT, lo que también incrementa el tiempo con respecto a esta red. Sin embargo, considerando bases de conocimiento de 50000 casos, R-CBR emplea 110 ms para clasificar un *lightpath* (más de un orden de magnitud más rápido que la *Q-Tool*). Si además se utiliza FixE-CBR el tiempo se reduce a 32 ms ya que el tamaño de la base de conocimiento se decrementa hasta 12633 casos. Por tanto, el estimador cognitivo utilizando una base de conocimiento optimizada, es aproximadamente dos órdenes de magnitud más rápido que la *Q-Tool*.

Tal y como se indicó previamente, el estimador cognitivo es una aproximación híbrida que primero utiliza la longitud del *lightpath* para clasificarlo y después si es necesario

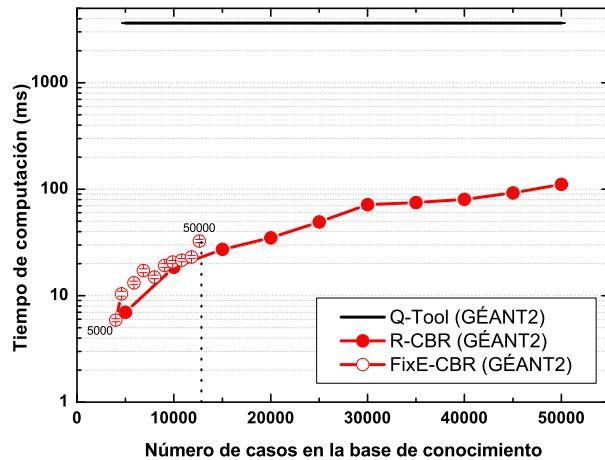


Figura 7.8: Tiempo de computación para la clasificación de un *lightpath* para la red GÉANT2 con 32 longitudes de onda utilizando la *Q-Tool*, R-CBR y FixE-CBR. Los números que aparecen al lado de los puntos de FixE-CBR se refieren a los tamaños iniciales de la base de conocimiento antes de ejecutar el proceso de optimización.

utiliza el sistema CBR. Esta aproximación híbrida reduce de forma significativa el tiempo de computación. Sin embargo, es interesante recalcar que incluso si el estimador cognitivo solo utilizase el CBR para clasificar los *lightpaths*, esto es, si se eliminase la primera fase de decisión por medio de la longitud del camino óptico, el sistema cognitivo seguiría siendo más rápido que la *Q-Tool*. Por ejemplo, considerando la red DT equipada con 32 longitudes de onda, el tiempo de computación obtenido en las simulaciones es de 6.6 ms (por *lightpath*) cuando se usa el sistema CBR frente a los 768.7 ms (por *lightpath*) que serían necesarios si se usa la *Q-Tool*. Por otro lado, incluso si se utilizase la versión de la *Q-Tool* en FPGAs (que reduce el tiempo de computación aproximadamente 28 veces comparada con la implementación de la *Q-Tool* en MATLAB [161]), el estimador cognitivo en MATLAB seguiría siendo ligeramente más rápido.

7.4.3. Puesta en marcha del estimador cognitivo de QoT en una nueva red

Una vez analizadas las prestaciones del estimador cognitivo, se va a estudiar cómo poblar la base de datos en una nueva red. Tal y como se ha visto anteriormente, un punto clave en el funcionamiento del estimador de QoT cognitivo es la base de conocimiento subyacente. Por este motivo, se van a describir a continuación dos métodos prácticos que han sido diseñados para poblar la base de datos antes de empezar la operación dinámica en una nueva red. Para ello, se va a utilizar la red DT, pero partiendo de una base de conocimiento previo vacía.

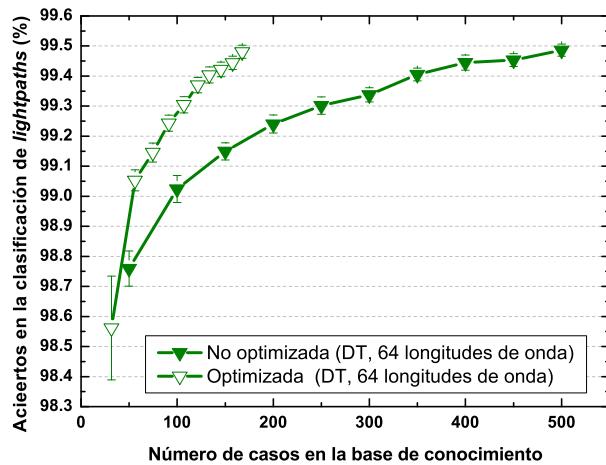


Figura 7.9: Porcentaje de aciertos para tamaños pequeños de la base de conocimiento con y sin optimización previa para la red DT con 64 longitudes de onda.

El primer método para poblar la base de conocimiento consiste en llevar a cabo una serie de simulaciones previas a su puesta en marcha, emulando diferentes configuraciones de la red y guardando el valor de la calidad de transmisión de los distintos *lightpaths* obtenidos a partir de la *Q-Tool* u otro tipo de herramientas. Puesto que este método puede ser lento y tedioso e incluso puede darse el caso de que no se recopilen muchos casos, se ha analizado el funcionamiento del estimador cuando la base de conocimiento consiste en un número reducido de casos (pero representando escenarios muy diversos). Por tanto, la Figura 7.9 representa el porcentaje de acierto en las clasificaciones de los *lightpaths* para la red DT cuando se considera una base de conocimiento sin optimizar poblada con un número pequeño de casos (menos de 500) y también para la versión optimizada de las bases de conocimiento (aplicando el algoritmo CRR explicado en el Apartado 7.3.4). Como puede observarse, incluso para una base de conocimiento pequeña, de unos 50 casos, el porcentaje de aciertos es mayor de 98.7 %.

Una segunda opción consiste en llenar la base de conocimiento reuniendo datos experimentales de la red óptica de forma previa a su operación dinámica. Para llevar a cabo este objetivo, el operador de red puede necesitar usar un sistema de gestión para testear un número reducido de configuraciones útiles o interesantes correspondientes a diferentes cargas de red (diferente número de *lightpaths* establecidos, y por tanto diferentes escenarios en términos de *lightpaths* copropagativos) que se espera vayan a ser tratadas por la red. Por tanto, para cada configuración, se establecen en la red un número de *lightpaths*, y sus niveles de calidad de transmisión son medidos mediante monitores de red. Esta información se utiliza para crear la base de conocimiento inicial. La Figura 7.10 muestra los resultados obtenidos para la red DT, equipada con 64 longitudes de onda, cuando se

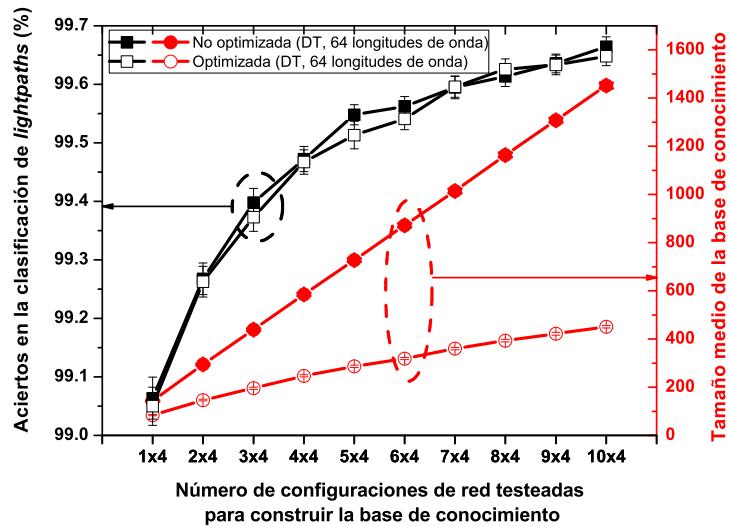


Figura 7.10: Porcentaje de aciertos y tamaño de la base de conocimiento para diferente número de configuraciones de red testeadas en el proceso de construcción de la base de conocimiento para la red DT con 64 longitudes de onda.

utiliza este procedimiento para construir la base de conocimiento. Para construir dicha base, se han establecido cuatro configuraciones aleatorias, cada una representativa de una carga de red diferente: una configuración con carga baja, una con carga media, otra con una carga alta y la última con una carga muy alta. Como se está usando un entorno de simulación y no se disponen de monitores reales, se han reemplazado las medidas de los monitores por los valores obtenidos mediante la *Q-Tool* bajo estas cuatro configuraciones. Entonces, se ha evaluado el funcionamiento del estimador cognitivo de QoT con esas bases de conocimiento subyacentes variando la carga de red desde la más baja hasta la más elevada, pero incluyendo también las cargas de tráfico intermedias que no han sido consideradas en el proceso de población de la base de conocimiento. Para obtener resultados estadísticos significativos, este procedimiento se ha repetido para 100 bases de conocimiento diferentes, construidas como se ha descrito previamente.

Por otro lado, se han construido también bases de conocimiento más grandes recopilando los resultados de más de una configuración aleatoria del estado de la red (esto es, de los *lightpaths* establecidos en distintos momentos) para cada una de las cuatro cargas de red (nombradas estas configuraciones en el eje x de la Figura 7.10 como 2x4, 3x4, y así sucesivamente para indicar que se han tomado dos configuraciones de las cuatro cargas, tres configuraciones, etc.) y se ha analizado el funcionamiento del estimador cognitivo. Por tanto la Figura 7.10 representa el porcentaje de acierto en función del número

de configuraciones de red testeadas en el proceso de construcción de la base de conocimiento. Además, el tamaño resultante de dicha base de conocimiento está representado en la misma figura. En la gráfica, existen dos familias de resultados: aquellas obtenidas utilizando una base de conocimiento no optimizada (símbolos rellenos) y aquellas obtenidas realizando una optimización previa de la base de conocimiento mediante la técnica CRR (símbolos vacíos).

Como puede observarse, el porcentaje de aciertos alcanza los 99.65 % cuando se emplean los resultados de $10 \times 4 = 40$ configuraciones de red para construir la base de conocimiento, dando lugar a una base con un tamaño de aproximadamente 1450 casos excluyendo aquellos que están fuera del área de incertidumbre. En ese escenario, si se utiliza una base de conocimiento optimizada, la tasa de acierto se mantiene pero el tamaño de la base de conocimiento se reduce significativamente (450 casos aproximadamente). Además, utilizando únicamente la información recopilada con 4 configuraciones de red (una configuración para cada una de las cuatro cargas de red), la base de datos subyacente contiene únicamente 84 casos y la tasa de acierto es superior al 99 %.

7.4.4. Validación experimental del estimador cognitivo de QoT

En el marco del proyecto Europeo CHRON (*Cognitive Heterogeneous Reconfigurable Optical Networks*) se ha llevado a cabo una demostración experimental del estimador cognitivo de QoT diseñado. En este sentido, la Universidad Técnica de Dinamarca (DTU) desarrolló un banco de pruebas con un enlace WDM PDM-QPSK (*Polarization Division Multiplexed-Quadrature Phase Shift Keying*) transmitiendo a 80 Gbit/s [164]. El transmisor está compuesto por 5 láseres equiespaciados en 50 GHz y combinados utilizando un multiplexor AWG (*Arrayed Waveguide Grating*) (Figura 7.11). La transmisión sobre la fibra se realiza a través de un máximo de 6 intervalos de 80 km, compuestos por fibra estándar monomodo seguidos de fibra compensadora de dispersión (DCF), alcanzando una distancia máxima de 480 km. Para poder emular diferentes configuraciones de red y *lightpaths*, se permite la modificación de los valores de una serie de parámetros. En concreto, el número de canales que pueden estar activos simultáneamente en el enlace, puede variar entre 2 y 5, la potencia de cada canal varía entre -4 y 4 dBm (en saltos de 2 dB) y el número de intervalos de fibra usados puede variar entre 3 y 6 (variación de 240 km a 480 km). Además, se han medido dos indicadores relacionados con la calidad de transmisión, en concreto, el EVM y la OSNR, para comprobar que el estimador cognitivo es lo suficientemente genérico como para ser aplicado con diferentes indicadores de la calidad de transmisión.

En primer lugar, se realizaron pruebas de clasificación utilizando únicamente el in-

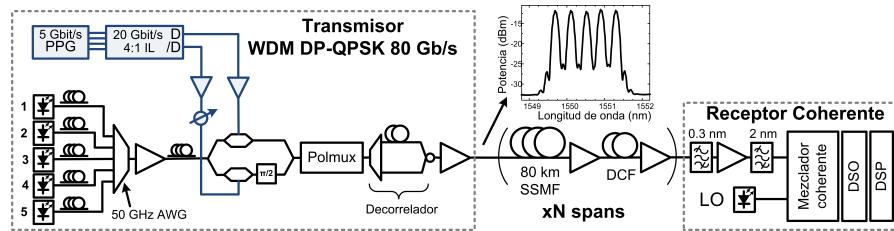


Figura 7.11: Esquema del banco de pruebas desarrollado por la Universidad de Dinamarca para la validación experimental del estimador cognitivo.

dicador EVM. La relación entre la QoT y el EVM es tal que un menor valor de EVM implica una mayor calidad de transmisión. Por tanto, fijando un valor umbral de EVM, si el valor EVM estimado para el *lightpath* es menor que dicho umbral, entonces se considera que dicho *lightpath* tiene una calidad alta, mientras que si el valor obtenido está por encima de dicho umbral la calidad es baja. A partir de 153 datos experimentales, se han utilizado 135 de ellos seleccionados aleatoriamente para poblar la base de conocimiento y 18 para ser clasificados por el estimador de calidad cognitivo. Además, las bases de conocimiento no se han optimizado. Este proceso se ha repetido 10 veces para realizar un análisis estadístico, obteniendo la media y los intervalos de confianza al 95 %. Al no disponer de otras herramientas que estimen el valor de EVM de una conexión antes de su establecimiento, se ha comparado el funcionamiento del estimador cognitivo con un estimador que siempre decide según la clase más probable para distintos valores umbrales de EVM. Esto es, este estimador predice que un *lightpath* tiene buena calidad si para el umbral fijado la mayoría de los *lightpaths* tienen un valor de EVM por debajo de dicho umbral. En caso contrario decidirá que la conexión óptica no tiene la calidad necesaria. En concreto, se ha analizado la tasa de aciertos en las clasificaciones en los dos estimadores. Los resultados demostraron que el estimador cognitivo basado en CBR mejoraba desde un 5.8 % hasta el 29 % la tasa de aciertos obtenida apostando por la clase más probable y dependiendo del valor umbral escogido. Así mismo, se estudió también la tasa de aciertos en función del tamaño de la base de conocimiento. En este sentido, el estimador aumenta la tasa de aciertos desde el 70 % cuando solo se consideran 20 casos, hasta el 83 % para una base de conocimiento de 152 casos.

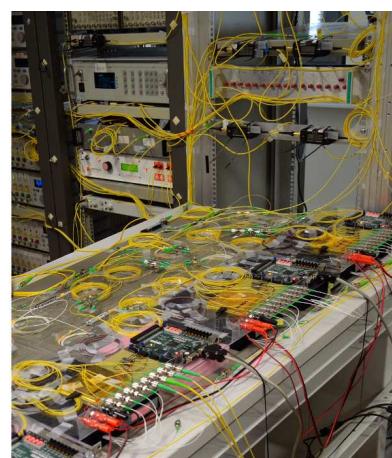
Para demostrar que el estimador cognitivo es lo suficientemente genérico para ser usado con otros indicadores de la calidad de transmisión, se llevó a cabo el mismo análisis considerando la OSNR. En este sentido, se fija un valor umbral de OSNR y si el valor de OSNR estimado es mayor que dicho umbral, se considera que el *lightpath* tiene buena calidad. Utilizando este parámetro, el estimador alcanza una tasa de aciertos del 100 % para todos los umbrales de OSNR con una base de conocimiento de 135 casos.

Además, la variación de la tasa de aciertos con el tamaño de la base de conocimiento, es de un 67 % con tan solo 20 casos, hasta el 100 % de aciertos para bases de conocimiento de 120 casos o más.

Así mismo, el estimador se ha validado experimentalmente en el banco de pruebas del proyecto CHRON (Figura 7.12) [177]. En este sentido, se demostraron los beneficios del uso de cognición con el estimador de QoT. En concreto, ante un escenario con fallos en enlaces, gracias al aprendizaje y la incorporación de nuevos casos a la base de conocimiento del estimador, se reduce en un 48 % el tiempo empleado en la recuperación de un camino óptico cuando se compara con un sistema sin cognición.



(a)



(b)

Figura 7.12: Banco de pruebas del proyecto CHRON.

7.5. Desarrollo de técnicas cognitivas para la sintonización de PIDs en redes de acceso LR-EPON

En esta segunda parte del capítulo, se va a analizar la inclusión de cognición en el segmento de la red de acceso. En concreto, se va a añadir un sistema de razonamiento basado en casos para llevar a cabo la sintonización adaptativa de un controlador PID. Como ejemplo, se va a incorporar dicho sistema cognitivo al algoritmo SPID, encargado de la asignación de ancho de banda en redes LR-EPON para garantizar distintos niveles de ancho de banda estipulados. En este sentido, y tal y como se ha visto en capítulos anteriores, ni la sintonización manual mediante el método de Ziegler-Nichols, ni la sintonización automática basada en algoritmos genéticos (GA-SPID), llevan a cabo una sintonización adaptativa. En el capítulo anterior, se desarrolló un método de sintonización adaptativa basada en redes neuronales (NN-SPID). Este algoritmo, modifica los parámetros de sintonización dinámicamente en tiempo real con el fin de adaptar la sintonización a las distintas circunstancias de la red, presentando por lo tanto un comportamiento más robusto y estable. Siguiendo esta idea, en este apartado se va a presentar un nuevo algoritmo que lleva a cabo una sintonización adaptativa, pero a partir del conocimiento adquirido en experiencias pasadas gracias al uso de un sistema CBR.

