

Long Memory Statistical Models in Telecommunication Networks¹

Modelli Statistici a Memoria Lunga in Reti di Telecomunicazioni

Livia De Giovanni
Dipartimento SEGeS
Università del Molise
degiovanni@unimol.it

Riassunto: Studi recenti basati su misurazioni empiriche mostrano che il traffico su reti a commutazione di pacchetti è auto-simile. In questo lavoro vengono presentati alcuni modelli statistici per il traffico a pacchetti. L'idea di base è quella di fornire una spiegazione “fisica” alla presenza di auto-similarità tramite risultati limite per processi ad alta variabilità. Il risultato principale, in questa direzione, è che la sovrapposizione di sorgenti *ON-OFF* con varianza infinita produce un processo di traffico aggregato di tipo auto-simile, con incrementi a memoria lunga. Nel lavoro vengono poi discusse le conseguenze dell'auto-similarità sulle prestazioni di sistemi di telecomunicazioni, e vengono descritte tecniche statistiche di stima in presenza di auto-similarità.

Keywords: Queues, Performance, Telecommunication, Long Range Dependence.

1. Introduction

In packet switching networks the information stream produced by a terminal equipment is transmitted in the form of packets. A packet is divided into a header and a data area. The header contains the source and destination addresses and the information to interpret the data area. Packets are routed through routers, which interconnect transmission input links with transmission output links through a switching fabric to deliver packets on the basis of the destination address (Fig. 1). Output links are equipped with buffers (queues) to store and schedule for transmission packets requiring simultaneously the same output link. Each connection source/destination is characterised by a Quality of Service (QoS) requirement. The most widespread packet switching technologies are IP (Internet Protocol) and ATM (Asynchronous Transfer Mode). IP is a protocol that provides the delivery of variable length packets, named datagrams. ATM is a technique based on fixed length packets (named cells). Standards and corresponding implementations already exist to run IP over ATM, *i.e.* for transmitting IP datagrams in the form of ATM *via* fast ATM switching fabrics (after assembly of ATM cells in the input port and fragmentation of IP datagrams in the output port).

¹ Financial support from the Italian Ministry of University and Scientific Research (MIUR) and from the National Research Council (CNR) is gratefully acknowledged.

Since packet switching networks are networks of queues, evaluating the performance of a single server queue (a multiplexer, composed by a transmission link with a buffer (Fig. 1)) is a fundamental step in assessing the performance of packet switching networks. This point is relevant, due to the recent growth of applications using Internet Protocol. One of the most important performance measure and QoS parameter in IP and ATM networks is Packet (Cell) Loss Probability, *i.e.* the probability that a packet is lost because it arrives when the buffer is full. It is usually approximated by the overflow probability, which is defined as the probability that the number of packets in an infinite buffer exceeds a given threshold. The overflow probability depends on both the packet arrival process (packet network traffic) and the packet service process. The last one is easy to be specified, as it depends on the capacity of the transmission link. The situation is different for the packet network traffic. As shown by recent measurements in telecommunication networks, the increments of this process are characterised by long range dependence. As a consequence, a general autocorrelation function for the process of the increments of the packet network process must be allowed.

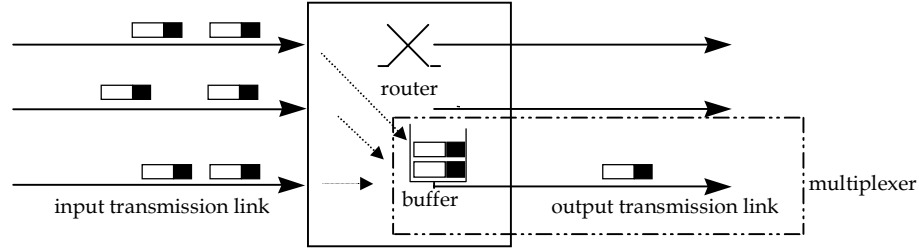


Figure 1: Router: output transmission link equipped with buffer (multiplexer)

Section 2 of the present paper contains some definitions of self similarity and its relationships with long range dependence. In section 3 statistical models for packet network traffic are presented. The impact of self similarity in the arrival process of packets on performance evaluation is studied in section 4. Statistical inference methods in the presence of self similarity are described in section 5. Conclusions are summarized in section 6.

