



Ingeniare. Revista Chilena de Ingeniería

ISSN: 0718-3291

facing@uta.cl

Universidad de Tarapacá

Chile

Millán Naveas, Ginno; Lefranc Hernández, Gastón

INCLUSIÓN DE CORRELACIONES TEMPORALES CON DEPENDENCIA A LARGO PLAZO Y
PATRONES AUTOSIMILARES EN MODELOS DE REDES IEEE 802.3

Ingeniare. Revista Chilena de Ingeniería, vol. 18, núm. 1, 2010, pp. 95-104

Universidad de Tarapacá

Arica, Chile

Disponible en: <http://www.redalyc.org/articulo.oa?id=77218811011>

- ▶ Cómo citar el artículo
- ▶ Número completo
- ▶ Más información del artículo
- ▶ Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica

Red de Revistas Científicas de América Latina, el Caribe, España y Portugal
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

INCLUSIÓN DE CORRELACIONES TEMPORALES CON DEPENDENCIA A LARGO PLAZO Y PATRONES AUTOSIMILARES EN MODELOS DE REDES IEEE 802.3

INCLUSION OF LONG-RANGE-DEPENDENT TEMPORARY CORRELATIONS AND SELF-SIMILAR PATTERNS IN IEEE 802.3 NETWORKS MODELS

Ginno Millán Naveas¹ Gastón Lefranc Hernández²

Recibido 1 de septiembre de 2009, aceptado 11 de diciembre de 2009

Received: September 1, 2009 Accepted: December 11, 2009

RESUMEN

En este trabajo se presentan los fundamentos de un proyecto de investigación sobre el modelado de redes de computadoras con mecanismo de control de acceso al medio, según el estándar IEEE 802.3-2005, empleando los postulados de la teoría de conjuntos autosimilares para establecer el nivel de impacto que poseen las correlaciones temporales con dependencia de largo alcance sobre el rendimiento de tales redes. Se postula una nueva forma de estimar grados de autosimilaridad basada en una variación del estimador de Whittle.

Palabras clave: Autosimilaridad, correlación, dependencia de largo alcance, estimador de Whittle.

ABSTRACT

The foundation of a research project about a model of computer networks with media access control mechanism based on the IEEE standard 802.3-2005 is presented. The model draws from the theory of self-similar sets for establishing the impact level that the long-range-dependent temporary correlations have on the performance of such networks. A new method for the estimation of self-similar levels based on a variation of the Whittle estimator is postulated.

Keywords: Self-similarity, correlation, long-range-dependent, Whittle estimator.

INTRODUCCIÓN

El posicionamiento y consolidación de Ethernet como el estándar predominante en el campo de las redes de computadoras, tanto a niveles de coberturas locales como extensas, frente a tecnologías tradicionales como Frame Relay, DQDB y ATM, son hechos que se explican a partir de sus principales características, a saber: compatibilidad e interoperatividad entre los equipamientos Ethernet de distintas velocidades, altas prestaciones, escalabilidad y capacidad de autoconfiguración, independencia de los esquemas de direccionamiento IP y, por supuesto, por su consabida economía de escala.

Ethernet, inicialmente a 2,94 Mb/s, ha evolucionado de 10 Mb/s a 10 Gb/s en veintidós años (ello sin considerar, además, el actual estándar IEEE 802.3ba que especifica Ethernet a 40 Gb/s y 100 Gb/s), y del uso de simples puentes ideados para la interconexión de redes con idénticos

protocolos a nivel de capas física y de acceso al medio, ha progresado hacia los commutadores con capacidades de N*10 Gb/s [1-3].

Insertos en esta continua evolución que experimentan las redes Ethernet, se encuentran dos aspectos de especial interés y criticidad. El primero, dice relación con el total abandono del medio compartido half-duplex originario, para dar paso a enlaces dedicados full-duplex, mientras el segundo aborda su extensión: Ethernet ha evolucionado desde las distancias del rango LAN a las coberturas del rango WAN [3]. Y aun cuando ambos cambios han sido graduales en el tiempo, son radicales desde el punto de vista de las arquitecturas Ethernet, por cuanto en su conjunto significan, por una parte, la desaparición del mecanismo de control de acceso al medio CSMA/CD en un primer término y, por otra, presuponen un drástico cambio en los medios de transmisión tendiente a un uso absoluto de fibras ópticas.

¹ Departamento de Ingeniería Eléctrica. Universidad de Santiago de Chile. Santiago, Chile. E-mail: ginno.millan@gmail.com

² Escuela de Ingeniería Eléctrica. Pontificia Universidad Católica de Valparaíso. Valparaíso, Chile. E-mail: glefranc@ucv.cl

Las redes de área local (LAN), en general, y Ethernet en particular, nacieron siendo, en esencia, redes de medio compartido de alta capacidad frente a tecnologías de redes de área extensa (WAN) basadas en conmutación y con caudales de transmisión inferiores, en general, a los disponibles en las redes LAN. Sin embargo, la evolución de las tecnologías utilizadas tanto en los entornos LAN como WAN, hoy converge en soluciones basadas en Ethernet y sus diversas especificaciones. De esta forma, las actuales redes LAN Ethernet son conmutadas, están compuestas de enlaces dedicados full-duplex, incorporan multiplexación según el estándar IEEE 802.1Q y soportan distancias de transmisión comparables, y en diversos casos idénticas, a las soportadas por los enlaces WAN convencionales.

