

NATURALEZA FRACTAL DEL TRAFICO INTERNET

Ricardo Lent, Peter Yamakawa

Universidad Nacional de Ingeniería

Facultad de Ingeniería Eléctrica y Electrónica, Sección de Post-Grado.

RESUMEN

En este artículo mostramos los resultados del análisis de 18 millones de datagramas Internet recolectados entre Abril de 1996 y Junio de 1997 en una de las redes de la Red Científica Peruana. Se demuestra que el flujo de paquetes TCP/IP presentan una naturaleza asintóticamente Auto-similar (fractal) análoga a la encontrada en redes de área local Ethernet [1] y que difiere de la prevista por los modelos tradicionales de tráfico basados en procesos de Markov y semi-Markov como Poisson.

ABSTRACT

This article shows the results of the analysis performed in 18 Millions Internet datagrams collected between April 1996 and June 1997 in a network of Red Científica Peruana. We demonstrate that the flow of TCP/IP packets present an asymptotic self-similar nature (fractal) like the one found in Ethernet local area networks [1]. Such a fractal behaviour differs from these derived by traditional traffic models based in Markov and Semi-Markov processes like Poisson.

INTRODUCCIÓN

Internet, la red global de paquetes TCP/IP, se ha convertido en un medio esencial en el intercambio mundial de información. Su crecimiento acelerado y el rápido desarrollo de nuevas aplicaciones han generado un aumento considerable en su tráfico interno [2], [3].

El conocimiento adecuado del comportamiento del tráfico en Internet es esencial para evaluar, predecir y controlar correctamente el rendimiento de la red. Usualmente se evalúa el flujo de paquetes para determinar el rendimiento de la red, el cual está inversamente relacionado con el tiempo de llegada entre paquetes. Una característica presente en las redes reales de datos y que no ha podido ser capturada por los modelos tradicionales de tráfico, es que los paquetes aparecen en forma de ráfagas 4 ó 5 veces mayores al promedio habitual de paquetes por unidad de tiempo, lo que provoca elevados retardos y en al-

gunos casos pérdida de datagramas en los routers.

Los estudios de datos tomados en redes de comunicaciones reales efectuados hasta el momento han sido realizados con medidas en las capas inferiores y superiores del modelo de referencia de redes. No existen investigaciones sobre el comportamiento del tráfico real en niveles intermedios como el caso del nivel de red.

En [4] Jain demuestra que el proceso de llegada de paquetes de una red Token Ring no puede ser modelado correctamente con procesos de Poisson comparando las funciones de densidad de probabilidad de muestras tomadas a nivel físico en una red de área local en anillo con modelos analíticos. Paxson [5] realiza un estudio sobre el comportamiento de las aplicaciones más comunes en Internet determinando que la mayoría de ellas son difícilmente modeladas con procesos de Poisson. En estudios recientes [1], [6] usando tráfico Ethernet en redes de área local (LAN), se demuestra que los modelos estadísticamente Auto-si-

milares (comúnmente llamados fractales) pueden representar mejor las características estadísticas del tráfico Ethernet LAN.

En este trabajo, a partir de mediciones en la capa de red en una red Internet TCP/IP demostramos que el tráfico Internet heterogéneo en cuanto a su composición interna de protocolos presenta también auto-similaridad estadística.

MEDIDAS DE TRAFICO

Conceptualmente el protocolo TCP/IP puede ser estudiado usando un modelo referencial de cuatro capas funcionales: aplicación, transporte, red y física. La capa de aplicación es la que generalmente origina la mayor parte del tráfico en una red porque es la que recibe la información del usuario (SMTP, HTTP, FTP, etc.). La información es pasada a la capa de transporte que agrupa los bytes en segmentos y les agrega información de control en una cabecera. El segmento de transporte se traslada a la capa de red donde se divide en paquetes con longitud variable (datagramas), dependiendo de la capacidad del medio físico. La capa de red puede además generar datagramas de control (ICMP). Finalmente el datagrama pasa a la capa física para su envío por el medio de transmisión.

Nuestras medidas fueron tomadas en el nivel de red, usando una estación de trabajo Sun Sparestation 4 usando Solaris 2.5.1 leyendo el kernel del sistema operativo [6]. El programa nos permite registrar el tiempo en que el datagrama fue visto por el kernel con una precisión relativa al reloj interno de 1 ms. En el proceso de medición se desprecia el tiempo que existe entre la llegada del paquete en el interface de red y su servicio por la interrupción "nuevo paquete" del kernel.

