

Estudio sobre la Distribución de Tráfico Autosimilar de Redes Wireless 802.11

Santiago Pérez, Higinio Facchini, Gustavo Mercado, Luis Bisaro
GRID TICs (Grupo UTN de Investigación y Desarrollo en TICs
Facultad Regional Mendoza, Universidad Tecnológica Nacional
Rodríguez 273, Mendoza, Argentina
0261-5244576

{santiagocp, higiniofac, gmercado, lbisaro}@frm.utn.edu.ar

Resumen

El análisis de colas ha sido y es de enorme utilidad para los diseñadores de redes y analistas de tráfico, a efectos de planificar las capacidades de los elementos de red y predecir su rendimiento. Además, los modelos exactos del tráfico Ethernet y Wireless 802.11 son importantes para modelar las aplicaciones de capa superior y los buffers de memoria de switches. Estos análisis dependen de la naturaleza Poisson del tráfico de datos. Sin embargo, muchos resultados predichos a partir del análisis de colas difieren significativamente del rendimiento observado en la realidad. Diversos estudios han demostrado que para algunos entornos el patrón de tráfico es autosimilar, en lugar de Poisson. Este concepto está relacionado con otros más conocidos como son los fractales y la teoría del caos. Desde principio de los años 90 se comenzaron a publicar documentos referidos a la autosimilitud del tráfico de Ethernet.

En esta investigación se analizará la distribución intertrama de secuencias de tráfico Wireless 802.11 de contenido diverso. Usando capturas de tráfico Wireless 802.11 realizadas en distinto momento, se analizará la distribución Pareto como posible modelo, y comparará la bondad de ajuste de este posible modelo de distribución, y de las muestras entre sí usando el estadístico Kolmogorov-Smirnov.

Palabras claves: tráfico, autosimilitud, Ethernet, Wireless 802.11, Pareto, bondad de ajuste

Contexto

La línea de investigación está inserta en el proyecto PID “Modelación del rendimiento de WLAN con diferentes tipos de nodos usando Redes de Petri”, en el ámbito del Grupo UTN GRID TICs (GRUPO UTN DE INVESTIGACION Y DESARROLLO EN TICs), del Departamento Ingeniería en Electrónica, de Facultad Regional Mendoza, de la Universidad Tecnológica Nacional, aprobado por el Consejo Directivo con resolución n° 898/2009, y bajo el número UTN1193 de la UTN, número 25/J070 del Ministerio de Educación, Ciencia y Tecnología para el programa de incentivos, y aprobado y financiado por resolución 71/2010 de la Secretaría de Ciencia y Técnica de Rectorado de la UTN, desde enero/2010 hasta la fecha.

Introducción

Para que las redes de comunicaciones puedan proveer un espectro creciente de servicios requiere técnicas eficientes de análisis, monitoreo, evaluación y diseño de las mismas. El análisis es tradicionalmente encarado de forma incompleta, y siempre, con crecientes demandas de los usuarios e incertidumbre acerca de la evolución de los sistemas de red. Por ello, para cumplir los requerimientos de los usuarios y para proveer garantías sobre la confiabilidad y adaptabilidad, deben desarrollarse modelos de sistemas que puedan capturar las características de la carga de red real y conducir a predicciones precisas de

performance del sistema, en una cantidad razonable de tiempo.

El análisis de tráfico es un componente vital para entender los requerimientos y capacidades de una red. En los años pasados han sido propuestos incontables modelos de tráfico para entender y analizar las características de tráfico de las redes. No obstante, no hay un simple modelo de tráfico que pueda eficientemente capturar las características de tráfico de todos los tipos de redes, bajo cada circunstancia posible. Consecuentemente, el estudio de los modelos de tráfico para identificar eventualmente el mejor modelo es una tarea crucial. La modelación correcta de tráfico es también un requerimiento básico para la planificación de la capacidad exacta.

Si los modelos de tráfico subyacente no capturan eficientemente las características del tráfico real, el resultado puede ser una subestimación o sobreestimación de la performance de la red. Esto perjudicaría totalmente el diseño de la red. Los modelos de tráfico son por lo tanto, un componente central de cualquier evaluación de performance de las redes, y ellos necesitan ser muy exactos.