2. Self similarity and Long Range Dependence

A continuous parameter stochastic process $\{Z(t); t \geq 0\}$ is *self similar*, with self similarity parameter $0 < H < 1$, if for any positive c the process $\{c^{-H}Z(ct); t \geq 0\}$ is equal in distribution to the original process $\{Z(t); t \geq 0\}$ (Beran, 1994). Self similar processes are invariant in distribution under scaling of time and space. A discrete parameter (second order) stationary process $\{U_j; j \geq 1\}$, with variance $\sigma^2 = E[U_j - E[U_j]]^2$ is *long range dependence (long memory)* if the correlation function is $r_k = \sigma^{-2} \text{Cov}(U_j, U_{j+k}) \approx k^{-\alpha}$ for some $0 < \alpha < 1$ when $k \rightarrow \infty$ (Beran, 1994; Samorodnitsky et al., 1994). If the original process $\{Z(t); t \geq 0\}$ is self similar with stationary increments and self similarity parameter H , then the correlation function r_k , $k \geq 1$, has the form (Beran, 1994):

$$r_k = \frac{1}{2}[(k+1)^{2H} - 2k^{2H} + (k-1)^{2H}] \equiv g(k) \approx k^{2(H-1)} \text{ as } k \rightarrow \infty \quad (1)$$

where $\{X_j; j \geq 1\}$ is the process defined as $\{X_j = (Z(j) - Z(j-1)); j = 1, 2, \dots\}$. The case $1/2 < H < 1$ corresponds to long range dependence, whilst for $H = 1/2$ r_k is zero for every $k \geq 1$. Relationship (1) clarifies that the increments of a self similar process with (second order) stationary increments form a stationary process with long range dependence ($H = 1 - \alpha/2$). The definition of self similarity involves infinitely many finite dimensional distributions. A practical definition of second order self similarity has been introduced (Cox, 1984), as equivalent characterization of long range dependence.

From the original process $\{X_j; j \geq 1\}$ an “averaged” second order stationary process $\{X_j^{(m)}; j \geq 1\}$ can be obtained by averaging the X_j ’s over adjacent, non overlapping blocks of size m :

$$X_j^{(m)} = m^{-1} \sum_{k=(j-1)m+1}^{jm} X_k \quad \forall j \geq 1. \quad (2)$$

If we indicate the associated correlation function by $r_k^{(m)}$, the process $\{X_j; j \geq 1\}$ is *exactly second order self similar* with Hurst parameter H ($1/2 < H < 1$), if $r_k \equiv g(k)$ and *asymptotically second order self similar* if $\lim_{m \rightarrow \infty} r_k^{(m)} \equiv g(k)$ (Cox, 1984).

Remark 1. By Taylor expansion $g(k)$ tends to $k^{-\alpha}$ ($0 < \alpha = (2-2H) < 1$) as $k \rightarrow \infty$.

Remark 2. From $r_k \equiv g(k)$ it follows $r_k^{(m)} = r_k$. Second order self similarity captures the property that the correlation structure is exactly (or asymptotically) preserved under time aggregation. The condition $r_k \equiv g(k)$ is equivalent to $\text{Var}(X^{(m)}) = \sigma^2 m^{-\alpha}$ ($0 < \alpha = (2-2H) < 1$) and to a spectral density function of the form $f(\lambda) \approx G \lambda^{1-2H} (\lambda \rightarrow 0^+, G \in (0, \infty))$ (Beran, 1994).

Remark 3. If the process $\{X_j; j \geq 1\}$ has a correlation function of the form (1), then $\text{Var}(X^{(m)}) = \sigma^2 m^{-\alpha}$ ($0 < \alpha = (2-2H) < 1$) as $m \rightarrow \infty$ and $r_k^{(m)} = g(k)$ as $m \rightarrow \infty$ (Tsybakov et al., 1997).

The above discussion clarifies two basic points. *i)* Asymptotic second order self similarity is equivalent to long range dependence. *ii)* Even if self similarity does not generally imply long range dependence, in the case $1/2 < H < 1$ and finite second moment the increment process of a continuous time self similar process is exactly (and asymptotically) second order self similar and hence long range dependence.