Se atribuye en gran parte esta evolución al alto nivel de desarrollo experimentado por los conmutadores Ethernet, puesto que desde el punto de vista operacional no sólo han aumentado su grado de transparencia y simplicidad en sus procesos de puesta en marcha, operación y gestión, sino que ello ha incidido directamente en la incorporación de diversas funcionalidades a la simple conmutación, lo cual, desde el punto de vista del estándar, se traduce hoy en la extensión del formato de la trama Ethernet original, establecida en IEEE 802.3-1985, con la incorporación de etiquetas para identificación de VLAN y establecimiento de prioridades para clases de servicio [4-6], en el aumento de tamaño de la señal portadora de CSMA/CD, en la incorporación de ráfagas de paquetes para compensar la pérdida de velocidad en las redes ocasionada por los bits de extensión de la portadora y, final y prioritariamente, en el abandono del esquema de contención y resolución de colisiones regido por el algoritmo de retroceso exponencial binario [7].

Además, esta tendencia migratoria hacia redes Ethernet sin medio compartido se confirma por la incorporación en los dispositivos activos de red de los estándares IEEE 802.1X, IEEE 802.1w (RSTP; actualmente IEEE 802.1D) e, IEEE 802.1s (MSTP; actualmente IEEE 802.1Q), todos los cuales establecen como requisito enlaces punto a punto para su operación. Cabe señalar que los enlaces dedicados no tan solo son necesarios para obtener las máximas prestaciones de la red, sino que además hacen efectiva la seguridad a nivel de enlace lógico de datos (LLC), simplifican sus protocolos y posibilitan los mecanismos de convergencia rápida en capa dos, como el empleado por el protocolo RSTP.

Un ejemplo categórico de todo lo anterior lo constituye el estándar IEEE 802.3ae (Ethernet a 10 Gb/s), el cual, en su especificación, no contempla el uso de enlaces half-

duplex como en IEEE 802.3z (Ethernet a 1.000 Mb/s), en cuya especificación se mantuvieron por estrictas razones de compatibilidad con las bases de equipamiento con anterioridad instaladas, respondiendo, como objetivo final, al de servir como plataforma en los procesos de migración o de transición tecnológicas. Cabe en este punto recordar que se trata de la última de las especificaciones Ethernet en ofrecer esta modalidad de comunicación entre dispositivos.

Un antecedente que resulta de especial atención e interés para el desarrollo de este trabajo es el hecho de que en las redes Ethernet tradicionales predominan la difusión y la inundación como mecanismos básicos y válidos al momento de establecer la presencia o la ausencia de estaciones y, por contraparte, hoy se busca la mínima difusión de las tramas por las mismas dos principales razones que se evitan en las redes WAN: la degradación del rendimiento y el control eficiente de tráfico.

De forma paralela las redes de mayores coberturas van incorporando tecnologías propias de los entornos LAN, debido a su robustez y buena relación precio/prestaciones [8], estando hoy bastante implantadas en entornos tanto metropolitanos como de acceso y de forma creciente en los entornos WAN propiamente tales. Esta tendencia es continua y con perspectivas de seguirlo siendo en el futuro, y se produce, principalmente, debido al aumento en el tamaño, la capacidad y la complejidad de las redes de campus. De hecho, las actuales funcionalidades de este tipo de redes incluyen: la agregación de enlaces según IEEE 802.3ad para el aumento de la fiabilidad y el ancho de banda, la desagregación y el balance de la carga, la posibilidad de ampliación gradual mediante la creación de VLANs y árboles múltiples de expansión por VLANs según el estándar IEEE 802.1s (MSTP; actualmente IEEE 802.1Q), la inspección IGMP en capa dos para provisión de transmisiones de multidifusión con acceso y desconexión rápida de clientes y fraccionamiento del ancho de banda para el tráfico intensivo de video únicamente para aquellos clientes que lo solicitan o bien lo reservan, y otras que hacen más compleja y crítica su gestión y administración [3].

El conjunto de todos los argumentos anteriores avalan un replanteamiento en la forma de estudiar las prestaciones del protocolo y mecanismo de control de acceso al medio CSMA/CD, en términos del impacto que su rendimiento provoca en los actuales entornos conmutados, entendiendo el rendimiento como la cantidad de información útil que la red es capaz de transportar, en relación con la cantidad de bits transportados realmente, además de caracterizar la naturaleza del tráfico sometido a estudio en términos

de la búsqueda de un patrón de comportamiento capaz de describir su evolución temporal y explicar sus efectos sobre el rendimiento anteriormente definido. Se justifica considerar el parámetro de rendimiento como una forma activa de medir las prestaciones de una red, pues es sin lugar a dudas uno de los aspectos de mayor interés dentro del análisis global de los sistemas de comunicaciones por la repercusión que tiene sobre los usuarios finales. Asimismo la caracterización del tráfico que en las redes se cursa, por cuanto de sus singularidades dependerá el comportamiento que este parámetro exhiba.

De igual forma se justifica un nuevo enfoque para llevar a cabo los procesos de modelado de redes Ethernet, puesto que, en términos de la evolución antes planteada, se infiere que el sucesor natural del estándar IEEE 802.3u, es el estándar IEEE 802.3z, el cual deberá ceder su lugar en los entornos de acceso y WAN, a los estándares IEEE 802.3ae e IEEE 802.3ba, respectivamente, y las repercusiones que consigo traerán asociadas estas migraciones tecnológicas, deben ser adecuadamente evaluadas, dimensionadas y categorizadas en términos de su impacto sobre las actuales bases de equipamiento instaladas bajo la óptica definida con anterioridad, antes de su adopción. Se aclara en este punto que en función de la definición de rendimiento antes planteada, tan solo se pretende entregar una idea intuitiva del aprovechamiento de la capacidad de canal, para posteriormente realizar un estudio en detalle sobre la base de los correspondientes modelos analíticos.