Los modelos de tráfico se analizan y comparan en base al número de parámetros requeridos para describir el modelo, la tratabilidad, la estimación de parámetros, y como mejor el modelo representa el tráfico real, conocido como bondad de ajuste. Por ello, para evaluar la bondad de ajuste, deberían definirse métricas que permitan cuantificar cuan cerca el modelo esta a los datos reales. Las métricas definidas deberían también estar directamente relacionadas a las medidas de performance que van a ser predichas desde el modelo de tráfico.

Modelos de Tráfico de Red

Los modelos de tráfico permiten a los diseñadores de red hacer suposiciones acerca de las redes a diseñar, basados sobre la experiencia pasada, y también la predicción de performance para requerimientos futuros. Los modelos de tráfico son usados de dos

formas fundamentales: 1) Como parte de un modelo analítico, ó 2) Para activar una Simulación de Eventos Discretos. Algunos modelos característicos son el Modelo de Distribución Poisson y el Modelo de Distribución Pareto.

Uno de los modelos de tráfico más ampliamente usado y antiguo es el Modelo de Poisson. La distribución de Poisson sin memoria es el modelo predominante usado para analizar tráfico en redes de telefonía tradicional, y por extensión en las redes de datos [1]. En un proceso Poisson, los tiempos interarribos están distribuidos exponencialmente con un parámetro frecuencia λ :

$$f(k; \lambda) = \frac{e^{-\lambda} \lambda^k}{k!},$$

La distribución tiene una media y una varianza igual al parámetro λ . La distribución Poisson puede visualizarse como una forma limitada de la distribución binomial, y también es usada ampliamente en los modelos de colas. La función de distribución de probabilidad y la función de densidad del modelo están dadas por:

$$F(t) = 1 - e^{-\lambda t}$$

$$f(t) = \lambda e^{-\lambda t},$$

El proceso de distribución Pareto, que es discutido en detalle posteriormente, produce tiempos interarribos independientes e idénticamente distribuidos. Una característica importante del modelo es que la distribución Pareto, del tipo hiperbólico, tiene valores grandes de varianza, respecto a Poisson, y varianza y media infinitas para ciertos valores de sus parámetros.

Aunque no hay un simple modelo que pueda ser usado efectivamente para modelar tráfico en todas las clases de redes, puede mostrarse que el modelo Poisson no responde bien para ciertos tipos tráfico. Por el contrario, por ejemplo, en casos de redes de alta velocidad, con demanda no esperada sobre las transferencias de paquetes, los modelos de

tráfico basados en Pareto son excelentes candidatos, dado que el modelo tiene en cuenta la correlación de largo alcance en los tiempos de arribo.

Trafico de Datos Ethernet Autosimilar con Distribución Pareto

El artículo fundamental del estudio de los datos de tráfico autosimilar es «On the Self-Similar Nature of Ethernet Traffic» (La naturaleza autosimilar del tráfico de Ethernet), que posteriormente sería corregido y aumentado en [2]. Este documento contradujo la idea de que un simple análisis de colas basado en la suposición de que el tráfico fuera de Poisson pudiera modelar adecuadamente todo tráfico de red. Empleando una masiva cantidad de datos y un cuidadoso análisis estadístico, el artículo manifiesta que, para el tráfico de Ethernet, se requiere un nuevo planteamiento de modelado y de análisis.

Esto es, el tráfico Ethernet muestra cierta cantidad de picos o ráfagas. De este modo, el tráfico de Ethernet posee un aspecto similar para escalas grandes (horas y minutos) y para escalas pequeñas (segundos y milisegundos). Este tráfico autosimilar es muy diferente de lo que se observa en el tráfico telefónico, y en los modelos estocásticos usando Poisson de los análisis y diseño de redes de datos.

En una simulación Poisson, con alta resolución, el tráfico tiene bastantes picos. Pero a medida que se van agregando los datos mediante escalas de tiempo cada vez mayores, el patrón de tráfico se suaviza. Por lo tanto, es de esperar que la varianza de los datos se reduzca por cada nivel, a diferencia de lo que sucede en un tráfico autosimilar, como es Ethernet.

Por ello, en las simulaciones se prefiere modelar los periodos de tiempo de tráfico, con distribuciones de varianza infinita, utilizando en particular la distribución de Pareto. Esto da como resultado una distribución de elevada varianza, con muchas ráfagas muy cortas, muchas ráfagas largas y algunas ráfagas muy largas. Esto ha permitido determinar el origen de las discrepancias, por

ejemplo, entre el tiempo real de espera y el tiempo estimado de espera obtenidos mediante el uso de la teoría de colas convencional usando Poisson.