7.5.1. Integración de la técnica CBR para la sintonización adaptativa de SPID. Algoritmo CBR-SPID

En este apartado se va a proceder a describir el proceso de sintonización cognitivo para llevar a cabo una sintonización adaptativa del controlador PID utilizado por SPID para la asignación de ancho de banda. El nuevo algoritmo, denominado CBR-SPID (*Case-Based Reasoning Service level agreement PID*), integra un sistema de razonamiento basado en casos para realizar este proceso. El esquema general del proceso es el que se muestra en la Figura 7.13. Tal y como se observa, la principal novedad frente a los algoritmos presentados anteriormente, es que los valores de los parámetros para la sintonización del PID (K_p , T_i y T_d) son proporcionados por el CBR. Para ello, el CBR consulta la base de conocimiento, donde se han almacenado experiencias previas obtenidas a partir del algoritmo GA-SPID.

En primer lugar, se ha construido una base de conocimiento que se ha poblado con diferentes casos que representan posibles escenarios de red. Cada uno de los casos está compuesto por una descripción del escenario y el individuo propuesto por el algoritmo genético de GA-SPID para sintonizar el PID bajo dichas circunstancias. Esto es, antes del funcionamiento dinámico de la red, se han realizado unas simulaciones previas

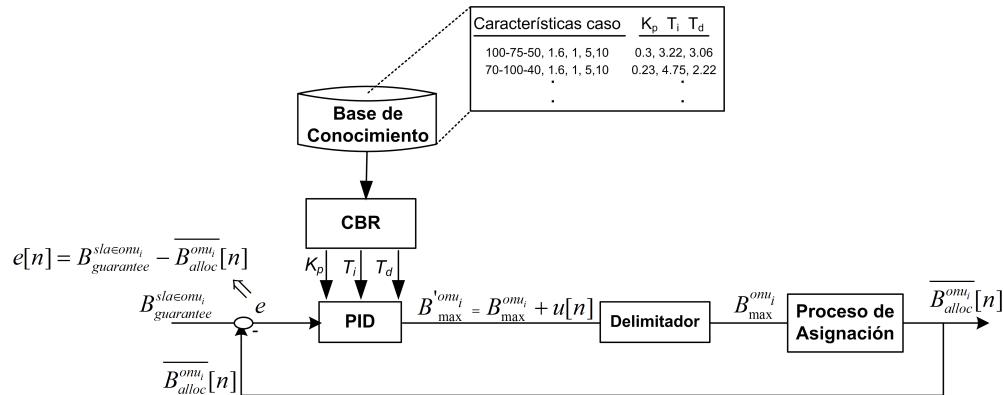


Figura 7.13: Diagrama de bloques propuesto para la sintonización adaptativa del controlador PID en SPID mediante un sistema CBR.

con el algoritmo GA-SPID para distintos escenarios de red. Las soluciones propuestas por el algoritmo genético para cada una de estas situaciones de red, son las que conforman los casos de la base de conocimiento. Específicamente, un caso está descrito por los siguientes parámetros: el nivel de ancho de banda garantizado para cada uno de los SLAs, la carga total de red y el número de ONUs asociadas a cada perfil. Por ejemplo, si se consideran tres perfiles de abonado, con unos niveles garantizados de 100, 75, y 50 Mbit/s, para la máxima carga de red de 1.6 y con 1, 5 y 10 ONUs por cada SLA, el caso vendría representado por el valor de estos siete parámetros (tal y como se aprecia en la primera línea de la base de conocimiento de la Figura 7.13). Además, junto a la descripción del caso se guardan los parámetros K_p , T_i y T_d del individuo óptimo proporcionado por el algoritmo genético para dicha configuración de red.

Por tanto, el funcionamiento del CBR es el siguiente. Cuando se produce un cambio en la red, por ejemplo una modificación en los anchos de banda a garantizar o en el número de ONUs pertenecientes a cada SLA, el CBR proporciona los valores óptimos de los parámetros del PID (K_p , T_i y T_d) para esa situación concreta. Así pues, el CBR en primer lugar recupera el caso más similar al nuevo escenario donde se pretende sintonizar el PID. Para ello, las características del nuevo escenario se comparan con las características de todos los casos de la base de conocimiento y se selecciona el caso más similar.

Para calcular la similitud entre el nuevo caso (x) y cada uno de los casos contenidos en la base de conocimiento (y), los valores de cada una de las características del caso se normalizan y se calcula la distancia Euclídea ponderada siguiendo la ecuación (7.1) (Apartado 7.2 de este capítulo). En este caso se ha supuesto que el nivel de ancho de banda garantizado y el número de ONUs de cada SLA tienen una mayor importancia a la hora de obtener el caso más similar que la carga de red. Por tanto, todas las características

tienen un peso asociado (W_a) igual a 1 excepto el peso de la carga total de red que se ha reducido a la mitad ($W_a = 0.5$). Una vez calculada la similitud y recuperado el caso más similar, éste se va a reutilizar. En concreto, los valores K_p , T_i y T_d propuestos para la sintonización del nuevo escenario son iguales que los del caso recuperado. Por último, en esta primera aproximación del sistema diseñado e implementado no se va a incorporar el nuevo caso a la base de conocimiento, esto es, no existe aprendizaje. Por tanto, el conocimiento no se incrementa y solo se disponen de los individuos propuestos por el algoritmo genético para cada situación.

7.5.2. Estudio de simulación del algoritmo CBR-SPID

Para comprobar el funcionamiento del algoritmo CBR-SPID, se va a proceder en primer lugar a la descripción de la construcción de la base de conocimiento del CBR. Después se analizarán las prestaciones del algoritmo mediante simulaciones en escenarios donde se producen variaciones en tiempo real del ancho de banda garantizado de cada perfil de abonado y los resultados se compararán con NN-SPID, que también realiza una sintonización adaptativa.

7.5.2.1. Escenario de simulación

Para llevar a cabo la simulación del algoritmo CBR-SPID, se ha utilizado el entorno de simulación genérico del algoritmo SPID (Apartado 4.4.1 del Capítulo 4).

La Tabla 7.3 recoge un resumen de los principales parámetros de red y los valores considerados. No obstante, para mostrar el proceso de sintonización adaptativo, los anchos de banda estipulados por el proveedor de servicios para cada SLA se modificarán a lo largo del tiempo.

7.5.2.2. Construcción de la base de conocimiento

Para la construcción de la base del conocimiento se ha partido de los individuos proporcionados como soluciones buenas del algoritmo genético para la combinación de distintos valores de ancho de banda. En concreto, se han utilizado las combinaciones realizadas en el Apartado 5.3 del Capítulo 5, esto es, 100/75/50 Mbit/s, 70/100/40 Mbit/s, 90/70/55 Mbit/s y 60/90/45 Mbit/s. Además se han incluido otros individuos óptimos obtenidos para estas configuraciones bajo distintas cargas de red. Para este ejemplo, solo se ha considerado una configuración del número de ONUs igual a 1, 5 y 10 para el SLA₀, SLA₁ y SLA₂, respectivamente. El número de casos incluidos en la base del conocimiento es un número reducido (un total de 8 casos). Esto es así porque el número

Parámetros de simulación	Valores
Número de ONUs	16 ONUs (N_{onu})
Tasa de transmisión canal ascendente	1 Gbit/s ($EPON_Line_rate$)
Tasa de transmisión ONU-abonado	100 Mbit/s ($User_Line_rate$)
RTT	1 ms
Tiempo de guarda	1 μ s
Tiempo de ciclo máximo	2 ms
Distribución de tráfico	Fuente de Pareto H=0.8
Tamaño de paquete	Variable entre 64-1500 bytes (más 38 bytes de cabeceras)
<i>Service Level Agreement (SLA)</i>	SLA ₀ : 1 ONU SLA ₁ : 5 ONUs SLA ₂ : 10 ONUs
Tiempo muestreo del PID (T)	1 s
Configuración pesos iniciales	$W^{sla_0} = W^{sla_1} = W^{sla_2} = 1$

Tabla 7.3: Parámetros de red considerados en el escenario de simulación de CBR-SPID.

de características a variar no es muy elevado, a diferencia de lo que ocurría en el CBR del estimador cognitivo de QoT presentado en la primera parte de este capítulo. Así mismo, tal y como se explica en la siguiente sección, el análisis de prestaciones se quiere llevar a cabo probando casos que no están contenidos en la base de conocimiento. Por todo ello, el número de posibles casos es reducido.

7.5.2.3. Análisis de prestaciones de CBR-SPID

Para comprobar la funcionalidad del algoritmo CBR-SPID, se van a llevar a cabo simulaciones donde los niveles de ancho de banda garantizados a cada SLA varían a lo largo del tiempo. En concreto, se han escogido dos escenarios, en los que el ancho de banda es distinto y además los niveles cambian cada 300 s. Además, los niveles garantizados son diferentes a los existentes en la base de conocimiento. De esta forma, se van a comprobar los beneficios del uso de la sintonización adaptativa del CBR, comparado con el método de Ziegler-Nichols y con el algoritmo NN-SPID, cuando no existe exactamente el mismo caso en la base del conocimiento. En concreto los valores escogidos para cada escenario aparecen recogidos en la Tabla 7.4.

En concreto, la Figura 7.14 (a), (b) y (c) muestra la evolución en tiempo real de la desviación del ancho de banda medio asignado frente al garantizado para el Escenario 1 y para el SLA₀, SLA₁ y SLA₂, respectivamente cuando se compara el algoritmo SPID (sintonizado con el método de Ziegler-Nichols) con NN-SPID y CBR-SPID. Tal y como puede observarse, la desviación obtenida mediante la sintonización adaptativa

Tiempo (s)	Nivel de ancho de banda garantizado a cada SLA (Mbit/s)					
	Escenario 1			Escenario 2		
	SLA ₀	SLA ₁	SLA ₂	SLA ₀	SLA ₁	SLA ₂
0 - 300	100	85	45	85	55	60
300 - 600	65	95	40	95	80	45
600 - 900	95	75	50	80	70	55
900 - 1200	75	65	60	100	90	40

Tabla 7.4: Niveles de ancho de banda garantizado considerados para cada SLA a lo largo del tiempo en los dos escenarios analizados.

con NN-SPID y CBR-SPID es menor que la obtenida con la sintonización mediante el método de Ziegler-Nichols. Esto se debe a la ventaja de la sintonización adaptativa que presentan ambos algoritmos y que modifica las constantes de sintonización en tiempo real y de acuerdo a las condiciones del nuevo escenario de red.

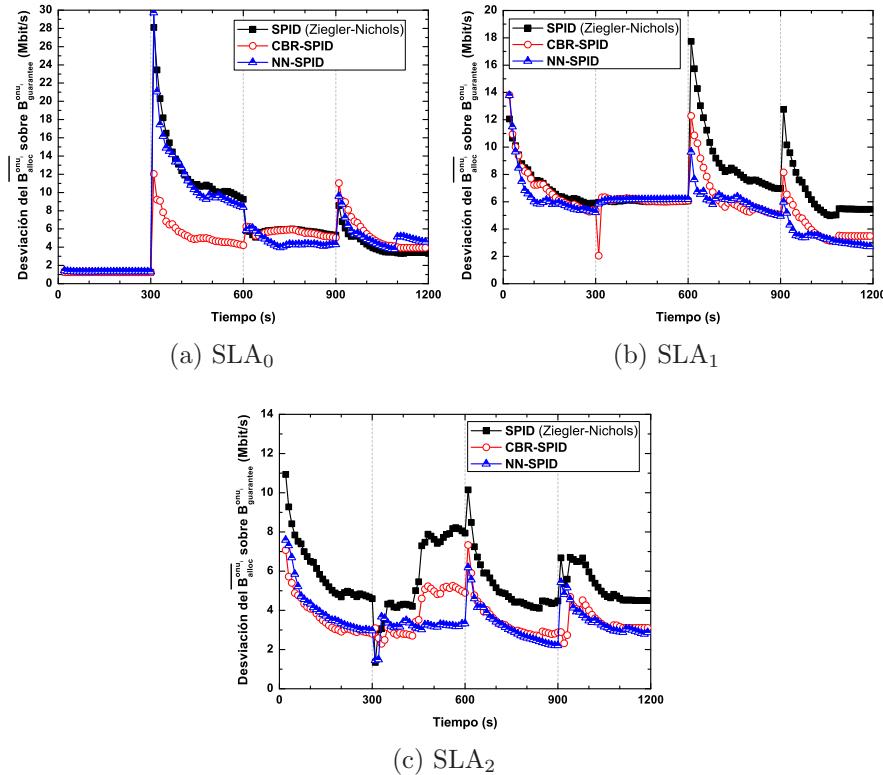


Figura 7.14: Variación en tiempo real de la desviación del ancho de banda medio asignado sobre el ancho de banda garantizado para las condiciones de red del Escenario 1. (a) SLA₀ (b) SLA₁ (c) SLA₂.

Tiempo	Desviación de B_{alloc}^{onui} sobre $B_{guarantee}^{sla \in onui}$ (Mbit/s)							
	SLA ₀		SLA ₁		SLA ₂		Total	
300 s	SPID	1.25	SPID	5.89	SPID	4.6	SPID	11.74
	CBR-SPID	1.25	CBR-SPID	5.35	CBR-SPID	2.80	CBR-SPID	9.4
	NN-SPID	1.32	NN-SPID	5.22	NN-SPID	2.96	NN-SPID	9.5
600 s	SPID	9.26	SPID	6.07	SPID	7.94	SPID	23.27
	CBR-SPID	4.21	CBR-SPID	6.08	CBR-SPID	4.89	CBR-SPID	15.18
	NN-SPID	8.4	NN-SPID	6.17	NN-SPID	3.32	NN-SPID	17.89
900 s	SPID	5.31	SPID	6.97	SPID	4.48	SPID	16.76
	CBR-SPID	5.03	CBR-SPID	5.06	CBR-SPID	2.87	CBR-SPID	12.96
	NN-SPID	4.29	NN-SPID	4.94	NN-SPID	2.21	NN-SPID	11.44
1200 s	SPID	3.4	SPID	5.43	SPID	4.49	SPID	13.32
	CBR-SPID	3.93	CBR-SPID	3.48	CBR-SPID	3.1	CBR-SPID	10.51
	NN-SPID	4.57	NN-SPID	2.72	NN-SPID	2.87	NN-SPID	10.16

Tabla 7.5: Desviación del ancho de banda medio asignado sobre el ancho de banda garantizado en 300 s, 600 s, 900 s y 1200 s para las condiciones del Escenario 1 comparando SPID, CBR-SPID y NN-SPID.

Para ver de forma más exacta las diferencias entre los resultados de las tres técnicas de sintonización, la Tabla 7.5 muestra la desviación del ancho de banda medio asignado sobre el ancho de banda garantizado para cada SLA y la desviación total, al finalizar el tramo correspondiente a cada configuración de ancho de banda (esto es, 300 s, 600 s, 900 s y 1200 s). Están recogidos los resultados obtenidos para la sintonización con el método de Ziegler-Nichols (SPID), con la sintonización realizada por el CBR (CBR-SPID) y con las redes neuronales (NN-SPID). Los resultados obtenidos refuerzan lo visualizado en las gráficas anteriores. Tal y como se observa, con una sintonización adaptativa, bien con CBR-SPID o NN-SPID, se obtienen menores niveles de desviación para todos los SLAs y por tanto una menor desviación total acumulada en todas las combinaciones de niveles garantizados consideradas.

En cuanto a las diferencias entre NN-SPID y CBR-SPID, se puede observar cómo éstas son muy pequeñas. En general, ambos algoritmos obtienen resultados similares. Esto se debe a que los dos algoritmos utilizan las soluciones del algoritmo genético, en el algoritmo CBR-SPID para poblar la base de conocimiento y en el algoritmo NN-SPID para entrenar la red neuronal. Por tanto, al tener ambos algoritmos la misma base, las variables de sintonización son muy parecidas. La principal diferencia entre ambas técnicas es que en CBR-SPID solo se produce un cambio en las variables de sintonización cuando se realiza un cambio en alguna de las características de sus casos (niveles garantizados, carga de red o número de ONUs de cada SLA). Por el contrario, NN-SPID puede cambiar periódicamente el valor de los parámetros de sintonización si se producen variaciones en el error cometido. Esto puede verse en la Figura 7.15, donde aparecen representados

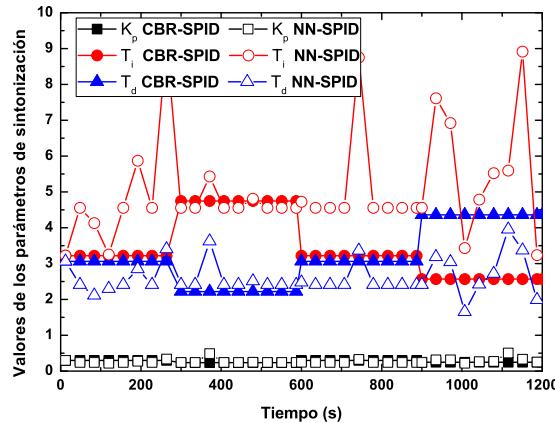


Figura 7.15: Evolución en tiempo real de los parámetros de sintonización en los algoritmos CBR-SPID y NN-SPID para el Escenario 1.

los valores de sintonización a lo largo del tiempo para NN-SPID y CBR-SPID bajo las condiciones de red del Escenario 1. Por tanto, en principio NN-SPID puede resultar más robusto, ya que reacciona a todos aquellos cambios que producen incrementos en el error cometido a la hora de asegurar los diferentes niveles de ancho de banda, mientras que CBR-SPID solo reacciona si se producen cambios en las características que definen los casos.

Respecto a las condiciones de red del Escenario 2 de la Tabla 7.4, en la Figura 7.16 (a), (b) y (c) está representada la evolución instantánea de la desviación del ancho de banda medio asignado sobre el nivel garantizado para el SLA₀, SLA₁ y SLA₂, respectivamente, comparando los tres algoritmos. Las conclusiones son similares a las obtenidas en el Escenario 1, esto es, la sintonización que peores resultados ofrece es la conseguida mediante el método de Ziegler-Nichols. Los métodos de sintonización adaptativa ofrecen mejores resultados, aunque, al igual que en el Escenario 1, las diferencias entre NN-SPID y CBR-SPID son muy pequeñas.

Del mismo modo que para el Escenario 1, la Tabla 7.6 recoge los resultados de la desviación acumulada de cada SLA y la suma total, justo antes de cambiar los niveles garantizados en cada tramo, para las tres estrategias de sintonización. En este caso, se observa también que los mejores resultados se alcanzan con la sintonización adaptativa, apreciándose pocas diferencias entre los métodos adaptativos, NN-SPID y CBR-SPID.