Por otra parte, la aceptación de la existencia de patrones de tráfico autosimilares, se realiza sobre el hecho empírico de que éstos se caracterizan por la presencia permanente de ráfagas de tráfico de paquetes, a través de diferentes escalas de tiempo y considerando el argumento de que las propiedades características de todos los procesos autosimilares es la dependencia de largo alcance (LRD), la cual se presenta al aumentar el nivel de agregación [1]. Luego, el comportamiento fractal de este tipo de tráfico no coincide con el comportamiento del tráfico modelado tradicionalmente a través de procesos de Poisson o Markov, los cuales se caracterizan por la ausencia de ráfagas y una baja variabilidad reflejada en la total independencia temporal entre las muestras. En resumen, se trata de procesos que con toda claridad exhiben dependencia temporal de corto alcance (SRD), la cual desprecia toda forma de relación entre procesos temporalmente distantes, es decir, se trata de procesos estocásticos de memoria nula, y no consideran válida la existencia de algún tipo de patrón mínimo representativo de la totalidad del tráfico cursado. Luego, considerando los mismos argumentos, además del hecho de que los fenómenos autosimilares presentan el mismo aspecto o comportamiento cuando se visualizan con distintos grados de ampliación o a

distintas escalas de una cierta dimensión, la cual puede ser el espacio o el tiempo, y que el objeto de interés en redes de computadoras son las series temporales y los procesos estocásticos que muestran autosimilaridad con respecto al tiempo, se debe establecer una medida, un grado de autosimilaridad para dichas series, el cual, por definición, se expresa utilizando únicamente un parámetro que representa la rapidez de decrecimiento de la función de autocorrelación, lo cual responde al hecho de que una serie de tiempo es autosimilar si la serie agregada tiene la misma función de autocorrelación que la serie original. Este parámetro de caracterización se denomina parámetro de Hurst, H , y puede ser estimado por diversos métodos, siendo el de Whittle el que ofrece el mayor rigor estadístico, y para cuya obtención existen diversos algoritmos, todos los cuales precisan necesariamente de la forma del proceso estocástico subyacente. En este trabajo se propone una nueva variante del estimador de Whittle, que permite obtener el grado de autosimilaridad de una serie temporal bajo una relación aceptable de compromiso entre el costo del modelo computacional, el mayor problema asociado con la obtención del estimador de Whittle y la calidad de las estimaciones.

Formalmente, la propuesta consiste en modificar el estimador local de Whittle (LWE) o estimador gaussiano semiparamétrico, del parámetro de memoria en procesos estándar a corto alcance expuesto en [9], en torno a un punto de operación susceptible de aproximaciones por bisección. Luego, con esta condición se esperan obtener las ventajas heredadas de la técnica original, que exhibe sus principales atributos como alternativa a la técnica de regresión del logaritmo del periodograma propuesta en [10]. En particular, se espera que bajo supuestos menos restrictivos se muestre una forma de ganancia de eficiencia asintótica. Para ello, primero se analiza el comportamiento asintótico del estimador gaussiano semiparamétrico original del parámetro de memoria en procesos de memoria cíclica o estacional permitiendo divergencias o ceros espectrales asimétricos, para luego obtener, por la vía de modificar el algoritmo, la consistencia y la normalidad asintótica necesaria para la caracterización de los flujos de tráfico de las muestras.

En este trabajo se presentan los fundamentos de un proyecto de investigación sobre el modelado de redes de computadoras que presentan un mecanismo de control de acceso al medio regido por el estándar IEEE 802.3 en su versión 2005, empleando postulados de la teoría de conjuntos autosimilares para determinar el nivel de impacto que poseen las correlaciones temporales de largo plazo sobre el rendimiento de tales redes, postulando una nueva forma de estimar grados de autosimilaridad basada en una variación del estimador de Whittle.

ANTECEDENTES DEL PROBLEMA

El análisis de colas ha resultado ser de enorme utilidad para quienes diseñan redes y para los analistas de sistemas, a efectos de realizar una planificación de capacidades y también para predecir rendimientos [11]. Sin embargo, existen muchos casos del mundo real en los cuales se ha observado que los resultados predichos a partir de un análisis de colas difieren de forma significativa de los rendimientos observados en la realidad [11]. En este sentido, se debe enfatizar en el hecho de que la validez de los análisis basados en la teoría de colas depende de la naturaleza de Poisson del tráfico de datos, y que al tratarse de procesos de Poisson, tanto la representación de la duración de cada arribo como de los tiempos entre arribos sucesivos de tramas son realizados por variables aleatorias independientes exponencialmente distribuidas. Por lo tanto se trata de modelos de memoria nula, y siendo así, de modelos en los cuales la probabilidad de llegada en un determinado instante, es por completo independiente de los instantes de llegadas anteriores y posteriores, propiedad que demuestra que no tan solo no se cumple en las redes de conmutación de paquetes, sino que además conlleva a modelos erróneos.

Con respecto al último planteamiento, se reconoce que el objetivo de estas suposiciones responde con rigurosidad a la obtención de modelos relativamente simples desde el punto de vista analítico. No obstante, por ejemplo, la variabilidad del tráfico multimedia coloca bajo fuertes cuestionamientos la validez de los modelos tradicionales y en concreto a las hipótesis de independencia sobre las cuales son planteados [12, 13].

A partir de los resultados obtenidos en [14], la existencia de la dependencia temporal es objeto de estudio por parte de numerosos autores, destacando el descubrimiento del enorme impacto que puede llegar a tener sobre las prestaciones de un sistema de colas. A la luz de este hecho existe abundante literatura sobre modelos de tráfico de entrada que exhiben estructuras de correlación más o menos complejas, aplicadas a casos en los cuales los modelos de los sistemas de comunicaciones bajo estudio permiten mantener un trato analítico adecuado. Pero en cualquier caso, estos modelos, fundamentalmente markovianos, desprecian la correlación a partir de una determinada separación temporal, aun cuando esta pueda ser arbitrariamente aumentada a costa de complicar el modelo con parámetros adicionales y disminuir con ello la plausibilidad de interpretarlos físicamente. Se informa a continuación un conjunto de resultados relevantes, referentes sin parangón de lo planteado.

En [1] se demuestra tras exhaustivas mediciones sobre una red Ethernet, que el tráfico, entendido como el número de tramas en la red por unidad de tiempo, presenta naturaleza autosimilar o fractal, lo cual pone al descubierto la existencia de una acusada correlación de largo alcance.

