



DEPARTAMENTO DE ELECTRÓNICA Y ELECTROMAGNETISMO

MINING AND CONTROL OF NETWORK TRAFFIC BY
COMPUTATIONAL INTELLIGENCE

MINERÍA DE DATOS Y CONTROL DE TRÁFICO DE RED
MEDIANTE INTELIGENCIA COMPUTACIONAL

Tesis Doctoral
Resumen en español

Federico Montesino Pouzols

Directores:

Ángel
Barriga Barros

Diego Rafael
Lopez García

Noviembre de 2008

**Departamento de Electrónica
y Electromagnetismo**

Universidad de Sevilla

**Instituto de Microelectrónica
de Sevilla**

Centro Nacional de
Microelectrónica, CSIC

**MINING AND CONTROL OF NETWORK TRAFFIC BY
COMPUTATIONAL INTELLIGENCE**

**MINERÍA DE DATOS Y CONTROL DE TRÁFICO DE RED
MEDIANTE INTELIGENCIA COMPUTACIONAL**

RESUMEN EN ESPAÑOL

**Memoria presentada para aspirar al
grado de Doctor por**

**Federico Montesino Pouzols
Ingeniero en Informática**

MINING AND CONTROL OF NETWORK TRAFFIC BY
COMPUTATIONAL INTELLIGENCE

MINERÍA DE DATOS Y CONTROL DE TRÁFICO DE RED
MEDIANTE INTELIGENCIA COMPUTACIONAL

RESUMEN EN ESPAÑOL

Memoria presentada por
Federico Montesino Pouzols
para aspirar al grado de Doctor

Los directores:

Dr. D. Ángel
Barriga Barros

Profesor Titular de Universidad
Dpto. Electrónica y Electromagnetismo

Dr. D. Diego Rafael
Lopez García

Doctor en Ciencias Físicas
RedIRIS, Red.es

Departamento de Electrónica y Electromagnetismo
Universidad de Sevilla

La estructura y el comportamiento de las redes de conmutación de paquetes son difíciles de modelar de manera análoga a muchos otros sistemas naturales y artificiales. No obstante, Internet es un caso extraordinario que plantea numerosos retos debido a su rápido desarrollo, elevada heterogeneidad y a la ausencia de mecanismos de medida y monitorización en su concepción esencial. En esta tesis aplicamos métodos de inteligencia computacional, especialmente aquellos basados en lógica difusa, a un conjunto de problemas actuales en torno a la medida, análisis y control de tráfico en Internet. Se diseñan y analizan nuevos métodos y se aborda la implementación hardware de algunos de ellos.

En primer lugar, se analizan, mediante técnicas de minería de datos, series temporales de tráfico de redes y medidas de flujos de red. Se desarrolla un método de predicción de series temporales mediante sistemas de inferencia difusa interpretables con ajuste automático, combinados con una técnica no paramétrica de estimación de la varianza residual. Las ventajas de este método respecto a otras técnicas se ilustran mediante series de referencia así como un extenso conjunto de series de tráfico de redes. Asimismo, se desarrolla un método de análisis y resumen lingüístico de medidas de flujos de red basado en resúmenes lingüísticos difusos. Se muestra que el método es rápido y proporciona resúmenes apropiados y concisos para un extenso conjunto de trazas de paquetes y colecciones de flujos de red.

A continuación abordamos el control de tráfico en Internet. Se desarrollan métodos tanto para control de congestión de extremo a extremo como para control de tráfico a nivel IP. En cuanto al control de extremo a extremo, se desarrolla un esquema de control de extremo a extremo que generaliza el mecanismo de ventana deslizante de TCP mediante lógica difusa. El método se evalúa en escenarios simulados, emulados y en producción, ilustrándose sus ventajas frente a las alternativas tradicionales. En cuanto al control a nivel IP, se diseñan controladores difusos para gestión activa de colas, y se realiza una detallada evaluación de estos en escenarios simulados y emulados, comparándose asimismo frente a alternativas tradicionales.

Por último, se estudia la implementación hardware de algunos de los métodos desarrollados en esta tesis. Proponemos una plataforma abierta basada en dispositivos FPGA, incluyendo una metodología de implementación para el desarrollo de componentes difusos en sistemas digitales complejos, especialmente en arquitecturas de routers modernas. Se muestra que los sistemas hardware de inferencia difusa analizados satisfacen los requisitos operacionales de routers de altas prestaciones, tanto actualmente como en un futuro anticipable, en términos de velocidad de inferencia y consumo de recursos.

Palabras clave: inteligencia computacional, tráfico de red, soft computing, lógica difusa, inferencia difusa, predicción de series temporales, estimación de la varianza residual, resúmenes lingüísticos, reglas de asociación, medida de redes, métricas de tráfico, prestaciones extremo a extremo, TCP, control de congestión, control de tráfico, gestión activa de colas, arquitectura de routers, hardware abierto, FPGA, módulos IP, VHDL, aprendizaje supervisado, aprendizaje no supervisado, máquinas de soporte vectorial, vecinos más cercanos, agrupamiento.

Contenido

1. Introducción	11
1.1. Contribuciones de esta tesis	13
1.2. Estructura de la tesis	14
2. Ciencia de Internet	17
3. Minería de datos sobre tráfico de red mediante inteligencia computacional	21
3.1. Modelos de series temporales basados en sistemas de inferencia difusa	22
3.2. Predicción de la carga de tráfico en enlaces de red	24
3.3. Análisis y resumen de registros de flujos de red	40
3.4. Conclusiones	44
4. Sistemas de inferencia difusa para control de tráfico de red	45
4.1. Escenarios de simulación	46
4.2. Control difuso de extremo a extremo para protocolos de transporte	48
4.3. Gestión activa de colas mediante sistemas de inferencia difusa	55
4.4. Conclusiones	57
5. Plataforma abierta de desarrollo basada en FPGA para sistemas de inferencia difusa	59
5.1. Arquitecturas de routers	60
5.2. Plataforma de desarrollo	62
5.3. Conclusiones	70
6. Conclusiones	71
Bibliografía	75

1. Introducción

La tesis que presentamos bajo el título de «minería de datos y control de tráfico de red mediante inteligencia computacional» aborda mediante técnicas de inteligencia computacional un conjunto de problemas actuales en los campos de medida, análisis y control de tráfico en redes de conmutación de paquetes.

Los problemas tratados en esta tesis incluyen el modelado predictivo de la carga de tráfico en enlaces, la generación de resúmenes de colecciones de flujos de red, la gestión activa de colas en routers y el control de congestión de extremo a extremo. Para ello, se utilizan diversas técnicas de inteligencia computacional, con especial énfasis en los sistemas basados en lógica difusa.

La estructura y el comportamiento de las redes de conmutación de paquetes es difícil de modelar de un modo similar a muchos sistemas naturales y artificiales. No obstante, Internet es a la vez un caso excepcional y un desafío debido a su rápido desarrollo, alta heterogeneidad y a la inherente ausencia de mecanismos de medida y monitorización en su diseño básico. De manera breve, las redes de conmutación de paquetes suponen un desafío para el modelado analítico.

De manera general, se acepta que la investigación en torno a protocolos y mecanismos de Internet requiere mejores modelos que los disponibles actualmente. Durante la última década, el creciente desarrollo de los sistemas e infraestructuras de medida de tráfico ha hecho posibles numerosos avances en la comprensión del funcionamiento de e interacciones entre los mecanismos básicos de Internet. En particular, varios trabajos en torno al área de medida de tráfico han dado lugar a los primeros resultados en lo que algunos autores llaman Ciencia de Internet, esto es, una ciencia experimental que estudia leyes y patrones en la estructura de Internet.

El estudio la dinámica del tráfico de red constituye un área de investigación de complejidad creciente, que plantea numerosos retos. Durante los pasados años se ha avanzado significativamente en el desarrollo de infraestructuras y sistemas de medida de tráfico. Gran parte de estas infraestructuras y sistemas están basados en medida pasiva mediante flujos de tráfico. Con una creciente diversidad de tecnologías, aplicaciones y patrones de tráfico, el análisis de flujos de red es un proceso cada vez más complejo. Actualmente la comprensión de todos los hechos relevantes derivables de una colección de datos sobre flujos de red está más allá de las posibilidades prácticas de los operadores, gestores y planificadores.

Un problema de investigación abierto de esencial importancia para las capas de red y transporte de los protocolos de Internet es el desarrollo de mecanismos de regulación de tráfico capaces de satisfacer los requisitos de una creciente diversidad de tecnologías, aplicaciones y servicios.

Las tendencias tecnológicas que durante los últimos años han venido afectando al diseño de routers troncales y equipos de comunicaciones de altas prestaciones en general han generado unas duras restricciones operacionales sobre los mecanismos de control de tráfico. En particular, las elevadas tasas de procesado de paquetes necesarias restringen el diseño de diversos mecanismos de análisis y control. Así, han surgido dos restricciones fundamentales: las implementaciones han de ser escalables y flexibles.

En este contexto, diversas técnicas de inteligencia computacional, y, en particular, las basadas en lógica difusa están ganando aceptación como herramientas para análisis y control de tráfico de red. Los sistemas difusos encuentran aplicaciones en un conjunto de áreas tales como regulación de tráfico en routers, soporte de servicios diferenciados dentro de la arquitectura DiffServ, evaluación de políticas y calidad de servicio, medida análisis y monitorización de tráfico en tiempo real, ahorro de consumo en redes inalámbricas, así como control de tráfico de extremo a extremo tanto en redes cableadas como inalámbricas.

Los sistemas de inferencia difusa (FIS o FRBS, por sus siglas en inglés) ocupan un papel central en esta tesis. Se utilizan para realizar tareas tales como evaluación de prestaciones, modelado predictivo y control. No obstante, además de técnicas basadas en lógica difusa, se aplican otras técnicas de inteligencia computacional tales como métodos estadísticos no paramétricos, operadores OWA (Ordered Weighted Averaging), algoritmos de minería de reglas de asociación, least squares support vector machines y ciertos tipos específicos de redes neuronales.

La lógica difusa es una lógica precisa de la imprecisión, basada en el concepto de conjunto difuso. Los sistemas de inferencia difusa son sistemas de modelado basados en reglas. Una razón fundamental que ha motivado el uso de mecanismos de inferencia difusa en múltiples aplicaciones es el hecho que que las reglas difusas se pueden expresar de forma lingüística, siendo por tanto altamente interpretables por los seres humanos. Así, es posible usar el conocimiento a priori para construir modelos. Además, los sistemas basados en inferencia difusa permiten la evaluación e interpretación de modelos por parte de expertos. Así, se han propuesto numerosos métodos para generar diferentes tipos de modelos de inferencia difusa manteniendo un cierto equilibrio entre interpretabilidad y precisión.

Una característica adicional de los sistemas de inferencia difusa es el hecho de que sean aproximadores universales bajo restricciones asumibles en la práctica. Asimismo, los llamados sistemas neuro-difusos combinan los FIS con las capacidades de aprendizaje de las redes neuronales artificiales (ANN), utilizando a menudo los mismos algoritmos de aprendizaje que fueron desarrollados inicialmente para las ANN.

El enfoque general seguido para la realización de esta tesis ha sido abordar problemas concretos en el área del análisis y control de tráfico mediante técnicas específicas basadas en lógica difusa. El conjunto de problemas tratados se ha elegido tomando como base su interés práctico en las redes de computadores actuales, y con objeto de proporcionar un enfoque unificado al área de análisis y control de tráfico. Por supuesto, en esta tesis no tratamos todos los problemas abiertos en el área mencionada. No obstante, el conjunto de métodos propuestos y aplicados aquí constituye un enfoque considerablemente completo para tratar los problemas actuales en este contexto. Este conjunto de métodos queda asimismo abierto a numerosas ampliaciones para afrontar problemas adicionales, tanto actuales como futuros.

En particular, abordamos problemas de minería de datos y control. En la primera clase incluimos dos problemas: el modelado predictivo de la carga de tráfico, y la generación de resúmenes y análisis inductivo de medidas de flujos de red. En la segunda clase incluimos otros dos problemas: la gestión activa de colas en routers de Internet así como el control de congestión de extremo a extremo.

Todos los experimentos incluidos en esta tesis se han realizado utilizando conjuntos de datos y software disponible públicamente. A lo largo del texto se proporcionan las corres-

pondientes referencias y enlaces.

El resto de esta tesis aborda los problemas mencionados. Si bien se describen algunos desarrollos teóricos, hacemos mayor énfasis en la evaluación de modelos utilizando datos reales por medio de simulaciones e implementaciones experimentales. En el siguiente epígrafe enumeramos de manera breve las contribuciones de esta tesis. A continuación, resumimos el contenido de los capítulos siguientes.

1.1. Contribuciones de esta tesis

En esta tesis tratamos los siguientes problemas en torno al modelado, análisis y control del tráfico de redes:

- Predicción de la carga de tráfico.
- Generación de resúmenes y análisis inductivo de medidas de flujos de red.
- Control de tráfico de extremo a extremo.
- Control de tráfico en routers.
- Obtención de implementaciones prácticas y eficientes de sistemas difusos para los problemas anteriores, que sean integrables en equipos actuales.

En estas áreas, las contribuciones desarrolladas en el presente trabajo se pueden enumerar como sigue:

- Analizamos sistemáticamente la predecibilidad del tráfico de red en diferentes escalas temporales. Se sigue un enfoque cuantitativo basado en técnicas no paramétricas para estimación de la varianza residual. Con un extenso conjunto experimental de trazas de tráfico de red, se muestra en qué grado es posible predecir la carga de tráfico en diversos escenarios.
- Proponemos una metodología para analizar series temporales de tráfico mediante FIS. Combinando técnicas estadísticas así como técnicas neuro-difusas, se desarrolla un marco metodológico para derivar de modo automático modelos predictivos basados en FIS como aproximadores no lineales. La metodología propuesta se aplica a un extenso conjunto experimental de trazas de tráfico y se implementa en una herramienta integrada en el entorno de desarrollo Xfuzzy.
- Se desarrolla un método para extraer resúmenes lingüísticos concisos a partir de medidas de flujos de red. El método se implementa en una herramienta capaz de operar en tiempo real. Además, se desarrolla un procedimiento para realizar minería de datos sobre colecciones de medidas de flujos de red. Ambos métodos son robustos ante técnicas de muestreo de medidas y se aplican igualmente sobre un extenso conjunto experimental de trazas de tráfico.

- Proponemos un esquema de control de tráfico de extremo a extremo mediante FIS. Se lleva a cabo una evaluación mediante simulación e implementación, comparando la propuesta frente a controladores tradicionales en un amplio conjunto de escenarios realistas.
- Se propone un esquema de control de tráfico en routers basado en FIS. El esquema se evalúa asimismo en un amplio conjunto de escenarios realistas a fin de obtener una propuesta robusta en el mayor rango de situaciones. La propuesta se compara frente a un conjunto de esquemas tradicionales, evaluándose las prestaciones atendiendo a diversas métricas.
- Se desarrolla una plataforma abierta basada en dispositivos FPGA para la implementación de FIS eficientes en el área de análisis y control de tráfico. Analizamos la factibilidad de la integración de los controladores propuestos en las arquitecturas de routers actuales y previsibles a medio plazo.

1.2. Estructura de la tesis

Este trabajo consta de tres partes. En primer lugar, se desarrollan métodos de inteligencia computacional para realizar minería de datos sobre tráfico de red. En particular, se desarrollan métodos automáticos para extraer modelos predictivos e interpretables de carga de tráfico, resumir medidas de flujos de red y extraer reglas lingüísticas de asociación acerca de registros de flujos de red.

A continuación, definimos dos esquemas de control basados en FIS para control de tráfico. El primer método es de aplicación en control de la tasa de transferencia y congestión de extremo a extremo, mientras que el segundo está orientado a gestión activa de colas en routers.

Por último, en la tercera parte abordamos el problema de la implementación práctica de algunos de los problemas propuestos, que plantean ciertos retos de cara a su implementación en equipos de comunicaciones. Se consideran los requisitos tanto arquitecturales como operacionales. Como resultado de este análisis, proponemos una plataforma hardware abierta basada en dispositivos FPGA para la implementación eficiente de FIS en problemas relacionados con el análisis y control de tráfico de red.

Se han empleado diversas técnicas de modelado no lineal, incluyendo sistemas difusos, ciertos tipos de redes neuronales y máquinas de soporte vectorial. En esta tesis obviamos los detalles de estas técnicas. Para una visión general de los sistemas difusos, el lector interesado puede consultar, entre otros muchos, el siguiente artículo de Bezdek [8], mientras que para una discusión clara de las aplicaciones y relaciones entre las tres técnicas mencionadas, se puede consultar el siguiente libro de Kecman [60].

Asimismo, Jang [54] proporciona en su libro una revisión de muchas de las técnicas de soft computing empleadas en esta tesis, desde una perspectiva práctica, incluyendo conjuntos difusos, sistemas de inferencia difusa, técnicas de regresión y optimización comunes basadas en mínimos cuadrados y derivadas, métodos de optimización sin derivadas, aprendizaje supervisado y por refuerzo en redes neuronales, aprendizaje no supervisado y modelado neuro-difuso.

En el capítulo 2, introducimos los elementos básicos de un conjunto de resultados recientes que algunos autores engloban bajo el término Ciencia de Internet. Se dan descripciones concisas de algunos de estos elementos básicos así como otros aspectos de las redes de comunicaciones que serán clave en los siguientes capítulos. Además, se describen brevemente algunas de las técnicas de predicción de series temporales aplicadas hasta la fecha al modelado del tráfico de redes, así como otras técnicas cuyo empleo se propone en los siguientes capítulos.

En el capítulo 3, nos centramos en la minería de datos sobre tráfico de red. Se abordan dos problemas relacionados: la predicción de la carga de tráfico en enlaces, y el análisis de medidas de flujos de red.

El capítulo 4 trata sobre el control de tráfico de red desde dos perspectivas. Por una parte, se aborda el control de tráfico en routers mediante esquemas de gestión activa de colas. Por otra parte, se introduce un nuevo esquema de control de congestión de extremo a extremo mediante mecanismos de ventana deslizante.

En el capítulo 5, abordamos la implementación práctica de algunos de los sistemas de inferencia difusa propuestos en los capítulos anteriores. Se consideran las restricciones impuestas por las arquitecturas de routers, así como las restricciones operacionales de estos sistemas en equipos de altas prestaciones.

Por último, hacemos algunos comentarios finales y delineamos algunas líneas de investigación abiertas de cara al futuro en el capítulo 6. En este último capítulo, hacemos brevemente hincapié en las ventajas aportadas por los métodos desarrollados en esta tesis.

2. Ciencia de Internet

Del mismo modo que otros sistemas complejos objeto de estudio de las ciencias sociales y naturales o la ingeniería, el funcionamiento de Internet es difícil de comprender desde un punto de vista técnico. La estructura y el comportamiento de las redes de conmutación de paquetes es difícil de modelar de un modo similar a muchos sistemas naturales y artificiales. No obstante, Internet es a la vez un caso excepcional y un desafío debido a su rápido desarrollo, alta heterogeneidad y a la inherente ausencia de mecanismos de medida y monitorización en su diseño básico. De manera breve, las redes de conmutación de paquetes suponen un desafío para el modelado analítico.

El objetivo del capítulo 2 de esta tesis es introducir y proporcionar descripciones concisas de algunos de los bloques elementales de lo que algunos autores llaman Ciencia de Internet [12, 64], es decir, el estudio de leyes y patrones en la estructura de Internet. Se discuten asimismo algunos aspectos adicionales que serán clave en los siguientes capítulos.

Se definen brevemente los conceptos más relevantes sobre la medida y evaluación de prestaciones de redes que se usan a lo largo de los siguientes capítulos. Sin embargo, no entramos en detalles acerca de todos los conceptos relacionados con las redes de ordenadores relacionados con esta tesis. En el libro de Crovella y Krishnamurthy [27] se puede encontrar un claro y completo análisis en profundidad de los sistemas de medida y análisis de prestaciones de redes de ordenadores. Asimismo, varios artículos de investigación detallan con claridad algunos aspectos más específicos. Entre estos, destacamos el artículo en el que Broido y colaboradores [12] analizan algunos conceptos matemáticos clave para el análisis del tráfico de Internet.

Del mismo modo, queda fuera del ámbito de esta tesis realizar un análisis detallado de los aspectos matemáticos de la mayoría de los conceptos tratados, en particular aquellos relacionados con el control de tráfico. En este ámbito recomendamos los siguientes libros [104, 10]. Además, se puede consultar algunos de los artículos más relevantes e influyentes en este área [93, 92, 91, 113, 46].

En el primer apartado del capítulo 2 se introduce el problema del modelado general de Internet. La complejidad del modelado de Internet en su funcionamiento actual y en un futuro previsible se puede comprender si consideramos el continuo crecimiento exponencial del volumen de tráfico y el número de nodos observado durante los últimos años [43], así como la rápida evolución de los protocolos y aplicaciones de red. Actualmente, la captura de cabeceras de paquetes en enlaces de alta velocidad durante unos minutos u horas puede producir volúmenes de datos del orden de cientos de GBs o incluso varios TBs de datos.

El reciente desarrollo de hardware de altas prestaciones para captura de tráfico IP a velocidades de hasta 10 Gb/s [34] ha hecho posible registrar trazas de tráfico en enlaces troncales de las redes de altas prestaciones actuales. Sin embargo, no es posible utilizar tales volúmenes de información de manera directa para tareas de investigación y operación. En este contexto, se necesitan métodos de preprocesado y filtrado. A menudo, los volúmenes de datos han de ser reducidos en unos 12 órdenes de magnitud, desde 10^{12} bytes a un informe de apenas 10 líneas de texto [35]. Asimismo, es una práctica común reducir grandes volúmenes de medidas de tráfico a unas pocas tablas y gráficas [101].

En el primer apartado del capítulo 2 de esta tesis se comentan asimismo los tres aspectos básicos de modelado de internet: el tráfico, las topologías y los efectos de los protocolos sobre tráfico y topologías.

A continuación, se analizan los sistemas e infraestructuras de medida de Internet, considerándose los sistemas activos, pasivos y los repositorios públicos de medidas.

Los sistemas activos funcionan mediante la inyección de tráfico sonda desde nodos de la red, con objeto de medir parámetros tales como el tiempo de retorno de paquetes desde diversos puntos o el porcentaje de pérdida de paquetes en determinados caminos de extremo a extremo [71, 75, 86, 94, 70]. Estos sistemas introducen paquetes sonda en la red y analizan la respuesta recibida. Siguiendo un determinado modelo de la red, pueden llegar a estimar algunas características, tales como el retardo de propagación y algunas métricas relacionadas con el ancho de banda. Un ejemplo de estos sistemas es la herramienta thrulay [100], que mide el retardo de retorno de paquetes y la capacidad disponible entre dos nodos mediante conexiones TCP o flujos UDP.

Por otra parte, los sistemas pasivos se basan en el registro de datos en un nodo de la red, esto es, no se envían paquetes sonda. Si bien los sistemas pasivos no requieren la cooperación entre nodos finales, la relevancia y calidad de los datos obtenidos depende en gran medida de la localización del punto de medida. Por ello, las infraestructuras de medida basadas en sistemas pasivos requieren una estrecha cooperación entre diferentes operadores de red [75, 23].

A continuación, se analizan diversos modelos de tráfico, prestándose atención a las limitaciones de la teoría de teletráfico tradicional y a las alternativas propuestas hasta la fecha. En particular, se analizan algunos modelos de nivel de transporte para conexiones TCP y las métricas de prestaciones más comunes. También se tratan brevemente algunos modelos de tráfico desde la perspectiva de las capas de red y transporte así como su impacto sobre métricas genéricas de prestaciones percibidas por usuarios finales [75], desde el punto de vista de aplicaciones específicas, tales como el tráfico Web [83], de sistemas peer-to-peer [76] o multimedia [84]. Asimismo, se estudia el uso de simuladores de red de diferentes características.

Además, se introduce el problema de la congestión de redes de conmutación de paquetes y los mecanismos de control de tráfico que se pueden utilizar para prevenir y tratar adecuadamente situaciones de congestión. Estos mecanismos se pueden clasificar de manera general en dos categorías. Los primeros son aquellos que controlan el tráfico de extremo a extremo, esto es, entre dos nodos finales de la red que establecen conexiones de nivel de transporte entre sí. La segunda categoría incluye aquellos mecanismos implementados en nodos intermedios de la red o routers, que generalmente se basan en la gestión activa de colas de paquetes. Se analizan asimismo las arquitecturas de provisión de calidad de servicio más ampliamente aceptadas, definiéndose brevemente las clases de servicio estándar definidas dentro de la arquitectura DiffServ.

En el apartado 2.5 de esta tesis se analizan diversos modelos de series temporales para tráfico de redes. El tráfico de redes de conmutación de paquetes exhibe de manera general un comportamiento altamente no lineal. Se han descrito muchos tipos de comportamientos no lineales, incluyendo pseudoaleatoriedad, comportamiento regular predecible, comportamientos impredecibles, caos transitorio e intermitente, caos en banda estrecha y banda

ancha, así como diferentes modos de superposición de varios patrones básicos [58]. La gran mayoría de estos patrones de comportamiento son observables en series temporales de carga de tráfico en enlaces de redes de ordenadores, de lo cual damos abundantes evidencias en el capítulo 3 de esta tesis.

En muchos casos, el tráfico de redes muestra en determinados enlaces patrones que sugieren que, en cuanto a su comportamiento dinámico, se puede clasificar en la categoría de los fenómenos regulares y predecibles. En este sentido, el análisis del tráfico de red se puede abordar desde el punto de vista del análisis de series temporales. En estos casos, la teoría de sistemas dinámicos no lineales proporciona un marco adecuado para el análisis, identificación y predicción de series de tráfico de red.

La posibilidad de predecir la carga de tráfico en enlaces de redes tiene importantes implicaciones en numerosos campos de aplicación, tales como el control de congestión y admisión, o la realización de mecanismos adaptativos para aplicaciones y sistemas de gestión de redes. Iniciativas como el servicio meteorológico de Internet (<http://nws.cs.ucsb.edu>) están adquiriendo relevancia recientemente por su utilidad en la implementación de diversos servicios adaptativos. En particular, los sistemas de computación grid pueden mejorar significativamente sus prestaciones si disponen de medidas y predicciones de las condiciones de la red. Esto ha motivado el desarrollo de servicios de predicción de la tasa de transferencia de extremo a extremo mediante TCP, orientados a servir de base a la computación grid [68].

En el capítulo 2 de esta tesis se describen diversos modelos de series temporales aplicables al tráfico de redes. Se consideran modelos con memoria tanto a corto como a largo plazo, paramétricos y no paramétricos. En primer lugar, se revisan los modelos estocásticos previamente propuestos para modelar el tráfico de Internet. A continuación, se describen brevemente varias técnicas basadas en inteligencia computacional, incluyendo los operadores OWA inducidos por reglas de vecinos más próximos, o IOWA-NN [117], los basados en máquinas de soporte vectorial de mínimos cuadrados, o LS-SVM [105], y las máquinas de aprendizaje extremo truncadas de forma óptima (optimally pruned extreme learning machines), o OP-ELM [69]. Estos modelos se aplicarán en el capítulo 3. Asimismo, se revisan algunas de las métricas de error o prestaciones de modelos predictivos que se utilizarán en el capítulo 3.

3. Minería de datos sobre tráfico de red mediante inteligencia computacional

La comprensión de la dinámica y las prestaciones de las redes de conmutación de paquetes tomando medidas como base permite optimizar el uso de recursos. A medida que la investigación en sistemas de medida de redes avanza y, con ello, surgen nuevas herramientas e infraestructuras, la tarea de operación de redes se hace cada vez más compleja.