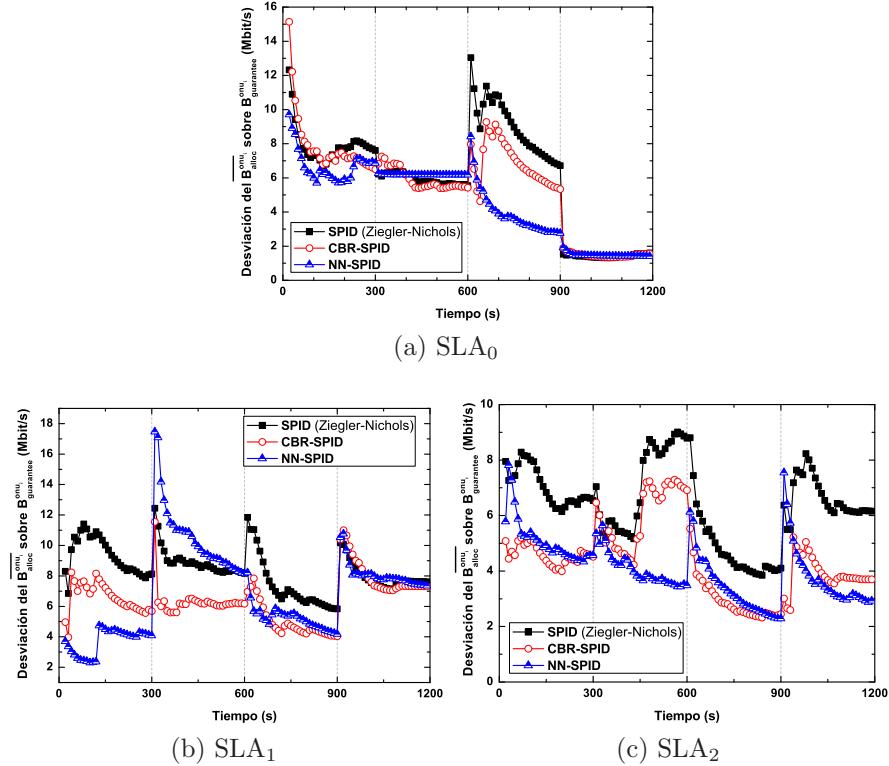


Figura 7.16: Variación en tiempo real de la desviación del ancho de banda medio asignado sobre el ancho de banda garantizado para las condiciones de red del Escenario 2. (a) SLA₀ (b) SLA₁ (c) SLA₂.

Tiempo	Desviación de $B_{alloc}^{onu_i}$ sobre $B_{garantee}^{onu_i}$ (Mbit/s)							
	SLA ₀		SLA ₁		SLA ₂		Total	
300 s	SPID	7.6	SPID	8.13	SPID	6.53	SPID	22.26
	CBR-SPID	6.5	CBR-SPID	5.86	CBR-SPID	4.51	CBR-SPID	16.87
	NN-SPID	6.86	NN-SPID	4.08	NN-SPID	4.56	NN-SPID	15.5
600 s	SPID	5.54	SPID	8.16	SPID	8.79	SPID	22.49
	CBR-SPID	5.42	CBR-SPID	6.17	CBR-SPID	6.91	CBR-SPID	18.5
	NN-SPID	6.18	NN-SPID	8.16	NN-SPID	3.48	NN-SPID	17.82
900 s	SPID	6.71	SPID	5.83	SPID	4.1	SPID	16.64
	CBR-SPID	5.33	CBR-SPID	4.03	CBR-SPID	2.44	CBR-SPID	11.8
	NN-SPID	2.76	NN-SPID	4.18	NN-SPID	2.28	NN-SPID	9.32
1200 s	SPID	1.62	SPID	7.61	SPID	6.12	SPID	15.35
	CBR-SPID	1.58	CBR-SPID	7.31	CBR-SPID	5.01	CBR-SPID	13.9
	NN-SPID	1.41	NN-SPID	7.35	NN-SPID	2.92	NN-SPID	11.68

Tabla 7.6: Desviación del ancho de banda medio asignado sobre el ancho de banda garantizado en 300 s, 600 s, 900 s y 1200 s para las condiciones del Escenario 2 comparando SPID, CBR-SPID y NN-SPID.

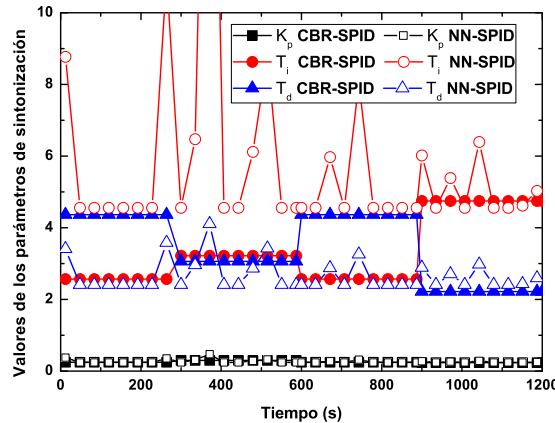


Figura 7.17: Evolución en tiempo real de los parámetros de sintonización en CBR-SPID y NN-SPID para las condiciones del Escenario 2.

Por último, la Figura 7.17, muestra la evolución de los valores de los parámetros de sintonización para los algoritmos CBR-SPID y NN-SPID en el Escenario 2. Al igual que para el Escenario 1, en el CBR solo se producen cambios en los parámetros cuando se produce un cambio en alguna de las características del caso, mientras que en NN-SPID los valores se modifican periódicamente siempre que se produzcan modificaciones en el error cometido.

Por tanto, se ha demostrado que la sintonización adaptativa ofrece un mejor comportamiento ya que adecua los parámetros de sintonización al escenario. En cuanto a las técnicas de sintonización adaptativa desarrolladas en esta Tesis, ambas obtienen resultados similares debido principalmente a que utilizan como base de conocimiento y casos de entrenamiento datos óptimos obtenidos mediante la simulación el algoritmo genético.

7.5.3. Análisis de prestaciones de CBR-SPID bajo un patrón de tráfico dinámico

Ya que el algoritmo CBR-SPID es un algoritmo adaptativo que modifica los parámetros de sintonización a lo largo del tiempo, se va a analizar su comportamiento bajo un patrón de tráfico dinámico más realista en un contexto de red real. Para ello, se ha considerado la variación dinámica de la carga de la ONU utilizada en capítulos anteriores, y que viene dada por la ecuación (3.9) del Capítulo 3. Además, se han considerado dos frecuencias de cambio de carga, cada 1800 s y otra más rápida cada 900 s.

En concreto, la Figura 7.18 (a), (b) y (c) representa la variación del ancho de banda medio asignado a lo largo del tiempo para una ONU del SLA₀, SLA₁ y SLA₂, respectivamente, cuando la frecuencia de los cambios en la carga es de 1800 s. En azul

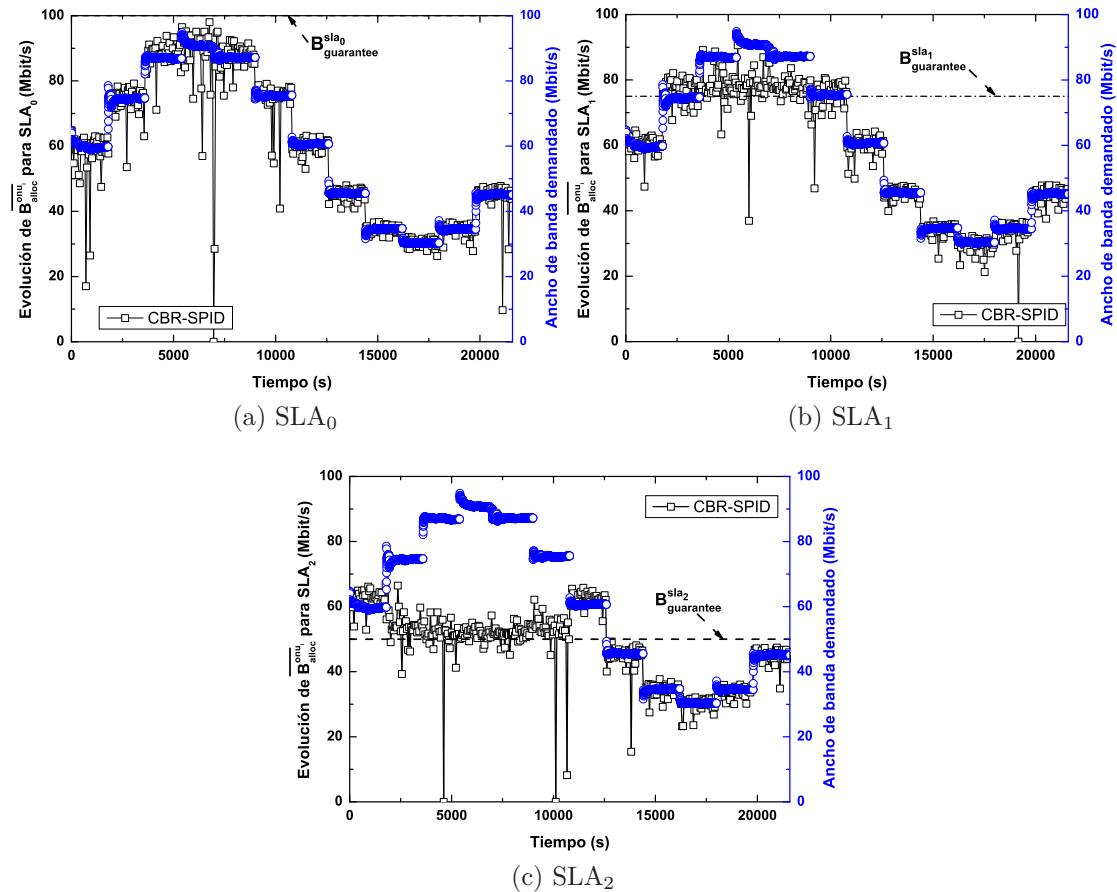


Figura 7.18: Ancho de banda demandado para una ONU de cada SLA (azul) y variación en tiempo real del ancho de banda medio asignado (negro) considerando una frecuencia de cambio de 1800 s. (a) SLA₀ (b) SLA₁ (c) SLA₂.

está representada la variación de la carga de ONU en Mbit/s referida al eje de la derecha, también en azul. En negro aparece la variación en tiempo real del ancho de banda medio asignado en CBR-SPID (eje de la izquierda). Tal y como puede observarse, para los tres SLAs el ancho de banda medio asignado se corresponde con el ancho de banda demandado, siempre que existan recursos para satisfacer la demanda total. En el caso de que la demanda total del ancho de banda sea mayor que la capacidad del canal, CBR-SPID asigna de forma óptima el nivel mínimo garantizado a cada SLA (100, 75 y 50 Mbit/s respectivamente).

Por otro lado, para ver la adaptación en tiempo real de la sintonización, la Figura 7.19 muestra la variación de los parámetros de sintonización K_p , T_i y T_d a lo largo del tiempo al considerar cambios de la carga cada 1800 s. Así mismo, referido al eje de la derecha en

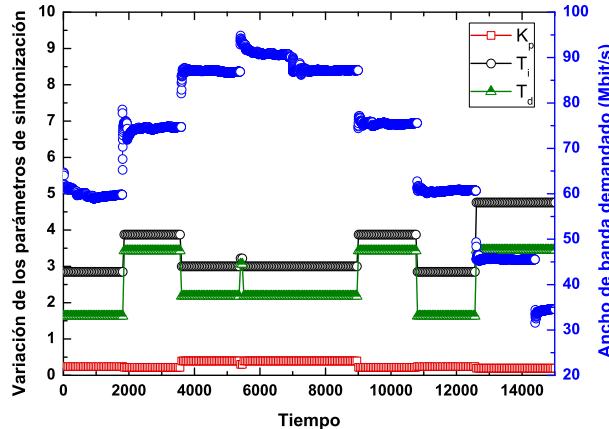


Figura 7.19: Ancho de banda demandado por cada ONU (azul) y variación en tiempo real de los parámetros de sintonización K_p , T_i y T_d considerando una frecuencia de cambio de 1800 s.

azul, aparece el ancho de banda demandado por cada ONU. Para facilitar la visualización de los datos, solo se han representado los primeros 15000 s de simulación. Tal y como se aprecia, los valores de los distintos parámetros varían a lo largo del tiempo para ofrecer una sintonización óptima en función de la carga.

Para analizar la adaptación que realiza el algoritmo CBR-SPID con cambios de la carga más rápidos, la Figura 7.20 (a), (b) y (c) representa la evolución del ancho de banda medio asignado (negro) según la variación del ancho de banda demandado por una ONU (azul) para el SLA₀, SLA₁ y SLA₂, respectivamente cuando se consideran cambios en la carga cada 900 s. Tal y como se aprecia, CBR-SPID adapta de forma óptima el ancho de banda medio asignado al nivel garantizado por el proveedor de servicios para cada SLA cuando no se puede ofrecer todo el ancho de banda demandado.

Así mismo, en la Figura 7.21 donde está representada la variación de los parámetros de sintonización con el ancho de banda demandado, se puede apreciar una modificación de las tres variables similar a la obtenida para una frecuencia de cambio de 1800 s. Esto es, cada 900 s el CBR proporciona nuevos parámetros de sintonización adaptados a las nuevas circunstancias de red.

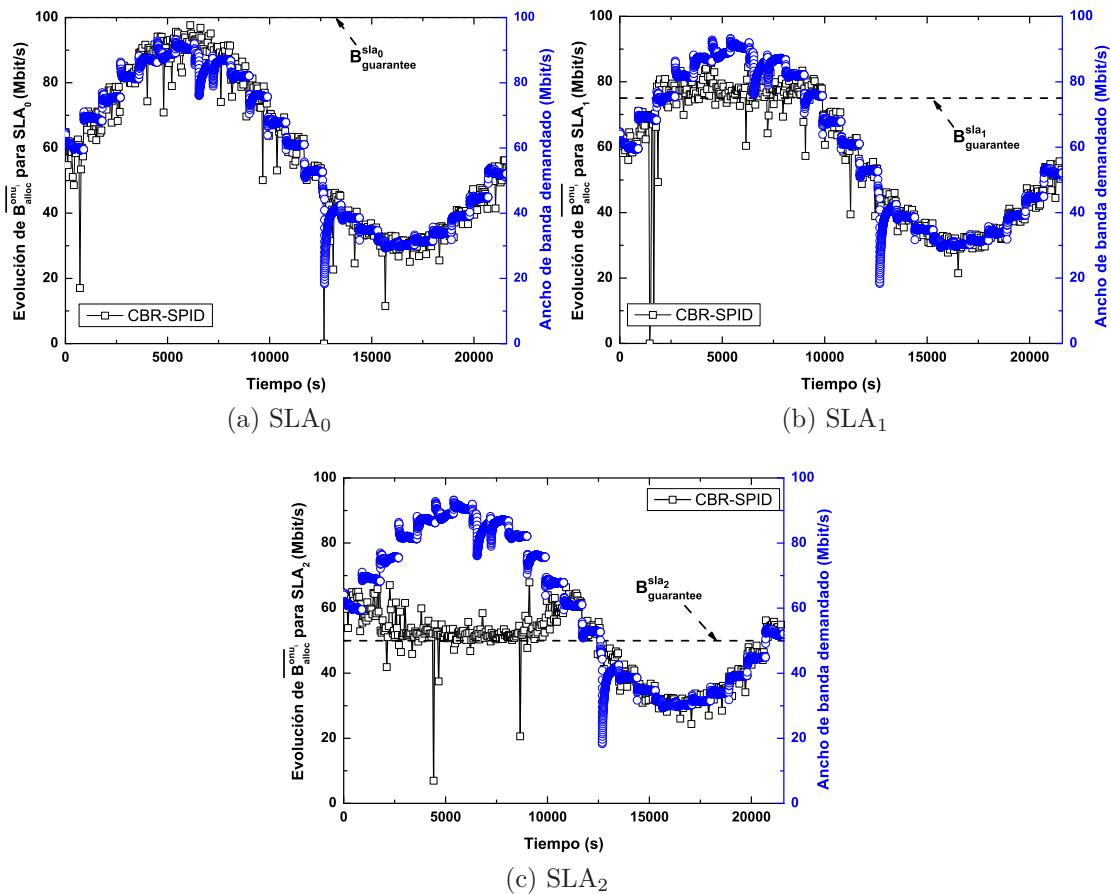


Figura 7.20: Ancho de banda demandado para una ONU de cada SLA (azul) y variación en tiempo real del ancho de banda medio asignado (negro) considerando una frecuencia de cambio de 900 s. (a) SLA₀ (b) SLA₁ (c) SLA₂.

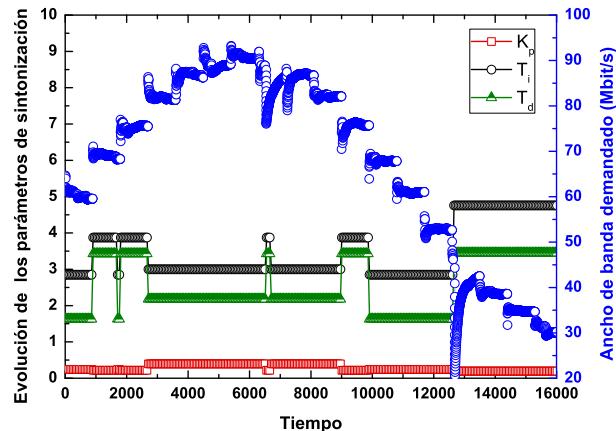


Figura 7.21: Ancho de banda demandado por cada ONU (azul) y variación en tiempo real de los parámetros de sintonización K_p , T_i y T_d considerando una frecuencia de cambio de 900 s.

7.6. Conclusiones

En este capítulo se han diseñado e implementado técnicas cognitivas en la gestión de recursos y de calidad de servicio en redes ópticas troncales y en las redes de acceso EPON y LR-EPON. En concreto, se ha utilizado la técnica de razonamiento basado en casos (CBR), para tomar decisiones a partir de experiencias pasadas.