3. Long Memory Statistical Models: Self Similarity through high variability

Structural models based on physical features of the underlying communication network structure are presented. Continuous time and Discrete time models are described below.

3.1 Continuous time models

Continuous time models are based on a fluid flow approximation. It consists in replacing the discontinuous stochastic arrival process of packets by its average value as a function of time. As an example (without loss of generality referring to fixed length cells) if a source transmits 100 cells/second (constant rate), based on the fluid flow approximation in each one second length time interval it transmits 100 cells, in each half second length time interval it transmits 50 cells, and so on. The same holds for the transmission link. Cells are lost only if the sum of the cells transmitted by the connected sources exceeds the transmission capacity of the transmission link. If three sources share a 100 cells/second transmission link, two of which transmitting 40 cells/second and the other 20 cells/second, no cell will be lost because the transmission link receives in a cell time interval (the transmission time of a single cell) 0.4 cells from two sources and 0.2 cells from one source.

The ON-OFF source model

It is shown in (Willinger et al., 1998) that the superposition of many (strictly alternating) independent and identically distributed (i.i.d.) *ON/OFF* sources, each of which exhibits high variability, results in self similar aggregate traffic. The two effects are referred to (Mandelbrot et al., 1968) as “Noah Effect” and “Joseph effect” alluding, respectively, to the biblical figure Noah who lived through an exceptionally heavy flood and to the biblical figure Joseph who was faced with long periods of plenty followed by long periods of famine. Intuitively, the Noah Effect for an individual *ON/OFF* source model results in *ON* and/or *OFF* periods that can be very large with non-negligible probability. Heavy tailed distributions with infinite variance (Pareto, truncated stable distributions and others) are used to account for the Noah Effect. A simple relation holds between the parameter describing the “heaviness” of the tail of such distributions and the Hurst parameter H measuring the degree of self similarity (the “Joseph Effect”) of the aggregate traffic. The root cause underlying long range dependence in network traffic is the persistence of heavy tailed distribution: opposite to the exponential distribution, where prediction of activity is not enhanced by conditioning on ever longer periods of observed activity, in the heavy tailed case the longer the period of activity the more certain it will persist into the future.

Suppose first that in the network there are a number of users/sources communicating with each other. An individual source is modeled according to an *ON/OFF* alternating renewal process, that alternates active (or *ON*) states, where it sends packets at a constant (unitary) rate into the network, and inactive (or *OFF*) states, where it is idle. The *ON* periods are i.i.d., the *OFF* periods are i.i.d., and the *ON* and *OFF* period sequences are independent. The *ON* and *OFF* periods do not need to have the same distribution. Let F_{ON} and F_{OFF} denote the cumulative distribution function of the *ON* and *OFF* period respectively, and let $F^C = 1 - F$ denote a complementary cumulative distribution function. Let also μ_{ON} , σ_{ON} and μ_{OFF} , σ_{OFF} denote the respective means and variances. Assume that, as $x \rightarrow \infty$:

$$F_{ON}^C(x) \approx l_{ON} x^{-\alpha_{ON}}, \quad 1 < \alpha_{ON} < 2 \text{ or } \sigma_{ON} < \infty; \quad F_{OFF}^C(x) \approx l_{OFF} x^{-\alpha_{OFF}}, \quad 1 < \alpha_{OFF} < 2 \text{ or } \sigma_{OFF} < \infty \quad (3)$$

Here l_{ON} and l_{OFF} are constant. When $1 < \alpha_{ON} < 2$, the distribution of the *ON* periods is said to be heavy tailed with exponent α_{ON} . Since it has infinite variance, the *ON* periods can be very long with high probability. In the case $\sigma_{ON} < \infty$, referred to as $\alpha_{ON} = 2$, the *ON* periods are not likely to last very long. Similar remarks hold for the distribution of the *OFF* periods. Let $\{W(t), t \geq 0\}$ be the process where $W(t) = 1$ if time t is an *ON* period and $W(t) = 0$ if time t is an *OFF* period. In stationary state, $E(W(t)) = p_{ON} = p(W(t) = 1) = \mu_{ON} / (\mu_{ON} + \mu_{OFF})$. The interest is concentrated on the cumulative packet count at time t , namely $A(t) = \int_0^t W(u) du$, when t is “large”. In Taqqu et al. (1997) it is shown that for some constant $\sigma > 0$, for $\alpha_{min} = \min(\alpha_{ON}, \alpha_{OFF})$, and for covariance $\gamma(u) = EW(u)W(0) - E^2W(0)$ the relationship:

$$V(t) = Var(A(t)) = 2 \int_0^t \left(\int_0^v \gamma(u) du \right) dv \approx \sigma^2 t^{2H} \text{ as } t \rightarrow \infty; \quad H = \frac{(3 - \alpha_{min})}{2} \quad (4)$$

holds true. Suppose now that there are M i.i.d. sources. The aggregate packet count at time t is $\sum_{m=1}^M W^m(t)$, $W^m(t)$ denoting the *ON/OFF* process for the m -th source (Fig. 2).

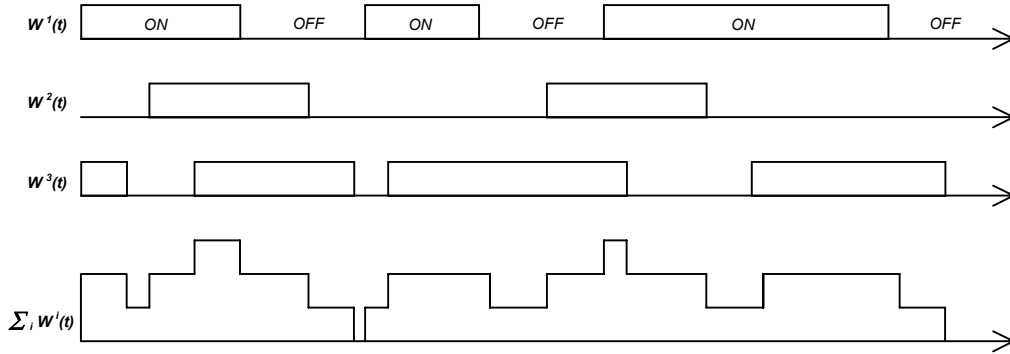


Figure 2: *ON/OFF sources and the corresponding aggregate process*

The normalized aggregate packet count $M^{-1/2} \sum_{m=1}^M (W^m(t) - E(W^m(t)))$ as $M \rightarrow \infty$ satisfies:

$$\lim_{M \rightarrow \infty} M^{-1/2} \sum_{m=1}^M (W^m(t) - E(W^m(t))) = G(t), \quad t \geq 0 \quad (5)$$

by the usual Central Limit Theorem. The process $\{G(t), t \geq 0\}$ is zero mean stationary (since the $W^m(t)$ are stationary) Gaussian with covariance function $\gamma(\cdot)$. After rescaling time by a factor T , the aggregate packet count satisfies the relationship:

$$\lim_{T \rightarrow \infty} T^{-H} \int_0^{Tt} G(u) du = \lim_{T \rightarrow \infty} \lim_{M \rightarrow \infty} M^{-1/2} T^{-H} \int_0^T \sum_{m=1}^M (W^m(t) - E(W^m(t))) = \sigma B_H(t), \quad t \geq 0. \quad (6)$$

The process $\{\sigma B_H(t), t \geq 0\}$, which appears in the limit, is Fractional Brownian Motion (FBM), *i.e.* a self similar (with Hurst parameter H) Gaussian process with stationary increments, mean 0, variance $\sigma^2 t^{2H}$ and covariance function $E[B_H(s)B_H(t)] = \sigma^2/2 \{s^{2H} + t^{2H} - |s-t|^{2H}\}$. Its increments $G_j = \{B_H(j) - B_H(j-1), j=1,2,\dots\}$ are called Fractional Gaussian Noise (FGN), and exhibit correlation function $E[G_H(j)G_H(j+k)] = (\text{const.})k^{2H-2}$ as $k \rightarrow \infty$ (Samorodnitsky et al., 1994).