Los sistemas de medida de redes actuales están creciendo en sofisticación, produciendo enormes cantidades de datos y estadísticas con complejas interrelaciones. En las redes actuales, es muy común generar informes basados en la tecnología NetFlow [22] mediante sistemas de medida que pueden producir GBs de datos diariamente. Además, estas medidas son frecuentemente difíciles de interpretar por parte de los operadores.

Uno de los retos fundamentales planteados por la complejidad del tráfico de red es el desarrollo de modelos predictivos de la carga de enlaces bajo diferentes escalas temporales. Aquí, abordamos este problema desde la perspectiva del campo de la predicción de series temporales. En este campo, la precisión de las predicciones no es el único objetivo importante; una mejor comprensión de la dinámica del proceso analizado es antes al contrario un objetivo esencial de los métodos de predicción [112].

En particular, los FIS exhiben una capacidad de modelado predictivo y descripción de forma combinada como consecuencia de su estructura basada en reglas. La aplicación de los FIS a predicción de series temporales se remonta a las primeras aplicaciones de la lógica difusa. Sin embargo, a pesar de los buenos resultados en términos de precisión e interpretabilidad, los FIS apenas se han aplicado en el campo de la predicción de series temporales en comparación con otras técnicas de inteligencia computacional, como las redes neuronales o las máquinas de soporte vectorial. En este trabajo proponemos un marco metodológico para predicción autoregresiva a largo plazo de series temporales por medio de FIS [78, 79].

Otro área de investigación en torno a los sistemas de medida de red es el desarrollo de herramientas de visualización que faciliten la comprensión de la dinámica y patrones de tráfico por parte de investigadores y profesionales. Existen numerosas herramientas de visualización de medidas de redes (véase [26] para una amplia lista), en su mayoría basadas en la representación de propiedades estadísticas mediante gráficos y mapas. Así, las herramientas empleadas actualmente producen informes consistentes en decenas de gráficas, mapas y tablas. Por ello, no es fácil extraer resúmenes simples acerca del comportamiento de la red. En efecto, la complejidad de las herramientas de monitorización y generación de informes dificulta considerablemente su adopción en la actualidad.

A pesar de que existen algunas herramientas que generan resúmenes breves de estadísticas de red, tales como la utilidad `flow-report`, incluida en las herramientas conocidas como `flow-tools` [44], su legibilidad y concisión son muy limitadas. Asimismo, producen informes según plantillas previamente fijadas, por lo que no captan complejas relaciones normalmente ocultas, y requieren herramientas adicionales de visualización para comprender los informes.

En este trabajo, definimos resúmenes lingüísticos de registros de flujos de red basados en la teoría general de resúmenes lingüísticos [77]. La idea fundamental de los resúmenes

lingüísticos es utilizar términos del lenguaje natural para expresar información y conocimiento oculto en una colección de objetos generalmente de gran tamaño.

Así, en el tercer capítulo de esta tesis se describen dos métodos heterogéneos basados en técnicas difusas para realizar minería de datos sobre tráfico de redes. El primer método proporciona un mecanismo automático para modelar y predecir la carga de tráfico, y se puede clasificar por ello como un método de minería de datos predictiva. El segundo método se ha diseñado para generar resúmenes lingüísticos de medidas de tráfico de red así como para proporcionar un procedimiento de identificación de hechos ocultos en medidas de tráfico en forma de reglas de asociación lingüísticas. Por tanto, este método aborda un problema de minería de reglas de asociación.

Se muestra la aplicabilidad práctica de ambos métodos en funciones de operación de red e ingeniería de tráfico [3, 4], para las que pueden ayudar a abordar varios problemas actuales.

En el epígrafe 3.1 nos centramos en el modelado predictivo a largo plazo de series temporales no lineales de una variable. En primer lugar, se desarrolla un método de predicción a largo plazo de series temporales mediante sistemas de inferencia difusa combinados con técnicas de estimación de la varianza residual. El método se valida por medio de un conjunto de series de referencia. Aunque el objetivo fundamental de esta sección es el desarrollo de una metodología para construir sistemas de inferencia simples e interpretables, se muestra además que estos sistemas pueden proporcionar mejores prestaciones que algunas de las técnicas más precisas y comunes en el campo de la predicción de series temporales.

A continuación, en el apartado 3.2, aplicamos la metodología desarrollada a series de carga de tráfico de redes en diversos enlaces y subredes. Se realiza un extenso análisis de predecibilidad mediante la misma técnica de estimación de la varianza residual empleada dentro de la metodología de predicción mediante sistemas difusos. Tomando como base estas estimaciones de predecibilidad, se derivan sistemas de inferencia difusa que son tanto interpretables como precisos para un conjunto heterogéneo de series de tráfico de red.

En la sección 3.3, describimos un método para resumir y analizar registros de medidas de flujos de red. Como primer paso, se aplica la teoría de resúmenes lingüísticos difusos para analizar y extraer resúmenes concisos e interpretables a partir de colecciones NetFlow. A continuación, se desarrolla un método para extraer hechos ocultos en medidas de flujos de red en forma de reglas de asociación difusas. El método se aplica a un amplio y heterogéneo conjunto de flujos de red.

3.1. Modelos de series temporales basados en sistemas de inferencia difusa

Tanto la predicción de series temporales como su análisis en general son problemas recurrentes en todas las áreas de las ciencias naturales y sociales, así como en la ingeniería. En el campo de la predicción de series temporales, la precisión de las predicciones no es el único objetivo fundamental. En efecto, comprender el comportamiento de las series temporales y lograr visualizar u obtener una cierta percepción de su dinámica es una capacidad muy deseable para los métodos de predicción [112].

Algunas técnicas estadísticas convencionales, como los modelos AR y ARMA se han venido usando ampliamente durante décadas para predicción y pronóstico en diversos cam-

pos [11]. Sin embargo, estas técnicas tienen serias limitaciones para modelar series temporales derivadas de fenómenos complejos, por lo que más recientemente se han aplicado con éxito métodos más avanzados, no lineales, incluyendo las redes neuronales artificiales, en diversos problemas [18].

Las técnicas de modelado basadas en lógica difusa son atractivas en virtud de su potencial para abordar un amplio espectro de problemas así como su interpretabilidad. En particular, los sistemas de inferencia difusa exhiben una interesante combinación de capacidades descriptivas y predictivas como consecuencia de su estructura basada en reglas [110]. Las primeras aplicaciones de los sistemas de inferencia difusa al modelado predictivo de series temporales surgen con algunos de los primeros trabajos sobre identificación de sistemas mediante sistemas de inferencia difusa, y en particular en [111], en el que los autores desarrollan el conocido algoritmo de identificación a partir de ejemplos y lo aplican a la serie temporal de Mackey-Glass como caso de validación.

No obstante, a pesar de sus buenas prestaciones en términos de precisión e interpretabilidad, los sistemas de inferencia difusa no han sido aplicados ampliamente en el campo de la predicción de series temporales si los comparamos con otras técnicas de modelado no lineal, como las redes neuronales artificiales o las máquinas de soporte vectorial.

La metodología propuesta en esta tesis tiene como objeto de aplicación series temporales de valores reales. Esto es, se propone un marco metodológico para realizar predicción autoregresiva de series temporales numéricas reales por medio de sistemas de inferencia difusa [78]. Llamaremos a estos sistemas implementados mediante sistemas de inferencia difusa autoregresores difusos, por oposición a lo que en la literatura se denomina habitualmente regresión difusa [17] y series temporales difusas [103].

Al desarrollar sistemas de inferencia difusa para predicción de series temporales, se plantean numerosas cuestiones: ¿cuántas y qué entradas al sistema se deben definir? ¿En qué grado se puede lograr la teórica capacidad de aproximación universal de ciertos sistemas de inferencia difusa mediante las técnicas existentes en la actualidad? ¿Cuáles son las mejores técnicas difusas para estas tareas? ¿Es posible realizar predicciones a largo plazo?

En la práctica, la construcción de un modelo difuso para series temporales requiere abordar dos problemas frecuentes en el campo de la minería de datos. En primer lugar, se ha de elegir el conjunto de variables o entradas del sistema de inferencia difusa, así como la estructura del sistema, que en este caso viene dada por las etiquetas lingüísticas y la base de reglas. Una vez que se han dado estos pasos, el modelo difuso se puede ajustar mediante técnicas de aprendizaje supervisado. En esta tesis se propone un marco metodológico automático para abordar ambos problemas utilizando técnicas difusas junto con técnicas de estimación de la varianza residual.

El primer problema se puede abordar mediante técnicas de selección a priori de características basadas en estimadores de la varianza residual, los cuales proporcionan una estimación del error de predicción del modelo no lineal más preciso que se puede construir sin que se produzca sobreajuste. El segundo problema se aborda mediante técnicas de identificación de sistemas difusos a partir de ejemplos numéricos [48], tales como el algoritmo de Wang y Mendel (W&M) [110, 111], así como algoritmos de identificación de sistemas mediante técnicas de agrupamiento [72, 20].

En el apartado 3.1 de esta tesis se trata asimismo un nuevo reto afrontado solo de manera relativamente reciente en el campo de la predicción de series temporales: la predicción a largo plazo, como generalización de los modelos de predicción a corto plazo. En este caso, la falta de información y acumulación de errores plantean dificultades adicionales. Asimismo, se analizan tanto series sintéticas, de comportamiento caótico pero sin ruido, como series temporales de test extraídas de aplicaciones del mundo real. Con objeto de evaluar las prestaciones de una versión concreta del marco metodológico propuesto, comparamos sus resultados frente a los obtenidos mediante modelos basados en máquinas de soporte vectorial de mínimos cuadrados, o least squares support vector machines, LS-SVM [105], un método altamente preciso y de frecuente uso en el campo de la predicción de series temporales.

Tras describir brevemente la técnica de estimación de la varianza residual empleada, Delta Test [55], así como la forma analítica general de los sistemas de inferencia difusa que se van a desarrollar, se describe el marco metodológico propuesto. Este marco se puede resumir en dos etapas, como se muestra en la figura 1

Dentro de este marco, se propone una variante concreta basada en el algoritmo de identificación de sistemas de inferencia difusa de Wang y Mendel [110] así como el método de optimización de segundo orden de Levenberg-Marquardt [38, 6]. Esta variante se estudia detalladamente y se valida mediante la serie temporal utilizada para la competición del primer simposio europeo sobre predicción de series temporales [36], que describe la evolución del síndrome climático conocido como Oscilación del Sur El Niño.

A continuación, la metodología se evalúa para un conjunto de series temporales de test, comparándose los resultados frente a los obtenidos con modelos LS-SVM. Los resultados de esta evaluación comparativa se resumen en la tabla 1, que muestra los errores de entrenamiento y test promediados para los modelos LS-SVM y los basados en inferencia difusa. De la tabla se puede concluir que los modelos difusos proporcionan mejores prestaciones que los modelos LS-SVM en términos de precisión en el caso de las series temporales con ruido, para las cuales no parece factible encontrar modelos deterministas satisfactorios. Una capacidad notable de los modelos difusos desarrollados es su capacidad de generalización. La diferencia entre los errores de entrenamiento y de test está típicamente por debajo de un rango entre un 20 % y un 30 %, excepto en el caso de la serie Sunspots.

En cuanto a recursos computacionales, en el capítulo 3 de esta tesis se dan amplias evidencias de que la metodología propuesta para construir modelos predictivos basados en lógica difusa viene caracterizada por un coste computacional en términos de tiempo de computación entre uno y dos órdenes de magnitud menor que el requerido por los modelos LS-SVM.

3.2. Predicción de la carga de tráfico en enlaces de red

Modelar y predecir la carga de tráfico en enlaces de redes de conmutación de paquetes es una tarea especialmente difícil. En general, en el tráfico de Internet se observan la mayoría de los comportamientos dinámicos complejos conocidos, si no todos. Por ello, los métodos lineales tradicionales no parecen ser suficientemente flexibles como modelos predictivos en este ámbito. En su lugar, es necesario explorar el uso de modelos no lineales y no

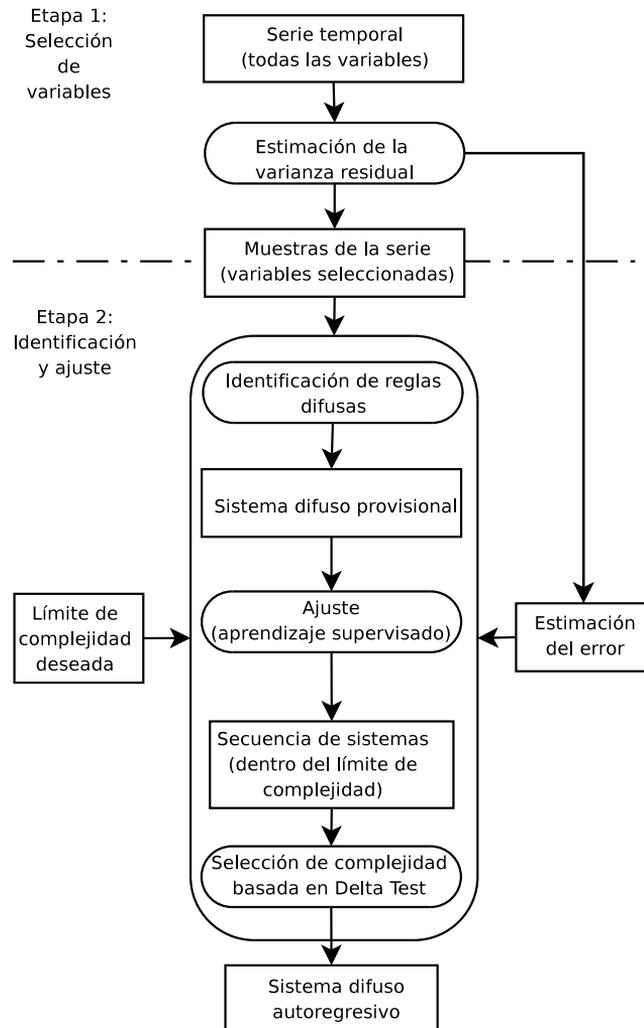


Figura 1: Marco metodológico para predicción de series temporales mediante sistemas de inferencia difusa.

Serie	LS-SVM		Inferencia difusa	
	Entrenamiento	Test	Entrenamiento	Test
ENSO (10)	$8,055 \cdot 10^{-3}$	$3,192 \cdot 10^{-2}$	$1,943 \cdot 10^{-2}$	$2,043 \cdot 10^{-2}$
PolElec (7)	$1,158 \cdot 10^{-2}$	$3,566 \cdot 10^{-2}$	$1,696 \cdot 10^{-2}$	$1,779 \cdot 10^{-2}$
PolElec (14)	$1,037 \cdot 10^{-2}$	$3,241 \cdot 10^{-2}$	$1,582 \cdot 10^{-2}$	$1,816 \cdot 10^{-2}$
Sunspots (9)	$1,338 \cdot 10^{-2}$	$3,284 \cdot 10^{-2}$	$1,691 \cdot 10^{-2}$	$2,623 \cdot 10^{-2}$
Sunspots (12)	$9,637 \cdot 10^{-3}$	$3,024 \cdot 10^{-2}$	$1,590 \cdot 10^{-2}$	$2,546 \cdot 10^{-2}$
AbileneI (7)	$8,587 \cdot 10^{-3}$	$2,476 \cdot 10^{-2}$	$1,448 \cdot 10^{-2}$	$1,732 \cdot 10^{-2}$
AbileneI (12)	$6,771 \cdot 10^{-3}$	$2,153 \cdot 10^{-2}$	$1,228 \cdot 10^{-2}$	$1,506 \cdot 10^{-2}$
SFL (10)	$1,481 \cdot 10^{-3}$	$6,578 \cdot 10^{-3}$	$1,020 \cdot 10^{-2}$	$1,285 \cdot 10^{-2}$
SFL (16)	$5,275 \cdot 10^{-4}$	$5,290 \cdot 10^{-3}$	$8,791 \cdot 10^{-3}$	$1,202 \cdot 10^{-2}$
MG (9)	$7,881 \cdot 10^{-4}$	$3,658 \cdot 10^{-3}$	$1,385 \cdot 10^{-2}$	$1,775 \cdot 10^{-2}$

Tabla 1: Errores de entrenamiento y test de los modelos difusos y LS-SVM promediados para los horizontes de predicción 1 a 50. Todos los errores se muestran como error cuadrático medio (MSE) normalizado frente a la varianza de la serie. Entre paréntesis se especifica el tamaño máximo de los regresores utilizados para cada caso.

paramétricos en general. En este contexto, se cuenta con pocos antecedentes en la literatura.

En el apartado 3.2 de esta tesis aplicamos al modelado de tráfico la metodología desarrollada para construir modelos predictivos de series temporales mediante sistemas de inferencia difusa combinados con un método de estimación de la varianza residual, método que llamaremos NRVE-FIS. Se analizan series temporales para el volumen de tráfico medido en ciertos puntos de determinadas redes durante un periodo de tiempo dado.

Así, se realiza un análisis en profundidad de un conjunto diverso de series temporales de tráfico. En este estudio se comparan los modelos NRVE-FIS, OP-ELM, IOWA-NN y ARIMA, o modelos autoregresivos integrados de media móvil. Cada una de las trazas de tráfico se analizan a varias escalas temporales, según la disponibilidad de medidas. Para cada serie se estudia su predecibilidad, los errores de modelado obtenidos con los cuatro modelos utilizados, la forma de las predicciones para algunos casos de ejemplo, así como el comportamiento de los errores tanto a corto como a largo plazo. La predecibilidad se estudia mediante el estimador Delta Test normalizado frente a la varianza de cada serie, o NDT.

Como ejemplo de aplicación, mostramos en este resumen los resultados obtenidos para la series correspondientes a la traza Abilene-I, que describen la cantidad total de tráfico agregado entrante en los routers de la red Abilene, la red troncal de Internet2 durante varios años. Los datos están disponibles en el Observatorio de Abilene e Internet2 [53]. Esta traza se analiza a las siguientes escalas temporales: un día, 12 horas, dos horas, 30 minutos y cinco minutos. La serie de medidas diarias cubre desde el 4 de enero de 2003 hasta el 31 de diciembre de 2006. Las series correspondientes a escalas más pequeñas cubren desde el 4 de agosto de 2005 hasta el 31 de diciembre de 2006.

En la figura 2, se muestran las series de tráfico correspondientes a la traza Abilene-I para cinco escalas de tiempo distintas. La figura 3 muestra la estimación de predecibilidad

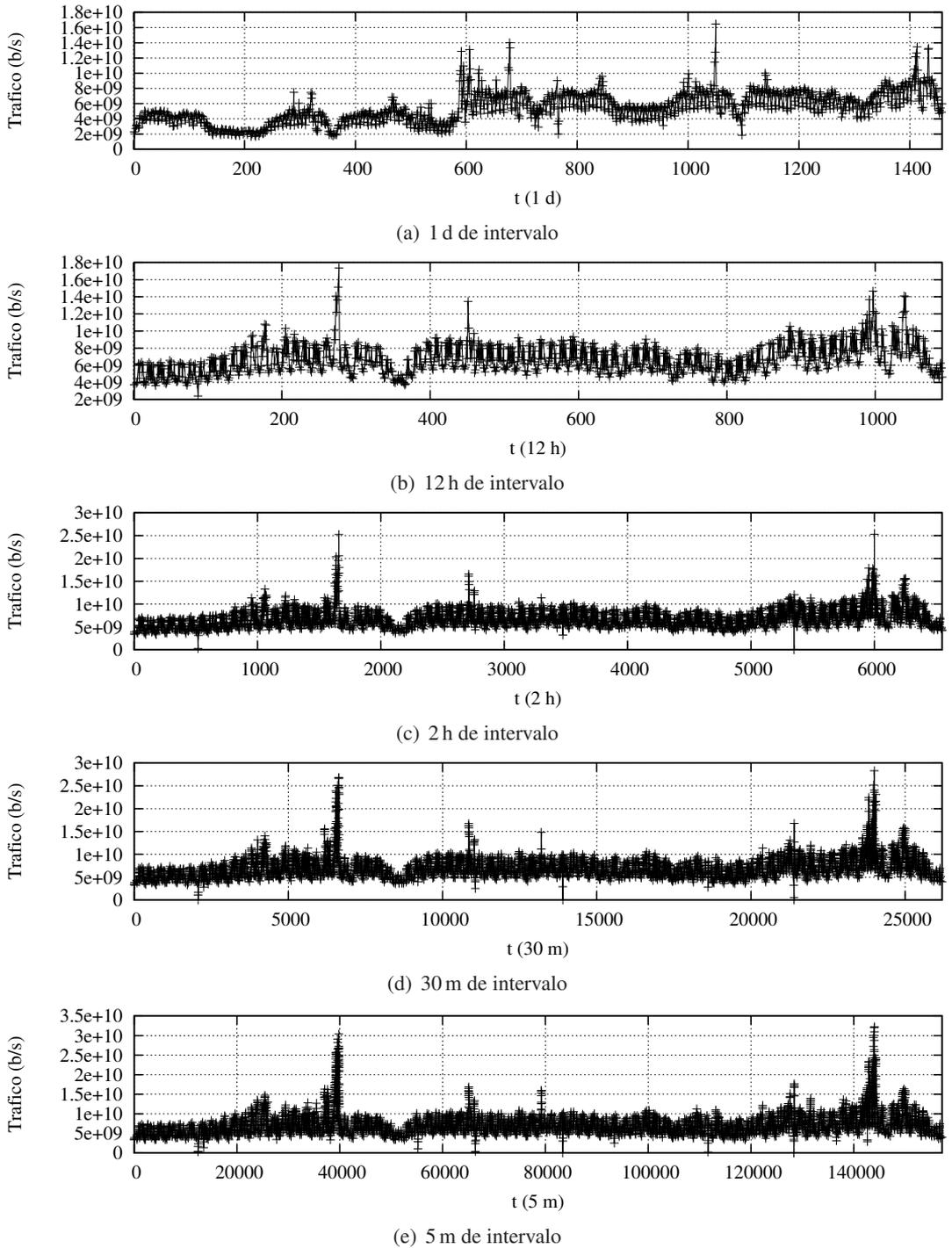


Figura 2: Abilene-I: carga de tráfico a diferentes escalas temporales.

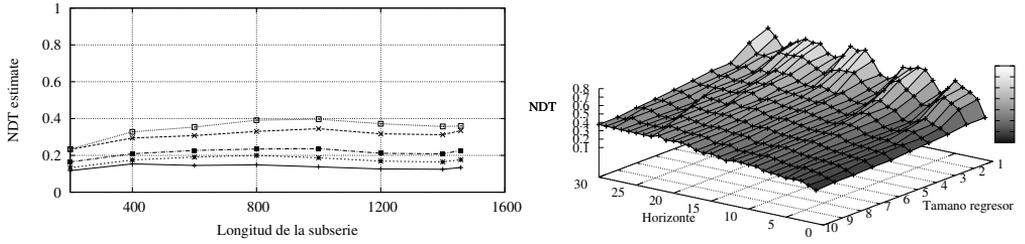
basada en Delta Test para las series derivadas de la traza Abilene-I a diferentes escalas. La columna de la izquierda muestra para cada escala la estimación NDT promediada para 100 subseries elegidas aleatoriamente, para los horizontes de predicción 1, 2, 5, 10 y 15, y utilizando regresores de tamaño máximo 10. La columna de la derecha muestra para cada escala temporal la estimación NDT para los horizontes entre 1 y 30, y regresores de tamaño máximo entre 1 y 10, en el caso de una de las subseries de longitud hasta 1000, elegida aleatoriamente.

En la figura 4, se representan los errores de entrenamiento para la traza Abilene-I. Se aplican cuatro modelos. En la columna de la izquierda se muestran los resultados para los sistemas difusos construidos con el método NRVE-FIS, y los modelos OP-ELM. En la columna de la derecha se muestran los resultados para los modelos IOWA-NN y ARIMA. El rango de horizontes de predicción comprende desde 1 hasta 30.

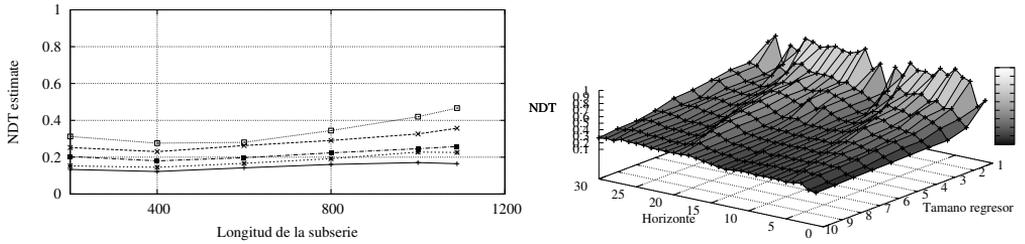
La figura 5 muestra un ejemplo de predicción realizada con los modelos difusos para cada una de las escalas temporales consideradas. Se utiliza una serie de entrenamiento de hasta 1000 valores, escogida aleatoriamente, y una serie de test de hasta 1000 valores. En la columna de la izquierda, se muestra el error absoluto porcentual simétrico (SAPE) para la serie de test. En la columna de la derecha se muestran las predicciones de los 30 valores posteriores a la serie de entrenamiento.

Los resultados del extenso análisis realizado se discuten en detalle en el capítulo 3 de esta tesis. En primer lugar, se consideran las siguientes conclusiones generales:

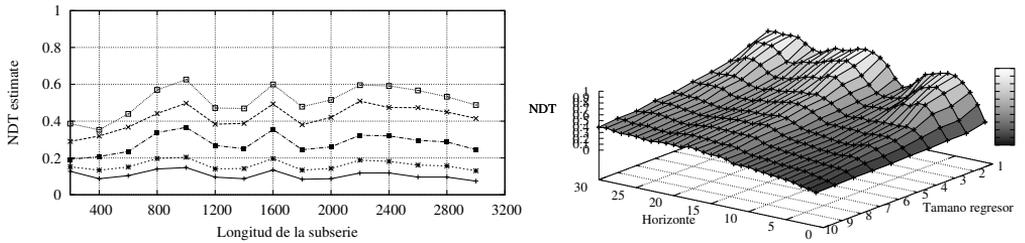
- La representación gráfica de las series temporales analizadas a diferentes escalas temporales (primera figura) confirma ampliamente el hecho frecuentemente comprobado de que la carga de tráfico es autosimilar para un amplio rango de escalas de tiempo. No obstante, se observan notables excepciones. Asimismo, el grado de autosimilitud es diverso, lo que encaja con resultados experimentales que muestran que el coeficiente de Hurst no tiene normalmente un valor alto para medidas de tráfico actuales [59]. Por ejemplo, el grado de autosimilitud es marcadamente más alto en el caso de las trazas Bellcore, correspondientes a tráfico Ethernet con un nivel de agregación bajo.
- El análisis de predecibilidad realizado mediante un estimador no paramétrico (segunda figura) nos lleva a extraer las siguientes conclusiones:
 - A partir de las gráficas bidimensionales se puede concluir que la estimación es estable para series de longitudes en el rango comprendido entre 200 y 3000 muestras, aunque se puede observar una ligera tendencia ascendente con el número de muestras en general. Esta tendencia es una consecuencia lógica de la no estacionariedad de las series analizadas. Existen sin embargo algunas excepciones para algunas series a ciertas escalas temporales, para las cuales la estimación puede oscilar hasta alrededor de un 100 %. No obstante, esto ocurre para series con un elevado grado de autosimilitud, típico en trazas con bajo grado de agregación, como las series Bellcore y DEC. Los valores de predecibilidad para subseries de longitud entre 1000 y 2000 son en la mayoría de los casos considerablemente estables, variando dentro de un rango del 10 % aproximadamente.