En la primera parte de este capítulo, se ha desarrollado un estimador cognitivo que puede utilizarse en redes ópticas troncales de encaminamiento por longitud de onda dinámico para determinar la calidad de transmisión. En particular, se ha propuesto en primer lugar un novedoso estimador cognitivo (R-CBR) que es capaz de determinar antes de su establecimiento si un *lightpath* cumple con los requisitos de transmisión establecidos. Se ha demostrado que es capaz de obtener un alto porcentaje de aciertos (más del 99 % en la red DT y alrededor del 98 % en la red GÉANT2) y además es mucho más rápido que la *Q-Tool* (cerca de tres órdenes de magnitud para la red DT y más de un orden de magnitud para la red GÉANT2). Esta mejora en el tiempo de computación supone una gran ventaja cuando es necesario decidir si se establece el nuevo *lightpath* en escenarios donde existen limitaciones de tiempo, ya que no solo se estima la calidad éste, sino también de todos aquellos *lightpaths* copropagativos cuya calidad puede verse afectada por el establecimiento de este nuevo camino óptico.

Además el estimador R-CBR se ha mejorado gracias a la incorporación de técnicas de aprendizaje y olvido para optimizar la base de conocimiento subyacente. Por tanto, se ha diseñado un nuevo estimador, FixE-CBR, que alcanza una tasa de acierto similar o ligeramente superior comparada con R-CBR, pero que consigue una reducción significativa del número de casos almacenados en la base de conocimiento, lo que se traduce en una importante reducción del tiempo de computación. En este sentido, FixE-CBR es aproximadamente un orden de magnitud más rápido que R-CBR para ambas redes y aproximadamente cuatro y dos órdenes de magnitud más rápido que la *Q-Tool* para las redes DT y GÉANT2, respectivamente. Aunque solo se ha llevado a cabo una optimización de la base de conocimiento previa a la puesta en marcha del estimador, las técnicas de aprendizaje y olvido presentadas pueden utilizarse para que el estimador cognitivo se auto-adapte, en tiempo real, a los cambios en el entorno. En este sentido, la base de conocimiento subyacente evolucionaría para reflejar cambios en la red tales como el envejecimiento de los componentes y su deterioro.

Por otro lado, se han descrito dos métodos pragmáticos para poblar la base de conocimiento del estimador de QoT cognitivo antes de la puesta en marcha de una nueva red. En concreto, se ha mostrado que testeando un número reducido de configuraciones de red antes de comenzar con su operación dinámica, se puede recoger suficiente

información para construir una base de conocimiento adecuada para obtener una tasa de aciertos elevada (mayor del 99 % en la red DT).

También se ha descrito una prueba experimental real del QoT diseñado, que demuestra que es lo suficientemente genérico para realizar la estimación de la calidad de transmisión a partir de diferentes parámetros. Así mismo, se han demostrado experimentalmente los beneficios de incorporar el estimador cognitivo con aprendizaje.

Por otro lado, en la segunda parte del capítulo, el sistema CBR se ha incorporado en redes de acceso LR-EAPON para llevar a cabo una sintonización adaptativa, en concreto orientada al algoritmo SPID. Por tanto, el nuevo algoritmo, denominado CBR-SPID, es capaz de modificar el valor de los parámetros de sintonización del PID (K_p , T_i y T_d) a lo largo del tiempo para asegurar una sintonización óptima para cada situación de red. Para ello, se utilizan como experiencias pasadas los individuos que el algoritmo genético ha considerado óptimos para cada situación de red. Por tanto, se combina el uso de un método de sintonización automático, como el ofrecido por GA-SPID, con el uso de algoritmos genéticos con el CBR para conseguir una sintonización óptima adaptativa.

En este sentido, el algoritmo CBR-SPID ha demostrado obtener una menor desviación del ancho de banda medio asignado frente al nivel de ancho de banda garantizado para todos los SLAs cuando los niveles estipulados por el proveedor de servicios varían a lo largo del tiempo al comparar con métodos no adaptativos. No obstante, los resultados son similares al método de sintonización adaptativa basado en redes neuronales (NN-SPID) desarrollado en el capítulo anterior.

Además, debido a la naturaleza adaptativa del controlador, se han llevado a cabo simulaciones bajo patrones de tráfico dinámico. De esta forma se ha comprobado que CBR-SPID consigue una asignación de ancho de banda adecuada de acuerdo a las modificaciones producidas en la red. En concreto, las constantes de sintonización del controlador se modifican según la carga de la red.

Por tanto, se ha demostrado en dos ámbitos diferentes, en redes troncales para la estimación de la calidad de transmisión y redes de acceso LR-EAPON para la sintonización adaptativa del PID en la asignación de ancho de banda, que la incorporación de técnicas cognitivas ayuda a realizar una gestión eficiente de los recursos disponibles. Por tanto, el uso de técnicas cognitivas mejora sustancialmente tanto la adaptabilidad a diferentes condiciones de red, como la eficiencia en la gestión de los recursos de las redes, consiguiendo de esta forma un funcionamiento óptimo.

Capítulo 8

Conclusiones y Líneas Futuras de Investigación

8.1. Conclusiones

La calidad de servicio y la gestión eficiente de los recursos de una red son dos aspectos muy importantes que los proveedores de servicio tienen que tener en cuenta. En este sentido, en esta Tesis Doctoral se han diseñado e implementado una serie de técnicas adaptativas y cognitivas, tanto para el segmento de la red de acceso como para la red troncal, con el fin de llevar a cabo una gestión eficiente de los recursos disponibles y poder ofrecer una calidad de servicio y de transmisión adecuada a los requisitos estipulados por los proveedores de servicio. A continuación, se presentan las conclusiones más relevantes en ambos tramos de red.

8.1.1. Red de acceso LR-EPON

En primer lugar, la Tesis se ha centrado principalmente en el segmento de la red de acceso. Este segmento, se convirtió en un cuello de botella de las redes de telecomunicaciones, debido al gran incremento del tráfico y de los tipos servicios, unido a la poca capacidad que ofrecían las técnicas comúnmente desplegadas hasta hace unos años. En este sentido, actualmente se están desplegando a gran escala a nivel mundial las redes ópticas pasivas (PON), que utilizan fibra óptica como medio de transmisión, solventando la falta de capacidad en este segmento de la red. En concreto, esta Tesis se ha centrado en el estudio de las redes ópticas pasivas de largo alcance (LR-EPON), una arquitectura de red muy prometedora ya que resultan más económicas en su despliegue que las redes PON típicas. Esto se debe al hecho de que al incrementar su alcance hasta 100 km,

simplifican la red de acceso y pueden dar cobertura a un mayor número de usuarios con mayores niveles de capacidad.

Así mismo, se ha realizado un análisis exhaustivo de diversas arquitecturas existentes en la literatura para el despliegue de las redes LR-PON. En este sentido e independientemente de la arquitectura elegida, las redes PON de largo alcance se pueden ver como un conjunto de redes PON independientes, y por tanto presentan una arquitectura punto-multipunto, donde en el canal ascendente es necesario un mecanismo de control de acceso al medio para gestionar eficientemente la transmisión de tráfico de las ONUs hacia el OLT. Los algoritmos de asignación dinámica de ancho de banda (DBA), son los que mejores prestaciones ofrecen, ya que realizan una distribución del ancho de banda dependiendo de la demanda actual de cada ONU y de los requisitos de calidad de servicio. No obstante, en las redes LR-PON los mecanismos de control de acceso al medio tienen que hacer frente a las dificultades añadidas que aparecen debido al incremento de la distancia extremo a extremo. Por ello, en esta Tesis se ha trabajado en el diseño de algoritmos que gestionen dinámicamente la asignación de ancho de banda en redes LR-EPON basadas en el estándar Ethernet, solventando no solo las dificultades derivadas del gran incremento de la distancia, sino que también sean capaces de proporcionar niveles de calidad de servicio adecuados según los requisitos estipulados por los proveedores de servicio. Además, para el estudio realizado en la Tesis se ha asumido la tecnología EPON basada en el estándar Ethernet debido a su gran nivel de despliegue.

El primer algoritmo propuesto para soportar calidad de servicio en redes LR-EPON ha sido DySGAB (*Dynamic Service levels with GuAranteed Bandwidth*). Este algoritmo, ha sido diseñado para realizar una diferenciación de usuarios mediante la consideración de diferentes perfiles de abonado, esto es SLAs. En concreto, DySGAB garantiza dinámicamente unos niveles mínimos de ancho de banda estipulados por los proveedores de servicio según la prioridad del perfil contratado. Para ello, a diferencia de otros algoritmos propuestos en la literatura, no distribuye el ancho de banda a cada SLA a partir de un sistema de pesos fijos, sino que los anchos de banda máximos permitidos se adaptan en tiempo real a los requisitos de calidad de servicio que se deben garantizar. Esta adaptación de los anchos de banda máximos dota a DySGAB de independencia frente a diferentes parámetros de red, tales como condiciones iniciales de ancho de banda o diferentes niveles garantizados. Además, el algoritmo es capaz de adaptarse de forma eficiente a cambios en tiempo real de la red o del tráfico, haciendo un uso adecuado de los recursos disponibles. En este sentido, el algoritmo se ha probado en redes con patrones de tráfico más realistas, tales como el tráfico dinámico y el tráfico asimétrico, así como en escenarios donde los requisitos de calidad varían en tiempo real. Así mismo, cabe

destacar que el algoritmo se puede utilizar en redes EPON y LR-EAPON de diferente cobertura, ofreciendo iguales prestaciones independientemente de la distancia.

Con el fin de mejorar las prestaciones del algoritmo DySGAB, se diseñó un sistema de control de procesos para la asignación dinámica del ancho de banda disponible. En concreto, se ha optado por el uso de controladores PID debido a su simplicidad, robustez y aplicabilidad en diferentes campos. En este sentido, inicialmente se diseñó e implementó el algoritmo SPID (*Service level agreement PID*). Este algoritmo garantiza a cada perfil de abonado su nivel de ancho de banda estipulado modificando dinámicamente y automáticamente los anchos de banda máximos, al igual que DySGAB, pero mediante un controlador PID. En este sentido, SPID presenta todas las ventajas de independencia y auto-adaptación de DySGAB pero, gracias al uso del controlador PID, que tiene en cuenta el error cometido en el presente, los errores pasados y una estimación del error futuro a la hora de ajustar el ancho de banda, resulta un algoritmo mucho más estable y robusto que DySGAB. Además, en SPID el ancho de banda medio converge de forma más rápida que en DySGAB hacia los niveles garantizados independientemente del escenario considerado. Por el contrario, en DySGAB aparte de reaccionar de forma más lenta, la velocidad de convergencia depende en gran medida del escenario de red considerado. Por otro lado, SPID presenta un comportamiento más independiente que DySGAB frente a parámetros comunes relacionados con el diseño de ambos algoritmos, tales como el tiempo de muestreo o el tamaño de la ventana deslizante. En este sentido, el comportamiento de DySGAB también presenta una fuerte dependencia con otro parámetro propio de su diseño, en concreto, de un término que marca el ancho de banda involucrado en el reajuste continuo de los máximos.

En el siguiente paso de la investigación y con la finalidad de desarrollar nuevas técnicas de QoS en redes LR-EAPON basadas en el uso de controladores PID, se propuso un nuevo algoritmo DBA que, de forma automática, adaptaba los recursos de ancho de banda disponible dependiendo de los requisitos de calidad de servicio del tráfico prioritario soportado por la red para diferentes tipos de perfiles de abonado. En este sentido, se desarrolló el algoritmo DaSPID (*Delay aware Service level agreement PID*). Este algoritmo, realiza una doble diferenciación, a nivel de usuarios y a nivel de servicios, para garantizar niveles de calidad de servicio de forma mucho más eficiente e integrada. En concreto, el algoritmo DaSPID controla, mediante el uso de un controlador P, que el retardo medio de las clases de servicio más prioritarias se sitúe por debajo de unas cotas máximas estipuladas dependiendo del tipo de tráfico y de la prioridad del perfil del abonado. De esta forma, DaSPID realiza una asignación de ancho de banda muy eficiente, consiguiendo una evolución dinámica del retardo medio hacia las cotas estipuladas de

forma más rápida y estable que otros algoritmos existentes en la literatura. Además, este algoritmo se puede aplicar de forma directa en redes EPON y LR-EPON de diferente distancia, sin que sus buenas prestaciones se vean afectadas. Del mismo modo, también se han llevado a cabo pruebas del algoritmo bajo distintos tipos de patrones de tráfico más realistas, tales como tráfico asimétrico y dinámico, comprobando así que DaSPID mantiene su buen comportamiento.

Por tanto, la inclusión de sistemas de control de procesos en la gestión de recursos en redes EPON y LR-EPON ha demostrado tener un gran potencial. En concreto, esta idea, que nunca antes se había aplicado en tareas relacionadas con la gestión de parámetros de red (ancho de banda, retardo), ha superado en gran medida las prestaciones de otros algoritmos existentes, consiguiendo una gran eficiencia, robustez, estabilidad y rapidez a nivel de convergencia.

No obstante, para asegurar un buen funcionamiento de los controladores, es necesario realizar una sintonización óptima del controlador diseñado. Esto es, hallar los valores adecuados de los parámetros que determinan su funcionamiento. Como primera aproximación, para los algoritmos SPID y DaSPID, se ha utilizado el método de sintonización manual de Ziegler-Nichols. Este método es muy utilizado en procesos donde no existe un modelo matemático, al igual que nuestra red, ya que es un método experimental basado en la aplicación de reglas sencillas. Aunque este método de sintonización ha dado lugar a buenos resultados en ambos algoritmos, es un método manual que puede resultar tedioso y que conlleva demasiado tiempo en el proceso de sintonización. Por tanto, en la Tesis se ha desarrollado un sistema de sintonización automática basado en algoritmos genéticos. Los nuevos algoritmos, denominados GA-SPID (*Genetic Algorithm Service level agreement PID*) y GA-DaSPID (*Genetic Algorithm Delay aware Service level agreement PID*) incorporan un algoritmo genético para la sintonización automática de los controladores PID y P encargados de la gestión de ancho de banda y retardo, respectivamente. Aunque el uso de algoritmos genéticos en el proceso de sintonización ha proporcionado mejoras en el funcionamiento y prestaciones de los algoritmos, las mayores ventajas vienen dadas por la automatización del proceso de sintonización, una reducción considerable en el tiempo de sintonización y una mayor precisión a la hora de proporcionar los parámetros óptimos de los controladores.

Sin embargo, tanto la sintonización con el método de Ziegler-Nichols como la sintonización con el algoritmo genético son métodos de sintonización previos a la ejecución de los controladores que gestionan los recursos de la red. Por tanto, estos métodos no modifican en tiempo real los parámetros del controlador una vez que los algoritmos están en ejecución. Para lograr una sintonización adaptativa en tiempo real que proporcione

los valores óptimos del controlador en cada situación, se han incorporado al proceso de sintonización de SPID técnicas basadas en redes neuronales. En el nuevo algoritmo desarrollado, denominado NN-SPID (*Neural Network Service level agreement PID*), se entrena la red neuronal diseñada para que, dependiendo de las circunstancias de la red, ofrezca los parámetros de sintonización más adecuados en cada momento. En primer lugar, se procedió al diseño de la red neuronal determinando su estructura (número de capas, funciones de activación, etc.) para después realizar el proceso de entrenamiento de la misma. En esta fase de entrenamiento, se utilizan como patrones de ejemplo individuos que ofrecen una buena sintonización y que se han obtenido a partir de soluciones óptimas del algoritmo genético. Una vez diseñada y entrenada la red neuronal, se analizaron las prestaciones de NN-SPID bajo distintos escenarios y patrones de tráfico, obteniendo unos resultados de comportamiento adecuados en tiempo real a las circunstancias de la red. De forma paralela y siguiendo esta adaptabilidad en tiempo real, se ha desarrollado otro método para la sintonización adaptativa de controladores PID mediante la introducción de técnicas cognitivas gracias a la implementación de un sistema de razonamiento basado en casos (CBR). En concreto, se ha aplicado dicha técnica cognitiva en SPID, dando lugar al algoritmo denominado CBR-SPID, diseñado para garantizar a cada perfil de abonado diferentes niveles mínimos de ancho de banda estipulados según la prioridad. Por tanto, a partir de la experiencia adquirida en la sintonización previa mediante el uso de algoritmos genéticos, el CBR consigue realizar una sintonización adaptativa en función de las circunstancias de la red. Ambos métodos de sintonización adaptativa (NN-SPID y CBR-SPID) ofrecen mejores resultados que los métodos de sintonización no adaptativos. En este sentido, gracias a la variación en los parámetros de sintonización, los métodos adaptativos consiguen una mayor estabilidad y robustez frente a posibles cambios en tiempo real de las condiciones de red y patrones del tráfico. No obstante, como ambos algoritmos parten del conocimiento obtenido previamente mediante algoritmos genéticos, bien para entrenar a la red neuronal en NN-SPID o para formar parte de la base de conocimiento en CBR-SPID, la diferencia de comportamiento entre ambos algoritmos es muy pequeña.

En la Tabla 8.1 se presentan las principales características de los algoritmos desarrollados en esta Tesis para las redes LR-EPON.

Algoritmo	Asegura ancho de banda	Asegura cota de retardo	Aplicable en redes EPON y LR-EAPON	Utiliza controlador P-PI-PID	Sintonización manual Ziegler-Nichols	Sintonización automática Algoritmos Genéticos	Sintonización adaptativa Redes Neuronales	Sintonización adaptativa cognición CBR
DySGAB	♦		♦					
SPID	♦		♦	♦	♦			
DaSPID		♦	♦	♦	♦			
GA-SPID	♦		♦	♦		♦		
GA-DaSPID		♦	♦	♦		♦		
NN-SPID	♦		♦	♦			♦	
CBR-SPID	♦		♦	♦				♦

Tabla 8.1: Características principales de los algoritmos propuestos en esta Tesis para redes LR-EAPON.