En [3] se describe un estudio realizado por los autores del presente artículo, que verifica tal comportamiento, y es usado como antecedente para el desarrollo del presente trabajo, pero sobre ambientes 802.11. Y además, para efectuar la comparación de los parámetros de distribución Pareto característicos de Ethernet y Wireless 802.11. El presente trabajo se desarrollará sucesivamente con los siguientes aspectos: 1) Introducción General, 2) Modelos de Tráfico de Red, 3) Tráfico de Datos Ethernet Autosimilar con Distribución Pareto, 4) Caso de Estudio experimental de tráfico Wireless 802.11, 5) Análisis con distribución Pareto, 6) Análisis de bondad de ajuste con la prueba de Kolmogorov-Smirnov, y 7) Comparación de los Parámetros y la curva de Distribución Pareto de tráfico Ethernet y Wireless 802.11.

Caso de Estudio Experimental de Tráfico Wireless 802.11

Como en el trabajo publicado previamente [3], para este caso de estudio se adopta la distribución Pareto, considerándola más apropiada para modelar el tráfico autosimilar wireless 802.11. Y se pretende demostrar que a partir de dos muestras Wireless 802.11 distintas, las mismas independientemente responden al modelo Pareto, y además, establecer el valor de los parámetros de dicha distribución. Por otro lado, que las muestras wireless definen comportamientos similares. Es decir, que el tráfico mantiene su comportamiento general más allá del tiempo.

Colección de las muestras o trazas

Los métodos de colección de tramas de red Ethernet y Wireless, son el punto de partida para el entendimiento del comportamiento del tráfico y de los nodos de red. Se ha avanzado en las primeras fases, utilizando el programa Wireshark [4] (ex Ethereal) combinado con la herramienta de captura AIRPCAP NX: USB

802.11 a/b/g/n [5]. Y simultáneamente, con fines de comprobación y para obtener otros datos complementarios, se ejecutó el sniffer wireless AirMagnet [6]. Ambos tienen una interfaz flexible con opciones muy ricas de filtrado (Figuras n° 1 y n° 2).

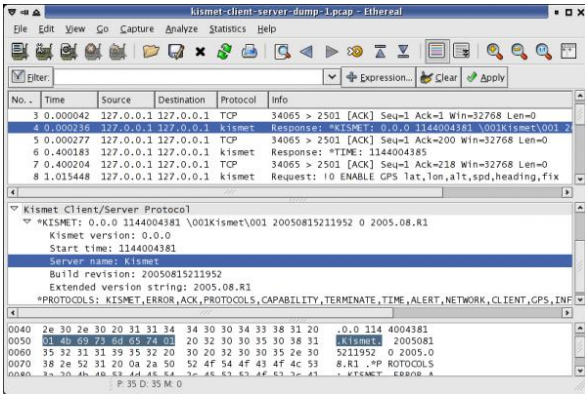


Figura n° 1 Sniffer Wireshark

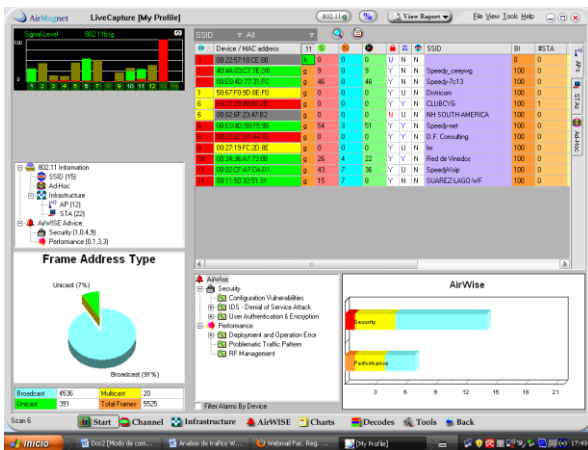


Figura n° 2 Sniffer AirMagnet

Con Wireshark, se tomaron 2 muestras wireless 802.11 de 10 minutos cada una, con una diferencia de 20 minutos entre sí, sobre un ambiente administrativo en producción, con varios APs, seleccionando uno de los canales habilitados, con el detalle dado en la Tabla n° 1.

Procesamiento de las muestras

En el procesamiento, las tramas se identificaron por su desplazamiento de tiempo entre sí, y sin los campos innecesarios de la muestra para este estudio. Luego, se exportaron como un vector al programa

Matchcad [7], para proceder a su análisis estadístico.