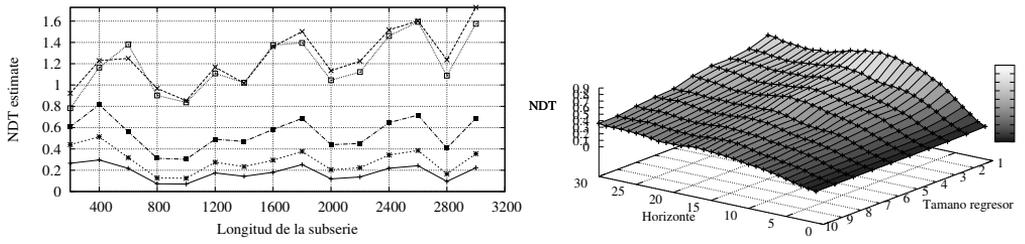
(a) 1 d de intervalo



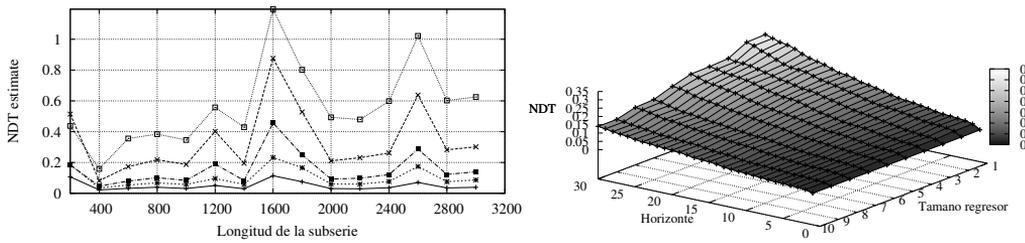
(b) 12 h de intervalo



(c) 2 h de intervalo

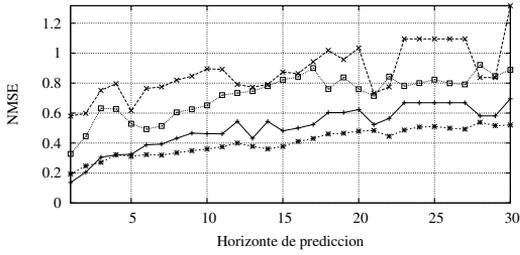


(d) 30 m de intervalo

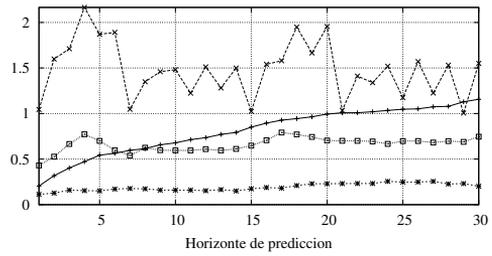
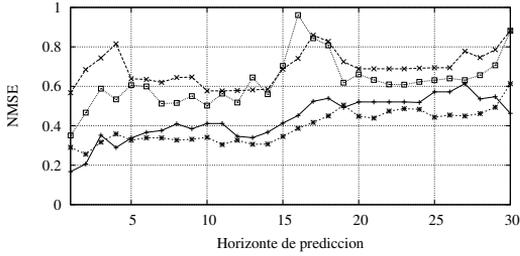
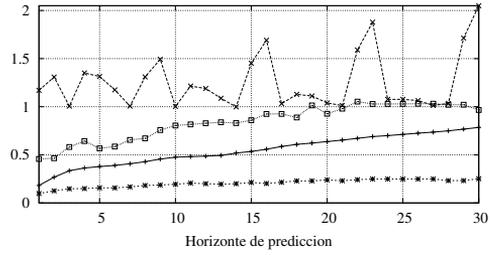


(e) 5 m de intervalo

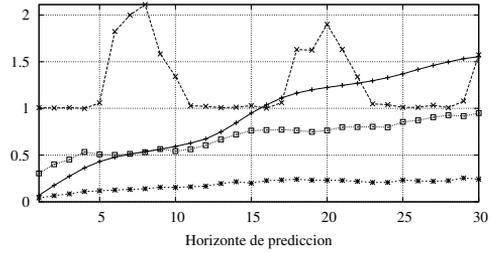
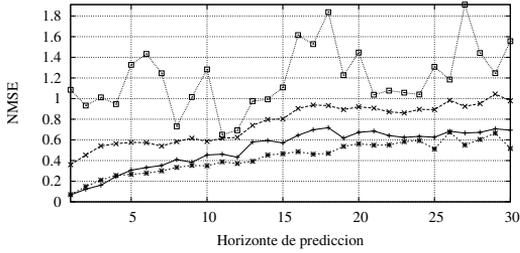
Figura 3: Abilene-I: estimación de predecibilidad para diferentes escalas temporales. Izquierda: NDT para los horizontes 1, 2, 5, 10, 15, para subseries de diferentes longitudes (con inicio aleatorio, promediado para 100 repeticiones). Derecha: NDT para una subserie de 1000 valores elegida aleatoriamente, horizontes 1-30 y tamaños de regresor 1-10.



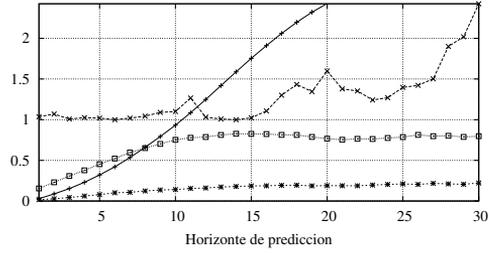
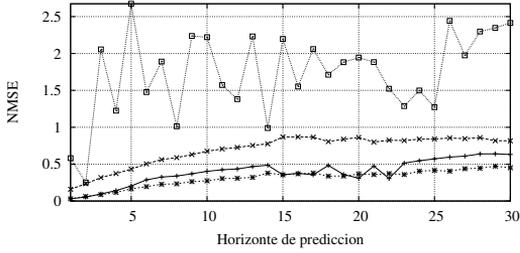
(a) 1 d de intervalo



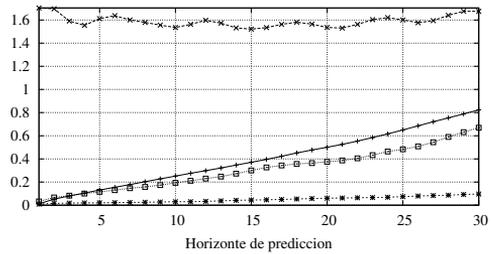
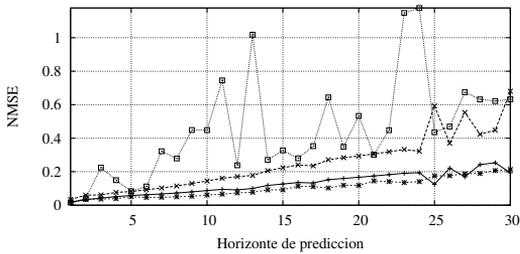
(b) 12 h de intervalo



(c) 2 h de intervalo



(d) 30 m de intervalo



(e) 5 m de intervalo

Figura 4: Abilene-I: errores de test. Izquierda: errores de los modelos NRVE-FIS (+, entrenamiento; x, test) y OP-ELM (*, entrenamiento; □, test). Derecha: errores de los modelos ARIMA (+, entrenamiento; x, test) e IOWA-NN (*, entrenamiento; □, test).

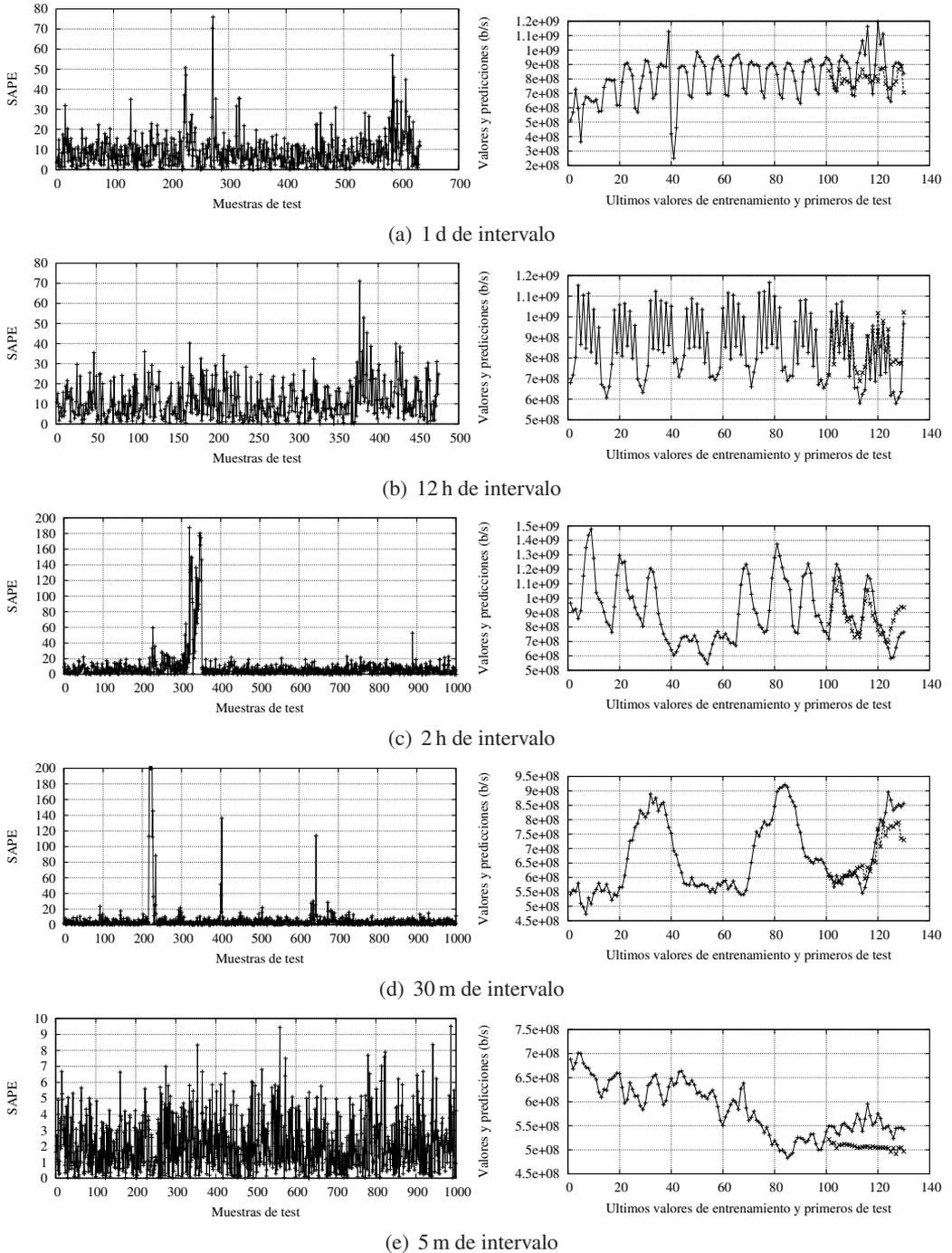


Figura 5: Abilene-I: predicciones usando NRVE-FIS para una serie de test elegida aleatoriamente. Izquierda: SAPE para predicciones del siguiente valor sobre la serie de test. Derecha: predicciones de los 30 valores posteriores a la serie de entrenamiento. Se muestran los últimos 100 valores de entrenamiento, los primeros 30 valores de la serie de test (ambos como +, línea continua) y las 30 predicciones correspondientes (x, línea discontinua).

Asimismo, la dependencia de la estimación de la varianza residual respecto a la longitud de las subseries es mayoritariamente invariante con el horizonte de predicción. Esto es, los incrementos y decrementos de la estimación que se observan para algunas longitudes de subseries ocurren de manera consistente para todos los horizontes de predicción. Por ejemplo, las estimaciones de predecibilidad para la serie AMPATH-OC12-200701-0 son claramente irregulares respecto a la longitud de la subserie seleccionada para las escalas de tiempo de 10 s, 1 s, 100 ms y 10 ms. Sin embargo, el patrón de variación es notablemente el mismo para todos los horizontes de predicción mostrados en las gráficas.

Esta primera parte de las gráficas de predecibilidad tiene una implicación importante: se pueden utilizar subseries relativamente cortas, de 1000 o incluso entre 400 y 1000 muestras en algunos casos, para calcular estimaciones aproximadas de la varianza residual. Se puede concluir que es posible construir modelos utilizando series de entrenamiento constituidas por 1000 muestras o incluso menos.

- La segunda gráfica de las figuras de predecibilidad muestra que la dependencia de la varianza residual respecto al tamaño máximo de los regresores así como el horizonte de predicción varía significativamente en función de la serie temporal analizada. En general, se confirma que el menor error de entrenamiento posible decrece a medida que el tamaño máximo del regresor crece. Respecto a la dependencia de la varianza residual frente al horizonte de predicción, la estimación es estrictamente creciente con el horizonte de predicción en algunos casos. Sin embargo, se puede observar un conjunto de series y escalas temporales para las cuales la estimación de varianza residual oscila, lo que indica un cierto grado de comportamiento cíclico. En estos casos, se puede esperar que la predecibilidad de la serie para algunos horizontes a medio o largo plazo sea mayor que para horizontes a más corto plazo.
- En la tercera y cuarta figura mostrada para cada traza de tráfico se comparan las prestaciones de los métodos de modelado usados. De entre ellos, es claro que los sistemas de inferencia difusa y los modelos OP-ELM son las dos mejores opciones en términos de precisión. En general, los modelos del primer tipo son ligeramente más precisos, aunque pueden llegar a tener una precisión notablemente mayor para predicción a largo plazo. Se pueden observar sin embargo algunas excepciones en las que los modelos OP-ELM son más precisos que los modelos difusos a corto plazo, con precisión ligera pero consistentemente mejor para un rango de horizontes de predicción. Por tanto, se puede concluir que los modelos difusos son normalmente, aun con algunas excepciones, ligeramente más precisos para predicción a corto plazo y claramente más precisos para predicción a largo plazo en general. No obstante, los dos métodos proporcionan resultados suficientemente similares como para que quepa esperar de ambos aproximadamente la misma capacidad de predecir la carga de tráfico en enlaces de red. Esto se puede observar en la columna izquierda de la figura 3.2. Nótese sin embargo que las gráficas de esta figura muestran los resultados globales para todas las trazas analizadas a todas las escalas temporales consideradas, y, por ello, se han de interpretar con cautela.

Los modelos de vecinos más próximos con IOWA así como los modelos ARIMA, como ejemplos de enfoques más tradicionales, son de forma consistente menos precisos que los modelos difusos y OP-ELM. Aunque los modelos ARIMA son de manera global menos precisos que los modelos IOWA-NN, como contrapartida son frecuentemente más robustos para predicciones a largo plazo. Esto se puede observar en la columna derecha de la figura 3.2. A lo largo de las gráficas mostradas para cada traza, se pueden observar asimismo varias excepciones en las que los modelos ARIMA proporcionan predicciones a corto plazo más precisas, como es el caso de las series WIDE-F-DITL-200701, especialmente a escala de 10 s.

Los resultados muestran que las técnicas basadas en inteligencia computacional son capaces de aprovechar la predecibilidad de las series analizadas. Esto es, son capaces de capturar la información sobre la dinámica del tráfico contenida en los conjuntos de datos de entrenamiento. Por ejemplo, si consideramos la serie correspondiente a la traza OC48-20030424-1 a escala de 100 ms, para un tamaño de regresor máximo de 10, la estimación de varianza residual crece con el horizonte de predicción hasta que éste llega aproximadamente a 10 y a 20, pero decrece para horizontes a más largo plazo. Este hecho revela cierto tipo de comportamiento periódico que hace que la serie sea más predecible a medio que a corto plazo.

Para las series temporales de carga de tráfico analizadas a diferentes escalas de tiempo, con una diferencia de hasta cinco órdenes de magnitud, la estimación mediante Delta Test de la varianza residual, o varianza del ruido, es notablemente estable para todas las longitudes de subseries consideradas. Este es el caso incluso para escalas temporales del orden del milisegundo, lo que sugiere que la predecibilidad de la carga de tráfico es invariante durante periodos de tiempo considerablemente prolongados. Por ejemplo, la varianza residual para la predicción del horizonte 1 oscila entre 0.45 y 0.55 para la traza Equinix-Chicago-DITL-2008 a intervalos de 1 ms. Teniendo en cuenta que esta traza comprende más de una hora, es decir, más de $3,6 \cdot 10^6$ intervalos de 1 ms, se puede concluir que la predecibilidad a escalas del orden del milisegundo es notablemente constante durante periodos cuatro órdenes de magnitud mayores.

A continuación listamos algunos casos de predicciones a corto plazo notablemente precisas:

- OC48-20020814-0, a escalas en el rango entre 10 s y 10 ms, así como a escala de 1 ms en menor medida.
- OC48-20020814-1, a todas las escalas temporales con algunas pocas excepciones para la escala de 1 ms.
- OC48-20030115-1 y OC48-20030115-0, especialmente a escalas en el rango comprendido entre 10 s, y 10 ms, así como a escala de 1 ms de forma más limitada.
- OC48-20030424-0 y OC48-20030424-1, especialmente a escalas comprendidas entre 10 s y 10 ms, así como a escala de 1 ms en menor medida. Es interesante observar que la carga de tráfico en la dirección 1 se puede predecir claramente mejor.

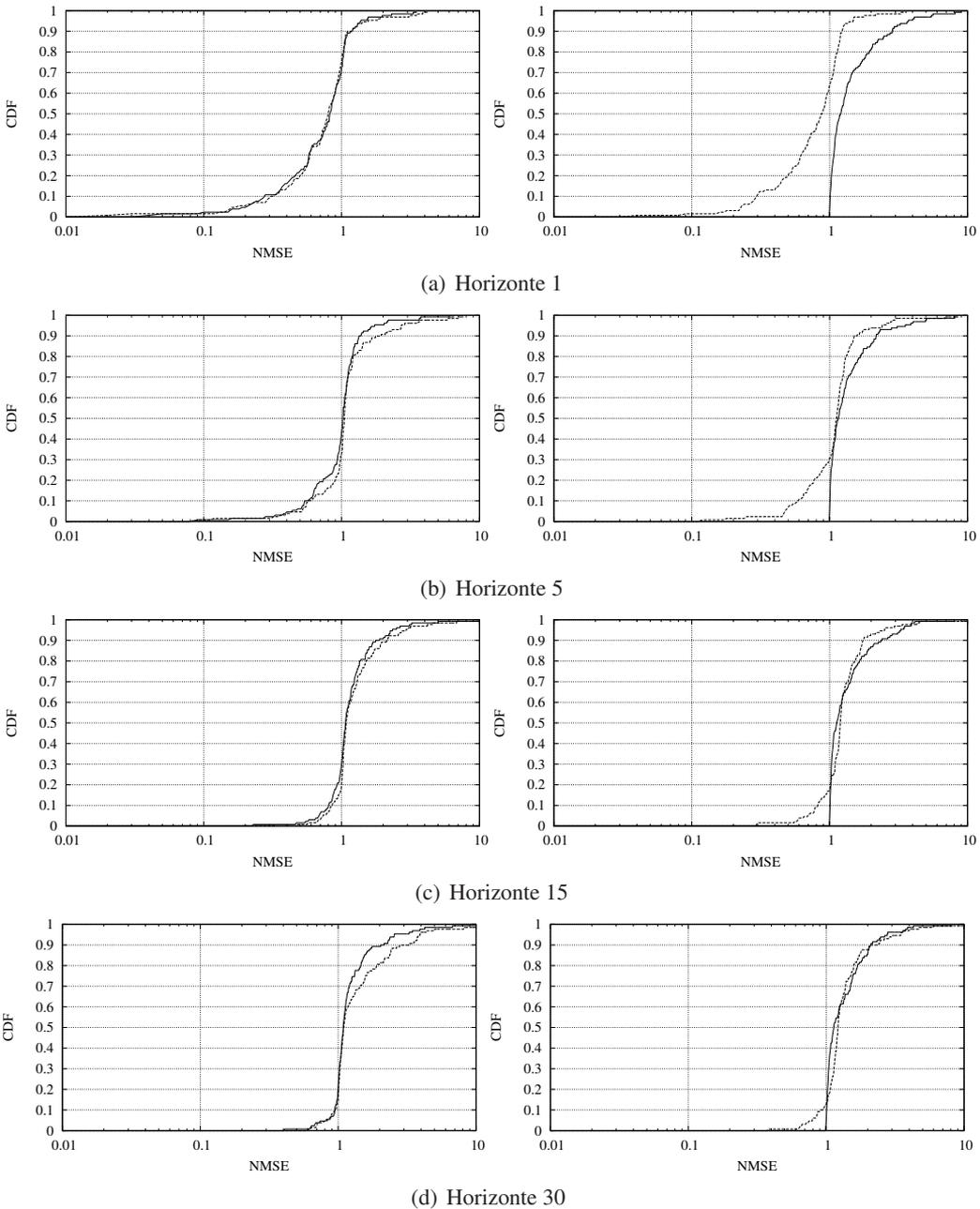


Figura 6: Función de distribución acumulada (CDF) de los errores de test normalizados (NMSE) de los modelos NRVE-FIS, OP-ELM, ARIMA y IOWA-NN para las predicciones de los horizontes 1, 5, 15 y 30, para el conjunto completo de trazas y todas las escalas temporales analizadas. Nótese la escala logarítmica-normal. Izquierda: NRVE-FIS (línea continua) y OP-ELM (línea a trazos). Derecha: ARIMA (línea continua) y IOWA-NN (línea a trazos).

En general, para las trazas CAIDA OC48 se puede observar que la predecibilidad a escalas del orden del segundo y superiores es notablemente mejor. Para escalas más pequeñas, la estimación de la varianza residual crece a medida que la escala decrece en aproximadamente 0.2 unidades normalizadas por cada orden de magnitud.

Sin embargo, en todos los casos mencionados, no existe la posibilidad de realizar predicciones a largo plazo precisas con los métodos empleados en esta tesis. Esto contrasta con lo observado para otras series, que exhiben claros patrones cíclicos, los cuales se traducen en una mayor predecibilidad, tanto en la estimación realizada mediante Delta Test como en los resultados finales obtenidos en la práctica. En particular, algunas subseries de la traza Equinix-Chicago-DITL-2008 a escalas de 100 ms y 10 ms muestran claros patrones cíclicos no lineales que se pueden predecir a largo plazo usando modelos NRVE-FIS.

Mediante un simple análisis de varianza (véase el capítulo 2 de esta tesis), el modelo NRVE-FIS para una subserie de la traza Equinix-Chicago-DITL-2008 a escala de 100 ms, explica el 84 % de la variabilidad total de la subserie, a pesar de la clara no linealidad. Debemos mencionar a este respecto que en general esta predecibilidad es ampliamente consecuencia de un determinismo no lineal, en lugar de correlaciones lineales, como se puede comprobar mediante un simple análisis de correlación de la series [58].

En algunos casos, los errores de los modelos ARIMA son considerablemente reducidos para predicciones a largo plazo si se comparan con los errores de los modelos basados en inteligencia computacional, que pueden llegar a ser del orden de varias unidades normalizadas. Este es el caso por ejemplo de la traza OC48-20020814-1 a escala de 10 s. No obstante, debemos hacer notar que en tales casos las series son en general predecibles en cierto grado sólo a muy corto plazo. Para predicciones a corto plazo, u horizontes menores o iguales a 5, tanto los modelos difusos como los OP-ELM aprovechan en mayor medida la predecibilidad de las series. Los modelos ARIMA muestran sin embargo mayor robustez en algunos casos para predicción a largo plazo. En cualquier caso, estas predicciones a largo plazo son muy poco precisas y por ello no parecen útiles a priori. Es interesante observar que, para aquellas series que son esencialmente impredecibles a largo plazo, los modelos ARIMA son más robustos frente a grandes cambios debidos a comportamientos no estacionarios. En tales casos, las técnicas basadas en inteligencia computacional parecen predecir errónea y parcialmente cambios importantes identificados en la serie de entrenamiento que no se repiten en la serie de test, lo que conlleva errores de predicción muy elevados. Este hecho es especialmente claro en el caso de la series OC48-20020814-1 a escala de 10 s.

La métrica de error utilizada para dirigir el proceso de aprendizaje de todos los métodos aplicados ha sido el error cuadrático medio, MSE por sus siglas en inglés. Se ha de tener en cuenta a este respecto que existen numerosas métricas alternativas definidas para distintos fines, y que las relaciones entre diversas métricas no son claras. Por ejemplo, un MSE normalizado de alrededor de 1, o entre 0.9 y 1.1, puede corresponder a un SMAPE que varía considerablemente, de forma aproximada entre 40 y 150 %. Asimismo, un NMSE de alrededor de 1 corresponde a un MxAE siempre por encima de 0.5, mayoritariamente entre 0.6 y 0.7. Es bien sabido que el modelado puede proporcionar resultados muy diferentes para distintas métricas de error. En esta tesis hemos limitado nuestro análisis a métodos de aprendizaje dirigidos por el error cuadrático medio. Por ello, un área importante de trabajo futuro es la identificación de métricas de error que puedan ser más adecuadas a la predicción

de variaciones extremas, entre otros propósitos específicos.

Es importante hacer notar que en algunos casos se puede observar un cierto grado de simetría en la predecibilidad del tráfico que atraviesa algunos enlaces en direcciones opuestas. En particular, las trazas Abilene-I y Abilene-O exhiben un alto grado de simetría para todas las escalas temporales analizadas. Además, las trazas CAIDA-OC48-20030424 y CAIDA-OC48-20030115 muestran una clara simetría a escalas de 10 ms y 1 ms, mientras que existen diferencias significativas de predecibilidad y prestaciones de los modelos predictivos para escalas de 100 ms y superiores, siendo el tráfico en la dirección 1 más predecible. Este hecho indica un cierto grado de persistencia de las características de la carga de tráfico en cuanto a predecibilidad en un mismo enlace. Comparando los resultados para las últimas dos trazas mencionadas en ambas direcciones, se puede observar asimismo que el hecho de que la predecibilidad sea mayor a una determinada escala temporal no implica necesariamente que lo mismo se cumpla para escalas temporales diferentes.

En el análisis que se realiza en el capítulo 3 de esta tesis, todas las series se han estudiado como procesos de una única variable, esto es, no se han considerado variables exógenas. Como primer paso para extender el análisis realizado en esta tesis, se podrían explorar las correlaciones existentes entre diferentes direcciones de un mismo enlace. Este análisis se podría ampliar muy posiblemente a un análisis más general de las correlaciones entre las series de carga de tráfico en enlaces relacionados topológicamente, lo que podría permitir hallar relaciones entre patrones de tráfico y topologías de redes, una de las áreas en que apenas se cuenta con resultados sólidos, como se discute en el capítulo 2. Esto conllevaría el análisis de trazas sincronizadas tomadas simultáneamente en diferentes puntos de determinadas redes, tales como los agregados de tráfico en la red Abilene, las dos direcciones de las trazas CAIDA OC48 y las trazas del archivo del proyecto NLANR PMA Special Traces Archive, incluyendo las trazas Abilene-I, Auckland-VI y Leipzig-II.

En este mismo área, otra línea de investigación futura consiste en el uso de variables exógenas como variables explicativas en general. En particular, para las trazas de tráfico disponibles para escalas de días y superiores, se puede esperar una mejora significativa de prestaciones si se consideran como entradas adicionales características tales como periodos de vacaciones y otros eventos predecibles. Estos y otros factores pueden tener un impacto significativo que se ha de analizar en cada caso particular, y que puede ser particularmente relevante en el contexto de los sistemas de inferencia difusa. En estos sistemas, muchas de tales características están descritas frecuentemente en forma de conocimiento lingüístico que se puede integrar en bases de reglas.

El objetivo fundamental de la metodología para predicción de series temporales desarrollada en el apartado 3.1 de esta tesis es proporcionar un método de identificación automática de modelos difusos que sean interpretables de forma lingüística. Sin embargo, se ha mostrado que, utilizando las técnicas adecuadas, estos sistemas pueden proporcionar prestaciones equivalentes o superiores a las de otras técnicas aplicadas ampliamente en el campo de la predicción de series temporales. No pretendemos con ello concluir que la metodología NRVE-FIS es en general más precisa que los otros métodos aplicados. Hemos de mencionar que la comparación realizada en esta tesis se puede ver como un estudio que muestra que para las series temporales analizadas, con las configuraciones concretas utilizadas, los sistemas de inferencia difusa son comparables o incluso superiores a los modelos LS-SVM

y OP-ELM en términos de capacidad de generalización.