Los paquetes registrados representan todo el tráfico TCP/IP de los servidores de la Red Científica Peruana con aplicaciones heterogéneas recibiendo (y generando) tráfico con Internet durante los tiempos de medición. Las condiciones en que se realizaron las observaciones están resumidas en la Tabla 1. En todos los casos el tráfico corresponde al horario diurno dentro de las horas de alto uso de la red (entre las 9am y 8pm).

La red en la que se realizaron las observaciones es

muy dinámica por lo que las tres muestras reflejan condiciones de tráfico diferentes por la presencia de nuevos servidores de aplicaciones.

	Fecha	Tiempo	#paq	Protocolos
ip1	18/4/96	6 h	3 M	SMTP/HTTP
ip2	16/9/96	6 h	5.8 M	SMTP/HTTP
ip3	9/6/97	4.5 h	10 M	SMTP/UUCP

Tabla 1: Resumen de las características de las mediciones.

En los dos primeros casos (ip1 e ip2) se tuvo una composición de tráfico similar: SMTP (49%), HTTP (12%), telnet (9%) y otros protocolos. Para el último caso (ip3) se obtuvo también predominancia de SMTP (31%) pero aparece una fuerte contribución de paquetes de origen UUCP (30%). El porcentaje restante estuvo compartido por FTP, X11, DNS, etc.

A pesar de que la distribución de densidad de probabilidad de las observaciones parece encajar con la de un proceso de Poisson (Fig. 1) existe una notable diferencia en el nivel de ráfagas que se presentan a diferentes escalas de tiempo.

La forma en que se presentan las ráfagas de paquetes en las tres muestras permanecen inalterables en diferentes escalas de tiempo. Los perfiles mostrados en la Fig. 2 corresponden a la secuencia de tráfico ip2.

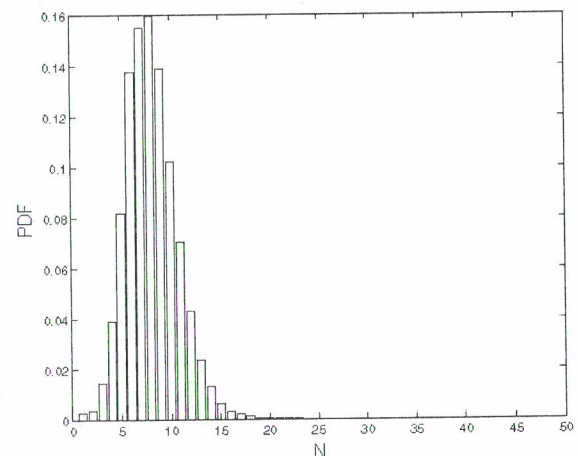


Fig. 1: Función de distribución de probabilidad de las muestras.

En un proceso de Poisson, la variación de las ráfagas tiende a disminuir a medida que se incrementa el tamaño de las unidades de tiempo cosa que no ocurre con las muestras reales [1].

PROCESOS AUTO-SIMILARES

Los procesos auto-similares son los modelos más simples de los procesos con dependencia en amplio rango (cuando un proceso estacionario tiene dependencia de amplio rango su función de autocorrelación $r(k)$ es no sumable y decae hiperbólicamente [7]).

En la Fig. 2 aparece una muestra de tráfico en cuatro escalas diferentes de tiempo (de arriba hacia abajo: unidades de tiempo correspondientes a 0.01s, 0.1s, 1s y 10s).

Para un proceso estocástico estacionario en el sentido amplio $X = (X_k : k=0,1, \dots)$ con media m , varianza σ^2 , y función de autocorrelación $r(k), k^3 \geq 0$; construimos $X^{(m)}$ una versión agregada, construida a partir de la división de X en bloques no superpuestos de m elementos secuenciales (con función de autocorrelación correspondiente $r_k^{(m)}$), donde cada elemento de $X^{(m)}$ se construye promediando m elementos:

$$X^{(m)}(k) = \frac{1}{m} X_{km-m+1} + \dots + X_{km} \quad ; t \geq 1 \quad (1)$$

El proceso X es exactamente Auto-similar de segundo orden con parámetro de Auto-similaridad (o persistencia) H ($0.5 < H \leq 1$) si los procesos agregados tienen la misma estructura de autocorrelación que X :

$$r^{(m)}(k) = r(k) \quad ; \forall m = 1, 2, \dots, k = 1, 2, \dots \quad (2)$$

Y la relación entre la varianza de la secuencia original con las nuevas secuencias se puede determinar por:

$$\sigma^2 = \text{Var}[X(k)] = \frac{1}{k^D} \text{Var}[X(k)] \quad (3)$$

donde:

$$H = 1 - \frac{D}{2} \quad (4)$$

Para un típico proceso estocástico ergódico, mientras m se incrementa, la autocorrelación de $X^{(m)}$ decrece con $D=1$ y la varianza del promedio en el tiempo cae a cero con tasa $1/k$. En cambio, para un proceso auto-similar, el proceso X y los procesos agregados $X^{(m)}$ se vuelven indistinguibles con respecto a sus funciones de autocorrelación, con una varianza que decae más lentamente que en un proceso ergódico ($0 \leq D < 1$).