En [15] y [16] se demuestra el carácter autosimilar del tráfico en redes WAN, mientras que en [17] se coloca de manifiesto la naturaleza fractal del flujo de los datos de los protocolos componentes del sistema de señalización número 7 (SS7) en redes de señalización de canal común, en [18] se demuestra la naturaleza autosimilar del tráfico debida al WWW y en [19] y [20], el tráfico de video de tasa variable (VBR).

En [21] se demuestra que la distribución de probabilidad que sigue el tamaño de la cola de un multiplexor, exhibe una caída asintótica del tipo Weibull al utilizar como tráfico de entrada cierto tipo de procesos autosimilares y en [22] que esta caída puede ser aún más lenta y de tipo hiperbólico utilizando otros procesos autosimilares.

En [23] se demuestra que el ancho de banda efectivo calculado sobre la base de modelos de Markov, en los cuales la distribución del tamaño de la cola presenta una caída exponencial, subestima con creces la tasa de pérdida de celdas en varios órdenes de magnitud.

En el análisis de un multiplexor de voz y datos expuesto en [24], se demuestra una excepcional correlación positiva de largo alcance en el tráfico agregado, obteniendo retardos muy superiores a los predichos con modelos de Poisson.

Finalmente, en [25] se demuestra que el número de llegadas en intervalos de tiempo adyacentes, resultante de la superposición de múltiples fuentes de voz independientes, homogéneas tratadas mediante modelos de procesos de renovación, es un proceso con fuertes correlaciones, que posee un impacto significativo sobre las prestaciones del sistema de comunicaciones bajo estudio.

Lo interesante de estos y otros muchos estudios es que colocan sobre relieve el impacto que la dependencia temporal de largo alcance, intrínseca a los más diversos tipos de tráfico, tiene sobre las prestaciones de las redes de comunicaciones frente a los modelos que no presentan correlación temporal por simplicidad analítica (como los procesos de renovación), o exhiben una estructura de correlación más complicada analíticamente (como los modelos de Markov o autorregresivos habituales), que pueden ser denominados como modelos de corto alcance. Además, el principal inconveniente que presentan la

totalidad de los modelos tradicionales es que requieren un número muy elevado de parámetros para modelar las fuertes correlaciones existentes entre los diferentes tipos de tráficos en una red. Y, por supuesto que en la medida que el número de parámetros aumenta, la complejidad analítica crece desmesuradamente, aparte de la dificultad añadida de dar una interpretación física plausible para todos los parámetros.

En pocas palabras, las características de autosimilaridad y fractalidad describen un fenómeno en el que una cierta propiedad de un objeto se preserva con respecto a un escalamiento temporal y/o espacial, y en un objeto autosimilar o fractal, sus partes magnificadas se asemejan a la forma del objeto completo, donde la semejanza se mide en algún sentido adecuado. Luego mediante la iteración de un cierto procedimiento, es posible obtener, por simple reconstrucción, la forma más sencilla de autosimilaridad. Si el procedimiento se repite indefinidamente para cada nuevo segmento, cualquier porción del objeto, por más pequeña que esta sea, puede ser magnificada para reproducir exactamente una porción mayor. A esta propiedad se la conoce con el nombre de “autosimilaridad exacta”. No se pretende observar en un proceso altamente aleatorio como el de la llegada de paquetes a una red de datos, características autosimilares exactas, pero si se considera el tráfico observado como una traza de muestras perteneciente a algún proceso estocástico y restringe la similaridad a ciertas estadísticas de las series temporales ajustadas en escala, se descubre autosimilaridad exacta en los objetos matemáticos abstractos y una autosimilaridad aproximada para cada una de las realizaciones específicas.

Aclarado lo anterior, no se espera, bajo ningún punto de vista, que los tráficos en las redes de computadoras bajo estudio exhiban características autosimilares exactas en sus comportamientos, pues ese supuesto sería tan o más atentatorio contra los intereses del trabajo como seguir validando los modelos tradicionales, sino que se esperan comportamientos del tipo autosimilares estocásticos, y en este respecto, para medir la autosimilaridad, pueden ser utilizados estadísticos de segundo orden que capturen la variabilidad de los procesos. De hecho, la invariancia a la escala puede ser definida en términos de la función de autocorrelación, puesto que el decrecimiento polinómico (en vez de exponencial) de esta función es la manifestación de su dependencia de largo alcance, concepto equivalente al de autosimilaridad y que constituye la óptica bajo la cual deben ser interpretados los procesos autosimilares durante el desarrollo de la investigación.

Por otra parte, con respecto al problema fundamental al análisis de procesos autosimilares, o más concretamente

de series temporales que exhiben LRD, es decir, el de la estimación del grado de autosimilaridad o parámetro de Hurst, los métodos propuestos en la literatura pueden ser clasificados en dos grandes grupos, a saber:

1. Métodos gráficos de regresión lineal. En ellos se calcula algún estadístico $T(x)$ que se comporta asintóticamente para un determinado conjunto de valores de x y, por lo tanto, están basados en obtener, por medio de regresión lineal de mínimos cuadrados, la recta que mejor se ajuste (para dicho conjunto de valores de x) a la representación de $\log(T(x))$ frente a $\log(x)$, obteniéndose así el valor del parámetro de Hurst (H) directamente a partir del valor de la pendiente de dicha recta.
2. Métodos basados en la utilización de estimadores de máxima verosimilitud (MLE) para H . El objetivo de este tipo de métodos es intentar minimizar las diferencias entre el periodograma de la serie y su espectro teórico.