No obstante, una de las ventajas esenciales de los sistemas de inferencia difusa, y la que motiva el desarrollo expuesto en la sección 3.1 de esta tesis es su interpretabilidad lingüística. Los sistemas de inferencia difusa son inherentemente comprensibles, especialmente cuando las reglas son definidas por expertos humanos. Cuando las reglas se identifican de forma automática a partir de datos y se aplican métodos de optimización, la interpretabilidad no se puede garantizar en general [16].

La metodología para predicción de series temporales mediante sistemas de inferencia difusa desarrollada en esta tesis se ha diseñado con objeto de abordar este problema. En primer lugar, la etapa de selección de entradas reduce el número de entradas significativamente para un tamaño máximo de regresor dado. En segundo lugar, la consistencia de la base de reglas está garantizada por el método de identificación empleado. En tercer lugar, como resultado de la reducción de dimensionalidad y la elección de un método de optimización adecuado, se obtienen un número reducido de términos lingüísticos por cada entrada así como bases de reglas compactas.

En el caso de las series de carga de tráfico, de manera similar a las series de test utilizadas para evaluar la metodología, en general, las bases de reglas son considerablemente compactas. Además, como regla general, los modelos para predicción a largo plazo constan de un número de reglas significativamente inferior, lo que se debe a que son modelos menos precisos. De este modo, los modelos NRVE-FIS son altamente interpretables en general. Desde nuestro punto de vista, no se debe esperar que los modelos NRVE-FIS proporcionen una explicación satisfactoria, completa y comprensible del comportamiento de las series de carga de tráfico en todos los casos. Sin embargo, todo ello es posible en algunos casos particulares, y es frecuente obtener detalles acerca de la dinámica de las series temporales expresados de manera simple. Por supuesto, el procedimiento necesario para dar una interpretación física a los modelos es, en gran medida, específico para cada caso.

En este resumen describiremos brevemente un caso que se detalla en esta tesis. El modelo NRVE-FIS de predicción del siguiente valor de la serie correspondiente a la traza AMPATH-OC12-200701-0 a escala de 1 s tiene cuatro entradas y una base de reglas compuesta por 16 reglas. Sin embargo, sólo se definen 15 valores de salida, por lo que dos reglas comparten el mismo valor de salida. Las cuatro entradas seleccionadas corresponden a la carga de tráfico en el último intervalo, y_{t-1} , dos segundos antes, y_{t-3} , cuatro segundos antes, y_{t-4} , y ocho segundos antes, y_{t-8} . La selección de estas variables indica que son las realmente relevantes para predecir la carga de tráfico en este enlace a escala de 1 s. Además, las relaciones entre las entradas y salidas se pueden interpretar de manera lingüística. Para cada entrada se definen dos términos lingüísticos, como se muestra en la figura 7, que se pueden interpretar como valores BAJOS (o B) y ALTOS (A), representados por sendas funciones de pertenencia gaussianas definidas en el rango de valores observados, $[7,6 \cdot 10^6, 3,0 \cdot 10^7]$. La base de reglas se muestra en la tabla 2, donde se especifican los centros de las salidas de tipo singleton.

En la tabla se puede observar, entre otros aspectos, que se puede esperar un incremento significativo de la carga de tráfico cuando se cumplen las condiciones especificadas por las reglas segunda y cuarta. En el proceso de interpretación, técnicas complementarias para simplificación de sistemas de inferencia difusa pueden ser de considerable ayuda. En

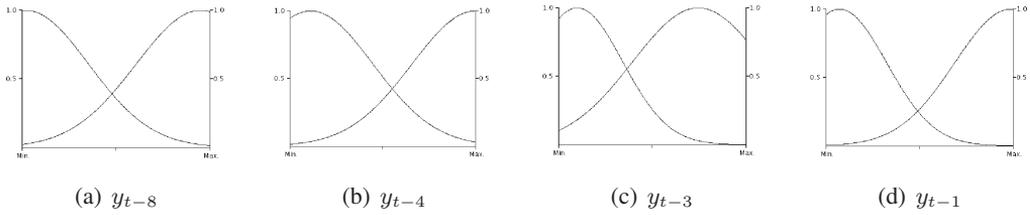


Figura 7: Términos lingüísticos para el modelo predictivo para el horizonte 1 de la serie AMPATH-OC12-200701-0 a escala de 1 s.

y_{t-8}	y_{t-4}	y_{t-3}	y_{t-1}	y_t
B	B	B	B	$8,15 \cdot 10^6$
B	B	B	A	$3,00 \cdot 10^7$
B	B	A	B	$1,12 \cdot 10^7$
B	B	A	A	$3,00 \cdot 10^7$
B	A	B	B	$1,48 \cdot 10^7$
B	A	B	A	$1,95 \cdot 10^7$
B	A	A	B	$1,22 \cdot 10^7$
B	A	A	A	$1,29 \cdot 10^7$
A	B	B	B	$1,52 \cdot 10^7$
A	B	B	A	$2,43 \cdot 10^7$
A	B	A	B	$2,44 \cdot 10^7$
A	B	A	A	$9,72 \cdot 10^6$
A	A	B	B	$1,63 \cdot 10^7$
A	A	B	A	$2,29 \cdot 10^7$
A	A	A	B	$1,77 \cdot 10^7$
A	A	A	A	$2,12 \cdot 10^7$

Tabla 2: Base de reglas del modelo NRVE-FIS de la serie correspondiente a la traza AMPATH-OC12-200701-0 a escala de 1 s.

este sentido, se pueden emplear los métodos de simplificación incluidos en el entorno de desarrollo Xfuzzy, en el que se engloba la herramienta xftsp. Por ejemplo, si el sistema se simplifica hasta obtener las seis mejores reglas, se obtiene un sistema que alcanza un error MSE sólo 2.03 % superior al del sistema original. En particular, se seleccionan las reglas primera, segunda, tercera, octava, decimoprimer y decimoquinta.

Los resultados mostrados en esta tesis se han obtenido utilizando los métodos de identificación y ajuste de sistemas difusos especificados en el primer apartado del capítulo 3. No obstante, la misma metodología se puede aplicar utilizando métodos de identificación alternativos, tales como el método basado en agrupamiento sustractivo propuesto en [72]. Obviamente, se obtendrán diferentes estructuras de sistemas dependiendo del método de identificación aplicado. Además de diferencias en la precisión de las predicciones, el número de reglas identificadas así como su forma pueden variar notablemente. Por ejemplo, en el caso descrito anteriormente para la traza AMPATH-OC12-200701-0, el método basado en agrupamiento sustractivo proporciona un modelo con un MSE 4.6 % mayor, mientras que se identifican 12 términos lingüísticos para cada entrada así como una base de reglas compuesta por 12 reglas.

Se puede argumentar que se ha aplicado un gran número de modelos al modelado de tráfico de redes sin que hayan aportado una mejor comprensión de la dinámica del tráfico. En este contexto, los modelos opacos que son descriptivos en un sentido puramente numérico y no proporcionan ninguna comprensión física son poco útiles. No obstante, en esta tesis se abordan aspectos que han permanecido esencialmente inexplorados hasta ahora: predicción a largo plazo a diferentes escalas temporales y predecibilidad mediante técnicas no paramétricas. Además, los modelos difusos desarrollados hacen énfasis en la interpretabilidad en un sentido lingüístico, siendo por tanto un nuevo enfoque que puede proporcionar descripciones fácilmente interpretables de la evolución temporal de la carga de tráfico en enlaces de red.

Como aspecto relacionado con el párrafo anterior debemos subrayar que el análisis realizado en esta tesis ha seguido un enfoque esencialmente neutro o ciego, en cuanto a que no se han aplicado técnicas exploratorias de análisis de datos previas a la aplicación de los métodos de predicción. Esto ha hecho posible analizar una extensa colección de trazas de tráfico heterogéneas. Sin embargo, la experiencia muestra que, mediante una etapa de análisis de datos inicial a medida de cada serie temporal, se pueden conseguir mejoras significativas en los resultados [112, 37].

Como último comentario al resumen de este apartado, hemos de mencionar que se ha mostrado que las series de tráfico a escalas de días, semanas y superiores son predecibles en la práctica en un grado considerable. Este hecho es una prueba de la aplicabilidad práctica de las técnicas de predicción actuales en tareas de operación y planificación de redes, confirmando trabajos previos realizados sobre medidas no disponibles públicamente [88], así como la asunción generalizada entre especialistas de que ciertos patrones diarios, semanales y anuales observados en muchas redes son predecibles. Sin embargo, se ha mostrado que las series analizadas a escalas más pequeñas son esencialmente impredecibles, con muy pocas excepciones. En efecto, sólo son posibles predicciones muy poco precisas y a muy corto plazo. Este hecho supone un reto para el desarrollo de esquemas de control de tráfico basados en predicciones.

3.3. Análisis y resumen de registros de flujos de red

Como se discute en el capítulo 2 de esta tesis, los sistemas de medida y monitorización son claves tanto para usuarios como para operadores y gestores. Durante los últimos años se han producido una gran cantidad de desarrollos en torno a infraestructuras y sistemas de medida [27, 9, 15, 101].

Actualmente, el enfoque dominante para desplegar sistemas de medida y monitorización está basado en medidas pasivas de flujos de red utilizando la tecnología NetFlow [22, 53]. Con una creciente diversidad de tecnologías, aplicaciones y patrones de tráfico, el análisis de flujos de tráfico es cada vez más complejo. Actualmente, interpretar de forma comprensiva los hechos relevantes contenidos en estas medidas queda en la práctica fuera de las posibilidades de los operadores y gestores.

A medida que la complejidad de las tareas de operación y gestión de redes crece, el desarrollo de los sistemas de medida ha avanzado significativamente durante los últimos años. Así, los sistemas actuales producen enormes cantidades de datos que se han de presentar en forma de informes y estadísticas. Estos informes pueden ser muy difíciles de interpretar. En particular, las redes que implementan la tecnología NetFlow [22] pueden generar varios GBs de datos diariamente incluso a pesar del empleo de técnicas de muestreo con objeto de reducir los volúmenes de datos generados. Por otra parte, la medida de precisión de tráfico en los nodos troncales actuales implica la generación de decenas de GBs de datos por hora.

El objetivo fundamental de los sistemas de medida de tráfico es hacer posible la comprensión del funcionamiento de las redes. Sin embargo, la distancia entre los sistemas de medida y la comprensión de los usuarios está creciendo. En particular, el tiempo necesario para analizar los datos lleva a que los experimentos y estudios sobre medida del tráfico quedan obsoletos desde el momento en que se terminan.

Existen diversas herramientas de visualización para medidas de redes, véase [26], que en la mayoría de los casos se fundamentan en gráficas de evaluación de propiedades estadísticas de series temporales, propiedades a diferentes escalas y el comportamiento de diferentes protocolos [98]. Las herramientas de visualización y generación de informes empleadas en la actualidad proporcionan informes compuestos por decenas de gráficas y tablas. Por ello, no es fácil para los expertos extraer resúmenes simples. Además, la complejidad de estas herramientas incide negativamente sobre su nivel de adopción.

Si bien existen varias herramientas que generan resúmenes breves de estadísticas de red, tales como las herramientas de análisis proporcionadas por los principales fabricantes de routers, o la conocida *flow-report*, incluida en la herramientas *flow-tools* [44], se puede aducir que sólo consiguen cierta interpretabilidad desde la perspectiva humana en un grado muy limitado. Asimismo, las relaciones complejas subyacentes en las medidas de flujos de red pasan desapercibidas y se requiere un considerable esfuerzo de investigación para su identificación. Se reconoce en general que son necesarias herramientas complementarias de visualización para comprender estos informes.

Es común que los sistemas de medida de flujos se utilicen para tareas de monitorización en tiempo real así como para tareas generales de operación. Debido a las grandes cantidades de datos que pueden producir los sistemas de medida de flujos, el análisis en tiempo real de las redes de altas prestaciones actuales requiere métodos rápidos, a menudo simplifica-

dos, e implementaciones optimizadas. Este hecho ha motivado el desarrollo de hardware específico para clasificación y monitorización de flujos de tráfico [120].

A lo largo de los últimos años se han desarrollado múltiples métodos para analizar datos de diversas medidas de Internet. Por ejemplo, algunas técnicas ampliamente extendidas en el ámbito de la minería de datos, tales como el análisis de componentes principales, PCA, se han aplicado para analizar flujos de tráfico desde un punto de vista estructural [65]. No obstante, la mayoría de este tipo de técnicas son cuantitativas, adecuadas para ciertos tipos de datos y, en general, están diseñadas para propósitos específicos. Existe así una carencia de herramientas de propósito general para la exploración cualitativa y el análisis de medidas de Internet, lo que constituye un primer paso conducente al análisis, la validación y el descubrimiento de hipótesis [102].

El enfoque que se propone en esta tesis parte de las ideas previas mencionadas. Nuestra propuesta va en la línea de obtener resúmenes lingüísticos que extraigan la información relevante contenida en medidas de flujos de red y tengan una forma directamente legible por seres humanos. En este contexto, es cada vez más necesario extraer resúmenes *concisos*, lo que implica que sean de un tamaño varios órdenes de magnitud inferior a las medidas originales, y que expresen el comportamiento de la red en idealmente no más de unas pocas líneas de *texto*. [35].

En esta tesis, abordamos el problema de la generación de resúmenes de medidas de red en forma de informes breves legibles por seres humanos mediante la teoría de resúmenes lingüísticos difusos [77].

Los resúmenes lingüísticos mediante lógica difusa han mostrado ser un método de minería de datos simple, eficiente y consistente con el razonamiento humano. Esta técnica se puede usar como una herramienta de descubrimiento de conocimiento basada en el lenguaje natural. La idea fundamental de los resúmenes lingüísticos difusos es usar términos lingüísticos, pertenecientes al lenguaje natural, para expresar información y conocimiento oculto en colecciones de objetos potencialmente de gran tamaño. Los resúmenes lingüísticos en la forma introducida por Yager [115, 116] y posteriormente desarrollados por Kacprzyk y Yager [56] así como Kacprzyk y Zadrozny [57], son proposiciones cuantificadas lingüísticamente (como “La mayoría de los flujos tienen una duración breve”), para los que se define un grado de verdad. Este tipo de resúmenes lingüísticos es especialmente general y de rápida implementación, dos características claves para analizar colecciones de flujos.

Cabe mencionar que se han propuesto otros enfoques de generación de resúmenes lingüísticos basados en lógica difusa. Por ejemplo, el método empleado en [95] se basa en el uso de técnicas de agrupamiento para construir una jerarquía de resúmenes, mientras que la propuesta realizada en [96] se centra en el análisis de dependencias graduales y difusas. Otro enfoque alternativo que no hemos considerado en esta tesis podría ser el uso de la teoría de la evidencia de Dempster-Shafer para extraer reglas [1]. Las funciones de creencia y plausibilidad en este caso se podrían utilizar para definir el grado de soporte de una proposición [14]. Por ello, estas funciones se pueden utilizar para realizar un proceso inductivo de extracción de reglas basado en las asignaciones de masa para varias proposiciones. Además, en esta línea de desarrollo las funciones de creencia y plausibilidad se pueden ver como medidas pesimistas y optimistas de la fuerza de una regla [118].

En el apartado 3.3 de esta tesis, proponemos dos métodos para extraer resúmenes lingüísticos de colecciones de datos resultante de medidas realizadas mediante sistemas NetFlow. Los métodos tienen como objetivo, por una parte, obtener resúmenes concisos de medidas de red y, por otra, extraer reglas de asociación difusas. Ambos métodos encuentran aplicaciones en análisis y caracterización del tráfico de red.

El primer método propuesto es útil para monitorización en tiempo real y generación de informes concisos. Dado que los algoritmos empleados para implementar este método son computacionalmente simples, es posible la generación de informes completos en tiempo real. Sin embargo, este primer método sólo explota una pequeña parte del potencial de los resúmenes lingüísticos, dado que sólo evalúa un conjunto fijo de protoformas especificadas a priori. Por ello, proponemos un segundo método, basado en técnicas de minería de reglas de asociación. Así, desarrollamos un método para extraer relaciones ocultas en forma de resúmenes lingüísticos difusos a partir de medidas de flujos de red.

En el capítulo 3 de esta tesis se analiza la tecnología NetFlow, se describen los resúmenes lingüísticos basados en lógica difusa y se definen resúmenes lingüísticos específicos para registros de flujos de red. Se muestran asimismo resultados experimentales para un amplio conjunto de colecciones de flujos.

Así, se aborda el problema del análisis de medidas de tráfico de red al nivel de flujos desde la perspectiva de los resúmenes lingüísticos. Los métodos propuestos se plantean como herramientas complementarias a las técnicas estadísticas descriptivas tradicionales. El hecho de que ambas técnicas son complementarias es especialmente evidente si se considera que las etiquetas lingüísticas definidas en el capítulo 3 de esta tesis se han extraído a partir de observaciones experimentales obtenidas usando técnicas estadísticas tradicionales, tales como funciones de distribución e histogramas.

Se proponen dos enfoques para generar resúmenes lingüísticos de colecciones de flujos, tanto en tiempo real como a posteriori. En el primer caso, se identifican un conjunto de protoformas como condiciones de especial interés. En el segundo caso, se utiliza un algoritmo de minería de reglas de asociación rápido con objeto de identificar los resúmenes relevantes.

En el capítulo 3, se detalla además el funcionamiento de una herramienta que implementa ambos enfoques de generación de resúmenes, *flow-lsummary* [77]. Los resultados experimentales obtenidos para un conjunto de colecciones de flujos de test confirman que los resúmenes lingüísticos aportan una nueva perspectiva alternativa para analizar estadísticas de tráfico de red, siendo útil tanto para usuarios como para técnicos. Así, no es necesario que los operadores tengan un conocimiento profundo de varios campos técnicos, es posible proporcionar resúmenes simples a los usuarios, y las estadísticas resumidas pueden ayudar en las funciones de planificación y dimensionamiento.

El enfoque propuesto es por tanto una nueva técnica para generar informes simples y legibles por seres humanos, siendo útil para usuarios y técnicos, al tiempo que proporciona una técnica prometedora para encontrar invariantes en medidas de tráfico de red. Este hecho se puede ver como un primer paso para el desarrollo de herramientas de descubrimiento de conocimiento basadas en lenguaje natural aplicadas a la Ciencia de Internet.

En la figura 8 se muestra el número de flujos procesador por segundo con la herramienta *flow-lsummary* [77] para la colección de flujos Internet2-KANS-May08, véase el capítulo

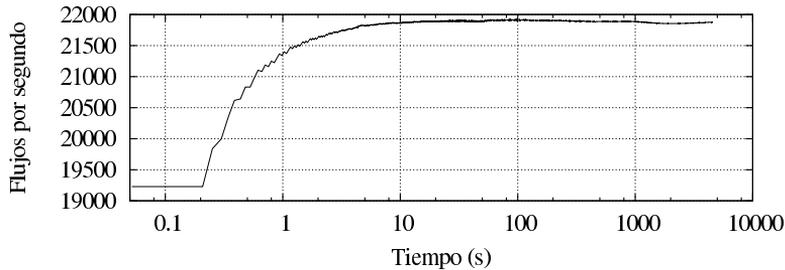


Figura 8: Flujos procesados por segundo con flow-lsummary para la colección de flujos Internet2-KANS-May08.

3 para más detalles. El proceso se ejecutó en un ordenador personal común sobre un procesador Intel® Core™ 2 Duo CPU E6550 a 2.33GHz con 4 MB de memoria cache L1 y 2 GB de memoria RAM, usando el sistema operativo GNU/Linux y un paquete estándar de Perl versión 5.10.0. El consumo de memoria es constante y menor de 7 MB, puesto que el proceso de generación de resúmenes sólo requiere un procedimiento de una pasada sobre los datos sin requisitos de memoria significativos.

En la figura se puede comprobar que, tras una breve etapa de inicio, el número de flujos procesado por segundo es superior a $20 \cdot 10^3$. Esta tasa de procesado es notablemente alta y constante debido a la naturaleza puramente estática del algoritmo aplicado, que emplea sólo procedimientos de una pasada sobre los datos. El procesado de 10^9 flujos lleva aproximadamente 4571 segundos, o, lo que es lo mismo, una hora, 16 minutos y 11 segundos. Esto es, una colección de flujos de red correspondiente a un mes completo de tráfico se puede resumir en aproximadamente tres horas y seis minutos. No obstante, la colección de flujos Internet2-KANS-May08 se generó utilizando una tasa de muestreo de paquetes del 1 %.

Si no se hubiese aplicado ninguna técnica de muestreo, y asumiendo como peor caso un número de flujos 100 veces mayor, una colección de flujos de red correspondiente a un mes completo de tráfico en un enlace de altas prestaciones se podría procesar en menos de 13 días. Por tanto, es posible generar resúmenes en tiempo real con la herramienta flow-lsummary sin optimizaciones específicas incluso para enlaces de muy altas prestaciones. En redes con cargas de tráfico más moderadas, el proceso de generación de resúmenes tiene unos requisitos de cálculo muy bajos. Por ejemplo, generar un resumen simple de la colección de flujos Darmouth-Fall03, correspondiente a 15 días de tráfico, toma aproximadamente tres minutos y 50 segundos con la misma configuración anterior.

Por último, hemos de mencionar que se ha estudiado el análisis y la generación de resúmenes de medidas de tráfico de red a nivel de flujos, en contraposición a medidas a nivel de paquetes. Así, mientras que en el apartado 3.2 de esta tesis se analizan series temporales generadas a partir de medidas de nivel de paquetes, en la sección 3.3 se describen métodos de análisis de medidas a nivel de flujos de red. El segundo enfoque de medida y análisis está mucho más generalizado en la práctica, encontrando multitud de aplicaciones, dado que en la actualidad hay disponibles funciones de medida de flujos en virtualmente todos los routers de gama media y alta. Sin embargo, la generación de resúmenes lingüísticos de medidas de tráfico a nivel de paquetes es también una opción plausible que creemos merece ser

explorada como ampliación futura del trabajo presentado. Con este fin, los procedimientos y algunos de los términos lingüísticos que se definen en esta tesis se pueden extrapolar de manera directa desde el análisis de tráfico a nivel de flujos de red al análisis a nivel de paquetes.

3.4. Conclusiones

Los resultados presentados en el capítulo 3 constituyen la primera contribución principal de esta tesis. En este capítulo se describen dos técnicas de modelado heterogéneas basadas en métodos difusos. Ambas se han aplicado al modelado de tráfico de red. Estas técnicas se han desarrollado con dos objetivos complementarios. Por una parte, se pretende predecir la carga de tráfico en enlaces a diferentes escalas de tiempo; por otra parte, se generan resúmenes concisos de medidas de tráfico de red.

Respecto al primer objetivo, se presenta una metodología para predicción de series temporales a largo plazo mediante sistemas de inferencia difusa. La metodología se valida mediante un conjunto de series de test y se aplica a un amplio conjunto de trazas de tráfico de red. En esta metodología, el uso del Delta Test como estimador de la varianza residual proporciona dos ventajas: se utiliza para realizar una selección de entradas a priori, y además proporciona una estimación cuantitativa que se usa para dirigir el proceso de aprendizaje.

A continuación se realiza un análisis de predecibilidad sobre trazas de tráfico mediante Delta Test. Los resultados muestran un grado de predecibilidad satisfactorio para algunas aplicaciones, asimismo se pone de manifiesto que los modelos regresivos basados en inferencia difusa propuestos aquí pueden alcanzar una precisión muy próxima al límite teórico de predecibilidad estimado. Estos resultados tienen implicaciones prácticas importantes en muchas aplicaciones, tales como sistemas de control de tráfico basados en predicción. Se muestra además que los sistemas regresivos difusos mejoran en prestaciones a los modelos LS-SVM, OP-ELM y ARIMA en cuanto a precisión y capacidad de generalización.

A continuación, se describen métodos para extraer resúmenes lingüísticos a partir de estadísticas y medidas de flujos de red. Se aborda tanto la generación de resúmenes sencillos en tiempo real como la extracción y minería de reglas de asociación difusa a posteriori. Se propone asimismo una medida de la bondad o grado de información de los resúmenes lingüísticos. Los métodos se han implementado en una herramienta útil tanto para investigadores como para operadores. La herramienta, *flow-lsummary*, se aplica a un amplio conjunto de colecciones de medidas de flujos diversas, mostrándose que el método propuesto genera resúmenes concisos, informativos y autoexplicativos.

4. Sistemas de inferencia difusa para control de tráfico de red

Los modelos disponibles hasta la fecha sobre la dinámica de extremo a extremo de Internet son, en el mejor de los casos, incompletos [75]. El desarrollo de mecanismos de control de tráfico robustos ante una creciente diversidad de tecnologías, aplicaciones y servicios es un área de investigación fundamental en el diseño de protocolos para las capas de red y transporte de Internet.

Dos razones hacen que sea clave obtener una mejor comprensión de la dinámica del tráfico de extremo a extremo en Internet. En primer lugar, haría posible la optimización de los recursos de la red y, como consecuencia, de las prestaciones de extremo a extremo observadas por los usuarios finales. En segundo lugar, posibilitaría la provisión de clases de servicio y garantías de calidad de servicio mejoradas. En palabras de Keshav [61], *“el Santo Grial de las redes de computadores es diseñar una red que tenga la flexibilidad y bajo coste de Internet y aun así ofrezca las garantías de calidad de servicio de extremo a extremo que ofrece la red de telefonía clásica.”*

Es decir, en la Internet actual existe una necesidad generalizada de mecanismos que proporcionen ciertas garantías de calidad de servicio. Esto requiere la capacidad de proporcionar calidad de servicio mejorada, es decir, la capacidad de evitar degradaciones de las prestaciones y fallos debidos a congestión. En casos extremos, tales como puedan ser las aplicaciones de cirugía remota, llamadas de emergencia y otras aplicaciones de las redes de misión crítica, puede ser necesario evitar cualquier falta de fiabilidad de extremo a extremo.

Sin embargo, la provisión de garantías de calidad de servicio no es el único problema abierto en este contexto. De hecho, el desarrollo de mecanismos eficientes para control de tráfico best effort plantea numerosos retos. En particular, existe la necesidad de garantizar un equilibrio entre factores antagónicos, como la utilización del ancho de banda, el reparto equitativo de recursos, la minimización del retardo de extremo a extremo y la variabilidad del retardo, y las funciones de prevención de congestión.

En la práctica, la gran mayoría de las redes actuales, y especialmente las redes troncales, funcionan satisfactoriamente gracias a un frecuente sobredimensionamiento. Sin embargo, debido a la naturaleza compleja y altamente variable del tráfico en redes de conmutación de paquetes, los usuarios de servicios best effort pueden sufrir degradaciones en las prestaciones incluso cuando el ancho de banda disponible es suficiente y el uso medio del ancho de banda es relativamente bajo [15]. En particular, se ha comprobado que incluso en enlaces con muy baja utilización son frecuentes los episodios de microcongestión, o periodos de congestión localizada que ocurren a escalas inferiores a las habitualmente analizadas por los operadores de red [89].

En el capítulo 4 de esta tesis, abordamos mediante sistemas de inferencia difusa dos problemas estrechamente relacionados en torno al control de tráfico: control de congestión de extremo a extremo y gestión activa de colas en routers. El capítulo se centra en el análisis de estos problemas con tráfico best effort. Se evalúan implementaciones software utilizando un simulador de nivel de paquetes, así como herramientas software experimentales y sistemas de emulación de redes.