8.1.2. Redes troncales

Por otro lado, enmarcado en el desarrollo del proyecto Europeo CHRON (*Cognitive Heterogeneous Reconfigurable Optical Network*), financiado por la Comisión Europea dentro del 7º Programa Marco, en esta Tesis también se han diseñado e implementado métodos cognitivos para gestionar la calidad de servicio en las redes ópticas troncales actuales. En concreto, mediante la incorporación de cognición a través de la técnica de razonamiento basado en casos (CBR), se ha desarrollado un estimador de la calidad de transmisión de los caminos ópticos que se pueden establecer en una red troncal con encaminamiento por longitud de onda. A partir de datos de calidad de transmisión de circuitos ópticos establecidos anteriormente, el estimador desarrollado predice la calidad de transmisión de un nuevo camino óptico antes de que éste sea establecido. Esto es, se utilizan datos de conexiones ópticas establecidas previamente, donde el factor de calidad se ha obtenido bien mediante otras herramientas o bien mediante monitores reales en la red, con el objetivo de poblar la base de conocimiento y, a partir de dicha base, el estimador predice la calidad de una nueva conexión. Además, se comprueba que el establecimiento de dicho camino óptico no influye negativamente en la calidad de los caminos ya establecidos. Para ello, se ha desarrollado un método híbrido que utiliza una estimación por longitud y un sistema de razonamiento basado en casos (CBR). De esta forma, el estimador cognitivo desarrollado presenta una probabilidad de acierto muy elevada y un tiempo de computación mucho menor que otras técnicas existentes. Así mismo, el análisis de prestaciones se ha llevado a cabo en redes de diferente tamaño y se ha validado experimentalmente en dos bancos de pruebas.

8.2. Líneas Futuras

La evolución de esta Tesis Doctoral en un campo tan emergente y potente como la gestión de recursos en redes ópticas, propicia tanto la optimización de aspectos abordados en esta Tesis, como la aparición de nuevas y prometedoras líneas de investigación.

En primer lugar, respecto a los métodos desarrollados en esta Tesis para las redes de acceso EPON y LR-EPON, se plantea como una línea de investigación muy interesante la modificación del algoritmo NN-SPID para ofrecer una adaptación en tiempo real más eficiente y óptima de los parámetros de sintonización. En este sentido, se podría realizar un proceso de aprendizaje o entrenamiento periódico de las redes neuronales con los últimos parámetros de sintonización y los resultados de comportamiento obtenidos, para que de este modo las redes neuronales infieran los parámetros óptimos a partir de los últimos resultados obtenidos y reflejar así posibles cambios de las condiciones de

red. O bien, se puede plantear utilizar otro tipo de métodos de aprendizaje donde no sea necesario el uso de ejemplos previos. De esta forma, las redes neuronales auto-aprenderían a partir de los errores cometidos en las continuas sintonizaciones y adecuándose de forma dinámica y en tiempo real a las nuevas condiciones de la red. En línea con estas técnicas auto-adaptativas, también resultaría de gran interés la incorporación de una fase de aprendizaje y olvido en el sistema CBR diseñado para la sintonización de controladores PID. Así pues, se podrían incrementar y actualizar los casos almacenados en la base de conocimiento para mejorar la selección del caso más similar y poder reflejar los cambios que se producen en la red.

Por otro lado, siguiendo la línea del uso de controladores PID para la gestión eficiente de recursos, resultaría de gran interés profundizar aún más en el análisis de los parámetros de calidad de servicio que debe soportar una red LR-EPON. En esta Tesis nos hemos centrado en el control por un lado del ancho de banda garantizado y por otro del retardo medio de las clases de servicio prioritarias. No obstante, sería muy interesante plantear un sistema de gestión conjunta de ambos parámetros. Esto es, no solo garantizar un nivel mínimo de ancho de banda, sino también asegurar de forma integrada y conjunta que el retardo medio de las clases prioritarias esté por debajo de las cotas establecidas dependiendo del perfil contratado. Así mismo, también sería relevante ampliar este control hacia otros parámetros importantes como la probabilidad de pérdida de paquetes o el *jitter* en aquellas clases de servicio que presenten restricciones en ambos parámetros de red.

Respecto a otras nuevas líneas de investigación que se pueden seguir, surge de un modo muy potente la combinación de redes ópticas de acceso PON y LR-PON con redes inalámbricas. De hecho, ambas tecnologías pueden considerarse como complementarias. La fibra óptica no consigue llegar a todos los lugares, pero ofrece un gran ancho de banda. Por el contrario, las redes inalámbricas potencialmente pueden llegar casi a todos los lugares pero el ancho de banda que ofrecen está restringido. Por tanto, la combinación de la alta capacidad de las redes ópticas PON y LR-PON junto con la ubicuidad y movilidad de las redes inalámbricas, forman las denominadas redes de acceso FiWi (*Fiber Wireless*). Esta integración de red da lugar a una potente tecnología capaz de soportar todas las aplicaciones emergentes y futuras, así como de proporcionar niveles adecuados de calidad de servicio a los usuarios, manteniendo al mismo tiempo unos costes de despliegue relativamente bajos. Además, hay que tener en cuenta que los perfiles de abonado tienen cada vez más un carácter inalámbrico por el uso extensivo de dispositivos móviles. Por tanto, debido al enorme interés y cualidades que presentan estas redes, se plantea la necesidad de diseñar e implementar algoritmos que controlen eficientemente

parámetros como el ancho de banda o el retardo medio extremo a extremo en redes FiWi, con el fin de ofrecer una calidad de servicio adecuada a las necesidades de sus usuarios. Esto es así, ya que en general, los niveles de retardo en la parte inalámbrica no son comparables a los obtenidos en las redes de fibra, fundamentalmente por las diferencias de propagación en ambos medios de transmisión. Además, es necesario diseñar algoritmos de encaminamiento en la parte inalámbrica que tengan en cuenta estos requisitos de calidad de servicio, con el fin de poder ofrecer una asignación de recursos eficiente extremo a extremo.

Por otro lado, otra línea de investigación a desarrollar está relacionada con las redes TDM-WDM PON. En estas redes, que se perfilan como la tecnología que se usará en las NG-PON2, también son necesarios mecanismos de asignación de ancho de banda y de longitud de onda de forma conjunta, para así poder explotar las ventajas de la multiplexación por división en longitud de onda. Por lo tanto, se abre una línea de investigación muy prometedora en el desarrollo de nuevos algoritmos que realicen una gestión eficiente de los recursos disponibles ofreciendo la calidad de servicio requerida.

Por último y referido al estimador cognitivo de calidad de transmisión en las redes troncales, sería muy interesante incorporar una fase de aprendizaje y olvido paralela a la ejecución del estimador y no de forma previa como la que se ha desarrollado en esta Tesis. De esta forma, el estimador cognitivo se adaptaría a cambios en tiempo real en la red, tales como la incorporación o caída de nuevos enlaces, situaciones de saturación en la red o el propio envejecimiento de los componentes, para ofrecer así una estimación de la calidad de transmisión más precisa. Incluso, también se podría considerar la extensión de este estimador de calidad óptico de la capa física a la capa de acceso, para decidir si se establece un nuevo *lightpath*, no solo en función de la calidad de parámetros físicos, sino también de otros parámetros de calidad de servicio como puede ser el retardo extremo a extremo, ancho de banda o pérdida de paquetes.

8.3. Publicaciones

A continuación se incluye un listado cronológico de las publicaciones derivadas de esta Tesis Doctoral hasta el momento de la finalización de la redacción de la presente memoria. Así, en primer lugar se muestran las contribuciones en revistas científicas con índice de impacto, en segundo lugar las contribuciones a capítulos de libro y finalmente las aportaciones en congresos nacionales e internacionales.

Contribuciones en revistas:

- N. Merayo, **T. Jiménez**, R.J. Durán, P. Fernández, R.M. Lorenzo, I. de Miguel, E.J. Abril, “Adaptive polling algorithm to provide subscriber and service differentiation in a Long-Reach EPON,” *Photonic Network Communications*, vol. 19, no. 3, pp. 247-264, Junio 2010.
- N. Merayo, **T. Jiménez**, P. Fernández, R.J. Durán, R.M. Lorenzo, I. de Miguel, E.J. Abril, “A Wavelength and Bandwidth Assignment Algorithm to Support Guaranteed Bandwidth Levels in Hybrid Time Division Multiplexing/Wavelength Division Multiplexing Ethernet Passive Optical Networks,” *Fiber and Integrated Optics*, vol. 29, no. 5, pp. 394-419, Septiembre 2010.
- N. Merayo, **T. Jiménez**, P. Fernández, R.J. Durán, R.M. Lorenzo, I. de Miguel, E.J. Abril, “A bandwidth assignment polling algorithm to enhance the efficiency in QoS Long-Reach EPONs,” *European Transactions on Telecommunications*, vol. 22, no. 1, pp. 35-44, Enero 2011.
- **T. Jiménez**, N. Merayo, P. Fernández, R.J. Durán, I. de Miguel, R.M. Lorenzo, E.J. Abril, “Self-adapted algorithm to provide Multi-Profile Bandwidth Guarantees in PONs with Symmetric and Asymmetric Traffic Load,” *Photonic Network Communications*, vol. 24, no. 1, pp. 58-70, Enero 2012.
- **T. Jiménez**, N. Merayo, P. Fernández, R.J. Durán, I. de Miguel, R.M. Lorenzo, E.J. Abril, “Implementation of a PID Controller for the Bandwidth Assignment in Long-Reach PONs,” *Journal of Optical Communications and Networking*, vol. 4, no. 5, pp. 392-401, Mayo 2012.
- A. Caballero, J.C. Aguado, R. Borkowski, S. Saldaña, **T. Jiménez**, I. de Miguel, V. Arlunno, R.J. Durán, D. Zibar, J.B. Jensen, R.M. Lorenzo, E.J. Abril, I. Tafur, “Experimental Demonstration of a Cognitive Quality of Transmission Estimator for Optical Communication Systems,” *Optics Express*, vol. 20, no. 26, pp. 64-70, Diciembre 2012.
- **T. Jiménez**, J.C. Aguado, I. de Miguel, R.J. Durán, M. Angelou, N. Merayo, P. Fernández, R.M. Lorenzo, I. Tomkos, E.J. Abril, “A Cognitive Quality of Transmission Estimator for Core Optical Networks,” *IEEE/OSA Journal of Lightwave Technology*, vol. 31, no. 6, pp. 942-951, Marzo 2013.
- I. de Miguel, R.J. Durán, **T. Jiménez**, N. Fernández, J.C. Aguado, R.M. Lorenzo, A. Caballero, I. Tafur, Y. Ye, A. Tymeki, I. Tomkos, M. Angelou, D. Klonidis,

- A. Francescon, D. Siracusa, E. Salvadori. “Cognitive Dynamic Optical Networks [Invited],” *Journal of Optical Communications and Networking*, vol. 5, no. 10, pp. A107-A118, Octubre 2013.
- **T. Jiménez**, N. Merayo, R.J. Durán, P. Fernández, I. de Miguel, J.C. Aguado, R.M. Lorenzo, E.J. Abril, “A PID-based algorithm to guarantee QoS delay requirements in LR-PONs,” *Optical Switching and Networking*. Aceptado para publicación. D.O.I: <http://dx.doi.org/10.1016/j.osn.2014.01.005>
 - A. Caballero, R. Borkowski, I. de Miguel, R.J. Durán, J.C. Aguado, N. Fernández, **T. Jimenez**, I. Rodríguez, D. Sánchez, R.M. Lorenzo, D. Klonidis, E. Palkopoulou, N. Diamantopoulos, I. Tomkos, D. Siracusa, A. Francescon, E. Salvadori, Y. Ye, J. López Vizcaíno, F. Pittala, A. Tymecki, I. Tafur, “Cognitive, Heterogeneous and Reconfigurable Optical Networks: the CHRON Project,” Aceptado para publicación en *IEEE/OSA Journal of Lightwave Technology*.

Contribuciones en capítulos de libro:

- N. Merayo, R.M. Lorenzo, **T. Jiménez**, R.J. Durán, P. Fernández, I. de Miguel, E.J. Abril, “Medium Access Control Protocols in Passive Optical Networks based on Ethernet (EPONs).” *Computer Science Research and Technology*. Hauppauge Editorial Nova Science Publishers, Inc. 2011. ISBN 978-1-61728-688-9.

Contribuciones en congresos nacionales e internacionales:

- **T. Jiménez**, N. Merayo, R.J. Durán, P. Fernández, R.M. Lorenzo, “Adaptive allocation algorithm to support Multi-Service Level Agreements in a Long-Reach EPON,” en *13th European Conference on Networks and Optical Communications (NOC 2008)*, Krems, Austria, Julio 2008.
- **T. Jiménez**, N. Merayo, R.J. Durán, P. Fernández, R.M. Lorenzo, I. de Miguel, M. Ramírez, E.J. Abril, “Polling algorithm with adaptive cycle to enhance efficiency in QoS Long-Reach EPONs,” en *14th European Conference on Networks and Optical Communications (NOC 2009)*, Valladolid, España, Junio 2009.
- **T. Jiménez**, N. Merayo, R.J. Durán, P. Fernández, R.M. Lorenzo, I. de Miguel, N. Fernández, E.J. Abril, “Interleaved polling algorithm with inserted cycles to support service level agreement in Long-Reach EPONs”, en *11th International Conference on Transparent Optical Networks (ICTON 2009)*, Ponta Delgada, Azores, Portugal, Julio 2009.

- **T. Jiménez**, N. Merayo, P. Fernández, R.J. Durán, R.M. Lorenzo, N. Fernández, I. de Miguel, E.J. Abril, “Self-adjustment bandwidth algorithm to ensure bandwidth levels in Multi-profile LR-EPONs under heterogeneous traffic load,” en *15th European Conference on Networks and Optical Communications (NOC 2010)*, pp. 353-358, Faro, Portugal, Junio 2010.
- **T. Jiménez**, N. Merayo, P. Fernández, R.J. Durán, R.M. Lorenzo, I. de Miguel, E.J. Abril, “LR-EPON Algorithm with Automatic Bandwidth Adaptation to Provide Multi-profiles Bandwidth Levels,” en *Advanced Photonics and Renewable Energy, 2010 OSA Optics & Photonics Congress. Access Networks and In-House Communications (ANIC 2010)*, Karlsruhe, Alemania, Junio 2010.
- I. de Miguel, R.J. Durán, **T. Jiménez**, N. Fernández, J.C. Aguado, N. Merayo, P. Fernández, R.M. Lorenzo, E.J. Abril, “Cognition and heterogeneity in optical networks”, en *1st Workshop of Future Internet: Efficiency in High-Speed Networks (W-FIERRO 2011)*, paper 09. Cartagena, España, Julio 2011.
- **T. Jiménez**, I. de Miguel, J.C. Aguado, R.J. Durán, N. Merayo, N. Fernández, D. Sánchez, P. Fernández, N. Atallah, E.J. Abril, R.M. Lorenzo, “Case-based reasoning (CBR) to estimate the Q-factor in optical networks: an initial approach,” en *16th European Conference on Networks and Optical Communications (NOC 2011)*, pp. 181-184, Newcastle-Upon-Tyne, Reino Unido, Julio 2011.
- **T. Jiménez**, J.C. Aguado, I. de Miguel, R.J. Durán, N. Fernández, M. Angelou, D. Sánchez, N. Merayo, P. Fernández, N. Atallah, R. Lorenzo, I. Tomkos, E.J. Abril, “Enhancing optical networks with cognition: Case-Based Reasoning to estimate the quality of transmission,” en *IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA 2012)*, pp. 166-169, Nueva Orleans, LA, Estados Unidos, Marzo 2012.
- **T. Jiménez**, J.C. Aguado, I. de Miguel, R.J. Durán, N. Fernández, M. Angelou, D. Sánchez, N. Merayo, P. Fernández, N. Atallah, R.M. Lorenzo, I. Tomkos, E.J. Abril, “A Cognitive System for Fast Quality of Transmission Estimation in Core Optical Network,” en *OFC/NFOEC 2012*, paper OW3A.5, Los Ángeles, Estados Unidos, Marzo 2012.
- **T. Jiménez**, J.C. Aguado, I. de Miguel, R.J. Durán, D. Sánchez, M. Angelou, N. Merayo, P. Fernández, N. Fernández, R.M. Lorenzo, I. Tomkos, E.J. Abril, “Optimization of the Knowledge Base of a Cognitive Quality of Transmission Esti-

mator for Core Optical Networks,” en *16th Optical Network Design and Modeling (ONDM 2012)*, Colchester, Reino Unido, Abril 2012.

- **T. Jiménez**, N. Merayo, I. de Miguel, R.J. Durán, P. Fernández, J.C. Aguado, N. Fernández, R.M. Lorenzo, E.J. Abril, “PID strategies for the dynamic allocation of resources in LR-Ethernet Passive Optical Networks,” en *16th Optical Network Design and Modeling (ONDM 2012)*, Colchester, Reino Unido, Abril 2012.
- **T. Jiménez**, N. Merayo, R.J. Durán, P. Fernández, I. de Miguel, J.C. Aguado, R.M. Lorenzo, E.J. Abril, “Auto-Tuning PID controller based on Genetic Algorithms for the Bandwidth Allocation in LR-PONs,” en *Advanced Photonics. OSA & Photonic Congress. Access Networks and In-house Communications (ANIC 2012)*, paper AW2A.2, Colorado, Estados Unidos, Junio 2012.
- N. Merayo, I. de Miguel, **T. Jiménez**, N. Fernández, R.J. Durán, J.C. Aguado, P. Fernández, R.M. Lorenzo, E.J. Abril, “OPNET Modeler and SP Guru Transport Planner. Overview and Utilization by the Optical Communications Group at University of Valladolid,” en *2nd Workshop of Future Internet: Efficiency in High-Speed Networks (W-FIERR 2012)*, Cartagena, España, Julio 2012.
- R.J. Durán, I. de Miguel, D. Sánchez, N. Fernández, **T. Jiménez**, J.C. Aguado, V.K. Yedugundla, M. Angelou, N. Merayo, P. Fernández, N. Atallah, R.M. Lorenzo, A. Francescon, E.J. Abril, “A Cognitive Decision System for Heterogeneous Reconfigurable Optical Networks,” en *Future Networks & Mobile Summit 2012*, Berlin, Alemania, Julio 2012.
- A. Caballero, J.C. Aguado, R. Borkowski, S. Saldaña, **T. Jiménez**, I. de Miguel, V. Arlunno, R.J. Durán, D. Zibar, J.B. Jensen, R.M. Lorenzo, E.J. Abril, I. Tafur, “Experimental demonstration of a cognitive quality of transmission estimator for optical communication systems,” en *European Conference and Exhibition on Optical Communication (ECOC 2012)*, Amsterdam, Holanda, Septiembre 2012.
- E. Palkopoulou, I. Stiakogiannaki, D. Klonidis, **T. Jiménez**, N. Fernández, J.C. Aguado, J. López, Y. Ye, I. Tomkos, “Cognitive Heterogeneous Reconfigurable Optical Network: A Techno-Economic Evaluation,” en *Future Networks & Mobile Summit 2013*, Lisboa, Portugal, Julio 2013.
- C. Kachris, D. Klonidis, A. Francescon, D. Siracusa, E. Salvadori, N. Fernández, **T. Jiménez**, R.J. Durán, I. de Miguel, J.C. Aguado, R.M. Lorenzo,

R. Borkowski, A. Caballero, I. Tafur, Y. Ye, A. Tymecki, I. Tomkos, “Experimental Demonstration of a Cognitive Optical Network for Reduction of Restoration Time,” en *OFC/NFOEC 2014*, paper W2A.28, San Francisco (CA), Estados Unidos, Marzo 2014.