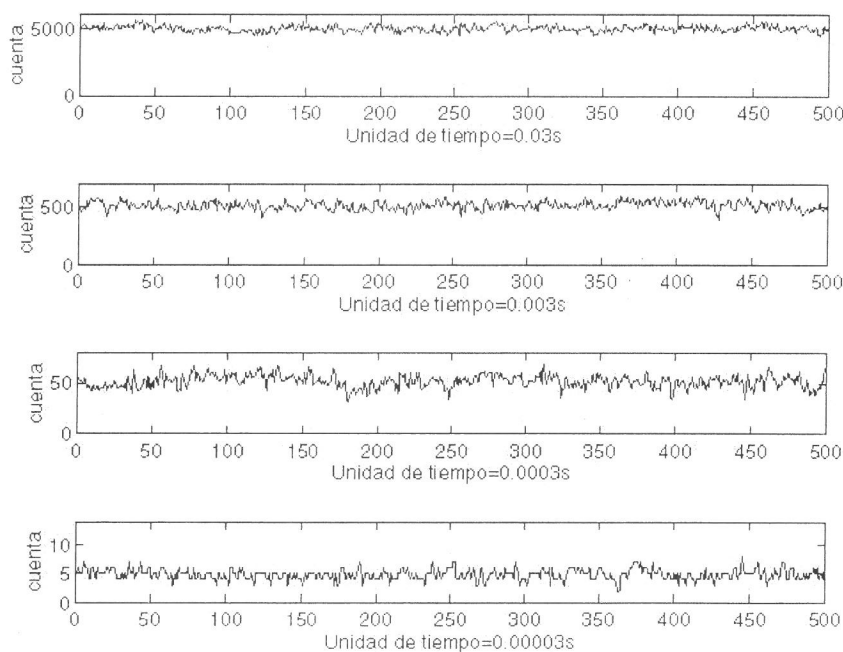


Fig. 2: El tamaño de las unidades de tiempo no modifica el "perfil" del tráfico real TCP/IP.

Una condición menos estricta es la siguiente: un proceso X es asintóticamente auto-similar si para todo k lo suficientemente grande:

$$r^{(m)}(k) \rightarrow r(k) \quad ; \quad \forall k \rightarrow \infty \quad (5)$$

Así, con esta condición de auto-similaridad, la autocorrelación del proceso agregado tiene la misma forma que el proceso original. Esto sugeriría que el grado de variabilidad, o grado de ráfagas, podría ser el mismo a diferentes escalas de tiempo.

ANÁLISIS DE LAS MUESTRAS

Podemos establecer que los tiempos de observación usados para medir el tráfico en la red TCP/IP son lo suficientemente grandes para afirmar que el proceso estudiado corresponde a un proceso estacionario en sentido amplio (WSS) dentro de los tiempos de medición. Se puede detectar esta propiedad analizando el comportamiento de la media y varianza con el número de muestras. En la Fig. 3 y Fig. 4 se aprecia la conducta de la media y la varianza para la muestra ip3. Ambos momentos estadísticos llegan a un nivel estable dentro del rango de las muestras tomadas. Para las otras dos muestras, el comportamiento es similar.