En un primer apartado, se analizan dos escenarios de simulación. En el apartado 4.2, desarrollamos mecanismos inteligentes para control de congestión de extremo a extremo

mediante sistemas de inferencia difusa [81]. En primer lugar se describe una generalización de los principios de control de congestión del protocolo TCP basada en lógica difusa. A continuación, se diseña un controlador de congestión siguiendo un esquema basado en mecanismo de ventana similar al de TCP. Se sigue una metodología sistemática para el diseño e implementación en forma de herramienta experimental de sistemas de inferencia difusa. Asimismo se realiza un estudio comparativo de prestaciones del sistema propuesto frente a controladores tradicionales. En esta tesis se muestra que, además de ser un enfoque de modelado útil, el esquema de control de congestión basado en reglas propuesto puede proporcionar mejoras de prestaciones frente a otras alternativas tradicionales.

Un problema relacionado y de gran importancia en el ámbito de las capas de red y transporte de la pila de protocolos de Internet es el desarrollo de mecanismos de control de tráfico capaces de responder adecuadamente a las necesidades de una creciente diversidad de tecnologías, aplicaciones y servicios. En el apartado 4.3 de esta tesis abordamos este problema y desarrollamos mecanismos para control inteligente de colas en routers por medio de sistemas de inferencia difusa [82]. Estos mecanismos se desarrollan siguiendo una línea de investigación iniciada por trabajos independientes, haciendo en nuestro caso especial énfasis en la evaluación en un amplio rango de escenarios así como la consideración de las restricciones prácticas de implementación. La propuesta desarrollada se compara extensivamente frente a algunos mecanismos tradicionales.

Para ello, se emplea una metodología sistemática y se lleva a cabo una evaluación comparativa mediante simulación así como una implementación experimental. Los resultados obtenidos para la propuesta de controlador presentada en esta tesis se comparan frente a un conjunto de controladores tradicionales, entre los cuales se encuentran las variaciones más ampliamente aceptadas y desplegadas en la actualidad.

4.1. Escenarios de simulación

Como se discute de manera pormenorizada en [43], la simulación de escenarios de red es una herramienta genérica útil para lograr comprender la dinámica, ilustrar ciertos efectos, o explorar comportamientos inesperados. Aun así, se deben tener precauciones al comparar diversas opciones y particularmente cuando se pretende obtener medidas de prestaciones comparativas. Es bien conocido que pequeños cambios, a veces unas pocas simples líneas de código, en protocolos y aplicaciones populares utilizadas en Internet tienen un enorme impacto potencial sobre los patrones de tráfico así como las infraestructuras y sus usuarios. En consecuencia, cualquier comparación numérica entre dos protocolos realizada mediante simulación está sujeta a dos cuestiones esenciales. En primer lugar, ¿podrían cambiar significativamente los resultados debido a un pequeño cambio en el modelo? En segundo lugar, ¿cual sería el impacto sobre los resultados de un cambio en un detalle de la implementación software de los modelos?

Por ello, en las secciones 4.2 y 4.3 de esta tesis, se explora mediante simulación el comportamiento de los esquemas propuestos para control de congestión de extremo y para gestión activa de colas, respectivamente. En particular, se ha definido un escenario de simulación complejo con objeto de analizar el comportamiento de los sistemas bajo un rango de

condiciones amplio y realista. En el apartado 4.1 de esta tesis se detallan los dos escenarios de simulación que se utilizan en los dos siguientes apartados.

El objetivo principal de este estudio es facilitar el diseño de los esquemas de control de tráfico mediante sistemas de inferencia difusa propuestos en los siguientes apartados. En estos apartados se comprobará mediante simulaciones la robustez y las posibles mejoras de prestaciones de las nuevas propuestas bajo diferentes condiciones.

De manera más específica, para gestión activa de colas, se analiza la robustez de diferentes algoritmos frente a tráfico con ráfagas con diferentes grados de carga global de tráfico sobre la red. Para control de congestión de extremo a extremo, se considera la robustez y las prestaciones bajo diferentes niveles de carga en un camino de extremo a extremo en particular y diferentes niveles de carga de tráfico competitivo. Se comparan las propuestas frente a dos alternativas tradicionales y se muestran las mejoras aportadas por las primeras en diversos escenarios de simulación. No obstante, hemos de hacer notar que este análisis comparativo está limitado por dos factores:

- La ingente diversidad de propuestas de algoritmos de control de congestión de extremo a extremo, entre las cuales las que se mencionan en el apartado 4.2 de esta tesis constituyen sólo un pequeño subconjunto. Este hecho hace casi imposible realizar una evaluación minuciosa de un conjunto representativo de propuestas bajo un conjunto de escenarios de simulación suficientemente amplio.
- Todo análisis de prestaciones de protocolos, ya se realice mediante simulación o sobre implementaciones experimentales, está afectado por un cierto grado de parcialidad creciente a medida que Internet evoluciona.

En esta tesis se ilustra el comportamiento de las dos propuestas mencionadas mediante simulación, al tiempo que se comparan con otras alternativas. Sin embargo, esta comparación tiene sentido en la medida en que los resultados de cualquiera de las evaluaciones cualitativas y, especialmente, cuantitativas se interpreten bajo la posición de cautela formulada en los párrafos anteriores.

Las simulaciones realizadas se complementan con tests mediante redes emuladas así como implementaciones. Si bien los escenarios de emulación e implementación proporcionan resultados de los cuales cabe esperar un alto grado de aproximación a los resultados de implementaciones desplegadas en redes en producción, las pruebas de simulación nos permiten analizar la dinámica de los algoritmos y esquemas de control de manera mucho más flexible y en un conjunto de condiciones más amplio.

En esta tesis se realizan simulaciones a nivel de paquetes. Este nivel de detalle es necesario para poder realizar una evaluación eficaz de los dos esquemas de control mencionados anteriormente. En el caso del control de extremo a extremo, el complejo comportamiento resultante de los algoritmos de control de congestión basados en mecanismo de ventana, y especialmente cuando existe tráfico de fondo, no se puede modelar completamente salvo que las simulaciones se realicen a nivel de paquetes. Lo mismo se puede aplicar a los esquemas de gestión activa de colas, para los cuales los modelos analíticos existentes son incompletos en el mejor de los casos. Por ello, hemos usado el simulador ns-2 [52], un estándar de facto dentro de la comunidad de investigación en protocolos de Internet.

Se utilizan dos escenarios a lo largo de los tests de simulación: dumbbell y GREN, por las siglas en inglés de la red global de redes de investigación y educación. Por una parte, el primer escenario cumple el papel de un escenario simple con el que se ilustran los principios de funcionamiento básico. Por otra parte, el segundo escenario está caracterizado por una topología extensa y compleja así como por esquemas de generación de patrones de tráfico complejos [43].

4.2. Control difuso de extremo a extremo para protocolos de transporte

La dinámica de extremo a extremo del tráfico de paquetes en Internet es un problema complejo para el cual los modelos disponibles hasta la fecha son, en el mejor de los casos, incompletos [42]. En este área, dos problemas concretos vienen recibiendo especial atención por parte de la comunidad investigadora durante los últimos años: el diseño de protocolos para redes de muy altas prestaciones, y mecanismos de control de congestión adecuados a nuevos servicios y aplicaciones. Los problemas encontrados en estos campos han llevado a que surja un cierto interés en el seno del End-to-End Research Group de la Internet Research Task Force en la aplicación de técnicas de inteligencia artificial y computacional a problemas de tráfico de extremo a extremo.

Como se muestra en la figura 9, el tráfico de extremo a extremo está sujeto a acciones de control llevadas a cabo no sólo en los nodos finales de la red sino también en todos y cada uno de los nodos que conforman el camino de extremo a extremo. Esto incluye nodos de acceso, nodos de enlaces troncales y puntos de intercambio, con la consecuente variedad de tecnologías y condiciones, desconocidas en principio para los mecanismos de la capa de transporte. Los esquemas de control de tráfico desplegados actualmente en Internet, así como las alternativas propuestas más sólidas, se pueden englobar en alguno de los dos enfoques siguientes, véase también la figura 10 en la página 50:

- Control distribuido, con funcionalidad distribuida entre los nodos finales de la red, e implementado mediante protocolos de transporte de extremo a extremo. Los procesos transmisores y receptores situados en los nodos finales actúan de modo que se realizan funciones de control de congestión y reparto de recursos de forma cooperativa.
- Control de colas en nodos intermedios o routers. Estos mecanismos pueden tratar de forma diferenciada algunos flujos de tráfico e imponer ciertas condiciones de reserva y distribución de recursos.

Por ello, el control de flujos de paquetes desde un emisor hasta el correspondiente receptor puede involucrar a todos los nodos de la red que forman parte del camino de extremo a extremo. Este control se realiza así de extremo a extremo al tiempo que en cada uno de los nodos del camino. Un esquema de este tipo conforma un sistema de control constituido por múltiples bucles cerrados con interacciones complejas, como se muestra en la figura 10.

Los flujos de extremo a extremo que atraviesan los routers corresponden a un rango muy amplio de requisitos de usuario y características dinámicas. Por ejemplo, el número de enlaces de un camino de extremo a extremo típico puede estar entre unos pocos y hasta 20; el tiempo de retorno de paquetes puede estar comprendido entre unos pocos milisegundos y

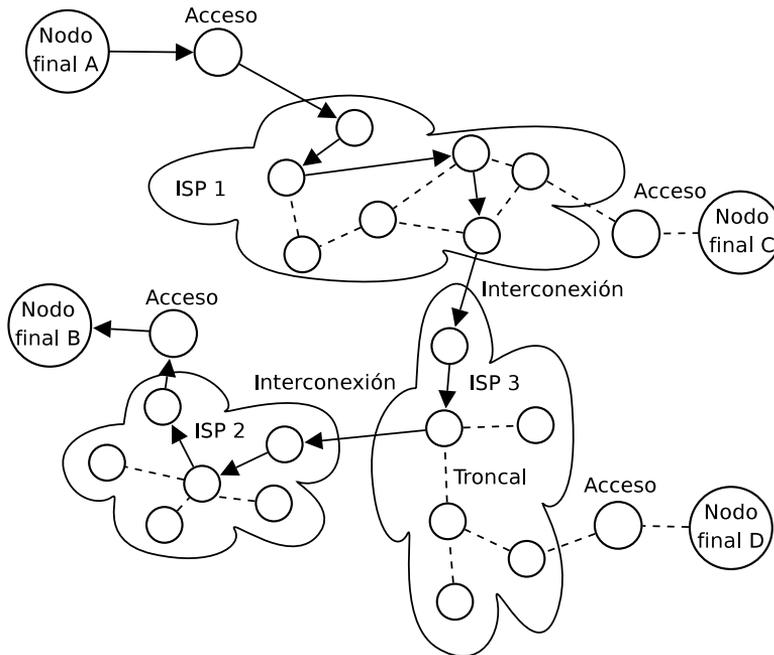


Figura 9: Camino de extremo a extremo de ejemplo entre dos nodos, y los elementos que pueden formar parte de este camino.

varios segundos; la duración de los flujos puede ser tan breve como unos pocos milisegundos o tan extensa como varias horas, y cada flujo puede transferir volúmenes comprendidos entre unos pocos KBs hasta varios GBs.

En una red como la esquematizada en la figura 9, los caminos posibles entre los nodos A y B pueden ser muy diferentes a los caminos posibles entre los nodos A y C o A y D, así como otros dos nodos finales cualquiera. En ello influyen factores tan diversos como las distancias físicas, los acuerdos y métodos de interconexión entre proveedores de servicio o ISP, la topología y dimensión de las redes troncales, las tecnologías y características de los enlaces de acceso, el uso de reguladores de tráfico, la aparición de tasas de pérdidas no despreciables debido a problemas de configuración, o la coexistencia de flujos de tráfico de características muy distintas. Asimismo, es difícil modelar la influencia de otros factores de mayor escala, tales como los efectos de los protocolos de encaminamiento, actualizaciones de subredes, y otros muchos.

Los requisitos de calidad de servicio así como los patrones de tráfico característicos de aplicaciones y servicios emergentes son difíciles de caracterizar, y demandan profundos avances en los esquemas de control de tráfico actuales. Debido a la naturaleza de estos problemas, una alternativa plausible para abordarlos es el empleo de sistemas inteligentes basados en lógica difusa y posiblemente otras técnicas de soft computing complementarias.

Los dos enfoques de control citados anteriormente se puede reformular en términos de sistemas difusos. En el apartado 4.2 de esta tesis nos centramos en el problema de control de congestión de extremo a extremo. Así, analizamos métodos de control basados en me-

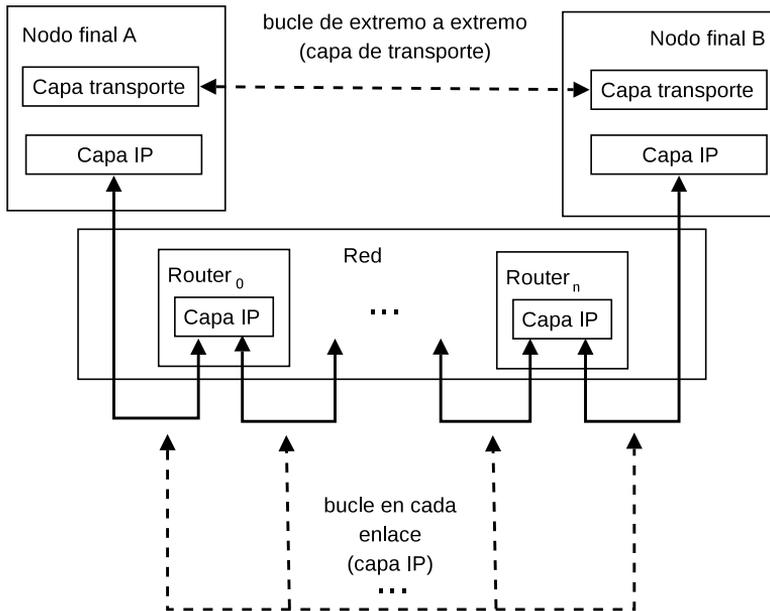


Figura 10: Bucles de realimentación en los sistemas de control de tráfico de Internet.

canismos de ventana similares a los empleados en el protocolo TCP estándar. El objeto de este apartado es reinterpretar el mencionado problema desde la perspectiva de los sistemas de inferencia difusa.

En esta tesis, se resumen brevemente los trabajos previos publicados en la literatura, se describe un procedimiento de generalización de los esquemas de control de congestión basados en mecanismos de ventana mediante lógica difusa, y se presenta el diseño de un controlador de congestión de extremo a extremo. Para este diseño se emplea una metodología de desarrollo de sistemas difusos adaptada al diseño de métodos de control de tráfico. Asimismo, se muestran resultados de experimentos realizados mediante simulación así como implementación experimental, mostrándose las ventajas del esquema propuesto en términos de prestaciones e interpretabilidad.

Para ilustrar brevemente los resultados obtenidos con la propuesta de control de congestión en el protocolo TCP mediante sistemas difusos, que llamaremos TCP difuso, mostramos en este resumen un ejemplo de funcionamiento de esta propuesta frente a dos alternativas tradicionales: TCP HighSpeed y TCP SACK.

Las figuras 11 y 12 comparan la evolución de la ventana de congestión de TCP, $cwnd$, así como la tasa de transferencia para tres tests de implementación. En el escenario utilizado dos nodos transfieren un archivo durante 30 segundos, la conexión TCP entre los nodos se establece a través de un camino formado por 5 enlaces. La capacidad máxima de estos enlaces es de 1 Gb/s, mientras que el cuello de botella tiene una capacidad de 100 Mb/s. El tiempo de retorno de paquetes en la dirección de la transferencia es de 20 ms aproximadamente. De entre varios tests de implementación realizados para corroborar los tests de simulación, se muestran tres casos.

Los resultados confirman las mejoras que TCP difuso aporta en términos de robustez

y tasa de transferencia final. La mejora en la tasa de transferencia final es de aproximadamente un 16 % y 12 % frente a TCP SACK y TCP HighSpeed, respectivamente, para los tests uno y dos. Estos dos tests se realizaron en un camino de extremo a extremo en el que la capacidad de los enlaces estaba mayoritariamente disponible debido a que el tráfico de fondo era escaso. Para realizar el test 3, se establecieron cuatro flujos de TCP SACK en competencia con los flujos de test. Dos de estos flujos se generan en un sentido, mientras que los otros dos se generan en el sentido contrario. Hemos de mencionar que todos los enlaces del camino son de tipo full-duplex y simétricos. En este test, véanse las figuras 11(c) y 12(c), se puede observar que la presencia de tráfico de fondo y en competencia, así como la presencia de flujos en el sentido inverso pueden degradar significativamente las prestaciones de las variantes tradicionales de TCP. En contraste con este hecho, la variante difusa consigue mantener una mayor estabilidad y tasa final de transferencia, con una mejora de esta última de alrededor del 50 %.

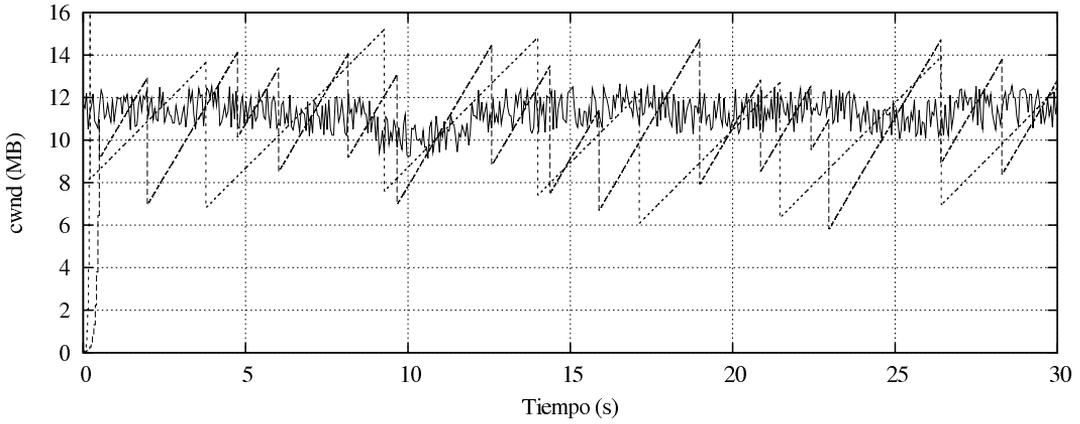
Con el análisis realizado mediante simulación y pruebas con una implementación experimental, se muestra que es posible reinterpretar y extender los mecanismos de control de congestión de TCP mediante lógica difusa. El modelo difuso descrito proporciona una perspectiva basada en reglas lingüísticas sobre los esquemas de control de tráfico de extremo a extremo. Esta propuesta es asimismo un primer resultado en la inferencia del estado de la red desde nodos finales mediante sistemas difusos.

De forma resumida, los experimentos realizados mediante simulación confirman que el esquema de control de congestión mediante sistemas difusos propuesto aporta mejoras significativas y puede lograr una mejor gestión de los recursos de la red en función del estado actual. Mediante pruebas de simulación, se analiza bajo condiciones controladas el comportamiento de la propuesta en un escenario realista con tráfico de fondo complejo, presentando entre otras características una alta variabilidad de tiempo de retorno de paquetes. Estos resultados se confirman asimismo mediante experimentos con una implementación real.

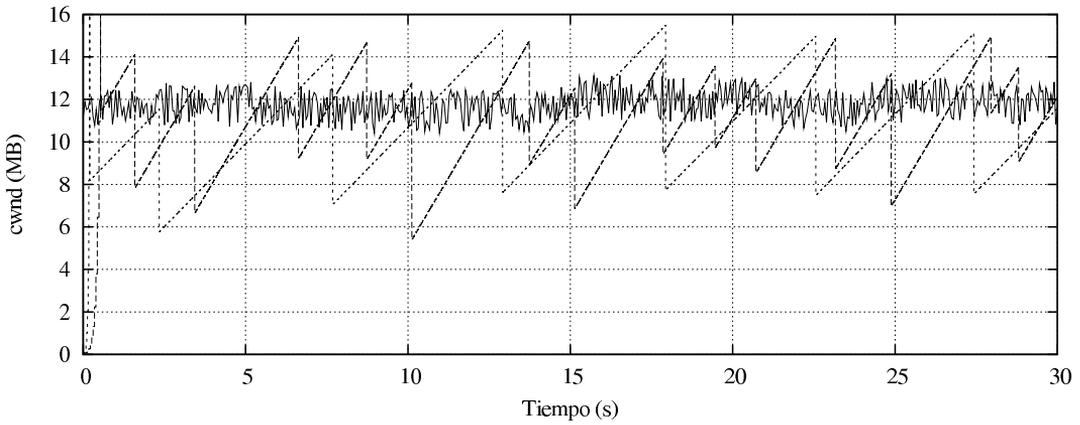
Así, los resultados de las etapas de simulación e implementación muestran que la propuesta presentada en el capítulo 4 de esta tesis puede mejorar las prestaciones de los sistemas de control de tráfico de extremo a extremo en términos de un conjunto de criterios, incluyendo una convergencia más rápida hacia la tasa de transferencia posible, una mayor tasa de transferencia final y menores oscilaciones en torno a la tasa de transferencia media para transferencias prolongadas. El sistema propuesto facilita el diseño de esquemas adaptados a diferentes requisitos.

No obstante, el trabajo presentado en el epígrafe 4.2 de este tesis abre numerosas líneas de trabajo para extender los sistemas propuestos. En particular, en futuras modificaciones cabría considerar la identificación de nuevas reglas, la exploración del espacio completo de reglas posibles, así como la introducción de nuevas entradas. Asimismo, los siguientes temas son de especial interés:

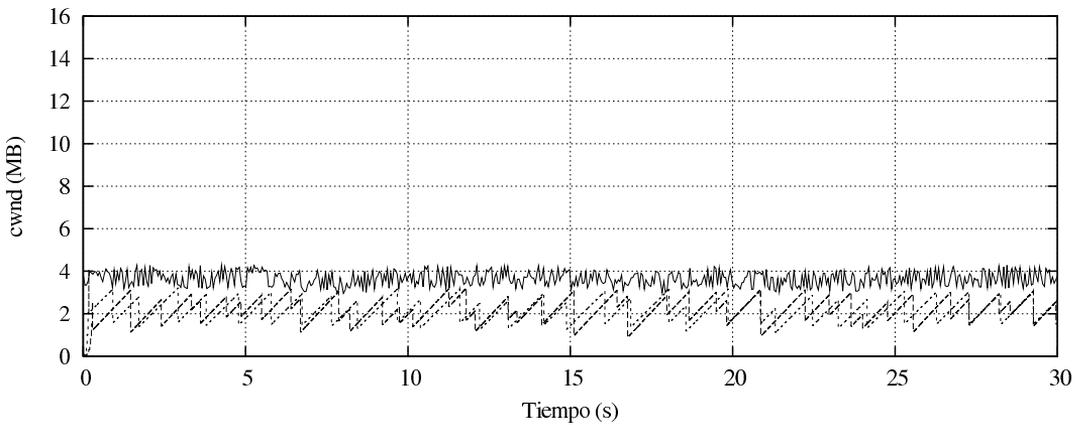
- Aplicación de técnicas automáticas de aprendizaje y ajuste en escenarios particulares, lo que permitiría analizar la dinámica del tráfico en estos casos.
- Análisis de extensiones de los mecanismos actuales de control de congestión propuestos con carácter experimental, tales como cambios en el valor inicial del tamaño de la



(a) Test 1



(b) Test 2



(c) Test 3

Figura 11: Evolución de la ventana de congestión en un escenario real para tres tests. Línea continua: TCP difuso; línea a trazos largos: HighSpeed TCP; línea a trazos cortos: TCP SACK.

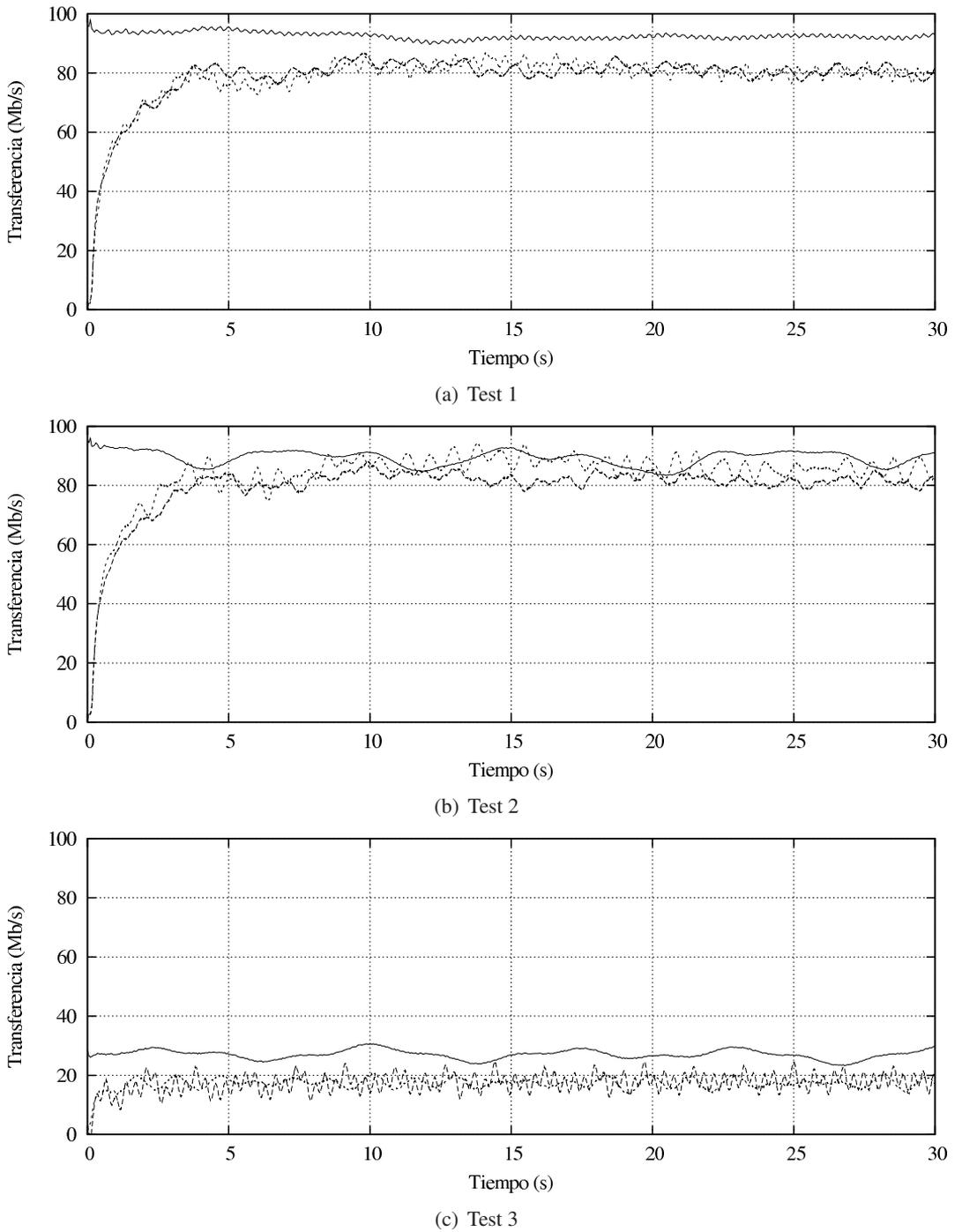


Figura 12: Evolución de la tasa de transferencia en un escenario real para tres tests. Línea continua: TCP difuso; línea a trazos largos: HighSpeed TCP; línea a trazos cortos: TCP SACK.

ventana de congestión, así como cambios en los procedimientos de aproximación del retardo de retorno de paquetes y el temporizador de retransmisión.

En este trabajo se han realizado tests sobre escenarios simples y complejos. Se han utilizado dos topologías. En primer lugar, se ha usado una topología dumbbell típica para simulaciones simples. Asimismo se ha utilizado un escenario complejo con una topología que reproduce de forma aproximada algunas redes reales con miles de nodos y patrones de tráfico complejos. El objetivo del uso del escenario complejo es generar tráfico de fondo con una dinámica compleja, alta variabilidad y propiedades estadísticas globales realistas.