Los métodos del primer grupo son relativamente sencillos y rápidos algorítmicamente de implementar, sin embargo su principal inconveniente reside en que primero debe estimarse un comportamiento asintótico a partir de un número finito de muestras, lo que hace que la estimación del parámetro H dependa considerablemente de la correcta elección del conjunto de valores de x . Por tal motivo, son fundamentales las representaciones gráficas para, por una parte, verificar que el conjunto elegido de valores de x corresponda con la zona de comportamiento lineal y, por otra, para comprobar que la recta es un buen ajuste de los puntos representados. Asimismo es importante destacar que estos métodos sólo permiten obtener una estimación puntual del parámetro H , ya que la obtención de intervalos de confianza trae aparejado un alto costo en recursos computacionales y tiempos de procesamiento. No obstante, son muy útiles para detectar la presencia de LRD y crearse una idea de su magnitud. Este grupo da cabida a todos los métodos basados en el análisis temporal de las series agregadas, y al método basado en la regresión lineal sobre el periodograma de [10]. Por el contrario, los métodos basados en estimadores de máxima verosimilitud (MLE), aun cuando sean más complejos e involucren un costo computacional mayor, son más flexibles y eficientes desde el punto de vista de la inferencia estadística, ya que permiten obtener intervalos de confianza para los valores estimados de H . Por esta razón, son los métodos habitualmente más utilizados.

Luego, para el caso de esta investigación se tratará con métodos basados en estimadores de máxima verosimilitud

a partir de los cuales serán resueltos los intervalos de confianza para llegar a las primeras aproximaciones de un valor para H , el cual luego será ajustado empleando el método propuesto. No obstante lo anterior, los métodos del primer grupo serán detenidamente analizados y los resultados que con ellos se obtengan serán contrastados con los obtenidos mediante el uso de MLE y el ajuste que propone el método desarrollado. Luego, por cuanto el análisis comienza empleando MLE, es necesario recordar que se trata de métodos que se emplean para minimizar las diferencias entre el periodograma de la serie y el modelo paramétrico supuesto para su densidad espectral teórica. Además, se hace hincapié en que el cálculo exacto de los MLE es computacionalmente costoso, por lo que deben emplearse funciones estimadoras de máxima verosimilitud gaussianas (en adelante, MLE gaussianos). Sin embargo, incluso en este caso, el costo computacional es elevado, por lo que en la práctica se utilizan aproximaciones basadas en dichos MLE gaussianos. Un punto trascendental es que la aproximación más ampliamente usada es la de Whittle, y en ella se basa el método propuesto.

VARIANTES DEL ESTIMADOR DE WHITTLE

Sea $f(\lambda, \theta)$ la forma paramétrica de la densidad espectral de un proceso estacionario gaussiano X_t , donde θ es el vector de parámetros que se desea estimar, además se verifica que $\theta = (\theta_1, \dots, \theta_M)$.

Sea el periodograma de muestras $I(\lambda)$ definido por

$$I(\lambda) = \frac{1}{2\pi N} \left| \sum_{t=1}^N X_t e^{j\lambda t} \right|^2 \quad (1)$$

El MLE aproximado de Whittle es el vector dado por la expresión siguiente

$$\hat{\theta} = (\hat{\theta}_1, \dots, \hat{\theta}_M) \quad (2)$$

que minimiza la función

$$\mathcal{Q}(\theta) \triangleq \frac{1}{2\pi} \left\{ \int_{-\pi}^{\pi} \frac{I(\lambda)}{f(\lambda, \theta)} d\lambda + \int_{-\pi}^{\pi} \log[f(\lambda, \theta)] d\lambda \right\} \quad (3)$$

En la práctica, el cálculo del estimador de Whittle se realiza eligiendo un parámetro de escala adecuado, θ_1 , tal que

$$f(\lambda, \theta) = \theta_1 f(\lambda, \theta^*) = \theta_1 f^*(\lambda, \eta) \quad (4)$$

de forma que anule el segundo sumando de (2), es decir

$$\int_{-\pi}^{\pi} \log[f(\lambda, \theta^*)] d\lambda + = \int_{-\pi}^{\pi} \log[f^*(\lambda, \eta)] d\lambda = 0 \quad (5)$$

donde $\eta = (\theta_2, \theta_3, \dots, \theta_M)$ y $\theta^* = (1, \eta)$

En [26] se demuestra que el parámetro de escala es:

$$\theta_1 = \frac{\sigma_e^2}{2\pi} \quad (6)$$

donde σ_e^2 es el error mínimo de predicción (MSPE).

En [10] se propone que la versión discreta del estimador de Whittle aproxima a (3) mediante una suma de Riemann en el rango de frecuencias dado por

$$\lambda_k = \frac{2\pi}{N} k \quad (7)$$

donde $k = 1, 2, \dots, N^*$ (N^* es la parte entera de $(N-1)/2$) con lo cual, la función a minimizar es la siguiente

$$\tilde{\mathcal{Q}}(\theta_1, H) = \frac{4\pi}{N} \left[\sum_{k=1}^{N^*} \frac{I(\lambda_k)}{\sum_{k=1}^{N^*} f(\lambda_k, \theta_1, H)} + \sum_{k=1}^{N^*} \log[f(\lambda_k, \theta_1, H)] \right] \quad (8)$$

Luego, con la elección del parámetro de escala adecuado se obtiene el parámetro H estimado, \hat{H} , valor que minimiza la expresión

$$\tilde{\mathcal{Q}}^*(H) = \tilde{\mathcal{Q}}(1, H) = \sum_{k=1}^{N^*} \frac{I(\lambda_k)}{\sum_{k=1}^{N^*} f(\lambda_k, 1, H)} = \sum_{k=1}^{N^*} \frac{I(\lambda_k)}{f^*(\lambda_k, H)} \quad (9)$$

donde se verifica que

$$f^*(\lambda, H) = \frac{1}{\theta_1} f(\lambda, \theta_1, H) = \frac{2\pi}{\sigma_e^2} \quad (10)$$

Luego, resaltan las siguientes desventajas del estimador de Whittle en sus formas convencionales:

1. La necesidad de conocer la forma paramétrica de la densidad espectral.
2. El alto tiempo de cálculo derivado de la ampliación de los métodos gráficos.