- **T. Jiménez**, N. Merayo, J.C. Aguado, R.J. Durán, I. de Miguel, P. Fernández, R.M. Lorenzo, E.J. Abril, “Design of auto-Tuning PID controller methods based on Genetic Algorithms for LR-PONs,” Aceptado para publicación en *5th International Conference on Access Networks*, Sevilla (España), Junio 2014.

Bibliografía

- [1] Cisco, “Cisco Visual Networking Index: Forecast and Methodology, 2012-2017,” 2013.
- [2] A. Banerjee, G. Kramer, Y. Ye, S. Dixit, B. Mukherjee, “Advances in Passive Optical Networks (PONs),” *Emerging optical network technologies: Architectures, Protocols and Performance*. K. M. Sivalingnan, S. Subramaniam, Eds. Spring Street, Nueva York: Springer, 2005.
- [3] G. Keiser, *FTTx concepts and Applications*. Ed. John Wiley & Sons, Inc., 2006.
- [4] I. Tomkos, “Techno-economic Comparison of Next Generation Optical Access Network Arcitecture,” en *50th FITCE Congress*, Palermo (Italia), Septiembre 2011.
- [5] IDATE Consulting & Research, “FTTx 2014. Markets & Trends. Facts & Figures. White Paper,” 2014.
- [6] C. Lange, T. Monath, E. Weis, J. Preuschaft, M. Kind, P. Bockelmann, M. Zerson, “FTTB or FTTH: An operator’s view,” en *14th European Conference on Networks and Optical Communications (NOC09)*, pp. 183-190, Valladolid (España), Junio 2009.
- [7] Ovum, “Brief Review of FTTx Market. Focus on China: Subscribers, Vendors, Network Deployments,” en *FTTH/ODN/OTN China Conference 2011 (FOOCC 2011)*, Shenzhen (China), Septiembre 2011.
- [8] R. Montagne, V. Chaillou, “FTTH/B Panorama. Europe (EU39) at December 2013,” en *FTTH Council Europe Conference*, Estocolmo (Suecia), Febrero 2014.
- [9] IDATE Consulting & Research, “FTTx 2013. Markets & Trends. Facts & Figures. White Paper,” 2013.
- [10] G. Kramer, *Ethernet Passive Optical Networks*. Ed. McGraw Hill, 2005.

- [11] A. Shami, M. Maier, C. Assi, *Broadband Access Networks. Technologies and Deployments*. Ed. Springer, 2009.
- [12] N. Ansari, J. Zhang, *Media Access Control and Resource Allocation For Next Generation Passive Optical Networks*. Ed. Springer, 2013.
- [13] E. Wong, “Next-Generation Broadband Access Networks and Technologies,” *IEEE/OSA Journal of Lightwave Technology*, vol. 30, no. 4, pp. 597-608, Febrero 2012.
- [14] IDATE Consulting & Research, “World FTTx Markets. FTTx Market Report,” Julio 2009.
- [15] G. Kramer, B. Mukherjee, A. Maislos, “Multiprotocol over WDM: Building the Next Generation Optical Internet,” *Ethernet Passive Optical Networks*, S. Dixit, Ed. John Wiley and Sons, 2003.
- [16] M. McGarry, M. Reisslein, M. Maier, “Ethernet Passive Optical Network Architectures and Dynamic Bandwidth Allocation Algorithms,” *IEEE Communications Surveys & Tutorials*, vol. 10, no. 3, pp. 46-60, tercer trimestre de 2008.
- [17] Y. Luo, N. Ansari, “Bandwidth allocation for multiservice access on EPONs,” *IEEE Communications Magazine*, vol. 43, no. 2, pp. 16-21, Diciembre 2005.
- [18] D. Shea, J.E. Mitchell, “Long Reach Optical Access Technologies,” *IEEE Network*, vol. 21, no. 5, pp. 5-11, Septiembre-Octubre 2007.
- [19] H. Song, B.W. Kim, B. Mukherjee, “Long-Reach Optical Access Networks: A Survey of Research Challenges, Demonstrations and Bandwidth Assignment Mechanisms,” *IEEE Communications Surveys & Tutorials*, vol. 12, no. 1, pp. 112-122, primer trimestre 2010.
- [20] ITU-T Recomendación G.983, “Broadband Optical Access Systems Based on Passive Optical Networks (PON).” Telecommunication Standardization Sector of ITU, 2001.
- [21] ITU-T Recomendación G.984, “Gigabit-Capable Passive Optical Networks (G-PON).” Telecommunication Standardization Sector of ITU, 2004.
- [22] ITU-T Recomendación G.987, “10-Gigabit-Capable Passive Optical Network (XG-PON) Systems: Definitions, Abbreviations and Acronymus.” Telecommunication Standardization Sector of ITU, 2009.

- [23] Y. Luo, X. Zhou, F. Effenberger, X. Yan, G. Peng, Y. Ma, "Time- and Wavelength-Division Multiplexed Passive Optical Network (TWDM-PON) for Next-Generation PON Stage 2 (NG-PON2)," *IEEE/OSA Journal of Lightwave Technology*, vol. 31, no. 4, pp. 587-593, Febrero 2013.
- [24] IEEE Standard 802.3ah - IEEE Standard for Information technology. Local and metropolitan area networks. Part 3: CSMA/CD Access Method and Physical Layer Specifications Amendment: Media Access Control Parameters, Physical Layers, and Management Parameters for Subscriber Access Networks, 2004.
- [25] IEEE Standard 802.3av - IEEE Standard for Information technology. Telecommunications and information exchange between systems. Local and metropolitan area networks. Specific requirements Part 3: Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications. Amendment 1: Physical Layer Specifications and Management Parameters for 10 Gb/s Passive Optical Networks, 2009.
- [26] G. Kramer, L. Khermosh, F. Daido, A. Brown, H. Yoon, K. Suzuki, W. Bo, "The IEEE 1904.1 Standard: SIEPON Architecture and Model," *IEEE Communications Magazine*, vol. 50 no. 9 pp. 98-108, Septiembre 2012.
- [27] Ovum, "FTTx Around the Globe: subscribers, shipments and regional deployments, trends business models, the supply chain," Mayo 2012.
- [28] K. Tanaka, "IEEE 802.3av 10G-EPON Standardization and its Research and Development Status," *IEEE/OSA Journal of Lightwave Technology*, vol. 28, no. 4, pp. 651-661, Febrero 2010.
- [29] M. Fujiwara, T. Imai, K. Taguchi, K. Suzuki, H. Ishii, N. Yoshimoto, "Field Trial of 100-km Reach Symmetric-Rate 10G-EPON System Using Automatic Level Controlled Burst-Mode SOAs," *IEEE/OSA Journal of Lightwave Technology*, vol. 31, no. 4, pp. 634-640, Febrero 2013.
- [30] ZTE Notas de Prensa. "ZTE Retains Leading Vendor Position in China Telecom 2013 PON Equipment Tenders," Julio 2013.
- [31] A.M. Ragheb, H. Fathallah, "Performance analysis of next generation-PON (NG-PON) architectures," en *High Capacity Optical Networks and Enabling Technologies (HONET 2011)*, pp. 339-345, Riyadh (Arabia Saudí), Diciembre 2011.

- [32] C.W. Chow, C.H. Yeh, "Technology advances for the 2nd stage next-generation passive-optical-network (NG-PON2)," en *6th International Conference on Advanced Infocomm Technology (ICAIT)*, pp. 83-84, Hsinchu (Taiwan), Julio 2013.
- [33] J. Escribano, C. García, C. Seldas, J.I. Moreno. "Diffserv como solución a la provisión de QoS en Internet," en *II Congreso Iberoamericano de Telemática (CITA 2002)*, Mérida (Venezuela), Septiembre 2002.
- [34] ITU-T Recomendación G.1010, "End-user multimedia QoS categories", Telecommunication Standardization Sector of ITU, 2001.
- [35] ITU-T Recomendación Y.1541, "Objetivos de calidad de funcionamiento de red para servicios basados en el protocolo Internet", Telecommunication Standardization Sector of ITU, 2006.
- [36] M. Ruffini, D. Mehta, B. O'Sullivan, L. Quesada, L. Doyle, D.B. Payne, "Deployment case studies of an energy efficient protected LR-PON architecture," en *16th Optical Network Design and Modeling (ONDM 2012)*, Colchester (Reino Unido), Abril 2012.
- [37] M. Ruffini, D. Mehta, B. O'Sullivan, L. Quesada, L. Doyle, D.B. Payne, "Deployment Strategies for Protected Long-Reach PON," *Journal of Optical Communications and Networking*, vol. 4, no. 2, pp. 118-129, Febrero 2012.
- [38] M. Ruffini, L. Wosinska, M. Achouche, J. Chen, N. Doran, F. Farjady, J. Montalvo, P. Ossieur, B. O'Sullivan, N. Parsons, T. Pfeiffer, X. Qiu, C. Raack, H. Rohde, M. Schiano, P. Townsend, R. Wessaly, X. Yin, D. Payne, "DISCUS: An End-to-End Solution for Ubiquitous Broadband Optical Access," *IEEE Communications Magazine*, vol. 52, no. 2, pp. S24-S32, Febrero 2014.
- [39] I. Voorde, C. Martin, J. Vandewege, X. Qiu, "The SuperPON demonstrator: an exploration of possible evolution paths for optical access networks," *IEEE Communications Magazine*, vol. 38, no. 2, pp. 74-82, Febrero 2000.
- [40] D. Shea, J. Mitchell, "A 10 Gbit/s 1024-Way Split 100 km Long-Reach Optical Access Network," *IEEE/OSA Journal of Lightwave Technology*, vol. 25, no. 3, pp. 685-693, Marzo 2007.
- [41] G. Talli, P.D. Townsend, "Hybrid DWDM-TDM Long-Reach PON for Next-Generation Optical Access," *IEEE/OSA Journal of Lightwave Technology*, vol. 24, no. 7, pp. 2827-2834, Julio 2006.

- [42] R. Davey, P. Healey, I. Hope, P. Watkinson, D. Payne, O. Marmur, J. Ruhmann, Y. Zuiderveld, "DWDM reach extension of a GPON to 135 km," *IEEE/OSA Journal of Lightwave Technology*, vol. 24, no. 1, pp. 29-31, Noviembre 2006.
- [43] D. Shea, J.E. Mitchell, "Architecture to integrate multiple PONs with long reach DWDM backhaul," *Journal on Selected Areas in Communications*, vol. 27, no. 2, pp. 126-133, Febrero 2009.
- [44] D. Shea, J.E. Mitchell, "Experimental Upstream demonstration of a Long-Reach Wavelength-Converting PON with DWDM Backhaul," en *OFC/NFOEC 2007*, Anaheim (Estados Unidos), pp. 1-3, Marzo 2007.
- [45] S. Smolorz, H. Rohde, P. Ossieur, C. Antony, P.D. Townsend, T. De Ridder, B. Baeckelandt, X.Z. Qui, S. Appathurai, H.G. Krimmel, D. Smith, A. Poustie, "Next generation access networks: PIEMAN and beyond," en *Photonics in Switching 2009*, Pisa (Italia), Septiembre 2009.
- [46] H. Rohde, S. Smolorz, E. Gottwald, K. Kloppe, "Next Generation Optical Access: 1 Gbit/s for Everyone," en *35th European Conference and Exhibition on Optical Communication (ECOC 2009)*, Viena (Austria), Septiembre 2009.
- [47] H. Song, B. Kim, B. Mukherjee, "Multi-Threat Polling: A Dynamic Bandwidth Distribution Scheme in Long-Reach PON," *Journal on Selected Areas in Communications*, vol. 27, no. 2, pp. 134-142, Febrero 2009.
- [48] F.T. An, K.S. Kim, D. Gutierrez, S. Yam, E. Hu, K. Shrikhande, L.G. Kazovsky, "SUCCESS: A Next-Generation Hybrid WDM/TDM Optical Access Network Architecture," *IEEE/OSA Journal of Lightwave Technology*, vol. 22, no. 11, pp. 2557-2569, Noviembre 2004.
- [49] F.T. An, D. Gutierrez, K.S Kim, J.W. Lee, L.G. Kazovsky, "SUCCESS-HPON: A next-generation optical access architecture for smooth migration from TDM-PON to WDM-PON," *IEEE Communications Magazine*, vol. 43, no. 11, pp. S40-S47, Noviembre 2005.
- [50] M. Rasztovits-Wiech, A. Stadler, K. Kloppe, "Realization of an XL-PON prototype," en *Broadband Europe*, Antwerp (Bélgica) , Diciembre 2007.
- [51] B.W. Kim, "Introduction to WDM-PON and WE-PON". Documento de trabajo. ETRI 2007.

- [52] J.A. Lazaro, J. Prat, P. Canclou, G.M. Tosi Beleffi, A. Teixeira, I. Tomkos, R. Soila, V. Koratzinos, "Scalable Extended Reach PON," en *OFC/NFOEC 2008*, California (Estados Unidos), Febrero 2008.
- [53] J.H. Yu, B.W. Kim, N. Kim, "WDM/TDMA Hybrid-PON: Wx-PON system," en *11th International Conference on Advanced Communication Technology (ICACT 2009)*, vol. 3, pp. 1707-1711, Gangwon-Do, (Corea del Sur), Febrero 2009.
- [54] ITU-T Recomendación G.984.3, "Gigabit-capable Passive Optical Networks (G-PON): Transmission convergence layer specification," Telecommunication Standardization Sector of ITU, 2008.
- [55] B. Kantarci, H.T. Mouftah, "Bandwidth Distribution Solutions for Performance Enhancement in Long-Reach Passive Optical Networks," *IEEE Communications Surveys & Tutorials*, vol. 14, no. 3, pp. 714-733, tercer trimestre 2012.
- [56] C. Chang, N. Álvarez, P. Kourtessis, R. Lorenzo, J. Senior, "Full-Service MAC Protocol for Metro-Reach GPONs," *IEEE/OSA Journal of Lightwave Technology*, vol. 28, no. 7, pp. 1016-1022, Abril 2010.
- [57] C.H. Chang, N.M. Álvarez, P. Kourtessis, J.M. Senior. "Dynamic Bandwidth assignment for Multi-service access in Long-Reach GPON," en *33th European Conference and Exhibition on Optical Communication (ECOC 2007)*, Berlin (Alemania), Septiembre 2007.
- [58] B. Skubic, J. Chen, J. Ahmed, B. Chen, L. Wosinska y B. Mukherjee, "Dynamic bandwidth allocation for Long-Reach PON: overcoming performance degradation," *IEEE Communications Magazine*, vol. 48, no. 11, pp. 100-108, Noviembre 2010.
- [59] J. Ahmed, J. Chen, L. Wosinska, B. Chen, B. Mukherjee, "Efficient Inter-Thread Scheduling Scheme for Long-Reach Passive Optical Networks," *IEEE Communications Magazine*, vol. 51, no. 2, pp. S35-S43, Febrero 2013.
- [60] A. Helmy, H. Fathallah, H. Mouftah, "Interleaved Polling versus Multi-Thread Polling for Bandwidth Allocation in Long-Reach PONs," *Journal of Optical Communications and Networking*, vol. 4, no. 3, pp. 210-218, Marzo 2012.
- [61] B. Kantarci, H.T. Mouftah, "Periodic GATE optimization (PGO): A new service scheme for long-reach passive optical networks," *IEEE Systems Journal*, vol. 4, no. 4, pp. 440-448, Diciembre 2010.

- [62] A. Dixit, G. Das, B. Lanoo, D. Colle, M. Pickavet, P. Demeester, "Adaptive Multi-Gate Polling with Void Filling for Long-Reach Passive Optical Networks," en *13th International Conference on Transparent Optical Networks (ICTON 2011)*, pp. 1-4, Estocolmo (Suecia), Junio 2011.
- [63] D. Murayama, N. Oota, K.I. Suzuki, N. Yoshimoto, "Low-Latency Dynamic Bandwidth Allocation for 100 km Long-Reach EPONs," *Journal of Optical Communications and Networking*, vol. 5, no. 1, pp. 48-55, Enero 2013.
- [64] N. Merayo, T. Jiménez, R.J Durán, P. Ferández, I. de Miguel, R.M. Lorenzo, E.J. Abril, "Adaptive polling algorithm to provide subscriber and service differentiation in a Long-Reach EPON," *Photonic Network Communications*, vol. 19, no. 3, pp. 257-264, Junio 2010.
- [65] N. Merayo, T. Jiménez, P. Fernández, R.J. Durán, R.M. Lorenzo, I. de Miguel, E.J. Abril, "A bandwidth assignment polling algorithm to enhance the efficiency in QoS long-reach EPONs," *European Transactions on Telecommunications*, vol. 22, no. 1, pp. 35-44, Enero 2011.
- [66] J.R. Ferguson, M. Reisslein, M.P. McGarry, "Online excess bandwidth distribution for Ethernet passive optical networks," *Journal of Optical Networking*, vol. 8, no. 4, pp. 358-369, Abril 2009.
- [67] A. Mercian, M.P. McGarry, M. Reisslein, "Offline and Online Multi-Thread Polling in Long-Reach PONs: A Critical Evaluation," *IEEE/OSA Journal of Lightwave Technology*, vol. 31, no. 12, pp. 2018-2028, Junio 2013.
- [68] A. Banerjee, G. Kramer, B. Mukherjee, "Fair Sharing Using Dual Service-Level Agreements to Achieve Open Access in a Passive Optical Network," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 8, pp. 32-44, Agosto 2006.
- [69] M. Ma, Y. Zhu, T.H Cheng, "A bandwidth guaranteed polling MAC protocol for Ethernet passive optical networks," en *22th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2003)*, pp. 22-31, San Francisco (Estados Unidos), 2003.
- [70] C.H. Chang, P. Kourtessis, J.M. Senior. "GPON Service Level Agreement based dynamic bandwidth assignment protocol," *IET Electronic Letters*, vol. 42, no. 20, pp. 1173-1174, Septiembre 2006.