	Horario de toma	N° de Tramas	% Uni cast	% Broad cast	% Multi cast
A	10:00 hs	95970	49%	49%	2%
B	10:20 hs	196400	55%	39%	4%

Tabla n° 1 Muestras de Tráfico

Las tramas se agruparon en un vector A y en un vector B según las muestras, de 61 elementos para cada caso, en intervalos de 0.9 mseg, y de 1.8 mseg, respectivamente. Y se ordenaron en forma decreciente según la cantidad de tramas.

Posteriormente, estos elementos se normalizaron dividiendo el vector A y B por el número total de tramas de cada muestra, dando origen a nuevos vectores a los que se le llamará ObservA y ObservB (vectores de los valores observados), a los fines del resto del proyecto (Figura n° 3).

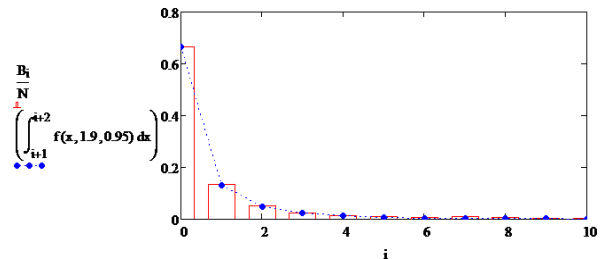


Figura n° 3 Vector ObservA

Resultados y Objetivos

En este trabajo, se pretende relacionar los temas de autosimilitud, con el tráfico Wireless 802.11, la distribución de Pareto y la prueba de Kolmogorov-Smirnov, con el mismo criterio que se hizo sobre tráfico Ethernet. El volumen de los trabajos y literatura sobre tráfico de datos es creciente, y el tema de la autosimilitud ha significado el principio de un nuevo examen del rendimiento del tráfico de datos, las técnicas de modelado, control de tráfico, dimensionamiento de buffers en dispositivos activos, entre otros. En el trabajo

se buscar verificar a través de un estudio experimental y usando la prueba de Kolmogorov-Smirnov, que el tráfico de datos Wireless 802.11 responde efectivamente a la distribución Pareto, estableciéndose los valores de los parámetros de dicha distribución. Y además, demostrar que dos muestras de tráfico Wireless 802.11, tomadas en distintos momentos tienen distribución Pareto con parámetros prácticamente similares.

Finalmente, se pretende efectuar, en base al estudio realizado sobre ambientes Ethernet [3], la comparación de los parámetros Pareto en Ethernet y en Wireless 802.11.

Líneas de investigación y desarrollo

El proyecto está direccionado hacia el Análisis de Tráfico de Red, como eje temático. El énfasis es en el análisis del comportamiento de las redes Ethernet y Wireless 802.11, para favorecer la mayor exactitud posible en los modelos de simulación de tráfico de red. Los autores experimentan actualmente un modelo de nodo Wireless 802.11e sobre el simulador Möbius [9].

Formación de Recursos Humanos

El equipo de trabajo está integrado por docentes investigadores, y becarios graduados y alumnos del Grupo GRID TICs (Grupo UTN de Investigación y Desarrollo en TICs) de la Universidad Tecnológica Nacional, Facultad Regional Mendoza. Entre los integrantes, hay en curso un tesis doctoral, una de magister y una de grado, relacionadas con la línea general de investigación.

Referencias

[1] Frost V. S. Frost y Melamed B., "Traffic Modeling for Telecommunications Networks", IEEE Communications, Mar. 1994, <http://ieeexplore.ieee.org/iel1/35/6685/00267444.pdf>

[2] Leland, W; Taqqu, M; Willinger, W. y Wilson, D; "On the Self-Similar Nature of Ethernet Traffic", IEEE/ACM Transactions on Networking, Febrero 1994

[3] Perez, S; Facchini, H; Mercado, G. y Bisaro, L.; "Análisis del Comportamiento Autosimilar del tráfico Ethernet de las Redes de Datos", CACIC 2010, 2010.

[4] <http://www.wireshark.org>

[5] http://www.cacetech.com/products/airpcap_nx.html

[6] <http://www.flukenetworks.com/enterprise-network/wireless-network/airmagnet-wifi-analyzer>

[7] <http://www.ptc.com/products/mathcad>

[8] <http://www.um.edu.ar/math/estadis/programa.htm>

[9] <https://www.mobius.illinois.edu>