Note that the limit must be Gaussian with mean 0 and have stationary increments since the integral of G has these properties. Moreover, due to $V(t)$, the limiting variance must be $\sigma^2 t^{2H}$ for fixed t . These properties uniquely characterize FBM (Samorodnitsky et al., 1994). It follows that if the distribution of the *ON* or of the *OFF* periods is heavy tailed ($1 < \alpha_{ON} < 2$ or $1 < \alpha_{OFF} < 2$, Noah effect), then $1/2 < H < 1$ and the increments of the aggregate cumulative packet process exhibit long range dependence (Joseph effect), providing the link between self similar characteristics at the macroscopic level (aggregate traffic) and heavy tailed phenomem at the microscopic level (traffic pattern displayed by source/destination pairs).

The infinite source Poisson model

An alternative to the superposition of *ON/OFF* sources is the infinite source Poisson model. In this model, sources arrive at a link at a rate of λ , each source transmits at a constant (unitary) rate (measured in packets) for a random time with complementary cumulative distribution function $F_{ON}^C(x)$ and, upon completing transmission, the source remains inactive. Opposite to the *ON/OFF* model, each source has only one period of activity (one *ON* period). The transmission times are assumed i.i.d. Let $W(t)$ be the number of active sources, or, equivalently, the number of packets generated from active sources, in the infinite source Poisson model at time $t > 0$. Furthermore let $\{X_j = W(j) - W(j-1), j=1,2,\dots\}$ be the corresponding increment process and let $A(t) = \int_0^t W(u) du$ be the cumulative arrival process of packets from time 0 to time t . Assuming $F_{ON}^C(x) \approx l_{ON} x^{-\alpha}$, $x \rightarrow \infty$ ($1 < \alpha < 2$), this model views traffic as generated by connection requests arriving in a Poisson fashion and whose lifetime is heavy tailed with infinite variance. The model arises as a limit of models that superimpose a finite number of independent *ON/OFF* sources (Resnick et al., 2002). $F_{ON}^C(x) \approx l_{ON} x^{-\alpha}$ implies $\text{Cov}(W(t), W(s)) = (\text{const.})(t-s)F_{ON}^C(t-s) \approx (\text{const.})(t-s)^{1-\alpha}$; and the relationship:

$$\gamma(k) = \text{Cov}(X_k, X_1) = (\text{const.}) \int_k^{k+1} \int_0^1 (t-s)^{1-\alpha} ds dt \approx (\text{const.})k^{1-\alpha} \quad (7)$$

holds true (Cox et al., 1980). In Resnick et al. (2002) it is shown that, after rescaling time by a factor T , and under a fast grow condition for the rate λ as a function of T , one obtains:

$$\lim_{T \rightarrow \infty} T^{-(3-\alpha)/2} \lambda^{-1/2} (A(Tt) - EA(Tt)) = cB_H(t), \quad t \geq 0; \quad H = \frac{(3-\alpha)}{2}. \quad (8)$$

The high variability in transmission times ($1 < \alpha_{OFF} < 2$, Noah effect) causes long range dependence in the process of the increments of the aggregate cumulative packet arrival process ($1/2 < H < 1$, Joseph effect).

3.2 Discrete time models

Discrete time models explicitly refer to the discrete nature of the packet arrival and service processes. Time is divided into consecutive intervals (time slots) and events (arrival, service) can only occur at discrete time instants, *i.e.* in time slots. In the ATM technology the natural elementary unit of time (the time slot) is the cell time, *i.e.* the transmission time of a single cell. Arrivals of cells can only occur at discrete time instants (*i.e.* in cell times) and the service time is constant (in a cell time only one cell can be transmitted). Opposite to the fluid flow approximation considered in the continuous time case, in the discrete time case cells can be lost even if the sum of the cells transmitted by the connected sources does not exceed the transmission capacity of the transmission link. Following the above example for the continuous time, if the three sources sharing a 100 cells/second transmission link (two of which transmitting 40 cells/second and the other 20 cells/second) transmit a cell in the same cell time interval of 0.01 second length two cells out of three will be lost because the transmission link receives three cells in a cell time interval.