- [71] B. Kantarci, H.T. Mouftah, "Periodic GATE Optimization with QoS-awareness for Long-Reach Passive Optical Networks," en *IEEE Symposium on Computers and Communications (ISCC 2010)*, pp. 879-884, Riccione (Italia), Junio 2010.
- [72] A. Shami, X. Bai, N. Ghani, C.M. Assi, "QoS Control Schemes for Two-Stage Ethernet Passive Optical Networks," *IEEE Journal of Selected Areas in Communications*, vol. 23, no. 8, pp. 1467-1478, Agosto 2005.
- [73] C. Assi, M. Maier, A. Shami, "Toward Quality of Service Protection in Ethernet Passive Optical Networks: Challenges and Solutions," *IEEE Network*, vol. 21, no. 5, pp. 12-19, Septiembre-Octubre 2007.
- [74] A. Dixit, G. Das, B. Lannoo, D. Colle, M. Pickavet, P. Demeester, "Jitter performance for QoS in Ethernet Passive Optical Networks," en *37th European Conference and Exhibition on Optical Communication (ECOC 2011)*, pp. 1-3, Génova (Suiza), Septiembre 2011.
- [75] B. Kantarci, H.T. Mouftah, "Delay-Constrained Admission and Bandwidth Allocation for Long-Reach EPON," *Journal of Networks*, vol. 7, no. 5, pp. 812-820, Mayo 2012.
- [76] B. Kantarci, H.T. Mouftah, "On SLA constraints in dynamic bandwidth allocation for long-reach passive optical networks," en *12th International Conference on Transparent Optical Networks (ICTON 2010)*, pp. 1-7, Munich (Alemania), Julio 2010.
- [77] T. Berisa, A. Bazant, V. Mikac, "Bandwidth and delay guaranteed polling with adaptive cycle time (BDGPACT): a scheme for providing bandwidth and delay guarantees in passive optical networks," *Journal of Optical Networking*, vol. 8, no. 4, pp. 337-345, Abril 2009.
- [78] A. Dixit, B. Lanoo, G. Das, D. Colle, M. Pickavet, P. Demeester, "Dynamic Bandwidth Allocation With SLA Awareness for QoS in Ethernet Passive Optical Networks," *Journal of Optical Communications and Networking*, vol. 5 no. 3, pp. 240-253, Marzo 2013.
- [79] N. Merayo, R.J. Durán, P. Fernández, R.M. Lorenzo, I. de Miguel, E.J. Abril, "EPON bandwidth allocation algorithm based on automatic weight adaptation to provide client and service differentiation," *Photonic Network Communication*, vol. 17, no. 2, pp. 119-128, Abril 2009.

- [80] T. Jiménez, N. Merayo, P. Fernández, R.J. Durán, I. de Miguel, R.M. Lorenzo, E.J. Abril, “Self-adapted algorithm to provide Multi-Profile Bandwidth Guarantees in PONs with Symmetric and Asymmetric Traffic Load,” *Photonic Network Communication*, vol. 24, no. 1, pp. 58-70, Enero 2012.
- [81] T. Jiménez, N. Merayo, P. Fernández, R.J. Durán, R.M. Lorenzo, I. de Miguel, E.J. Abril, “LR-EPON Algorithm with Automatic Bandwidth Adaptation to Provide Multi-profiles Bandwidth Levels,” en *Advanced Photonics and Renewable Energy, 2010 OSA Optics & Photonics Congress. Access Networks and In-House Communications (ANIC 2010)*, Karlsruhe (Alemania), Junio 2010.
- [82] T. Jiménez, N. Merayo, P. Fernández, R.J. Durán, R.M. Lorenzo, N. Fernández, I. de Miguel, E.J. Abril. “Self-adjustment bandwidth algorithm to ensure bandwidth levels in Multi-profile LR-EPONs under heterogeneous traffic load,” en *15th European Conference on Networks and Optical Communications (NOC 2010)*, pp. 353-358, Faro (Portugal), Junio 2010.
- [83] G. Kramer, B. Mukherjee, G. Pesavento. “IPACT: a dynamic protocol for an Ethernet PON (EPON),” *IEEE Communications Magazine*, vol. 40, no. 2, pp. 74-80, Febrero 2002.
- [84] C. Assi, Y. Ye, S. Dixit, and M. Ali, “Dynamic bandwidth allocation for quality-of-service over Ethernet PONs,” *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 9, pp. 1467-1477, Noviembre 2003.
- [85] OPNET Modeler. Disponible en: <http://www.opnet.com>.
- [86] IEEE 802.3. Ethernet in the First Mile Study Group. Disponible en: <http://www.ieee802.org/3/efm/public/public/index.html>.
- [87] Plan de servicio de NTT EAST. Disponible en: <http://flets.com/english/opt/index>.
- [88] A. Gençata, B. Mukherjee, “Virtual-Topology Adaptation for WDM Mesh Networks Under Dynamic Traffic,” *IEEE/ACM Transactions on Networking*, vol. 11, no. 2, pp. 236-247, Abril 2003.
- [89] R.J. Durán, R.M. Lorenzo, N. Merayo, I. de Miguel, P. Fernández, J.C. Aguado, E.J. Abril, “Efficient reconfiguration of logical topologies: Multiobjective design algorithm and adaptation policy,” en *5th International Conference on Broadband*

- Communications, Networks and Systems, (BROADNETS 2008)*, pp. 544-551, Londres (Reino Unido), Septiembre 2008.
- [90] C. Knospe, “PID Control,” *IEEE Control Systems Magazine*, vol. 26, no. 1, pp. 30-31, Febrero 2006.
 - [91] K.J. Aström, T. Hägglund, “The future of PID control,” *Control Engineering Practice*, vol. 9, no. 11, pp. 1163-1175, Noviembre 2001.
 - [92] K.H. Ang, G. Chong, Y. Li, “PID Control System Analysis, Design and Technology,” *IEEE Transactions on Control Systems Technology*, vol. 13, no. 4, pp. 559-576, Julio 2005.
 - [93] P. Cominos, N. Munro, “PID controllers: recent tuning methods and design to specification,” *IEEE Control Theory and Applications*, vol. 149, no. 1, pp. 46-53, Enero 2002.
 - [94] N. Xiong, X. Jia, L.T. Yang, A.V. Vasilakos, Y. Li, Y. Pan, “A Distributed Efficient Flow Control Scheme for Multirate Multicast Networks,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 21, no. 9, pp. 1254-1266, Septiembre 2010.
 - [95] W. Tao, Z. Shun-Yi, “Active Queue Management Based on Single Neural Adaptive PID Algorithm,” en *International Conference on Computer Science and Software Engineering*, vol. 4, pp. 923-926, Wuhan (China), Diciembre 2008.
 - [96] S. Ryu, C. Rump, C. Qiao, “A Predictive and Robust Active Queue Management for Internet Congestion Control,” en *8th IEEE International Symposium on Computers and Communication (ISCC03)*, Julio 2003.
 - [97] T. Tachibana, K. Kogiso, K. Sugimoto, “Dynamic Management of Computing and Network Resources with PID Control in Optical Grid Networks,” en *IEEE International Conference on Communications, (ICC'08)*, pp. 396-400, Pekin (China), Mayo 2008.
 - [98] T. Tachibana, K. Sugimoto, “Lightpath establishment with PID control for effective wavelength utilization in all-optical wavelength-division multiplexing networks,” *Journal of Optical Networking*, vol. 8, no. 4, pp. 383-392, Abril 2009.
 - [99] H.J. Byun, J.M. Nho, and J.T. Lim, “Dynamic bandwidth allocation algorithm in Ethernet passive optical networks,” *Electronics Letters*, vol. 39, no. 13, pp. 1001-1002, Junio 2003.

- [100] S. Yin, N. Ansari, "Nonlinear Predictor-Based Dynamic Resource Allocation over Point-to-Multipoint (P2PM) Networks: A Control Theoretical Approach," *Journal of Optical Communications and Networking*, vol. 2, no. 12, pp. 1052-1062, Diciembre 2010.
- [101] S. Yin, Y. Luo, N. Ansari, T. Wang, "Non-Linear Predictor-Based Dynamic Bandwidth Allocation over TDM-PONs: Stability Analysis and Controller Design," en *IEEE International Conference on Communications (ICC08)*, pp. 5186-5190, Beijing (China), 2008.
- [102] K.J. Aström, T. Hägglund, *PID Controllers: Theory, Design and Tuning*, Ed. Research Triangle Park, NC, Instrumet Soc. Amer., 2nd edition, 1995.
- [103] K.J. Aström, T. Hägglund, *Advanced PID control*. Ed. Research Triangle Park, NC: ISA-The Instrumentation, Systems, and Automation Society, 2006.
- [104] R. Isermann, *Digital Control Systems Vol I: Fundamentals, Deterministic Control*. Ed. Springer-Verlag, 2nd edition, 1989.
- [105] J. Dorsey, *Sistemas de Control Continuos y Discretos*. Ed. McGraw-Hill Interamericana Editores S.A., 2005.
- [106] B.C. Kuo, *Sistemas de control automático*. Ed. Prentice Hall Hispanoamericana, S.A., 1996.
- [107] T. Jiménez, N. Merayo, P. Fernández, R.J. Durán, I. de Miguel, R.M. Lorenzo, E.J. Abril, "Implementation of a PID Controller for the Bandwidth Assignment in Long-Reach PONs," *Journal of Optical Communications and Networking*, vol. 4, no. 5, pp. 392-401, Mayo 2012.
- [108] T. Jiménez, N. Merayo, I. de Miguel, R.J. Duran, P. Fernández, J.C. Aguado, N. Fernández, R.M. Lorenzo, E.J. Abril, "PID strategies for the dynamic allocation of resources in LR-Ethernet Passive Optical Networks," en *16th Optical Network Design and Modeling (ONDM 2012)*, Colchester (Reino Unido), Abril 2012.
- [109] M.T. Ho, Y.W. Tu, "Position control of a single-link flexible manipulator using H_∞ -based PID control," en *IEEE Control Theory Applications*, vol. 153, no. 5, pp. 615-622, Septiembre 2006.
- [110] V. Dharmalingam, S. Sundaram, M. Arumugam, "Conductivity Based PID tuning for long time delay process using simulation technique," en *SICE 2003 Annual Conference*, vol. 1, pp. 191-196, Fukui (Japón), Agosto 2003.

- [111] T. Jiménez, N. Merayo, R.J. Durán, P. Fernández, I. de Miguel, J.C. Aguado, R.M. Lorenzo, E.J. Abril, “A PID-based algorithm to guarantee QoS delay requirements in LR-PONs,” *Optical Switching and Networking*. Aceptado para publicación. D.O.I:<http://dx.doi.org/10.1016/j.osn.2014.01.005>
- [112] S. Sherif, A. Hadjiantonis, G. Ellinas, C. Assi y M. Ali, “A novel decentralized Ethernet-Based PON Access Architecture for Provisioning Differentiated QoS,” *IEEE/OSA Journal of Lightwave Technology*, vol. 22, no. 11, pp. 2483-2497, Noviembre 2004.
- [113] J. Stribling, V. Arunarthi, C. Knittle, D. Murayama, M. Emmendorfer, “Implementing QoS in SIEPON (Service Interoperability in Ethernet Passive Optical Networks),” *IEEE Communications Magazine*, vol. 50, no. 9, pp. 128-135, Septiembre 2012.
- [114] S.I. Choi, J. Huh, “Dynamic bandwidth allocation algorithm for multimedia services over ethernet PONs,” *ETRI Journal*, vol. 24, no. 6, pp. 465-468, Diciembre 2002.
- [115] ITU-T Recomendation G.114, “One-way transmission time. Series G: Transmission Systems and Media, Digital Symstems and Networks”, Telecommunication Standardization Sector of ITU, 2000.
- [116] D.E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*. Ed. Addison-Wesley, 1989.
- [117] K.F. Man, K.S. Tang, S. Kwong, *Genetic algorithms*. Ed. Springer-Verlag. London Limited, 2001.
- [118] M. Gestal, D. Rivero, J.R. Rabuñal, J. Dorado, A. Pazos, *Introducción a los Algoritmos Genéticos y a la Programación Genética*. Ed. Universidade da Coruña. Servizo de Publicacións, 2010.
- [119] G. Lin, G. Liu, “Tuning PID Controller Using Adaptive Genetic Algorithms,” en *5th International Conference on Computer Science & Education*, Hefei (China), Agosto 2010.
- [120] L. Fan, E.M. Joo, “Design for Auto-tuning PID Controller Based on Genetic Algorithms,” en *4th IEEE Conference on Industrial Electronics and Applications, (ICIEA 2009)*, Xi'an (China), Mayo 2009.

- [121] P.J. Rensburg, I.S. Shaw, J.D. Wyk, "Adaptive PID-control using a Genetic Algorithm," en *2nd International Conference on Knowledge-Based Intelligent Electronic System*, Adelaide (Australia), Abril 1998.
- [122] J.C. Shen, "New tuning method for PID controller," *ISA Transactions*, vol. 41, no. 4, pp. 473-484, Octubre 2002.
- [123] P. Wang, D.P. Kwok, "Optimal design of PID process controllers based on Genetic Algorithms," *Control Engineering Practice*, vol. 2, no. 4, pp. 641-648, Agosto 1994.
- [124] C. Li, J. Lian, "The Application of Immune Genetic Algorithm in PID Parameter Optimization for Level Control System," en *IEEE International Conference on Automation and Logistics*, Jinan (China), Agosto 2007.
- [125] C.K. Chen, H.H. Kuo, J.J Yan, T.L. Liao, "GA-based PID active queue management control design for a class of TCP communication networks," *Expert Systems with Applications*, vol. 36, no. 2 parte 1, pp. 1903-1913, Marzo 2009.
- [126] T. Jiménez, N. Merayo, R.J. Durán, P. Fernández, I. de Miguel, J.C. Aguado, R.M. Lorenzo, E.J. Abril, "Auto-Tuning PID controller based on Genetic Algorithms for the Bandwidth Allocation in LR-PONs," en *Advanced Photonics. OSA & Photonic Congress. Access Networks and In-house Communications (ANIC 2012)*, paper AW2A.2, Colorado (Estados Unidos), Junio 2012.
- [127] T. Jiménez, N. Merayo, J.C. Aguado, R.J. Durán, I. de Miguel, P. Fernández, R.M. Lorenzo, E.J. Abril, "Design of auto-Tuning PID controller methods based on Genetic Algorithms for LR-PONs," Aceptado para publicación en *5th International Conference on Access Networks*, Sevilla (España), Junio 2014.
- [128] K.F. Man, K.S. Tang, S. Kwong, "Genetic Algorithms: Concepts and Applications," *IEEE Transactions on Industrial Electronics*, vol. 43, no. 5, pp. 519-534, Octubre 1996.
- [129] M. Wall, GAlib: A C++ Library of Genetic Algorithm Components (version 2.4.7), Massachusetts Institute of Technology. Disponible en: <http://lancet.mit.edu/ga>.
- [130] P. Isasi, I. M. Galván, *Redes de Neuronas Artificiales: Un enfoque práctico*. Ed. Pearson Educación, 2004.
- [131] E.N. Sánchez, A.Y. Alanis, *Redes Neuronales: Conceptos fundamentales y aplicaciones a control automático*. Ed. Pearson Educación, 2006.

- [132] R. Flórez, J.M. Fernández, *Las Redes Neuronales Artificiales. Fundamentos teóricos y aplicaciones prácticas.* Ed. Netbiblo, S.L., 2008.
- [133] R. López, PhD Thesis Neural Networks for Variational Problems in Engineering. Department of Computer Languages and Systems Technical University of Cataluña, Septiembre 2008.
- [134] B. Robitaille, B. Marcos, M. Veillette, G. Payre, “Modified Quasi-Newton Methods for Training Neural Networks,” *Computers & Chemical Engineering*, vol. 20, n. 9, pp. 1133-1140, Septiembre 1996.
- [135] O.T.C. Chen, B.J. Sheu, “Optimization Schemes for Neural Network Training,” en *IEEE World Congress on Computational Intelligence*, vol. 2, pp. 817-822, Orlando (Estados Unidos), Julio 1994.
- [136] A. Likas, A. Stafylopatis, “Training the random neural network using quasi-Newton methods,” *European Journal of Operational Research*, vol. 126, no. 2, pp. 331-339, Octubre 2000.
- [137] T.M. Yang, S.B. Xiong, “Neural network based online self-learning adaptive PID control for automatic ranging-cutting height of shearer,” en *International conference on Machine Learning and Cybernetics, 2002*, vol. 2, pp. 857-859, Pekin (China), Noviembre 2002.
- [138] J. Cao, B. Cao, W. Chen, P. Xu, “Neural network self-adaptive PID control for driving and regenerative braking of electric vehicle,” en *IEEE International Conference on Automation and Logistics*, pp. 2029-2034, Jinan (China), Agosto 2007.
- [139] M.G. Zhang, Z.G. Wang, P. Wang, “Adaptive PID decoupling control based on RBF neural network and its application,” en *International Conference on Wavelet Analysis and Pattern Recognition, 2007 (ICWAPR '07)*, vol. 2, pp. 727-731, Pekin (China), Noviembre 2007.
- [140] B. Guo, H. Liu, Z. Luo, F. Wang, “Adaptive PID controller based on BP neural network,” en *International Joint Conference on Artificial Intelligence, 2009 (JCAI '09)*, pp. 148-150, Hainan (China), Abril 2009.
- [141] B. Guo, H. Liu, Z. Luo, F. Wang, “Study of PID neural network for hydraulic system,” en *IEEE International Conference on Automation and Logistics 2009 (ICAL '09)*, pp. 228-232, Shenyang (China), Agosto 2009.