Se acepta que si se desconoce la forma exacta del espectro, el estimador puede tener un elevado sesgo que, además, puede no ser detectado al no estar disponible una salida

gráfica. Sin embargo, este problema puede ser aliviado mediante el uso del estimador de Whittle agregado.

Desde el punto de vista de la aplicación del estimador de Whittle a procesos donde no es posible asegurar nada acerca de su densidad espectral, resulta de enorme interés el Teorema del límite central para procesos autosimilares, producto que es una buena aproximación para series no gaussianas, hecho que permite aplicar todos los resultados obtenidos para procesos gaussianos autosimilares puros a series agregadas [27].

Resulta de gran interés el hecho de que este teorema permita suponer que para una serie temporal de tamaño N , que presenta una función de autocorrelación que exhibe caída hiperbólica con LRD, si m y N/m son lo suficientemente grandes y la varianza es finita, un proceso FGN sea una buena aproximación para las secuencias agregadas de la serie, inclusive en el caso que no sean gaussianas [27].

Precisamente en todo lo anterior se basa la modificación del estimador de Whittle conocida como estimador de Whittle agregado, la cual provee una forma de hacer más robusto y menos sesgado el estimador de Whittle cuando no se dispone de información acerca de la forma paramétrica exacta de la densidad espectral. Este método puede ser usado si la serie es lo suficientemente larga y para ello se agregan los datos para obtener una serie más corta $X_k^{(m)}$. De esta forma, se tiene que,

$$X_k^{(m)} = \frac{1}{m} \sum_{i=km}^{(k+1)m-1} X_i \quad 0 \leq k \leq [N/m] \quad (11)$$

y luego se aplica sobre la serie anterior dada por (11) el estimador de Whittle, tomando como modelo paramétrico de su densidad espectral el de FGN.

Sin embargo, y a pesar del hecho de que al usar una serie más corta se reduce el costo computacional asociado, el inconveniente que se presenta ahora es que aumenta la varianza del estimador y con ello también la del grado de autosimilaridad y, consecuentemente, disminuye el grado de representatividad del patrón.

Otro inconveniente del método es la imposibilidad de conocer *a priori* el valor de m apropiado. Pero, en este sentido, en [1] se expone un método para representar las estimaciones de H obtenidas para distintos valores de m , y buscar una región donde la gráfica se muestre aproximadamente plana. En [28] se muestran resultados para series de más de 10.000 muestras, y se estima que el valor $m=N/200$ parece ser el que entrega el mejor resultado en términos del compromiso entre el sesgo y la varianza del estimador.

En [29] se expone otra visión. Se propone el estimador local de Whittle, que a diferencia del estimador de Whittle, es un estimador semiparamétrico, puesto que especifica sólo la forma paramétrica de la densidad espectral para frecuencias próximas a cero, es decir,

$$f(\lambda) \sim G |\lambda|^{1-2H} \quad (12)$$

cuando $\lambda \rightarrow 0$

El método comparte características de otros estimadores basados en periodogramas. Sin embargo, sólo supone la forma del espectro para frecuencias próximas a cero, lo cual en la práctica se traduce en minimizar una versión modificada de la función $Q(H)$ dada por (2). Luego, de (12) sustituyendo $f(\lambda)$ por $f(\lambda, H)$ en (2), e integrando sólo hasta la frecuencia $2\pi M/N$, donde $1/M+M/N \rightarrow 0$ cuando $N \rightarrow \infty$, se obtiene que

$$Q(G, H) \triangleq \frac{1}{M} \sum_{j=1}^M \left(\frac{I(\lambda_j)}{G \lambda_j^{1-2H}} + \log(G \lambda_j^{1-2H}) \right) \quad (13)$$

reemplazando la constante G por su estimación, dada por

$$\hat{G} = \frac{1}{M} \sum_{j=1}^M \frac{I(\lambda_j)}{\lambda_j^{1-2H}} \quad (14)$$

se obtiene la función a minimizar, es decir,

$$R(H) \triangleq Q(\hat{G}, H) - 1 \quad (15)$$

De forma equivalente

$$R(H) \triangleq \log \left(\frac{1}{M} \sum_{j=1}^M \frac{I(\lambda_j)}{\lambda_j^{1-2H}} \right) - (2H-1) \frac{1}{M} \sum_{j=1}^M \log(\lambda_j) \quad (16)$$

En [29] se demuestra que el valor estimado para H , que se obtiene con este método, converge en probabilidad al valor real de H , y que el estimador es asintóticamente normal.

Sin embargo, el problema de fondo persiste, por cuanto la elección del valor de M resulta fundamental, y de él dependen el sesgo y la varianza, y nuevamente aparece el amplio margen de compromiso entre sesgo y varianza. En la medida en que se aumente M , el valor estimado para H converge más rápidamente al valor real de H pero, en cambio, la forma del espectro se apartará cada vez más de (9) y los efectos de la SRD serán mayores, por lo que el sesgo aumentará. También, al igual que en el método

anterior se debe optar por representar el valor estimado de H frente a M y encontrar la región más aproximadamente plana que se muestre en la gráfica. En [28], empleando series de 10.000 muestras, se informa sobre un valor de M tal que $M=N/32$ se considera como una buena opción para alcanzar un compromiso aceptable entre sesgo y varianza.

Como último comentario cabe señalar que a diferencia de Whittle agregado, en el caso de Whittle local, encontrar la zona plana requiere de una mayor cantidad de pruebas de ensayo y error, resultando por tanto imposible extraer conclusiones sobre un intervalo concreto que dependa en gran medida de la presencia de SRD.