Si bien se ha definido un escenario considerablemente complejo, muchos aspectos fundamentales de simulación de redes se encuentran aún en etapas iniciales de desarrollo desde nuestro punto de vista. En particular, no se tiene una respuesta clara a la cuestión fundamental de cuál debe ser la mínima escala de un escenario para que los resultados sean suficientemente aproximados a los que se obtendrían en un escenario que modelase por completo la Internet global o una parte significativa. Esta y otras cuestiones requieren la exploración de escenarios de simulación complejos como el desarrollado en esta tesis.

En este sentido, áreas de investigación tales como la definición de topologías, patrones de tráfico, métricas y criterios de evaluación están experimentando un considerable desarrollo recientemente. La generación de tráfico de fondo es un área en el que aún son necesarios desarrollos significativos. En particular, se tiene escaso conocimiento sobre el grado en que modelos de red simples pueden ser representativos del impacto sobre la dinámica del tráfico que resulta de las complejas topologías existentes en la Internet real [109].

Por tanto, se pueden considerar un gran número de aspectos de investigación futura sobre la línea abierta en el epígrafe 4.2 de esta tesis. No obstante, estos aspectos están más directamente relacionadas con desarrollos recientes o previsibles a corto y medio plazo en las áreas de simulación y emulación de redes, tales como métodos de generación de tráfico basados en mediciones, que puedan reproducir patrones estadísticos y características específicas de ciertas aplicaciones a partir de trazas de tráfico [108].

Hemos de hacer notar asimismo que el esquema de control de congestión de extremo a extremo propuesto en esta tesis se ha diseñado y estudiado en escenarios estándar. Por ello, un posible área de trabajo futuro conllevaría el estudio de mecanismos de control de tráfico en dos entornos de interés particular. Estos entornos comprenden, por una parte, redes de alta velocidad y grandes distancias con elevados anchos de banda disponible, y, por otra, redes con enlaces con porcentajes de pérdidas elevados, como pueden ser ciertas redes basadas en conexiones inalámbricas.

Otra área de investigación futura que consideramos de especial interés y como ampliación natural de la propuesta presentada en esta tesis consiste en la descripción de esquemas de control de congestión de extremo a extremo mediante mecanismos de ventana utilizando máquinas de estados difusos y cadenas de Markov.

Finalmente, cabe mencionar que independientemente de las prestaciones de la propuesta de control de tráfico de extremo a extremo presentada, la principal contribución de este trabajo consiste en la formulación de un esquema basado en reglas lingüísticas para realizar control de congestión de extremo a extremo basado en mecanismo de ventana.

4.3. Gestión activa de colas mediante sistemas de inferencia difusa

Un problema de investigación de importancia fundamental en las capas de transporte y red de la arquitectura de Internet es el desarrollo de mecanismos de control de tráfico que resulten adecuados a los requisitos de una creciente diversidad de tecnologías, aplicaciones y servicios. De manera más general, la dinámica del tráfico de internet es un tema de investigación de creciente complejidad [42, 98].

Los requisitos de calidad de servicio así como los patrones de tráfico de servicios y aplicaciones emergentes son difíciles de caracterizar y demandan profundos avances en los esquemas actuales de control de flujo y congestión. Debido a la naturaleza de estos problemas, su complejidad, la dificultad de solucionarlos de manera analítica, y al hecho de que la información disponible para estos mecanismos es incompleta e imprecisa, el empleo de sistemas inteligentes basados en lógica difusa y otras técnicas de soft computing es a priori una solución atractiva.

La dinámica resultante de las complejas interacciones entre mecanismos de control de extremo a extremo y controladores de routers es difícil de comprender y plantea numerosos retos. El diseño de nuevos algoritmos para uso general en Internet requiere la consideración de un conjunto de argumentos prácticos [41], principios de control de congestión [39] y métricas para la evaluación de esquemas de control de congestión [40].

En esta tesis se diseña, siguiendo una metodología a medida, un controlador difuso para gestión activa de colas para tráfico de tipo best effort, FAQMBestEffort. Aquí describiremos brevemente las características más generales. Se definen dos entradas, de tipos Te_i y Te_{i-1} , y una salida. La entrada e_i es una medida de la desviación entre el número de paquetes de la cola y un valor de referencia, normalizado entre 0 y 1. La entrada e_{i-1} mide la misma desviación en el intervalo de tiempo inmediatamente anterior, también normalizada entre 0 y 1. Para ambas entradas se definen siete etiquetas lingüísticas, que conforman una partición uniforme del dominio de las entradas. Por simplicidad, se emplean funciones de pertenencia triangulares.

La salida del sistema, p_i , se define como un valor de probabilidad de marcado del siguiente paquete disponible en la cola. En este caso, de manera equivalente a los esquemas de gestión activa de colas más aceptados actualmente, la acción de marcado implica descartar el paquete. Se definen asimismo siete términos lingüísticos para la salida, utilizándose funciones de pertenencia de tipo singleton. La superficie de control del sistema se muestra en la figura 13.

En el capítulo 4 de esta tesis se muestran los resultados de una evaluación comparativa del controlador FAQMBestEffort propuesto, considerándose un conjunto de características de los patrones de tráfico reales así como una topología compleja. Se argumenta que la definición de un conjunto completo de escenarios y métricas de evaluación para optimizar esquemas de gestión activa de colas es actualmente un problema para el que no existe una solución clara. Por ello, la optimización del controlador para escenarios específicos se ha dejado abierta como línea de investigación futura.

La sección 4.3 de esta tesis está centrada en la evaluación de un controlador simple que se puede considerar como una forma básica y general de una clase de controladores para gestión activa de colas. Se sugiere que, en lugar de métodos de optimización a posteriori, se

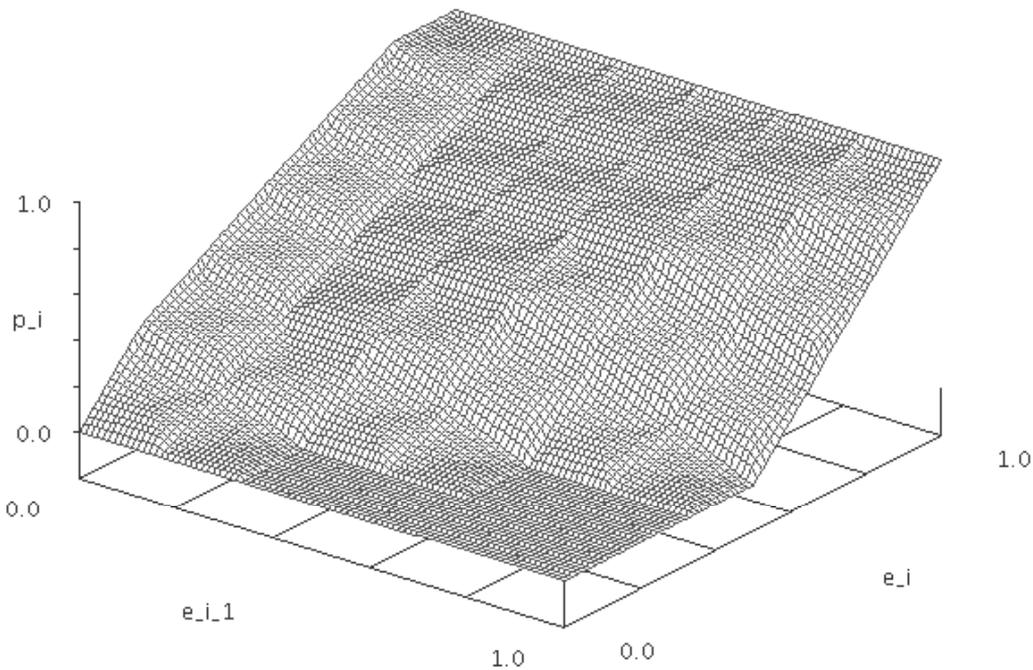


Figura 13: Superficie de control del sistema FAQMBestEffort.

exploren métodos en tiempo real tales como los utilizados por los sistemas neuro-difusos evolutivos. Sin embargo, esto requeriría un análisis en profundidad de las posibilidades de utilizar la información disponible en los routers para hacer evolucionar el sistemas de gestión activa de colas de modo que se optimicen las prestaciones en términos de medidas que no se conocen directamente.

Se explora asimismo las prestaciones de los sistema FAQMBestEffort, y RED mediante experimentos de simulación, confirmándose el comportamiento correcto de nuestra propuesta. Los resultados de tests realizados mediante emulación encajan con los obtenidos mediante simulación excepto en pequeñas variaciones numéricas debidas a detalles de implementación. Como conclusión general, FAQMBestEffort es más robusto frente a tráfico a ráfagas y proporciona mejores prestaciones que RED tanto para transferencias de grandes volúmenes de datos como para tráfico de tiempo real, mostrando mejor comportamiento en términos de la longitud y estabilidad de la cola de paquetes, utilización del ancho de banda e impacto sobre el retardo de extremo a extremo.

Hacemos notar que, aunque el sistema FAQMBestEffort se ha desarrollado para tráfico best effort, su mayor grado de robustez y velocidad de respuesta ante ráfagas de tráfico conlleva una mejora de las prestaciones de extremo a extremo que experimentarían las aplicaciones y servicios con restricciones temporales. FAQMBestEffort es por tanto una solución de compromiso práctica para los routers desplegados actualmente.

4.4. Conclusiones

El capítulo 4, junto con el capítulo anterior y el quinto describen las contribuciones principales de esta tesis. En este capítulo 4 se presentan sendas propuestas para realizar control de congestión de extremo a extremo así como en routers mediante sistemas de inferencia difusa.

En primer lugar se desarrolla un escenario de simulación complejo y realista sobre la base de varios estudios recientes sobre topologías y medidas de tráfico. A continuación, se diseña y prueba un nuevo esquema para realizar control de congestión de extremo a extremo de manera compatible con flujos TCP estándar mediante sistemas difusos basados en reglas. Este nuevo enfoque se prueba y compara frente a las variantes TCP SACK y HighSpeed TCP del protocolo TCP. Los resultados de los experimentos realizados mediante simulación e implementación muestran que la propuesta presentada aquí puede mejorar las prestaciones globales bajo un amplio conjunto de condiciones.

A continuación, se presenta un esquema para gestión activa de colas en routers basado en control difuso. En contraposición a otros controladores propuestos previamente, en este trabajo hemos centrado nuestra atención en la posibilidad de evaluar las prestaciones en escenarios realistas y variados. Este esquema de control se ha probado y comparado frente a los esquemas RED. Las simulaciones realizadas ponen de manifiesto que el controlador propuesto tiene mejores prestaciones bajo condiciones comunes y puede alcanzar prestaciones satisfactorias para un mayor rango de condiciones que los otros esquemas comparados. Estos resultados de simulación se han confirmado mediante una implementación experimental realizada en un escenario emulado.

5. Plataforma abierta de desarrollo basada en FPGA para sistemas de inferencia difusa

Como último bloque de este trabajo, se analizan las arquitecturas y los métodos de diseño propuestos hasta la fecha para routers de altas prestaciones. El objetivo de este análisis es identificar las posibilidades reales de integración de implementaciones hardware de sistemas difusos en las unidades de procesamiento de routers, tanto en arquitecturas actuales como en aquellas previsibles a medio plazo.

A lo largo de más de una década, se han desarrollado estrategias y metodologías para el desarrollo de implementaciones hardware de controladores basados en lógica difusa. Hasta la fecha, casi todo lo desarrollado se ha enfocado a diversas aplicaciones en el sector industrial [13]. Considerando los requisitos específicos así como el elevado coste y complejidad de los routers desplegados en la Internet actual, aquí proponemos una plataforma de desarrollo que facilita el diseño de sistemas de análisis y control de tráfico de red.

Como consecuencia del análisis realizado, se propone una plataforma abierta basada en dispositivos FPGA que facilita el diseño modular de componentes basados en lógica difusa integrables en equipos de comunicaciones complejos. En particular, la plataforma se ha empleado para desarrollar sistemas de análisis y control inteligente que alcanzan las tasas de procesamiento necesarias para operar en tiempo real en las arquitecturas actuales [73]. La plataforma utiliza el entorno Xfuzzy para el desarrollo de sistemas de inferencia difusa, propuesto anteriormente [85], así como una arquitectura específica optimizada para su implementación digital [7].

Las técnicas de inteligencia computacional, y, en particular, las basadas en lógica difusa están ganando aceptación como herramientas para modelado, análisis y control de tráfico de redes [80]. Sin embargo, la realización de implementaciones hardware de estos sistemas que puedan operar en tiempo real en equipos de comunicaciones de altas prestaciones así como en otros campos de aplicación altamente restrictivos es aún un problema abierto.

Las arquitecturas de routers actuales plantean dos retos fundamentales para el diseño de nuevos mecanismos, sus exigencias en términos de escalabilidad y flexibilidad. En esta tesis se desarrolla una plataforma y una metodología para desarrollar sistemas difusos que no sólo satisfacen los requisitos operacionales sino que además abordan los retos planteados actualmente por las arquitecturas de routers.

Para ello, se emplea una placa de desarrollo con interfaz PCI/PCI-E que soporta una plataforma abierta compuesta por herramientas de CAD así como módulos IP. Para el proceso de desarrollo establecemos una metodología y una cadena de herramientas de CAD que cubre desde la especificación inicial mediante un lenguaje de alto nivel hasta la implementación en dispositivos FPGA. Se implementan como sistemas-en-un-chip-programable, o SoPC, módulos de inferencia difusa compatibles con el bus PCI utilizando la arquitectura de interconexión abierta WISHBONE. Se analizan asimismo los resultados del diseño e implementación de sistemas difusos de análisis y control de tráfico, mostrándose que satisfacen los requisitos operacionales y arquitecturales de los routers de altas prestaciones actuales y previsibles para un futuro próximo.

En el apartado 5.1 de esta tesis describimos las prácticas y tendencias actuales en torno a las arquitecturas y el diseño de routers, discutiéndose cómo influyen en la integración

de mecanismos difusos. Se describen asimismo algunas arquitecturas y plataformas de investigación desarrolladas en el campo de las redes de ordenadores, lo cual nos sirve para motivar el desarrollo de la plataforma que es objeto del capítulo 5. En el apartado 5.2 de esta tesis se describen brevemente las tecnologías y arquitecturas propuestas para obtener implementaciones hardware eficientes de sistemas difusos, prestándose especial atención a la arquitectura digital empleada en los siguientes apartados. La sección 5.3 describe una plataforma basada en dispositivos FPGA para desarrollo rápido de prototipos de sistemas de inferencia difusa con aplicaciones en redes de comunicaciones, incluyendo funciones de análisis y control de tráfico. La plataforma se aplica en la implementación de algunos de los sistemas difusos descritos en los capítulos anteriores.

En los siguientes epígrafes resumimos el análisis de arquitecturas de routers realizado en el apartado 5.1 de esta tesis, así como las características generales de la plataforma presentada en el apartado 5.3 del texto completo de esta tesis.

5.1. Arquitecturas de routers

En el capítulo 5 de esta tesis se diseñan sistemas de inferencia difusa para tareas de control de tráfico que han de implementarse en routers de Internet. En lo que sigue resumimos las posibles opciones para integrar bloques de inferencia difusa en las arquitecturas de los routers de generaciones actuales y futuras.

Las tendencias tecnológicas que afectan al diseño de los routers de redes troncales de Internet y a los equipos de comunicaciones de altas prestaciones en general, véase la figura 14, conllevan restricciones estrictas especialmente en lo referente a las tasas de procesamiento de paquetes que se han de alcanzar. Durante los últimos años, el tráfico global de Internet ha crecido aproximadamente un 80 % anual, lo que se traduce en un crecimiento igual o incluso superior del volumen de tráfico en los enlaces troncales.

En general, el volumen del tráfico de la red crece a una velocidad que sobrepasa el avance de las tecnologías VLSI. En este contexto, han surgido dos restricciones fundamentales respecto a la escalabilidad y la flexibilidad de los componentes de routers de altas prestaciones. En primer lugar, las unidades de procesamiento han de ser capaces de procesar hasta millones de paquetes por segundo (Mpps o Mp/s). En segundo lugar, las implementaciones de estas unidades deben ser flexibles y reconfigurables, de modo que sea posible afrontar la creciente diversidad de protocolos y tecnologías que han de ser soportados [50, 107].

Durante los últimos años, desde los sectores académico e industrial se está promoviendo el desarrollo y despliegue de routers con diseño modular y distribuido. En estos diseños, los routers están compuestos por módulos funcionales que se pueden asignar a diferentes elementos de procesamiento. La comunicación entre los módulos se realiza mediante interfaces bien definidas y abiertas a través de una red interna [50]. El hardware de los sistemas de comunicaciones de altas prestaciones se ha desarrollado tradicionalmente con procedimientos a medida y sin estructura clara. No obstante, en la práctica la mayoría de los fabricantes emplean arquitecturas con cierto grado de reconfigurabilidad, y se persigue la implantación generalizada de metodologías de diseño que faciliten el proceso de desarrollo.

De manera general, los routers se pueden clasificar en tres clases en función del nivel en el que se despliegan dentro de Internet. Estas clases corresponden a los routers de acce-

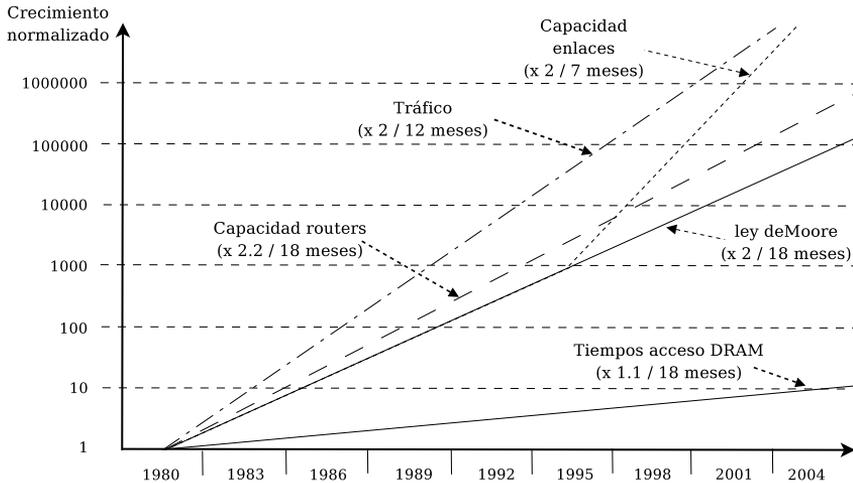


Figura 14: Tendencias tecnológicas en equipos de comunicaciones de altas prestaciones [50, 51, 63].

so, los de campus u organizaciones, y los routers centrales. En esta tesis nos centramos en los routers de la última clase, aquellos diseñados para enlaces troncales. En muchos aspectos, los routers de altas prestaciones actuales se asemejan a supercomputadores, teniendo restricciones operacionales muy estrictas.

Los routers conforman un todo unificado enlazando redes de la Internet global. Si bien la función principal de un router es retransmitir paquetes procedentes de un conjunto de enlaces entrantes a otro conjunto de enlaces salientes, esto no excluye que tengan que implementar algoritmos de encaminamiento distribuidos y complejos, gestionar diferentes tecnologías de enlace y proporcionar soporte a tareas de ingeniería de tráfico, servicios diferenciados así como esquemas de calidad de servicio de rápida evolución.

La arquitectura de los routers de Internet ha evolucionado rápidamente desde que aparecieron las primeras implementaciones [47, 99, 21, 62]. Esta evolución ha venido condicionada por un conjunto de factores tecnológicos y requisitos funcionales. Por una parte, la divergencia en el incremento de prestaciones experimentada por los diferentes componentes de un router, tales como elementos de memoria, enlaces de interconexión, dispositivos programables y procesadores, desafía el diseño de routers. Por otra parte, nuevos requisitos funcionales han surgido junto con las nuevas aplicaciones, servicios y tecnologías que se han desplegado durante los últimos años. Como consecuencia, afrontar estos retos de diseño requiere considerables esfuerzos de investigación y desarrollo.

En esta tesis se discuten las características distintivas de los routers de altas prestaciones así como la evolución esperada para estos sistemas en el futuro a medio plazo. La arquitectura de los equipos de redes de ordenadores ha cambiado dramáticamente durante las dos últimas décadas [24, 28, 66, 5]. Desde una perspectiva histórica, las arquitecturas existentes se pueden clasificar en tres generaciones en función del grado de centralización:

- Primera generación, correspondiente a los últimos años de la década de 1980 y los primeros años de la década de 1990. En estos routers, el software se ejecuta sobre un

procesador estándar, siendo común encontrar routers IP realizados añadiendo software especializado a computadores convencionales.

- Segunda generación, correspondiente a la década de 1990, en los que las funciones de clasificación así como algunas otras se ejecutan por medio de hardware de propósito específico. Asimismo, una red de conmutación de alta velocidad reemplaza al bus compartido habitualmente utilizado en los routers de primera generación.
- Tercera generación, surgida a finales de la década de 1990, y basada en un diseño completamente descentralizado. En esta generación se utilizan circuitos integrados de aplicación específica, o ASIC, además de un procesador dedicado para cada interfaz de red, que gestiona la retransmisión acelerada de paquetes, y mediante el cual se libera a la unidad de procesamiento de numerosas tareas.

5.2. Plataforma de desarrollo

En esta tesis se citan diversas referencias a la literatura que muestran que los sistemas difusos pueden aportar soluciones a los problemas actuales en torno al control de tráfico en Internet. En efecto, las técnicas de soft computing, y los sistemas de inferencia difusa en particular, están ganando aceptación como herramientas de modelado, análisis y control de tráfico. Los sistemas de inferencia difusa encuentran aplicaciones en diversas áreas, tales como control de tráfico en routers [32, 121], control de admisión [67], soporte a servicios diferenciados dentro de la arquitectura DiffServ [119], soporte a políticas y evaluación de calidad de servicio [97], medida de tráfico, análisis y monitorización [77], ahorro de energía en redes inalámbricas, así como control de tráfico de extremo a extremo [81] y control de tráfico de extremo a extremo en redes inalámbricas [31].

No obstante, mientras que existen numerosas aplicaciones industriales de sistemas difusos en una gran variedad de campos, los sistemas difusos para control de congestión no han llegado a desplegarse en aplicaciones reales. En particular, a pesar de las buenas prestaciones de los mecanismos basados en lógica difusa para análisis y control de tráfico, existe una carencia de arquitecturas y procedimientos de diseño para implementarlos de manera sistemática, al tiempo que se afrontan los retos actuales en el diseño de routers de altas prestaciones. Como resultado, aunque se han publicado importantes resultados sobre diversas aplicaciones de sistemas difusos en comunicaciones y redes de ordenadores desde hace más de una década [45], el despliegue de estos sistemas en el mundo real es aún un problema abierto.

En la Internet actual, las velocidades de los enlaces y con ello las tasas de procesamiento de paquetes requeridas están creciendo muy rápidamente. El ritmo al que crece la velocidad de los elementos de memoria así como otros componentes de las unidades de procesamiento de los routers es significativamente menor. Esto es, las tasas de aceleración de dos componentes tecnológicos clave divergen de manera creciente, esperándose que estas tendencias continúen. Como consecuencia, muchos elementos de almacenamiento y procesamiento de datos, o motores de procesamiento, de los routers actuales se implementan mediante motores específicos que utilizan arquitecturas hardware especializadas [28, 29, 30, 50, 25].

En este contexto, son necesarias arquitecturas hardware específicas que puedan afrontar las siempre crecientes tasas de procesamiento de paquetes en sistemas de medida, análisis y control de tráfico. Se han diseñado arquitecturas específicas para acelerar sistemas de medida [34], analizar flujos de red en tiempo real [120], e implementar ciertos motores de procesamiento comunes [28, 29, 30]. En particular, la implementación de esquemas de gestión activa de colas requiere una tasa de control de paquetes muy cercana a la máxima tasa de procesamiento de paquetes alcanzable. Esto es un requisito importante que se ve agravado por dos factores coincidentes:

- El tráfico de las redes de conmutación de paquetes es inherentemente irregular y las ráfagas de paquetes son frecuentes. Estas ráfagas, formadas frecuentemente por un gran número de paquetes que llegan a interfaces de entrada a la máxima velocidad del enlace, son comunes en un amplio espectro de escenarios de red [90].
- La longitud máxima de las colas de paquetes es por lo general reducida, alrededor de unas pocas decenas de paquetes [2].

Por tanto, los controladores de tráfico deben tener una granularidad de procesamiento de paquetes fina con objeto de tratar correctamente las ráfagas de paquetes y gestionar convenientemente colas de reducida longitud.

En el capítulo 5 de esta tesis se analiza la implementación e integración de sistemas de inferencia difusa en routers. Aunque las soluciones software serían obviamente más flexibles que las implementaciones hardware, es fácil comprobar que las primeras no pueden lograr las tasas de inferencia del orden del millón de inferencias difusas por segundo, o MFLIPS, de decenas de MFLIPS y superiores incluso a pesar de la utilización de CPUs de propósito general de altas prestaciones. Esto se muestra asimismo de forma breve en el apéndice B de la tesis. Además, si consideramos las tendencias existentes en diversos factores tecnológicos y que no se prevé que éstas cambien en un futuro previsible, se puede concluir que las tasas de inferencia necesarias seguirán creciendo a un ritmo superior al que es posible mediante unidades de procesamiento basadas en software.

Además, incluso si las implementaciones software fuesen suficientemente rápidas, el uso de una CPU de propósito general dedicada para implementar esquemas de gestión activa de colas en routers no es una opción factible. En las arquitecturas actuales, los bloques funcionales de análisis y control de tráfico han de replicarse para cada puerto de entrada en cada una de las llamadas colas de salida virtuales, o VOQ, tantas veces como puertos de salida estén activos. Como consecuencia, el número de VOQs y bloques de control asociados puede estar alrededor de varias decenas e incluso llegar a varias centenas en algunos casos. Este hecho impone fuertes limitaciones de consumo y coste sobre las implementaciones de los mecanismos de análisis y control de tráfico. Por ello, es muy improbable que los fabricantes pudieran asumir el coste de incluir un número considerable de CPUs de altas prestaciones por unidad.

En lugar de utilizar implementaciones software sobre hardware de propósito general, con el consecuente coste y limitación de prestaciones, las arquitecturas de routers actuales emplean motores de procesamiento hardware con arquitecturas especializadas para implementar procesos críticos e intensivos computacionalmente. Este es el caso por ejemplo de las memorias TCAM, empleadas para almacenar y extraer información de encaminamiento.

Aunque en el pasado se han utilizado implementaciones puramente ASIC, actualmente muchos de estos motores de procesamiento se implementan utilizando uno o varios dispositivos FPGA para cada placa de conexión, con objeto de incrementar la programabilidad y hacer más flexible el diseño. Algunas tareas específicas se implementan mediante subsistemas especializados, mientras que las unidades de procesamiento de propósito general realizan diversas funciones de coordinación así como funciones de gestión de alto nivel.

Por tanto, es necesario emplear esquemas de implementación hardware eficiente de sistemas difusos para cumplir los requisitos operacionales comunes de los mecanismos de análisis y control de tráfico de los routers actuales. Aunque los sistemas de inferencia difusa requieren normalmente una considerable capacidad computacional, sus prestaciones se pueden mejorar significativamente mediante implementaciones hardware basadas en arquitecturas optimizadas que explotan el paralelismo inherente de los sistemas de inferencia difusa, o simplifican el proceso de inferencia.

Como se describe en el epígrafe 5.2 de esta tesis, se han propuesto numerosas soluciones para la implementación hardware de sistemas de inferencia difusa dentro del campo de la microelectrónica [7]. Por ello, es factible aplicar este tipo de sistemas al procesamiento de volúmenes masivos de tráfico en tiempo real.