- [142] Z.S. Jiang , D.K. Li, Q.T. Meng, “PID controller based on BP neural network in the application of wind power generation and Matlab simulation,” en *2nd International Conference on Information Engineering and Computer Science (ICIECS 2010)*, pp. 1-4, Wuhan (China), Diciembre 2010.
- [143] S. Lizheng, G. Jun, W. Jianhua, “Pulp concentration control by PID with BP neural network in the production of light weight cardboard,” en *IEEE 11th International Conference on Computer-Aided Industrial Design & Conceptual Design (CAIDCD 2010)*, vol. 2, pp. 1217-1220, Yiwu (China), Noviembre 2010.
- [144] A. Abraham, “Meta learning evolutionary artificial neural networks,” *Neurocomputing*, vol. 56, no. 1-4, pp. 1-38, Enero 2004.
- [145] Open Neural Networks Library (OpenNN). Disponible en: <http://opennn.cimne.com>.
- [146] I. de Miguel, R.J. Durán, T. Jiménez, N. Fernández, J.C. Aguado, R.M. Lorenzo, A. Caballero, I. Tafur, Y. Ye, A. Tymecki, I. Tomkos, M. Angelou, D. Klonidis, A. Francescon, D. Siracusa, E. Salvadori, “Cognitive Dynamic Optical Networks,” *Journal of Optical Communications and Networking*, vol. 5, no. 10, pp. A107-A118, Octubre 2013.
- [147] W. Wei, C. Wang, J. Yu, “Cognitive Optical Networks: Key Drivers, Enabling Techniques, and Adaptive Bandwidth Services,” *IEEE Communications Magazine*, vol. 50, no. 1, pp. 106-113, Enero 2012.
- [148] G.S. Zervas, D. Simeonidou, “Cognitive Optical Networks: Need, Requirements and Architecture,” en *12th International Conference on Transparent Optical Networks (ICTON 2010)*, pp. 1-4, Munich (Alemania), Julio 2010.
- [149] R.W. Thomas, D.H. Friend, L.A. DaSilva, A.B. MacKenzie, “Cognitive networks: Adaptation and learning to achieve end-to-end performance objectives,” *IEEE Communications Magazine*, vol. 44, no. 12, pp. 51-57, Diciembre 2006.
- [150] A. Aamodt, E. Plaza, “Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches,” *Artificial Intelligence Communications*, vol. 7, no. 1, pp. 39-59, 1994.
- [151] I. Watson, *Applying Case-Based Reasoning: Techniques for Enterprise Systems*. Ed. Morgan Kaufmann Publishers, Inc., 1997.

- [152] S.K. Pal, S.C.K. Shiu, *Foundations of Soft Case-Based Reasoning*. Ed. John Wiley & Sons, Inc., 2004.
- [153] D. W. Aha, "Tolerating noisy, irrelevant and novel attributes in instance-based learning algorithms," *International Journal of Man-Machine Studies*, vol. 36, no. 2, pp. 267-287, Febrero 1992.
- [154] D. Wettscherek, D.W. Aha, T. Mohri , "A review and empirical evaluation of feature weighting methods for a class of lazy learning algorithms," *Artificial Intelligence Review*, vol. 11, no. 1-5, pp. 273-314, Febrero 1997.
- [155] R. Ramaswami, K. Sivarajan, G. Sasaki, *Optical Networks: A Practical Perspective*. Ed. Morgan Kaufmann, 3^a edición, 2009.
- [156] S. Azodolmolky, M. Klinkowski, E. Marin, D. Careglio, J. Solé, I. Tomkos, "A survey on physical layer impairments aware routing and wavelength assignment algorithms in optical networks," *Computer Networks*, vol. 53, no. 7, pp. 926-944, Mayo 2009.
- [157] C.V. Saradhi, S. Subramaniam, "Physical layer impairment aware routing (PLIAR) in WDM optical networks: Issues and challenges," *IEEE Communications Surveys & Tutorials*, vol. 11, no. 4, pp. 109-130, Diciembre 2009.
- [158] S. Azodolmolky, J. Perelló, M. Angelou, F. Agraz, L. Velasco, S. Spadaro, Y. Pointurier, A. Francescon, C.V. Saradhi, P. Kokkinos, E. Varvarigos, S.A. Zahr, M. Gagnaire, M. Gunkel, D. Klonidis, I. Tomkos, "Experimental demonstration of an impairment aware network planning an operation tool for transparent/translucent optical networks," *IEEE/OSA Journal of Lightwave Technology*, vol 29, no. 4, pp. 439-448, Febrero 2011.
- [159] S. Azodolmolky, P. Kokkinos, M. Angelou, E. Varvarigos, I. Tomkos, "DICO-NET NPOT: An impairments aware tool for planning and managing dynamic optical networks," *Journal of Network and Systems Management*, vol. 20, no. 1, pp. 116-133, Marzo 2012.
- [160] R. Schmogrow, B. Nevendahl, M. Winter, A. Josten, D. Hillerkuss, S. Koenig, J. Meyer, M. Dreschmann, M. Huebner, C. Koos, J. Becker, W. Freude, J. Leuthold, "Error vector magnitude as a performance measure for advanced modulation formats," *IEEE Photonics Technology Letters*, vol. 24, no. 1, pp. 61-63, Enero 2012.

- [161] Y. Qin, K.C.S. Cheng, J. Triay, E. Escalona, G.S. Zervas, G. Zarris, N. Amaya-Gonzalez, C. Cervelló-Pastor, R. Nejabati, D. Simeonidou, “Demonstration of C/S based hardware accelerated QoT estimation tool in dynamic impairment-aware optical network,” en *36th European Conference and Exhibition on Optical Communication (ECOC 2010)*, Turin (Italia), Septiembre 2010.
- [162] N. Fernández, R.J. Durán, I. de Miguel, N. Merayo, D. Sánchez, M. Angelou, J.C. Aguado, P. Fernández, T. Jiménez, R.M. Lorenzo, I. Tomkos, E.J. Abril, “Cognition to design energetically efficient and impairment aware virtual topologies for optical networks,” en *16th Optical Network Design and Modeling (ONDM 2012)*, Colchester (Reino Unido), Abril 2012.
- [163] P. Poggolini, “The GN model of non-linear propagation in uncompensated coherent optical systems,” *IEEE/OSA Journal of Lightwave Technology*, vol. 30, no. 24, pp. 3857-3879, Diciembre 2012.
- [164] A. Caballero, J.C. Aguado, R. Borkowski, S. Saldaña, T. Jiménez, I. de Miguel, V. Arlunno, R.J. Durán, D. Zibar, J.B. Jensen, R.M. Lorenzo, E.J. Abril, I. Tafur, “Experimental demonstration of a cognitive quality of transmission estimator for optical communication systems,” *Optics Express*, vol. 20, no. 26, pp. B64-B70, Diciembre 2012.
- [165] T. Jiménez, J.C. Aguado, I. de Miguel, R.J. Durán, M. Angelou, N. Merayo, P. Fernández, R.M. Lorenzo, I. Tomkos, E.J. Abril, “A Cognitive Quality of Transmission Estimator for Core Optical Networks,” *IEEE/OSA Journal of Lightwave Technology*, vol. 31, no. 6, pp. 942-951, Marzo 2013.
- [166] T. Jiménez, J.C. Aguado, I. de Miguel, R.J. Durán, N. Fernández, M. Angelou, D. Sánchez, N. Merayo, P. Fernández, N. Atallah, R.M. Lorenzo, I. Tomkos, E.J. Abril, “A cognitive system for fast quality of transmission estimation in core optical network,” en *OFC/NFOEC 2012*, paper OW3A.5, Los Ángeles (Estados Unidos), Marzo 2012.
- [167] T. Jiménez, J.C. Aguado, I. de Miguel, R.J. Durán, D. Sánchez, M. Angelou, N. Merayo, P. Fernández, N. Fernández, R.M. Lorenzo, I. Tomkos, E.J. Abril, “Optimization of the knowledge base of a cognitive quality of transmission estimator for core optical networks,” en *16th Optical Network Design and Modeling (ONDM 2012)*, Colchester (Reino Unido), Abril 2012.

- [168] T. Jiménez, J.C. Aguado, I. de Miguel, R.J. Durán, N. Fernández, M. Angelou, D. Sánchez, N. Merayo, P. Fernández, N. Atallah, R. Lorenzo, I. Tomkos, E.J. Abril, "Enhancing optical networks with cognition: Case-Based Reasoning to estimate the quality of transmission," en *IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA 2012)*, pp. 166-169, Nueva Orleans (Estados Unidos), Marzo 2012.
- [169] Proyecto Europeo CHRON. Entregable D2.1: "Specification of the network scenario and service requirements", Noviembre 2011. Disponible en: <http://www.ict-chron.eu>.
- [170] M. Angelou, Y. Pointurier, D. Careglio, S. Spadaro, I. Tomkos, "Optimized monitor placement for accurate QoT assessment in core optical networks," *IEEE/OSA Journal of Lightwave Technology*, vol. 4, no. 1, pp. 15-24, Enero 2012.
- [171] R. López, D. Mcsherry, D. Bridge, D. Leake, B. Smyth, S. Craw, B. Faltings, M. Lou, M.T. Cox, K. Forbus, M. Keane, A. Aamondt, I. Watson, "Retrieval, reuse, revision and retention in case-based reasoning," *The Knowledge Engineering Review*, vol. 30, no. 3, pp. 215-240, Cambridge University Press, 2006.
- [172] M. Salamó, M. López-Sánchez, "Adaptive case-base reasoning using retention and forgetting strategies," *Knowledge-Based Systems*, vol. 24, no. 2, pp. 230-247, Marzo 2011.
- [173] S.J. Delany, P. Cunningham, "An analysis of case-base editing in a spam filtering system," *Advances in Case-Based Reasoning*, Springer, pp. 128-141, 2004.
- [174] E. McKenna, B. Smyth, "Competence-guided editing methods for lazy learning," en *14th European Conference on Artificial Intelligence*, pp. 60-64, Berlín (Alemania), 2000.
- [175] A. Mokhtar, M. Azizoglu, "Adaptive wavelength routing in all-optical networks," *IEEE/ACM Transactions on Networking*, vol. 6, no. 2, pp. 197-206, Abril 1998.
- [176] A. Zapata-Beghelli, P. Bayvel, "Dynamic versus static wavelength-routed optical networks," *IEEE/OSA Journal of Lightwave Technology*, vol. 26, no. 10, pp. 3403-3415, Octubre 2008.
- [177] C. Kachris, D. Klonidis, A. Francescon, D. Siracusa, E. Salvadori, N. Fernández, T. Jiménez, R.J. Durán, I. de Miguel, J.C. Aguado, R.M. Lorenzo, R. Borkowski,

A. Caballero, I. Tafur, Y. Ye, A. Tymecki, I. Tomkos, “Experimental Demonstration of a Cognitive Optical Network for Reduction of Restoration Time,” en *OFC/NFOEC 2014*, paper W2A.28, San Francisco (Estados Unidos), Marzo 2014.

Lista de Acrónimos

APON	<i>Asynchronous Passive Optical Network</i>
ASE	<i>Amplified Spontaneous Emission</i>
AQM	<i>Active Queue Management</i>
AWG	<i>Arrayed Waveguide Grating</i>
BER	<i>Bit Error Rate</i>
BGP	<i>Bandwidth Guaranteed Pollling</i>
BPON	<i>Broadband Passive Optical Network</i>
CapEx	<i>Capital Expenditure</i>
CBR	<i>Case-Based Reasoning</i>
CBR-SPID	<i>Case-Based Reasoning Service level agreement Proportional Integral Derivative</i>
CHRON	<i>Cognitive Heterogeneous Reconfigurable Optical Network</i>
CO	<i>Central Office</i>
CoS	<i>Class of Service</i>
CRR	<i>Conservative Redundancy Reduction</i>
CS	<i>Coverage Set</i>
DaSPID	<i>Delay aware Service level agreement Proportional Integral Derivative</i>
DBA	<i>Dynamic Bandwidth Allocation</i>
DBRu	<i>Dynamic Bandwidth Report upstream</i>
DCF	<i>Dispersion Compensating Fibre</i>
DISCUS	<i>DIStributed Core for unlimited bandwidth supply for all Users and Services</i>
DSL	<i>Digital Subscriber Line</i>
DT	<i>Deutsche Telekom</i>
DWDM	<i>Dense Wavelength Division Multiplexing</i>

DySGAB	<i>Dynamic Service levels with GuAranteed Bandwidth</i>
DySLa	<i>Dynamic Service Level agreement</i>
EAM	<i>Electro Absorption Modulator</i>
EDFA	<i>Erbium Doped Fiber Amplifier</i>
EIS	<i>Enhanced Inter-Thread Scheduling</i>
EPON	<i>Ethernet Passive Optical Network</i>
EVM	<i>Error Vector Magnitude</i>
FEC	<i>Forward Error Correction</i>
FiWi	<i>Fiber Wireless</i>
FixE-CBR	<i>Fixed Error Case-Based Reasoning</i>
FPGA	<i>Field Programmable Gate Array</i>
FSAN	<i>Full Service Access Network</i>
FTTB	<i>Fiber To The Building</i>
FTTC	<i>Fiber To The Curb</i>
FTTH	<i>Fiber To The Home</i>
FTTN	<i>Fiber To The Neighborhood/Node</i>
FTTx	<i>Fiber To The x</i>
GA-SPID	<i>Genetic Algorithm Service level agreement Proportional Integral Derivative</i>
GA-DaSPID	<i>Genetic Algorithm Delay aware Service level agreement Proportional Integral Derivative</i>
GPON	<i>Gigabit Passive Optical Network</i>
HDTV	<i>High Definition Television</i>
IEEE	<i>Institute of Electrical and Electronic Engineers</i>
ILP	<i>Integer Linear Programming</i>
IP	<i>Internet Protocol</i>
ITU-T	<i>International Telecommunication Union</i>
KB	<i>Knowledge Base</i>
LAN	<i>Local Area Network</i>
LIPSA	<i>Long-Reach Interleaved Polling algorithm with Service level Agreement</i>
LOHEDA	<i>LOng reach Highly Efficient Dynamic bandwidth Assignment</i>
LR-EPON	<i>Long-Reach Ethernet Passive Optical Network</i>
LR-PON	<i>Long-Reach Passive Optical Network</i>
MAC	<i>Medium Access Control</i>
MAP	<i>Metro Access Points</i>
MPCP	<i>Multi-Point Control Protocol</i>

MT	<i>Multi-Thread polling</i>
MTP	<i>online Multi-Thread Polling</i>
NA+	<i>Newly Arrived Frames Plus</i>
NG-PON1	<i>Next Generation Passive Optical Network 1</i>
NG-PON2	<i>Next Generation Passive Optical Network 2</i>
NN-SPID	<i>Neural Network Service level agreement Proportional Integral Derivative</i>
NTT	<i>Nippon Telegraph and Telephone</i>
OADM	<i>Optical Add-Drop Multiplexer</i>
OEBD	<i>Online Excess Bandwidth Distribution</i>
OEO	<i>Optical-Electrical-Optical</i>
OFDM	<i>Orthogonal Frequency Division Multiplexing Passive Optical Network</i>
OLT	<i>Optical Line Terminal</i>
ONU	<i>Optical Network Unit</i>
OOK	<i>On-Off Keying</i>
OpEx	<i>Operational Expenditure</i>
OSNR	<i>Optical Signal to Noise Ratio</i>
P2P	<i>Point to Point</i>
PCBd	<i>Physical Control Block downstream</i>
PDM-QPSK	<i>Polarization Division Multiplexed - Quadrature Phase Shift Keying</i>
PGO	<i>Periodic Gate Optimization</i>
PIEMAN	<i>Photonic Integrated Extended Metro and Access Network</i>
PID	<i>Proportional-Integral-Derivative</i>
PON	<i>Passive Optical Network</i>
QoE	<i>Quality of Experience</i>
QoS	<i>Quality of Service</i>
QoT	<i>Quality of Transmission</i>
R-CBR	<i>Regular Case-Based Reasoning</i>
RF	<i>Radio Frequency</i>
RN	<i>Remote Node</i>
RSOA	<i>Reflective Semiconductor Optical Amplifier</i>
RTT	<i>Round Trip Time</i>
RWA	<i>Route and Wavelength Assignment</i>
SARDANA	<i>Scalable Advanced Ring Dense Access Network Architecture</i>

SDH	<i>Synchronous Digital Hierarchy</i>
SDTV	<i>Standard Definition Television</i>
SLA	<i>Service Level Agreement</i>
SMF	<i>Single Mode Fibre</i>
SNR	<i>Signal to Noise Ratio</i>
SOA	<i>Semiconductor Optical Amplifier</i>
SPID	<i>Service level agreement Proportional Integral Derivative</i>
SSFM	<i>Split-Step Fourier Method</i>
TCP	<i>Transmission Control Protocol</i>
TDMA	<i>Time Division Multiple Access</i>
TSD	<i>Two-State DMB</i>
TWDM-PON	<i>Time Wavelength Division Multiplexing Passive Optical Network</i>
UDWDM	<i>Ultra Dense Wavelength Division Multiplexing</i>
VDSL	<i>Very high speed Digital Subscriber Line</i>
VPN	<i>Virtual Private Network</i>
WC	<i>Wavelength Converter</i>
WDM	<i>Wavelength Division Multiplexing</i>
WDMA	<i>Wavelength Division Multiple Access</i>
WDM-PON	<i>Wavelength Division Multiplexing Passive Optical Network</i>
WE-PON	<i>Wavelength division multiplexing Ethernet Passive Optical Network</i>
WRON	<i>Wavelength-Routed Optical Networks</i>
XG-PON	<i>10 Gigabit Capable Passive Optical Network</i>
XL-PON	<i>eXtra-Large Passive Optical Network</i>