PROPIUESTA DE UN ESTIMADOR EFICIENTE

Todas las variantes del estimador de Whittle expuestas requieren la minimización de una expresión: (9) o (16). Una forma simple de realizar tales minimizaciones consiste en evaluar dichas expresiones para un cierto número de valores de H equidistantes, q , que dependerá, por cierto, de la resolución deseada. Sin embargo, puede observarse que basta un número de muestras moderadamente elevado para que el algoritmo implique un costo computacional muy elevado. Para reducir este costo, se propone una minimización por medio de un algoritmo que reduzca el número de puntos a evaluar. Para ello se aprovechará el hecho de que la función a minimizar es convexa en todo el dominio $[0.5, 1]$ y por lo tanto posee un mínimo único. Así, un método de búsqueda por biseción aplicado sobre la derivada de la función permite que el número de puntos evaluados esté sólo en torno a $\log_2(q)$. Considerando el supuesto de que el cálculo de la derivada en un punto H_i puede ser aproximado por un coeficiente de diferencias para un incremento suficientemente pequeño, h , es decir:

$$Q'(H_i) \approx \frac{Q(H_i + h) - Q(H_i)}{h} \quad (17)$$

para $h \rightarrow 0$

se deriva el hecho de que la principal hipótesis de trabajo dice relación con establecer un grado de autosimilitud basado en el estimador de Whittle, sin embargo, por las complicaciones que presenta el compromiso entre el sesgo y la varianza que subyace a todos los modelos, se plantea que en un espectro reducido de puntos es posible entregar una respuesta con un buen nivel de compromiso.

Luego, para la obtención de intervalos de confianza, partiendo del supuesto de que la función a minimizar es convexa en el dominio $[0.5, 1]$, es decir que posee un único mínimo, se tendrá en cuenta que si \hat{H} es el valor del parámetro H que minimiza la función $R(H)$ y H_0 es su valor real, entonces,

$$\hat{H} \rightarrow H_0 \quad (18)$$

$$(\hat{H} - H_0) \rightarrow N(0, \sigma_H) \quad (19)$$

cuando $N \rightarrow \infty$

siendo

$$\sigma_H^2 = \frac{4\pi}{N} \left\{ \int_{-\pi}^{\pi} \left[\frac{\partial \log[f(\lambda, H)]}{\partial H} \right]^2 d\lambda \right\}_{H=\hat{H}}^{-1} \quad (20)$$

Luego, el cálculo de la derivada dada por (17) puede ser aproximado mediante un coeficiente de diferencias para un incremento suficientemente pequeño, h . De esta forma

$$\sigma_H^2 = \left\{ \sum_{k=1}^{N^*} \left[\frac{\partial \log[f(\lambda_k, \theta_1, H)]}{\partial H} \right]^2 \right\}_{H=\hat{H}}^{-1} \quad (21)$$

de donde finalmente se verifica que,

$$\sigma_H^2 = \left\{ \sum_{k=1}^{N^*} \left\{ \frac{\log[f(\lambda_k, \theta_1, \hat{H} + h)]}{\partial H} - \frac{\log[f(\lambda_k, \theta_1, \hat{H})]}{\partial H} \right\}^2 \right\}_{H=\hat{H}}^{-1} \quad (22)$$

CONCLUSIONES

Se han presentado los fundamentos de un proyecto de investigación sobre modelado de redes de computadoras con mecanismo de control de acceso al medio regido por el estándar IEEE 802.3, empleando los postulados de la teoría de conjuntos autosimilares para establecer el nivel de impacto que poseen las correlaciones temporales con dependencia de largo alcance (LRD) sobre su rendimiento. Además, se ha postulado un nuevo método para estimar grados de autosimilaridad sobre la base de una variación del estimador local de Whittle.

Los modelos tradicionales basados en procesos de Poisson o, de manera más general, en procesos de dependencia de corto alcance (SRD) no son incapaces de describir el comportamiento de las actuales redes de datos de alta velocidad, en particular de las redes Ethernet

comutadas según el estándar IEEE 802.3 en su revisión del año 2005 (IEEE 802.3-2005). Consecuentemente, es necesario un replanteamiento del estudio de los sistemas de carga de tráfico, considerando procesos de entrada autosimilares, producto de que la demanda de tráfico autosimilar impone nuevos requerimientos en el diseño de las redes que no pueden ser caracterizados empleando modelos de tráfico que consideren procesos con memoria de corto alcance, en especial en lo que a estrategias de buffering se refiere.

Los métodos tradicionalmente empleados para evaluar el estimador de Whittle presentan desventajas derivadas de la necesidad de conocer la forma paramétrica de la densidad espectral y de un alto costo computacional, producto de la aplicación intensiva de métodos gráficos. Se estima que estos inconvenientes pueden ser salvados en la medida en que pueda ser factible introducir un algoritmo que reduzca el número de puntos a evaluar, lo cual no sólo incide en un menor costo de procesamiento computacional, sino que habilita una nueva alternativa para ser considerada en el estudio que posee sobre las prestaciones de una red, la consideración de tráfico con características autosimilares o fractales.

Se propone que una función para realizar la minimización de la función generatriz del estimador de Whittle que sea convexa en el dominio acotado de interés [0.5, 1], presenta la característica fundamental de poseer un solo mínimo, completamente individualizable, y por lo tanto la aplicación de un método de búsqueda por bisección sobre la derivada de la función permitiría determinar un punto en torno al cual fluctúen todos los valores. Y esto es, en definitiva, contar con una región plana en la cual el valor del parámetro H sea perfectamente aproximado por un coeficiente de diferencias, entregando una simple solución al problema fundamental de su valoración con un intervalo de confiabilidad, y todo dentro del dominio de interés.

Para el desarrollo de las simulaciones se está trabajando en la búsqueda, primeramente, de literatura que exponga de forma clara técnicas de simulación y análisis estadísticos para series con dependencia de largo alcance, pues se considera que no basta con inducir resultados a partir de agregaciones, sino que debe existir una formalización de los procedimientos para estudiar concretamente dichos fenómenos. Esta es una área que requiere atención urgente, pues de ella depende en gran medida no tan solo el costo operacional en materia de capacidad computacional, sino que la validez de muchos postulados futuros.