En esta tesis se introduce una plataforma abierta basada en dispositivos FPGA para el desarrollo de sistemas difusos como componentes modulares de sistemas complejos [74, 73]. La plataforma se ha aplicado a sistemas de comunicaciones, y, en particular, se ha empleado satisfactoriamente para desarrollar sistemas inteligentes de análisis y control de tráfico que pueden operar en tiempo real en routers actuales de Internet. La plataforma se ha desarrollado con un doble objetivo:

- Posibilitar la implementación automática y eficiente, en términos de prestaciones y esfuerzo de desarrollo, de un conjunto de sistemas de inferencia difusa propuestos a lo largo de los últimos años.
- Promover la investigación sobre soluciones a problemas de análisis y control de tráfico basadas en lógica difusa. Esta es una consecuencia que cabe esperar de la disponibilidad de una plataforma que permite la validación de prototipos hardware utilizando equipos de bajo coste. Así, se facilita la realización de pruebas en instalaciones de red para la realización de experimentos disruptivos, basadas en virtualización de recursos de la red. Se espera que estas instalaciones estén ampliamente disponibles en un futuro próximo.

A lo largo de más de una década, se han propuesto y aplicado estrategias y metodologías para el desarrollo de controladores basados en lógica difusa. Hasta el presente, la mayoría de los trabajos en este área han estado enfocados a aplicaciones industriales [13] y, más recientemente, a otras áreas tales como procesamiento de señal en general y procesamiento de imágenes en particular.

A la hora de evaluar sistemas de análisis y control de tráfico cuyas prestaciones dependen de la naturaleza del tráfico, sería deseable desplegar estos sistemas en redes reales y estudiar sus prestaciones. Sin embargo, esto es en general imposible, especialmente en el

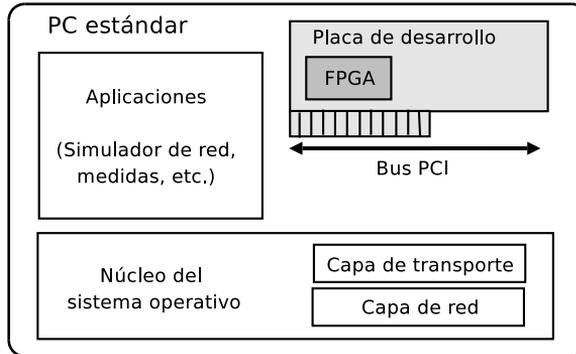


Figura 15: Esquema de la plataforma de desarrollo de prototipos.

caso de los esquemas de gestión activa de colas. Por ello, la disponibilidad de medios flexibles para evaluar este tipo de sistemas con diferentes patrones de tráfico, ya sea mediante simulación, emulación o implementación, es un aspecto clave.

Considerando todos estos requisitos específicos así como el alto coste y complejidad de los routers de altas prestaciones desplegados actualmente en Internet, en esta tesis hemos definido una plataforma de desarrollo flexible para realizar prototipos de sistemas de análisis y control de tráfico. La plataforma se ha definido de tal manera que los sistemas de inferencia difusa quedan integrados como módulos independientes en sistemas de redes de ordenadores complejos.

Algunos requisitos adicionales tomados en consideración incluyen la facilidad de integración en arquitecturas de routers actuales [63], la flexibilidad de los sistemas, y la escalabilidad de sus prestaciones hasta los requerimientos más elevados de las tecnologías de red actuales y previsibles. La plataforma proporciona un conjunto completo de herramientas y un entorno para facilitar el desarrollo y la validación de prototipos de sistemas difusos.

La arquitectura de la plataforma, esquematizada en la figura 15, se basa en un ordenador personal común equipado con una placa de desarrollo sobre dispositivos FPGA con interfaz PCI, con lo que se tiene una solución flexible y de bajo coste, sin requisitos de hardware específicos, que puede emular el comportamiento de equipos de comunicaciones complejos y de elevado coste. Así, es posible realizar experimentos de validación por medio de prototipos utilizando hardware de bajo coste.

Obviamente, estos prototipos estarán limitados en prestaciones y probablemente no puedan alcanzar una capacidad de transmisión de tráfico real de alrededor o mayor que el Gb/s. Esto es sin embargo una consecuencia de las limitaciones del ancho de banda de conmutación total alcanzable mediante una arquitectura de ordenador personal común, sin que por ello quede excluida la posibilidad de validar los módulos de inferencia difusa a velocidades superiores.

De manera general, se pueden distinguir dos bloques funcionales principales para la implementación de sistemas difusos: aquellos directamente relacionados con el proceso de inferencia y aquellos que se pueden clasificar como funciones auxiliares, tales como inicio, temporización, pre y postprocesado, etc. [13]. Para la implementación de prototipos de sistemas difusos aplicados al análisis y control de tráfico se ha utilizado el siguiente modelo:

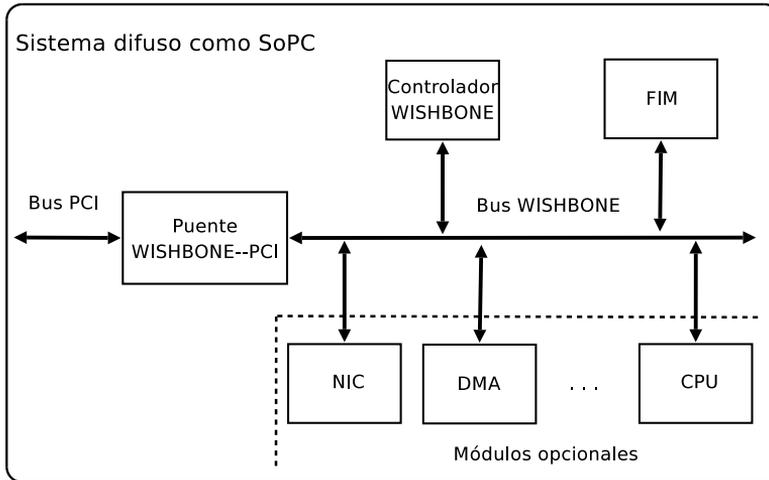


Figura 16: SoPC difuso como dispositivo PCI.

- Dado que los módulos de inferencia difusa (FIM) son el cuello de botella potencial del sistema, los FIM se implementan sobre dispositivos FPGA, describiéndose en VHDL de acuerdo a una arquitectura de procesamiento específico [7] adaptada para realizar inferencias de manera rápida y eficiente. La metodología y las herramientas empleadas para el desarrollo de los FIM se describe detalladamente en el capítulo 5 de esta tesis.
- En la configuración básica de la plataforma, todas las funciones auxiliares se implementan mediante software, que puede ejecutarse sobre el sistema operativo del PC así como sobre componentes opcionales implementados sobre el dispositivo FPGA de la placa de desarrollo.

Se ha definido una arquitectura abierta y flexible para implementar sistemas difusos sobre el dispositivo FPGA. En esta arquitectura, mostrada en la figura 16, los módulos FIM se integran como subsistemas de un sistema digital de inferencia difusa potencialmente complejo.

La interconexión entre el ordenador personal y el módulo FIM se realiza por medio de un bus PCI/PCI-E estándar. El bus interno del sistema difuso digital cumple la especificación del bus lógico WISHBONE [49], de dominio público. WISHBONE es una arquitectura de interconexión para sistemas-en-un-chip, desarrollada para módulos portables, que puede conectar un número variable de componentes.

Todas las interfaces de nivel superior son compatibles con la especificación WISHBONE. Tanto el controlador de bus WISHBONE como el puente PCI-WISHBONE [33] han sido desarrollados bajo licencias de libre distribución por la organización OpenCores [87] y otras entidades. Mientras que el puente WISHBONE-PCI, con soporte para el estándar PCI 2.2, y el controlador WISHBONE Conbus [19]) están implementados en lenguaje Verilog, proporcionados por la organización OpenCores, el módulo FIM se implementa en VHDL.

Los sistemas WISHBONE se pueden interconectar fácilmente con otros estándares de buses para sistemas-en-un-chip, tales como OPB [106], mediante un puente WISHBONE-

OPB. En su configuración básica, el sistema consta de tres módulos: el FIM, como dispositivo esclavo del bus WISHBONE, el controlador WISHBONE, y el puente PCI-WISHBONE, como dispositivo maestro. El módulo de control WISHBONE empleado permite la interconexión de hasta 8 dispositivos de tipo maestro y otros tantos de tipo esclavo.

Así, las tareas software se pueden definir utilizando lenguajes de programación comunes y se pueden ejecutar en el procesador de propósito general del PC, además de la posible utilización de unidades de procesamiento específicas implementadas en el dispositivo FPGA. Por ejemplo, la configuración descrita deja abierta la posibilidad de implementar aplicaciones de análisis de tráfico incorporando un módulo OpenRisc, procesador para el cual está disponible entre otras opciones el sistema operativo GNU/Linux. Por su parte, la interfaz PCI de los prototipos facilita la integración en arquitecturas de routers de los principales fabricantes [63, 47].

En las arquitecturas de routers desplegadas actualmente en Internet [51, 63], los sistemas de análisis y control de tráfico basados en lógica difusa se podrían integrar directamente como motores de procesamiento, ya sea en la unidad de procesamiento de red, NPU, o en las tarjetas de entrada y salida, en función de la arquitectura de calidad de servicio implementada en el router. Además, en aquellos casos en los que la placa de desarrollo incluye una interfaz de red, como en el caso de la placa utilizada para nuestra implementación, es posible implementar una aplicación completa de análisis de tráfico como sistema-en-un-chip-programable, SoPC, completamente autónomo sobre el dispositivo FPGA.

Como se detalla en el capítulo 5 de esta tesis, la configuración presentada permite implementar unidades de procesamiento para inferencia difusa que se pueden aplicar de manera genérica a problemas de análisis y control de tráfico. Además, se pueden incorporar módulos adicionales, tales como módulos de control de interfaces de red, dispositivos de acceso directo a memoria o unidades de procesamiento general, con objeto de desarrollar unidades de procesamiento difuso o completar estas con técnicas complementarias.

Dentro de la plataforma presentada, se sigue un flujo de diseño completamente automatizado. El flujo de diseño cubre desde la especificación inicial en un lenguaje de alto nivel hasta la implementación hardware mediante dispositivos FPGA, y utiliza herramientas incluidas en el entorno de desarrollo Xfuzzy así como herramientas pertenecientes al entorno ISE de Xilinx [114].

Las implementaciones hardware de sistemas de inferencia difusa analizadas en el capítulo 5 de esta tesis se han realizado utilizando dos dispositivos FPGA de Xilinx. El primer dispositivo es una FPGA Spartan-3 xc3s1500-fg456-5 (con 1.5 millones de puertas equivalentes) incluida en una placa de desarrollo ADS-XLX-SP3-EVL1500 de AvNet con interfaz PCI 2.0 estándar. Se trata de un dispositivo de coste bajo para esta aplicación, seleccionado con objeto de evaluar si el enfoque propuesto en esta tesis es factible con recursos limitados. El segundo dispositivo, con prestaciones más próximas a las de los dispositivos lógicos programables empleado actualmente en equipos de comunicaciones de altas prestaciones, es una FPGA Virtex-5 de Xilinx, modelo XC5VLX50T-1FF1136-1C-ES, incluida en una placa de desarrollo LXT FPGA ML505 de Xilinx, con interfaz de conexión PCI-E.

El flujo de diseño detallado y las herramientas empleadas se describen en el capítulo 5 de esta tesis. En el presente resumen mostramos en la tabla 3 las características generales de los sistemas. Los sistemas recogidos en la tabla implementan diversos mecanismos de

Sistemas	Entradas	Términos lingüísticos	Reglas
FAQMBestEffort	2	7,7,7	37
TCPSS	6	5,5,5,5,5,5	24
DSSelect	2	5,5,2	17
AQMDSAF	2	3,3,4	7
RxBufferSize	2	5,5,5	25
RTperf	4	5,5,5,5,5	27

Tabla 3: Sistemas de inferencia difusa y su complejidad en términos de entradas, términos lingüísticos y reglas.

análisis y control de tráfico FAQMBestEffort es el sistema de inferencia difusa para gestión activa de colas descrito en el apartado 4.3 de esta tesis. TCPSS implementa uno de los tres sistemas de inferencia que componen el sistema de control de congestión de extremo a extremo desarrollado en el epígrafe 4.2. DSSelect y AQMDSAF son controladores de tráfico adaptados a redes con soporte de calidad de servicio dentro de la arquitectura DiffServ. RxBufferSize implementa un sistema difuso para inferir el tamaño adecuado de un buffer de datos dinámico en función de las condiciones de la red, tomando como entradas el retardo unidireccional y el porcentaje de pérdida de paquetes. Por último, RTperf es un sistema de inferencia difusa para evaluación de prestaciones adaptado a aplicaciones y servicios de tiempo real.

Los resultados de implementación, analizados en detalle en el capítulo 5, se pueden resumir en las figuras 17 y 18. La figura 17 muestra un resumen de los resultados de implementación de los sistemas mencionados utilizando una precisión de 8 bits para las entradas, salidas y funciones de pertenencia. Esta configuración es suficiente, no habiéndose encontrado diferencias en la práctica en comparación con otras configuraciones de 16 bits.

La figura 18 el mismo resumen para el caso del dispositivo Virtex-5. Al igual que en el caso anterior, se usa una configuración de 8 bits para las entradas, salidas y funciones de pertenencia.

De los resultados mostrados aquí y los que se analizan con mayor detalle en el capítulo 5 de esta tesis se puede concluir que la variabilidad de la velocidad de inferencia con respecto a la complejidad de los sistemas, medida en términos de entradas, reglas, funciones de pertenencia y precisión, queda dentro de los márgenes aceptables para la arquitectura y las tecnologías utilizadas. Así, es posible alcanzar altas tasas de inferencia incluso para los sistemas más complejos considerados.

En el capítulo 5 de esta tesis se analiza asimismo el consumo estático y dinámico de estos sistemas, mostrándose que ambos quedan por debajo de los márgenes de consumo aceptables. En cuanto a la tasa de inferencia, los prototipos implementados sobre el dispositivo Spartan-3 pueden alcanzar por encima de los 60 millones de inferencias difusas por segundo, o MFLIPS. Como se detalla en el texto completo de esta tesis, los controladores de colas de paquetes deben tener una granularidad de procesamiento fina, con objeto de procesar adecuadamente ráfagas de paquetes y colas de paquetes de reducido tamaño.

La granularidad de procesamiento más fina es posible sólo si la tasa de procesamiento de paquetes es igual o mayor que el máximo número de paquetes por segundo aceptado por las colas.

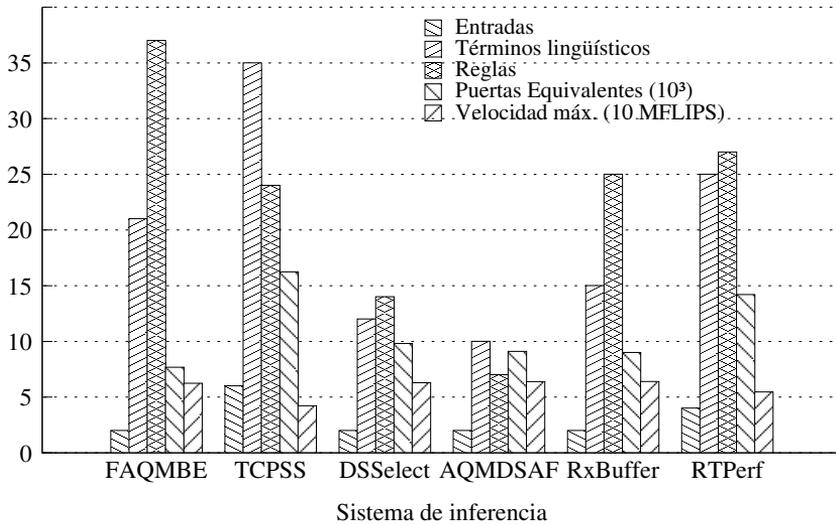


Figura 17: Resultados de implementación sobre el dispositivo FPGA Spartan-3. La complejidad de los sistemas se expresa en términos de número de entradas, términos lingüísticos y reglas. La complejidad de las implementaciones resultantes viene dada en términos de puertas equivalentes y tasa máxima de inferencia por segundo.

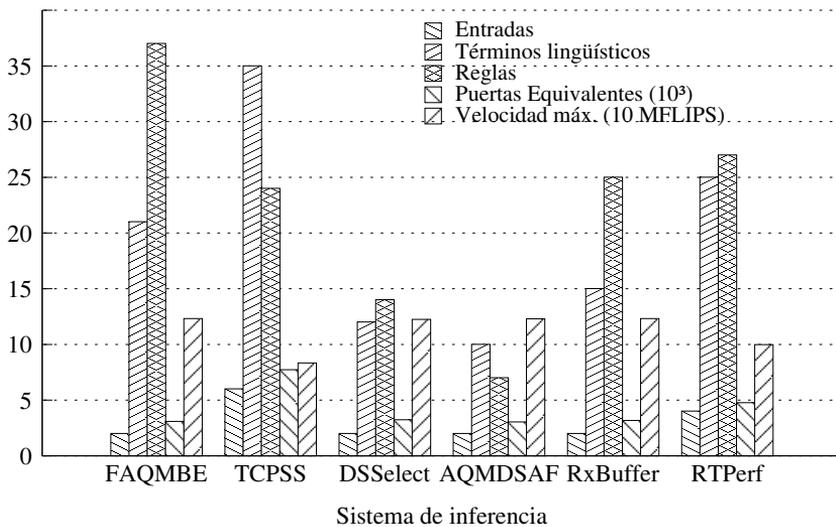


Figura 18: Resultados de implementación sobre el dispositivo FPGA Virtex-5. La complejidad de los sistemas se expresa en términos de número de entradas, términos lingüísticos y reglas. La complejidad de las implementaciones resultantes viene dada en términos de puertas equivalentes y tasa máxima de inferencia por segundo.

En los routers de Cisco de las series 12000 y CRS, así como los routers de Juniper de las series M y T, las tasas máximas de procesamiento de paquetes son de 25 Mp/s, 40 Mp/s, 24 Mp/s y 60 Mp/s, respectivamente, por cada cola de salida de las interfaces de conexión. Por tanto, incluso implementaciones prototipo realizadas utilizando dispositivos FPGA de bajo coste pueden alcanzar las tasas de inferencia necesarias para lograr la máxima granularidad de procesamiento de paquetes de las familias de routers de altas prestaciones actuales.

5.3. Conclusiones

El quinto capítulo de esta tesis describe la tercera contribución principal, que complementa las contribuciones descritas en los dos capítulos anteriores. En este capítulo se analizan los problemas más importantes que condicionan la implementación de sistemas de inferencia difusa para análisis y control de tráfico en routers, considerando las restricciones y tendencias tecnológicas actuales así como los factores arquitecturales. A continuación, se propone una plataforma que satisface tales restricciones para el desarrollo de implementaciones de sistemas difusos para análisis y control de tráfico basadas en dispositivos FPGA.

En primer lugar se describen los aspectos arquitecturales que restringen las implementaciones de métodos difusos en la práctica. Se diseña a continuación una arquitectura modular compatible con buses PCI/PCI-E para la integración de subsistemas difusos en arquitecturas de routers actuales. Se describe asimismo una plataforma abierta que facilita el desarrollo de sistemas difusos y su implementación como SoPC sobre dispositivos FPGA.

La plataforma integra herramientas abiertas así como módulos hardware. Además, se sigue una metodología de desarrollo junto con un flujo de diseño sistemático. Las herramientas incluidas en el entorno Xfuzzy automatizan el desarrollo desde la especificación inicial hasta la generación de VHDL sintetizable. En este esquema, los módulos de inferencia difusa se integran en una arquitectura SoPC compuesta de módulos abiertos que es adecuada para desarrollar sistemas difusos aplicados a redes de ordenadores entre otras posibles áreas. Estos sistemas se pueden integrar como unidades de procesamiento en arquitecturas de routers actuales.

A continuación, se analizan las prestaciones de estos sistemas en cuanto a velocidad de inferencia y consumo de recursos, con especial énfasis en la velocidad de inferencia. Se muestran resultados que avalan que los sistemas implementados satisfacen las restricciones operacionales características de los routers de altas prestaciones en la actualidad y en el futuro previsible, tanto en términos de velocidad de inferencia como de consumo de recursos.

Incluso algunos prototipos implementados utilizando dispositivos FPGA de bajo coste pueden proporcionar las velocidades de inferencia requeridas con un reducido coste en términos de área y energía. Además, los prototipos se han diseñado para que sean fácilmente integrables en las arquitecturas de routers actualmente desplegadas en Internet.

La plataforma de desarrollo presentada en este capítulo abre el camino para el futuro desarrollo de controladores de tráfico inteligentes y eficientes. Asimismo, esperamos que la disponibilidad de esta plataforma promueva el desarrollo de sistemas difusos en campos adicionales en los que son necesarios sistemas flexibles y eficientes, tales como identificación, filtrado y clasificación de paquetes y flujos, entre muchos otros.

6. Conclusiones

Como conclusión global, en esta tesis hemos aportado métodos y herramientas basados en inteligencia computacional para abordar un conjunto de problemas actuales en los campos de medida, modelado y control de tráfico de red. Además de desarrollar métodos, se han tenido en consideración un conjunto de aspectos de implementación práctica que condicionan la adopción de nuevos métodos y mecanismos para análisis y control de tráfico.

De manera más específica, se han abordado los siguientes problemas: modelado y predicción de la carga de tráfico a diferentes escalas temporales, generación de resúmenes e inducción de reglas de asociación a partir de registros de flujos de red, gestión activa de colas en routers, control de congestión de extremo a extremo, así como la implementación eficiente de sistemas de inferencia difusa dentro de las arquitecturas de routers actuales y previsibles para el futuro.

En general, como paso previo al desarrollo de métodos en las áreas mencionadas, se han considerado diversos factores determinantes y problemas prácticos. Así, hemos analizado los siguientes aspectos:

- Predecibilidad de la carga de tráfico a diferentes escalas temporales para un amplio conjunto de trazas de tráfico. Mientras que para algunos escenarios y escalas de tiempo sólo es posible realizar predicciones satisfactorias a muy corto plazo y con precisión limitada, para otros escenarios y escalas temporales se ha mostrado que es posible alcanzar una considerable precisión en las predicciones a corto y medio plazo.
- Cómo se puede utilizar el conocimiento previo para definir etiquetas lingüísticas para atributos de flujos de red y cómo estas se corresponden con medidas realizadas en redes en producción. Mediante técnicas de aprendizaje no supervisado, se ha mostrado la alta correspondencia existente entre las etiquetas lingüísticas a priori propuestas y los valores medidos en redes reales.
- Escenarios de red complejos para simular mecanismos de control bajo condiciones realistas. Se ha mostrado que estos escenarios son sustancialmente diferentes respecto a los que habitualmente se utilizan en la literatura para diseñar y optimizar métodos de control de tráfico basados en inteligencia computacional.
- Factores tecnológicos y arquitecturales que condicionan la implementación de métodos de inteligencia computacional en el contexto de las arquitecturas de routers. En particular, se ha mostrado que en la práctica las implementaciones software de sistemas de inferencia difusa presentan serios problemas de consumo de recursos computacionales, y apenas pueden proporcionar las tasas de inferencia requeridas en equipos de altas prestaciones.

Utilizando los puntos anteriores como elementos básicos, esta tesis ha aportado contribuciones en las siguientes áreas:

- Se ha desarrollado un método para predicción de series temporales a largo plazo mediante sistemas de inferencia difusa combinados con una técnica de estimación

de la varianza residual. El método, implementado en una herramienta incluida en el entorno de desarrollo Xfuzzy, proporciona mejores prestaciones que los modelos LS-SVM en términos de precisión, interpretabilidad y coste computacional para un conjunto heterogéneo de series de test. Se ha aplicado a un extenso conjunto de trazas de tráfico de nivel de paquetes con objeto de extraer modelos lingüísticos de series de carga de tráfico a diferentes escalas temporales.

- Asimismo se ha propuesto un método para resumir registros de flujos de red así como para inducir reglas difusas de asociación a partir de registros de flujos de red. Ambos métodos se han implementado en una herramienta experimental y se han aplicado a un amplio conjunto de registros de flujos de red. Se ha mostrado que el primero proporciona resúmenes concisos y apropiados con un coste computacional reducido. Se ha mostrado asimismo cómo el segundo método permite descubrir hechos relevantes ocultos en medidas de flujos de red.
- Hemos introducido un esquema para control de flujo y congestión de extremo a extremo mediante sistemas difusos basados en reglas, mostrándose que sus prestaciones mejoran otras propuestas actuales para control de congestión de extremo a extremo, incluyendo HighSpeed TCP y TCP SACK. Este método aporta una nueva perspectiva sobre los sistemas de control de tráfico basados en mecanismos de ventana deslizante.
- Asimismo, se ha abordado el problema del control de tráfico en routers, manteniendo la simplicidad del mecanismo de control y unas prestaciones satisfactorias para un amplio rango de escenarios y condiciones. El controlador difuso propuesto mejora las prestaciones de los controladores más ampliamente aceptados actualmente, incluyendo una variante del esquema RED, en términos de varios criterios en escenarios realistas, aun teniendo una configuración simple.
- Tomando como base los puntos anteriores, finalmente hemos abordado la realización hardware de aquellos sistemas cuyas restricciones de operación son críticas. Se ha mostrado que es factible implementar estos sistemas en la práctica y se ha desarrollado una plataforma abierta basada en dispositivos FPGA para el desarrollo automatizado de sistemas de inferencia difusa eficientes que se pueden integrar en arquitecturas de routers actuales y previsibles sin requerir elementos adicionales. Se ha mostrado que los sistemas de inferencia difusa analizados satisfacen las restricciones operacionales de los routers de altas prestaciones actuales y futuros, a medio plazo, en términos de velocidad de inferencia y consumo de recursos.

Como continuación natural al último punto cabe resaltar la aplicación de la plataforma de desarrollo en infraestructuras de comunicaciones experimentales para pruebas disruptivas de nuevos protocolos y mecanismos de red, cuya disponibilidad es previsible para el futuro próximo.

Por último, hemos de resaltar que en el contexto de esta tesis han quedado abiertas varias líneas de investigación a medio y largo plazo, entre las cuales podemos destacar las siguientes:

- Extender el análisis de predecibilidad de series de carga de tráfico realizado en el capítulo 3 a modelos de mayor dimensionalidad, lo que implicaría desarrollar modelos no lineales precisos para problemas de dimensionalidad muy alta. Esto, a su vez, conlleva nuevos desarrollos en un conjunto de áreas relacionadas con la predicción de series temporales, que pueden ser útiles para explotar la posible predecibilidad adicional derivada de las dependencias a largo plazo, frecuentes en las series de carga de tráfico.
- Ampliación del método de minería de reglas de asociación difusas aplicado en esta tesis con objeto de incluir información topológica y analizar las interacciones entre tráfico y topología. Esta línea de investigación puede conducir a resultados relevantes en un área que permanece aún esencialmente inexplorado debido a su complejidad.
- El diseño de los esquemas basados en reglas desarrollados en el capítulo 4 para control de congestión de extremo a extremo y gestión activa de colas permite adaptar, mediante un enfoque lingüístico, estos mecanismos para ajustarlos a los requisitos de aplicaciones específicas. Así, es posible, por ejemplo, adaptar gradualmente funciones de control de congestión para servicios con grados diversos de elasticidad de extremo a extremo.