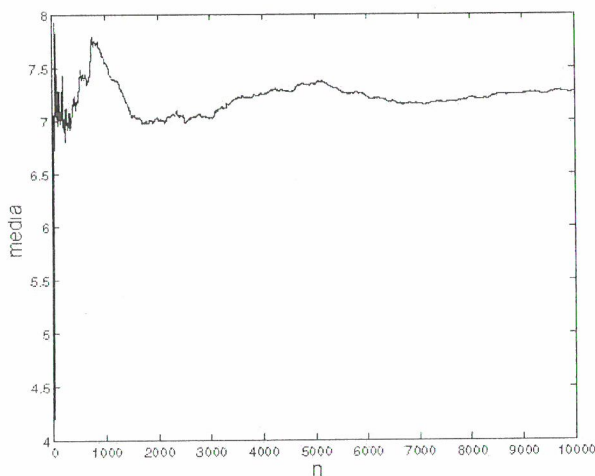


Fig. 3: Media de las muestras de tráfico TCP/IP

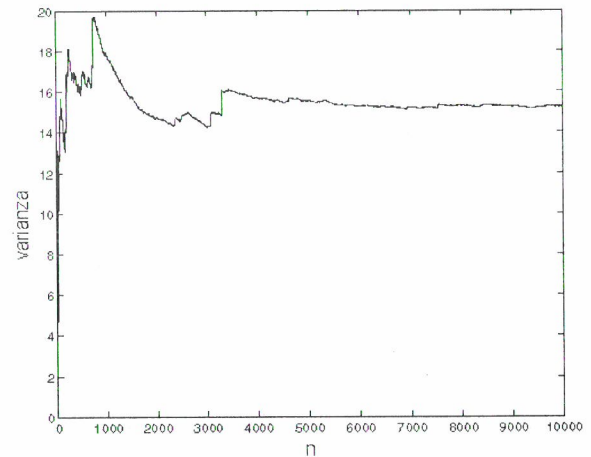


Fig. 4: Varianza de las muestras de tráfico TCP/IP.

Para establecer si las muestras obtenidas corresponden a procesos auto-similares aprovechamos la característica de los procesos estocásticos que presentan dependencia de amplio rango: la varianza de la media aritmética de estos procesos decrece más lentamente que el recíproco del tamaño de la muestra (ecuación 3). Si graficamos la varianza de los procesos agregados, esta decrece linealmente en un plano con ambos ejes logarítmicos y con pendiente inferior a -1.

$$\log(\text{Var}(X^{(m)})) \approx \log(c) - D \log(m) \quad (6)$$

A partir de la ecuación 6 puede determinarse que la pendiente de la relación entre la varianza y la agregación de las muestras corresponde a $-D$ usando ejes coordenados logarítmicos. Usando la relación entre D y H dada por la ecuación 4, puede estimarse el valor de la persistencia H .

La Fig. 5 muestra este caso para la muestra ip2. Ajustamos una línea con el método de los mínimos cuadrados a los puntos resultantes en el plano, ignorando los valores pequeños de k (período transitorio) así como los valores más altos, y obtenemos una pendiente de $D=0.27$, con lo que determinamos un parámetro de Hurst de $H=0.87$. Para fines comparativos, en la Fig. 5 también hemos graficado la relación dada para un proceso sintetizado de Poisson. La pendiente correspondiente en este caso se aproxima a $D=1$, con lo que

obtenemos un valor de 0.5 para la persistencia, lo que indica, la ausencia de correlación entre sus muestras. Esto era lo esperado dado que es una característica inherente a los procesos de Poisson,.

El estudio de las muestras ip1 e ip3 arrojan pendientes de 0.27 y 0.4, que corresponden a valores de 0.87 y 0.8 respectivamente para el parámetro de Hurst.

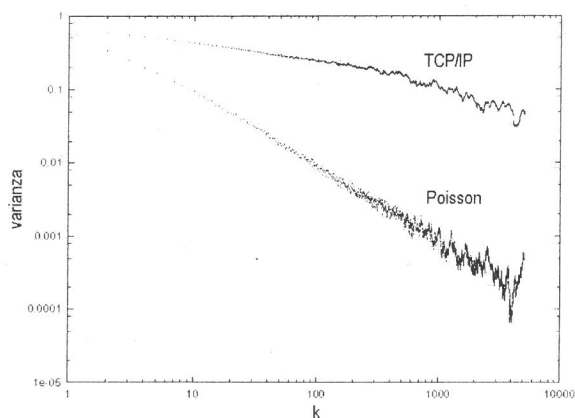


Fig. 5: Varianza vs. agregación de muestras para la muestra ip2 de tráfico real y un proceso de Poisson

Otra forma de estimar heurísticamente el grado de auto-similaridad de un proceso estocástico X es usar el análisis R/S propuesto por Hurst [8], que relaciona el rango del proceso con la desviación estándar de la muestra.