REFERENCIAS

- [1] W. Leland, M. Taqqu, W. Willinger and D. Wilson. "On the self-similar nature of ethernet traffic (extended version)". IEEE/ACM Trans. Netw. Vol. 2 Nº 1, pp. 1-15. February 1994.
- [2] P802.3ba 40 Gb/s and 100 Gb/s Ethernet Task Force. IEEE Std.802.3ba. 2007.
- [3] G. Ibáñez. "Contribución al diseño de redes de campus ethernet autoconfigurables". Ph.D. Thesis. Departamento de Ingeniería Telemática. Universidad Carlos III de Madrid. Madrid, España. 2005.
- [4] IEEE standard for information technology, telecommunications and information exchange between systems local and metropolitan area networks. IEEE Std. 802.3. 1985.
- [5] IEEE standard for local and metropolitan area networks. IEEE Std. 802.1Q. 2005.
- [6] J. García, S. Ferrando and M. Piattini. "Redes para proceso distribuido". Ra-Ma, pp. 127-160. Madrid, España. 1997.
- [7] C. Zacker. "Redes. Manual de referencia". McGraw-Hill, pp. 275-341. Madrid, España. 2002.
- [8] S. Halabi. "Metro Ethernet. The definitive guide to enterprise and carrier metro ethernet applications". Cisco Press. Indianápolis, EEUU, p. 1. 2003.
- [9] P.M. Robinson. "Gaussian semiparametric estimation of long-range dependence". Annals of Statistics. Vol. 23 Nº 5, pp. 1630-1661. October 1995.
- [10] J. Geweke and S. Porter-Hudak. "The estimation and applications of long memory time series models". Journal of Time Series Analysis. Vol. 4 Nº 4, pp. 221-238. July 1983.
- [11] W. Stallings. "Redes e Internet de alta velocidad: Rendimiento y calidad de servicio". Prentice Hall. 2nd Ed., pp. 224-225. Madrid, España. 2004.
- [12] K.W. Fendick, V.R. Saksena and W. Whitt. "Dependence in packet queues". IEEE Trans. Commun. Vol. 37 Nº 11, pp. 1173-1183. November 1989.
- [13] D.P. Heyman and T.V. Lakshman. "Source models for VBR broadcast-video traffic". IEEE/ACM Trans. Netw. Vol. 4 Nº 1, pp. 40-48. February 1996.
- [14] L. Kleinrock. "Communication Nets: Stochastic Message Flow and Delay". McGraw-Hill. New York. 1972.
- [15] V. Paxson and S. Floyd. "Wide-area traffic: The failure of Poisson modeling". Netw. IEEE/ACM Trans. Vol. 3 Nº 3, pp. 266-244. June 1995.
- [16] S. Klivansky, S. Mukherjee and C. Song. "Factor contributing to self-similarity over NFSNet". Georgia Institute of Technology. 1995.

- [17] D.E. Duffy, A.A. McIntosh, M. Rosenstein and W. Willinger. "Statistical analysis of CCSN/SS7 traffic data from working CCS subnetworks". IEEE J. Sel. Areas Commun. Vol. 12 Nº 3, pp. 544-551. April 1994.
- [18] M.E. Crovella and A. Bestavros. "Self-similarity in world wide web traffic: evidence and possible causes". IEEE/ACM Trans. Netw. Vol. 5 Nº 6, pp. 835-846. December 1997.
- [19] M.W. Garret and W. Willinger. "Analysis, modeling and generation of self-similar VBR video traffic". Proc. SIGCOMM, pp. 269-280. 1994.
- [20] J. Beran, R. Sherman, M. Taqqu and W. Willinger. "Long-range dependence in variable-bit-rate video traffic". IEEE Trans. Commun. Vol. 24 Nº 2, pp. 1566-1579. 1995.
- [21] I. Norros. "A storage model with self-similar input". Queueing Systems. Vol. 16 Nº 3-4, pp. 387-396. 1994.
- [22] N. Likhanov, B. Tsybakov and N. D. Georganas. "Analysis of an ATM buffer with self-similar ("fractal") input traffic". Proc. IEEE INFOCOM, pp. 985-992. Boston, Massachusetts, USA. April 1995.
- [23] A. Elwaid and D. Mitra. "Effective bandwidth of general markovian traffic sources and admission control of high-speed networks". IEEE/ACM Trans. Netw. Vol. 1 Nº 3, pp. 329-343. June 1993.
- [24] K. Sriram and W. Whitt. "Characterizing superposition arrival processes in packet multiplexers for voice and data". IEEE J. Sel. Areas in Commun. Vol. 4 Nº 6, pp. 833-846. September 1986.
- [25] H. Heffes and D. Lucantoni. "A Markov modulated characterization of packetized voice and data traffic and related statistical multiplexer performance". IEEE J. Sel. Areas Commun. Vol. 4 Nº 6, pp. 856-868. September 1986.
- [26] J. Beran. "Statistics for long-memory processes". In Generalized Additive Models (Monographs on Statistics and Applied probability), Chapman & Hall/CRC. New York, USA. 2000.
- [27] R.J. Adler, R.E. Feldman and M.S.Taqqu. "A practical guide to heavy tails: statistical techniques and applications". Birkhäuser. Boston, Massachusetts, USA. 1998.
- [28] M.S. Taqqu and V. Teverovsky. "Robustness of Whittle-type estimators for time series with long-range dependence". Stochastic Models. Vol. 13 Nº 4, pp. 723-757. 1997.
- [29] P. Robinson. "Log-periodogram regression of time series with long range dependence". Annals of Statistics. Vol. 23 Nº 3, pp. 1048-1072. June 1995.