Referencias

- [1] S. S. Anand, D. A. Bell, y J. G. Hughes, “EDM: A General Framework for Data Mining Based on Evidence Theory,” *Data & Knowledge Engineering*, vol. 18, no. 3, pp. 189–223, Abr. 1996.
- [2] G. Appenzeller, I. Keslassy, y N. McKeown, “Sizing Router Buffers,” en *ACM Special Interest Group on Data Communications (SIGCOMM) Conference*, Portland, OR, USA, Sep. 2004, pp. 281–292.
- [3] D. Awduche, A. Chiu, A. Elwalid, I. Widjaja, y X. Xiao, “Overview and Principles of Internet Traffic Engineering,” Internet Engineering Task Force, Network Working Group, RFC 3272, May. 2002, category: Informational.
- [4] D. Awduche, J. Malcolm, J. Agogbua, M. O’Dell, y J. McManus, “Requirements for Traffic Engineering Over MPLS,” Internet Engineering Task Force, Network Working Group, RFC 2702, Sep. 1999.
- [5] J. Aweya, “IP router architectures: an overview,” *International Journal of Communication Systems*, vol. 14, no. 5, pp. 447–475, May. 2001.
- [6] R. Battiti, “First and Second Order Methods for Learning: Between Steepest Descent and Newton’s Method,” *Neural Computation*, vol. 4, no. 2, pp. 141–166, Mar. 1992.
- [7] I. Baturone, A. Barriga, S. Sanchez-Solano, C. J. Jimenez, y D. R. Lopez, *Microelectronic Design of Fuzzy Logic-Based Systems*. CRC Press., 2000, ISBN: 0-8493-0091-6.
- [8] J. C. Bezdek, “Fuzzy models – what are they and why?” *IEEE Transactions on Fuzzy Systems*, vol. 1, no. 1, pp. 1–6, Feb. 1993.
- [9] J. W. Boote *et al.*, “Towards Multi-Domain Monitoring for the European Research Networks,” en *21th TERENA Networking Conference*. Poznan, Poland: Trans-European Research and Education Networking Association, Jun. 2005.
- [10] J.-Y. L. Boudec y P. Thiran, *Network Calculus: A Theory of Deterministic Queuing Systems for the Internet*, ser. Lecture Notes on Computer Science. Springer Verlag, Ago. 2001, vol. 2050, ISBN: 978-3540421849.
- [11] G. Box, G. M. Jenkins, y G. Reinsel, *Time Series Analysis: Forecasting & Control*. Prentice Hall; 3rd edition, 1994, ISBN: 0130607746.
- [12] A. Broido, Y. Hyun, R. Gao, y K. C. Claffy, “Their Share: Diversity and Disparity in IP Traffic,” en *5th Passive and Active Measurement Workshop (PAM)*, Antibes Juan-Les-Pins, France, Abr. 2004, pp. 113–125.
- [13] A. Cabrera, S. Sánchez-Solano, P. Brox, A. Barriga, y R. Senhadji, “Hardware/ Software Codesign of Configurable Fuzzy Control Systems,” *Applied Soft Computing*, vol. 4, no. 3, pp. 271–285, Dic. 2004.

- [14] D. Cai, M. F. McTear, y S. I. McClean, “Knowledge discovery in distributed databases using evidence theory,” *International Journal of Intelligent Systems*, vol. 15, no. 8, pp. 745–761, Jun. 2000.
- [15] P. Calyam, D. Krymskiy, M. Sridharan, y P. Schopis, “Active and Passive Measurements on Campus, Regional and National Network Backbone Paths,” en *14th IEEE International Conference on Computer Communications and Networks (ICCCN 2005)*, San Diego, California USA, Oct. 2005, pp. 537–542.
- [16] J. Casillas, O. Cordón, F. Herrera, y L. Magdalena, Eds., *Interpretability Issues in Fuzzy Modeling*, ser. Studies in Fuzziness and Soft Computing. Berlin, Germany: Springer Verlag, Ago. 2003, ISBN: 978-3-540-02932-8.
- [17] Y.-H. O. Chang y B. M. Ayyub, “Fuzzy regression methods - a comparative assessment,” *Fuzzy Sets and Systems*, vol. 119, no. 2, pp. 187–203, Abr. 2001.
- [18] C. Chatfield, *The Analysis of Time Series. An Introduction*. CRC Press, Jul. 2003, Sixth edition, ISBN: 1-58488-317-0.
- [19] J. Chi *et al.*, “WISHBONE Conbus IP Core,” May. 2008. [Online]. Disponible en: http://www.opencores.org/projects.cgi/web/wb_conbus/
- [20] S. L. Chiu, “A Cluster Estimation Method with Extension to Fuzzy Model Identification,” en *IEEE Conference on Fuzzy Systems, 1994. IEEE World Congress on Computational Intelligence*, Orlando, FL, USA, Jun. 1994, pp. 1240–1245.
- [21] Cisco Systems, Inc., “The Evolution of High-End Router Architectures,” Basic Scalability and Performance Considerations for Evaluating Large-Scale Router Designs, Tech. Rep., 2001. [Online]. Disponible en: <http://www.cisco.com/en/US/products/hw/routers/ps167/>
- [22] “Cisco IOS NetFlow,” Nov. 2007. [Online]. Disponible en: http://www.cisco.com/en/US/products/ps6601/products_ios_protocol_group_home.html
- [23] K. C. Claffy, “Measuring the Internet,” *IEEE Internet Computing*, vol. 4 (1), pp. 73–75, Ene. 2000.
- [24] D. E. Comer, “Network Processors: Programmable Technology for Building Network Systems,” *The Internet Protocol Journal*, vol. 7 (4), Dic. 2004. [Online]. Disponible en: http://www.cisco.com/en/US/about/ac123/ac147/archived_issues/ipj_7-4/
- [25] —, *Network Systems Design Using Network Processors: Intel 2XXX Version*. Upper Saddle River, NJ, USA: Pearson Prentice Hall, Jun. 2005, ISBN: 9780131872868.
- [26] Cooperative Association for Internet Data Analysis, “CAIDA Visualization Tools,” May. 2008. [Online]. Disponible en: <http://www.caida.org/tools/visualization/>

- [27] M. E. Crovella y B. Krishnamurthy, *Internet Measurement: Infrastructure, Traffic and Applications*. Wiley, Jul. 2006, ISBN: 978-0470014615.
- [28] P. Crowley, M. A. Franklin, H. Hadimioglu, y P. Z. Onufryk, Eds., *Network Processor Design: Issues and Practices*, ser. Computer Architecture and Design. San Francisco, CA, USA: Morgan Kaufmann Publishers, Oct. 2002, vol. 1, ISBN: 978-1558608757.
- [29] ———, *Network Processor Design: Issues and Practices*, ser. Computer Architecture and Design. San Francisco, CA, USA: Morgan Kaufmann Publishers, Nov. 2003, vol. 2, ISBN: 978-0121981570.
- [30] ———, *Network Processor Design: Issues and Practices*, ser. Computer Architecture and Design. San Francisco, CA, USA: Morgan Kaufmann Publishers, Feb. 2005, vol. 3, ISBN: 978-0120884766.
- [31] R. de Oliveira y T. Braun, “A Delay-based Approach Using Fuzzy Logic to Improve TCP Error Detection in Ad Hoc Networks,” en *IEEE Wireless Communications and Networking Conference*, Atlanta, USA, Mar. 2004.
- [32] G. Di Fatta, F. Hoffmann, G. L. Re, y A. Urso, “A Genetic Algorithm for the Design of a Fuzzy Controller for Active Queue Management,” *IEEE Transactions on Systems, Man and Cybernetics, Part C: Applications and Reviews*, vol. 33 (3), pp. 313–334, Ago. 2003.
- [33] M. Dolenc y T. Markovic, “PCI IP Core Specification,” OpenCores.Org Free Open Source IP Cores and Chip Design, Tech. Rep. Rev. 1.2, Jul. 2004.
- [34] Endace Limited., “DAG Network Monitoring Cards,” Jun. 2008. [Online]. Disponible en: <http://www.endace.com/our-products/dag-network-monitoring-cards/>
- [35] C. Estan, S. Savage, y G. Varghese, “Automatically Inferring Patterns of Resource Consumption in Network Traffic,” en *ACM SIGCOMM 2003*, Karlsruhe, Germany, Ago. 2003, pp. 137–148.
- [36] “ESTSP’07 European Symposium on Time Series Prediction: Prediction Competition,” Mar. 2008. [Online]. Disponible en: <http://www.estsp.org>
- [37] J. Faraday y C. Chatfield, “Time Series Forecasting with Neural Networks: A Comparative Study Using the Airline Data,” *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 47, no. 2, pp. 231–250, 1998.
- [38] R. Fletcher, *Practical Methods of Optimization, 2nd Edition*. New York, NY, USA: Wiley-Interscience, May. 2000, ISBN: 978-0-471-49463-8.
- [39] S. Floyd, “Congestion Control Principles,” Internet Engineering Task Force, Network Working Group, RFC 2914, Sep. 2000, category: Best Current Practice.

- [40] S. Floyd *et al.*, “Metrics for the Evaluation of Congestion Control Mechanisms,” Internet Engineering Task Force, Network Working Group, RFC 5166, Mar. 2008, category: Informational.
- [41] S. Floyd, J. Kempf *et al.*, “Concerns Regarding Congestion Control for Voice Traffic in the Internet,” Internet Engineering Task Force, Network Working Group, RFC 3714, Mar. 2004, category: Informational.
- [42] S. Floyd y E. Kohler, “Internet Research Needs Better Models,” *ACM SIGCOMM Computer Communication Review*, vol. 33, no. 1, pp. 29–34, Ene. 2003.
- [43] S. Floyd y V. Paxson, “Difficulties in Simulating the Internet,” *IEEE/ACM Transactions on Networking*, vol. 9, no. 4, pp. 392–403, Ago. 2001.
- [44] M. Fullmer *et al.*, “flow-tools,” Nov. 2007. [Online]. Disponible en: <http://www.splintered.net/sw/flow-tools/>
- [45] S. Ghosh, Q. Razouqi, H. J. Schumacher, y A. Celmins, “A Survey of Recent Advances in Fuzzy Logic in Telecommunications Networks and New Challenges,” *IEEE Transactions on Fuzzy Systems*, vol. 6, no. 39, pp. 443–447, Ago. 1998.
- [46] W. Gong, Y. Liu, V. Misra, y D. F. Towsley, “Self-similarity and long range dependence on the Internet: a second look at the evidence, origins and implications,” *Computer Networks*, vol. 48, no. 3, pp. 377–399, May. 2005.
- [47] W. J. Goralski, *Juniper and Cisco Routing. Policy and Protocols for Multivendor IP Networks*. Indianapolis, Indiana: Wiley Publishing Inc., 2002, ISBN: 0-471-21592-9.
- [48] S. Guillaume, “Designing fuzzy inference systems from data: An interpretability-oriented review,” *IEEE Transactions on Fuzzy Systems*, vol. 9, no. 3, pp. 426–443, Jun. 2001.
- [49] R. Herveille *et al.*, “WISHBONE System-on-Chip (SoC) Interconnection Architecture for Portable IP Cores.” OpenCores Organization, Tech. Rep. Revision B.3, Sep. 2002.
- [50] M. Hidell, “Decentralized Modular Router Architectures,” Ph.D. dissertation, KTH-Royal Institute of Technology, Sep. 2006.
- [51] M. Hidell, P. Sjödin, y O. Hagsand, “Control and Forwarding Plane Interaction in Distributed Routers,” Laboratory for Communication Networks, Department of Signals, Sensors, and Systems. KTH Royal Institute of Technology, Stockholm, Sweden, Tech. Rep. TRITA-S3-LCN-0501, Mar. 2005.
- [52] Information Sciences Institute. University of Southern California, Viterbi School of Engineering, “The Network Simulator – ns-2,” Ago. 2008. [Online]. Disponible en: <http://www.isi.edu/nsnam/ns/>

- [53] “The Internet2 Observatory,” Jul. 2008. [Online]. Disponible en: <http://www.internet2.edu/observatory/>
- [54] J.-S. R. Jang, C.-T. Sun, y E. Mizutani, *Neuro-Fuzzy and Soft Computing A Computational Approach to Learning and Machine Intelligence*. Upper Saddle River, New Jersey: Prentice Hall, Sep. 1997, ISBN 0-13-261066-3.
- [55] A. J. Jones, “New Tools in Non-linear Modelling and Prediction,” *Computational Management Science*, vol. 2, no. 1, pp. 109–149, Sep. 2004.
- [56] J. Kacprzyk y R. R. Yager, “Linguistic Summaries of Data Using Fuzzy Logic,” *International Journal of General Systems*, vol. 30, no. 2, pp. 133–1504, Ene. 2001.
- [57] J. Kacprzyk y S. Zadrozny, “Linguistic Database Summaries and Their Protoforms: Towards Natural Language Based Knowledge Discovery Tools,” *Information Sciences*, vol. 173, no. 4, pp. 281–304, Mar. 2005.
- [58] H. Kantz y T. Schreiber, *Nonlinear Time Series Analysis*, 2nd ed. Cambridge, UK: Cambridge University Press, 2004.
- [59] T. Karagiannis, M. Molle, M. Faloutsos, y A. Broido, “A Nonstationary Poisson View of Internet Traffic,” en *23th Annual Joint Conference of the IEEE Computer and Communications Societies (IEEE INFOCOM)*, vol. 3, Hong Kong, Mar. 2004, pp. 1558–1569.
- [60] V. Kecman, *Learning and Soft Computing. Support Vector Machines, Neural Networks, and Fuzzy Logic Models*. The MIT Press, Mar. 2001, ISBN 0-262-11255-8.
- [61] S. Keshav, *An Engineering Approach to Computer Networking: ATM Networks, the Internet, and the Telephone Network*, ser. Computing Series. Addison-Wesley Longman Publishing Co., Inc., May. 1997, ISBN: 978-0201634426.
- [62] S. Keshav y R. Sharma, “Issues and trends in router design,” *IEEE Communications Magazine*, vol. 36, no. 5, pp. 144–151, May. 1998.
- [63] A. K. Kloth, *Advanced Router Architectures*. CRC Press, Nov. 2005, ISBN: 0849335507.
- [64] A. Lakhina, J. W. Byers, M. E. Crovella, y I. Matta, “On the Geographic Location of Internet Resources,” *IEEE J-SAC, Special Issue on Internet and WWW Measurement, Mapping and Modeling*, vol. 21 (6), pp. 934–948, Ago. 2003.
- [65] A. Lakhina, K. Papagiannaki, M. E. Crovella, C. Diot, E. D. Kolaczyk, y N. Taft, “Structural analysis of network traffic flows,” en *Joint International Conference on Measurement and Modeling of Computer Systems (ACM SIGMETRICS)*, New York, NY, USA, Jun. 2004, pp. 61–72.
- [66] P. A. Lekkas, *Network Processors : Architectures, Protocols and Platforms*, 1st ed. McGraw-Hill Professional, Jul. 2003, ISBN: 978-0071409865.

- [67] Q. Liang, N. Karnik, y J. M. Mendel, “Connection Admission Control in ATM Networks Using Survey-Based Type-2 Fuzzy Logic Systems,” *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, vol. 30, no. 3, pp. 329–339, Ago. 2000.
- [68] B. B. Lowekamp, “Combining active and passive network measurements to build scalable monitoring systems on the grid,” *ACM SIGMETRICS Performance Evaluation Review. Special issue on grid computing*, vol. 30, no. 4, pp. 19–26, Mar. 2003.
- [69] Y. Miche, A. Sorjamaa, y A. Lendasse, “OP-ELM: Theory, Experiments and a Toolbox,” en *18th International Conference on Artificial Neural Networks (ICANN)*, ser. Lecture Notes in Computer Science, vol. 5163, Prague, Czech Republic, Sep. 2008, pp. 145–154.
- [70] F. Montesino-Pouzols, “Comparative Analysis of Active Bandwidth Estimation Tools,” en *5th Passive and Active Network Measurement*, ser. Lecture Notes in Computer Science, C. Barakat y I. Pratt, Eds., vol. 3015. Antibes Juan-Les-Pins, France: Springer, Abr. 2004, pp. 175–184.
- [71] F. Montesino-Pouzols y A. Barriga, “Medida y estimación activa de las prestaciones de la red,” *Boletín de la Red Nacional de I+D RedIRIS*, no. 74–75, pp. 28–31, Dic. 2005.
- [72] —, “Regressive fuzzy inference models with clustering identification: Application to the ESTSP08 competition,” en *2nd European Symposium on Time Series Prediction*, Porvoo, Finland, Sep. 2008, pp. 205–214.
- [73] F. Montesino-Pouzols, A. Barriga, D. R. Lopez, y S. Sánchez-Solano, “FPGA Based Implementation of Fuzzy Controllers for Internet Traffic,” en *XII IBERCHIP Workshop*, San José, Costa Rica, Mar. 2006, pp. 34–41.
- [74] —, “Open FPGA-Based Development Platform for Fuzzy Systems with Applications to Communications,” en *XXII Conference on Design of Circuits and Integrated Systems (DCIS’07)*, Seville, Spain, Nov. 2007, pp. 323–328.
- [75] —, *Encyclopedia of Networked and Virtual Organizations*. Hershey, New York, USA: Information Science Reference (an imprint of IGI Global), Mar. 2008, vol. II, cap. Performance Analysis of Computer Networks, pp. 1216–1222, ISBN: 978-1-59904-885-7.
- [76] —, *Encyclopedia of Networked and Virtual Organizations*. Hershey, New York, USA: Information Science Reference (an imprint of IGI Global), Mar. 2008, vol. II, cap. Performance Analysis of Peer-to-Peer Traffic, pp. 1210–1215, ISBN: 978-1-59904-885-7.
- [77] —, “Linguistic Summarization of Network Traffic Flows,” en *17th IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2008)*, *IEEE World Congress on Computational Intelligence*, Hong Kong, China, Jun. 2008, pp. 619–624.

- [78] F. Montesino-Pouzols, A. Lendasse, y A. Barriga, “Fuzzy Inference Based Autoregressors for Time Series Prediction Using Nonparametric Residual Variance Estimation,” en *17th IEEE International Conference on Fuzzy Systems (FUZZ-IEEE’08), IEEE World Congress on Computational Intelligence*, Hong Kong, China, Jun. 2008, pp. 613–618.
- [79] ———, “xftsp: a Tool for Time Series Prediction by Means of Fuzzy Inference Systems,” en *4th IEEE International Conference on Intelligent Systems (IS’08)*, Varna, Bulgaria, Sep. 2008, pp. 2–2–2–7.
- [80] F. Montesino-Pouzols, D. R. Lopez, A. Barriga, y S. Sánchez-Solano, “Sistemas Difusos para Control de Congestión y Calidad de Servicio en Internet,” en *XII Congreso Español Sobre Tecnologías y Lógica Fuzzy (ESTYLF’2004)*, Jaén, España, Sep. 2004.
- [81] ———, “Fuzzy End-to-End Rate Control for Internet Transport Protocols,” en *15th IEEE International Conference on Fuzzy Systems (FUZZ-IEEE’06)*, Vancouver, Canada, Jul. 2006, pp. 1347–1354.
- [82] ———, “Intelligent Scheduling of Aggregate Traffic in Internet Routers by Means of Fuzzy Systems,” en *Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU)*, Paris, France, Jul. 2006.
- [83] ———, *Encyclopedia of Networked and Virtual Organizations*. Hershey, New York, USA: Information Science Reference (an imprint of IGI Global), Mar. 2008, vol. II, cap. Performance Analysis and Models of Web Traffic, pp. 1196–1203, ISBN: 978-1-59904-885-7.
- [84] ———, *Encyclopedia of Networked and Virtual Organizations*. Hershey, New York, USA: Information Science Reference (an imprint of IGI Global), Mar. 2008, vol. II, cap. Performance Analysis of Multimedia Traffic, pp. 1204–1209, ISBN: 978-1-59904-885-7.
- [85] F. J. Moreno-Velo, I. Baturone, A. Barriga, y S. Sánchez-Solano, “Automatic Tuning of Complex Fuzzy Systems with Xfuzzy,” *Fuzzy Sets and Systems*, vol. 158, no. 18, pp. 2026–2038, Sep. 2007.
- [86] M. Murray y K. C. Claffy, “Measuring the immeasurable: Global internet measurement infrastructure,” en *Passive and Active Network Measurement (PAM) Workshop*, Amsterdam, Abr. 2001.
- [87] OpenCores Organization, “OpenCores.Org: Free Open Source IP Cores and Chip Design,” Sep. 2007. [Online]. Disponible en: <http://www.opencores.org>
- [88] K. Papagiannaki, N. Taft, Z.-L. Zhang, y C. Diot, “Long-Term Forecasting of Internet Backbone Traffic,” *IEEE Transactions on Neural Networks*, vol. 16, no. 5, pp. 110–1124, 2005.

- [89] K. Papagiannaki, D. Veitch, y N. Hohn, “Origins of Microcongestion in an Access Router,” en *5th Passive and Active Network Measurement*, ser. Lecture Notes in Computer Science, C. Barakat y I. Pratt, Eds., vol. 3015. Antibes Juan-Les-Pins, France: Springer, Abr. 2004, pp. 175–184.
- [90] K. Park y W. Willinger, Eds., *Self-Similar Network Traffic and Performance Evaluation*. New York, USA: Wiley Interscience, Sep. 2000, ISBN: 0-471-31974-0.
- [91] V. Paxson, “Empirically-Derived Analytic Models of Wide-Area TCP Connections,” *IEEE/ACM Transactions on Networking*, vol. 2, no. 4, pp. 316–336, Ago. 1994.
- [92] ———, “End-to-End Internet Packet Dynamics,” *IEEE/ACM Transactions on Networking*, vol. 7, no. 3, pp. 277–292, Jun. 1999.
- [93] V. Paxson y S. Floyd, “Wide Area Traffic: The Failure of Poisson Modeling,” *IEEE/ACM Transactions on Networking*, vol. 3 (3), pp. 226–244, Jun. 1995.
- [94] R. S. Prasad, M. Murray, C. Dovrolis, y K. C. Claffy, “Bandwidth estimation: metrics, measurement techniques, and tools,” *IEEE Network*, vol. 17, no. 6, pp. 27–35, Nov. 2003.
- [95] G. Raschia y N. Mouaddib, “SAINTETIQ: a fuzzy set-based approach to database summarization,” *Fuzzy Sets and Systems*, vol. 129, no. 2, pp. 137–162, Jul. 2002.
- [96] D. Rasmussen y R. R. Yager, “Finding fuzzy and gradual functional dependencies with SummarySQL,” *Fuzzy Sets and Systems*, vol. 106, no. 2, pp. 31–42, Sep. 1999.
- [97] R. A. Resende, N. A. Nassif, M. A. de Siquira, A. E. da Silva, y M. Lima-Marques, “Quality of Service Control in IP Networks Using Fuzzy Logic for Policy Condition Evaluation,” en *IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2005)*, Reno, USA, May. 2005, pp. 448–453.
- [98] D. Rolls, G. Michailidis, y F. Hernández-Campos, “Queueing Analysis of Network Traffic: Methodology and Visualization Tools,” *Computer Networks*, vol. 48 (3), pp. 447–473, Jun. 2005.
- [99] C. Semeria, “T-series routing platforms: System and packet forwarding architecture,” Juniper Networks, Inc., Tech. Rep. 200027-001, Abr. 2002. [Online]. Disponible en: <http://www.arl.wustl.edu/~jst/cse/577/readings/juniperTseries.pdf>
- [100] S. Shalunov, M. Swany, B. Lutzmann, y F. Montesino-Pouzols, “Reporting IP Performance Metrics to Users,” Internet Engineering Task Force, Network Working Group, IP Performance Metrics Working Group, Tech. Rep., Jul. 2008, Internet Draft, draft-ietf-ippm-reporting-02.
- [101] S. Shalunov y B. Teitelbaum, “TCP Use and Performance on Internet2,” en *ACM SIGCOMM Internet Measurement Workshop*, San Francisco, CA, USA, 2001, pp. 147–160.

- [102] J. Sommers, P. Barford, y W. Willinger, “SPLAT: A Visualization Tool for Mining Internet Measurements,” en *7th Passive and Active Network Measurement Workshop*, Mar. 2006, pp. 31–40.
- [103] Q. Song y B. Chissom, “Fuzzy time series and its models,” *Fuzzy Sets and Systems*, vol. 54, no. 3, pp. 269–277, Mar. 1993.
- [104] R. Srikant, *The Mathematics of Internet Congestion Control*, ser. Systems & Control: Foundations & Applications. Boston, MA, USA: Birkhauser, Dic. 2003, ISBN: 0-8176-3227-1.
- [105] J. A. K. Suykens, T. Van Gestel, J. De Brabanter, B. De Moor, y J. Vandewalle, *Least Squares Support Vector Machines*. Singapore: World Scientific, 2002, ISBN: 981-238-151-1.
- [106] R. Usselman *et al.*, “WISHBONE/OPB & OPB/WISHBONE Interface Wrapper: Overview,” Sep. 2004. [Online]. Disponible en: http://www.opencores.org/projects.cgi/web/opb_wb_wrapper
- [107] G. Varghese, *Network Algorithmics: An Interdisciplinary Approach to Designing Fast Networked Devices*. San Francisco, CA, USA: Morgan Kaufmann, Dic. 2004, ISBN: 978-0120884773.
- [108] K. V. Vishwanath y A. Vahdat, “Realistic and responsive network traffic generation,” en *ACM Special Interest Group on Data Communication (SIGCOMM) Conference*, Pisa, Italy, Sep. 2006, pp. 111–122.
- [109] ———, “Evaluating distributed systems: does background traffic matter?” en *ACM/USENIX 2008 Annual Technical Conference*, Boston, MA, USA, Jun. 2008, pp. 227–240.
- [110] L. X. Wang, “The WM Method Completed: A Flexible System Approach to Data Mining,” *IEEE Transactions on Fuzzy Systems*, vol. 11, no. 6, pp. 768–782, Dic. 2003.
- [111] L. X. Wang y J. M. Mendel, “Generating Fuzzy Rules by Learning from Examples,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 22, no. 4, pp. 1414–1427, Dic. 1992.
- [112] A. Weigend y N. Gershenfeld, *Times Series Prediction: Forecasting the Future and Understanding the Past*. Addison-Wesley Publishing Company, 1994, ISBN: 0201626020.
- [113] W. Willinger, V. Paxson, y M. S. Taqqu, *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*. Boston, USA: Brikhauser, 1998, cap. Self-Similarity and Heavy Tails: Structural Modeling of Network Traffic, pp. 27–53.

- [114] *Xilinx ISE 9.2i Software Manuals and Help - PDF Collection*, Xilinx©, 2007. [Online]. Disponible en: http://www.xilinx.com/support/sw_manuals/xilinx92/index.htm
- [115] R. R. Yager, “A New Approach to the Summarization of Data,” *Information Sciences*, vol. 28, pp. 69–86, 1982.
- [116] —, “Database Discovery Using Fuzzy Sets,” *International Journal of Intelligent Systems*, vol. 11, pp. 691–712, 1996.
- [117] —, “Using fuzzy methods to model nearest neighbor rules,” *IEEE Transactions on Systems, And, and Cybernetics—Part B: Cybernetics*, vol. 32, no. 4, pp. 512–525, Ago. 2002.
- [118] R. R. Yager, K. J. Engemann, y D. P. Filev, “On the Concept of Immediate Probabilities,” *International Journal of Intelligent Systems*, vol. 10, no. 4, pp. 373–397, 1995.
- [119] M. H. Yaghmaee, “Design and Performance Evaluation of a Fuzzy Based Traffic Conditioner for Differentiated Services,” *Computer Networks*, vol. 47 (6), pp. 847–869, Ene. 2005.
- [120] S. Yusuf, W. Luk, M. Sloman, N. Dulay, E. C. Lupu, y G. Brown, “Reconfigurable Architecture for Network Flow Analysis,” *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 16, no. 2, pp. 57–65, Ene. 2008.
- [121] R. Zhang, Y. A. Phillis, y V. Kouikoglou, *Fuzzy Systems for Queuing Control*. London, UK: Springer-Verlag, Ene. 2005, ISBN: 978-1-85233-824-4.