Para una muestra de n observaciones tomadas en el intervalo de tiempo entre t y t+n, denotemos X_k el valor de la muestra tomada en el tiempo t+k (o retardo k), $k=1,2 \dots n$, con media $E[X(n)]$ y varianza $S^2(n)$. El rango ajustado $R(n)/S(n)$ se calcula con [9], [10]:

$$\frac{R(n)}{S(n)} = \frac{1}{S(n)} [\max(0, W_1, W_2, \dots, W_n) - \min(0, W_1, W_2, \dots, W_n)] \quad (7)$$

donde $W_k = (X_1 + X_2 + \dots + X_k) - kE[X(n)]$, con $1 \leq k \leq n$. En la Fig. 6 graficamos la relación $\log(R(n)/S(n))$ vs. $\log(d)$ conocida como Diagrama Pox para la muestra ip3, donde d corresponde al número de particiones obtenidas de la secuencia original de N datos ($d=N/n$). Con el método de los mínimos cuadrados ajustamos

una línea, cuya pendiente corresponde directamente a un estimado del valor del parámetro de Hurst:

$$\text{Log}(Q) = a + H \log(k) \quad (8)$$

El diagrama típicamente comienza con una zona de transición representando la naturaleza de dependencia de corto rango de la muestra, pero eventualmente tiende a estabilizarse y fluctuar sobre una línea recta de pendiente H.

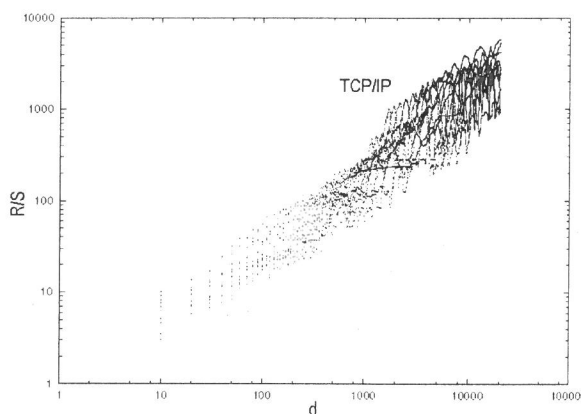


Fig. 6: Pox plot

Los valores encontrados para el parámetro de Hurst con este método en las muestras ip1, ip2 e ip3 corresponden a 0.85, 0.86 y 0.79 respectivamente.

CONCLUSIONES

Encontramos una marcada diferencia entre las muestras recolectadas de tráfico y los procesos de Poisson usados comúnmente para representar el tráfico de una red de datos. El tráfico Internet TCP/IP real presenta niveles no constantes que oscilan en el tiempo en períodos que se repiten en diferentes escalas (desde fracciones de segundo hasta horas), en forma de ráfagas de paquetes. Esta propiedad no es prevista correctamente por los procesos de Poisson.

Basados en el análisis estadístico de muestras de tráfico TCP/IP recolectadas en una red conectada a Internet concluimos que éste presenta una naturaleza asintóticamente auto-similar con un parámetro de Hurst aparentemente dependiente de la composición

de protocolos en el flujo de datagramas. Estos resultados abren un interesante campo de estudio en el modelamiento realístico del tráfico y en la revisión de sistemas de comunicaciones de datos, cuyos niveles de calidad de servicio han sido fijados usando parámetros determinados por modelos ergódicos en el modelamiento del tráfico de paquetes.

BIBLIOGRAFÍA

- [1] **W. Leland**, "On the self-similar nature of ethernet traffic (extended version)", IEEE/ACM Transactions on Networking, vol. 2, pp. 1-15, Feb. 1994.
- [2] **V. Paxson**, "Growth trends in wide-area tcp connections" tech. rep., Lawrence Berkeley Laboratory, University of California, May 1994.
- [3] **K. Claffy**, "Internet traffic characterization" PhD dissertation, University of California, San Diego, Department of Computer Science, 1994.
- [4] **Raj Jain, Shawn A. Routhier**, "Packet trains - measurements and a new model for computer network traffic", IEEE Journal on Selected Areas in Communications, vol. SAC-4, pp. 986-995, Sep. 1986.
- [5] **V. Paxson**, "Empirically derived analytic models of wide-area tcp connections" IEEE/ACM Transactions on Networking, vol. 2, pp. 316-336, Aug. 1994.
- [6] **V. Jacobson, C. Leres, S. McCanne**, "Tcpcdump" Disponible en <ftp://ftp.ee.lbl.gov>.
- [7] **R. Alexander**, "Monitoring, analysis and simulation of packed switched network traffic" Master's thesis, Auckland University, jan 1995.
- [8] **William Stallings**, "High-Speed Networks: TCP/IP and ATM Design Principles", Prentice Hall, ISBN 0-13-525965-7, 1998.
- [9] **P. Morin**, "The impact of self-similarity on network performance analysis" tech. rep., Carleton University, 1995.
- [10] **J. Beran**, "Long-range dependence in variable-bit-rate video traffic" IEEE Transactions on Communications, vol. 43, pp. 1566-1579, feb/mar/apr 1995.