Recent measurements in telecommunication networks (Willinger et al., 1995; De Giovanni et al., 2000; De Giovanni, 2000) have shown that the process of the increments of the arrival process is characterized by long range dependence. Several authors have proposed models in discrete time allowing any autocorrelation function for the increments of the arrival process. In some models (Tsybakov et al., 1997) the aggregate traffic is generated by a Poisson arrival of sources each of which generates cells at a constant (unitary) rate for an heavy tailed random time to account for long range dependence. In other models the amount of cells arriving in successive time intervals is jointly Gaussian with correlation function accounting for long range dependence (Montgomery et al., 1996; Addie, 1999).

4. Queueing with long range dependence input

Given that the arrival process of interest is self similar (and the corresponding increment process is long range dependence), a natural follow on question is whether this attribute is important for the queueing behaviour and hence for performance evaluation.

As an approximation of the cell loss probability (section 1), the overflow probability is generally used. As a consequence, the attention is focused on the unfinished work tail probability, which is equivalent to the overflow probability.

Let $A_M(t) = \int_0^t \sum_{m=1}^M W^m(t)$ be the aggregate (from M sources) cumulative packet count at time t offered to a transmission link with capacity c with an unlimited buffer, let $U(t)$ be the

unfinished work at time t , *i.e.* the amount of work remaining to be done at time t (Beneš, 1962), let U_∞ be the unfinished work in steady state and let K be the buffer size. The overflow probability is computed as $P(U_\infty > K)$ and can be defined, analogously to Reich's formula for the virtual waiting time in a queueing system, (Beneš, 1962) as:

$$P(U_\infty > K) = P(\sup_{t \geq 0} (A_M(t) - ct) > K). \quad (9)$$

Apart from special cases, the exact distribution of U_∞ is not known.

As far as the continuous time case is considered, approximations for the above probability have been obtained (Norros, 1994; Duffield et al., 1995; Husler et al., 1998) based on the assumptions that the process $\{A_M(t), t \geq 0\}$ is a self similar stationary Gaussian process using different techniques. These approximations have the following form:

$$P(U_\infty > K) \approx c(m, \sigma^2, H) \exp\{-K^{2(1-H)}\}, \quad (10)$$

m and σ^2 being the mean and variance of $A_M(t)$ and H is the Hurst parameter. As it appears from (10), the decay of the overflow probability is slower than the exponential form typical of arrival processes with independent increments (Markov processes, $H=1/2$ in (10)).

As far as the discrete time case is considered, the time is divided into consecutive intervals (time slots). Let U_n be the unfinished work of the queue at the beginning of n -th time interval, *i.e.* the amount of work remaining to be done at time n . The unfinished work obeys the recurrence formula $U_{n+1} = \max(0, U_n + Y_n)$, ($n \geq 1$), where $Y_n = (A_n - B_n)$ is the net input process, A_n , B_n being the number of cell arrivals in a time slot and the work that can be processed in a single time slot, respectively.

The random variate (r.v.) U_∞ represents the unfinished work in steady state. An upper bound for the tail probability of U_∞ , robust with respect to dependence among arrivals in time slots, is proposed in Kobayashi (1983). Moreover approximations of the tail probabilities of U_∞ have been proposed, based on the assumptions that the input process $(Y_n; n \geq 1)$ is a stationary Gaussian long range dependence process. They are obtained *via* Wiener-Hopf factorization technique (Addie, 1999) or using the large deviation theory (Montgomery et al., 1996), and behave as in the continuous time case.

5. Statistical Inference Methods for Long Range Dependence

Statistical inference on the overflow probability (via an estimation of the tail probability of the unfinished work) on the basis of sequences of interarrival and service times is of particular interest for packet switching networks. In fact, opposite to telephone networks, where a Connection Admission Control function checks if the network has enough resources to accept a new connection without infringing QoS requirements for all established connections as well as the new connection, packet switching networks operate without this preventive traffic control function. Therefore, the estimation of the unfinished

work tail probability is essential in assessing their performance. Performance evaluation is based either in the continuous time case or in the discrete time case on the assumption that the arrival process is Gaussian self similar, and that the corresponding increment process is Gaussian with long range dependence (section 4).

The Gaussian property is ensured by (space) aggregation obtained by superimposing independent traffic processes (independent sources) satisfying the usual conditions of the functional central limit theorem.

As far as the long range dependence property of the increments of the arrival process is concerned, we distinguish the structural and non structural modeling approach.

In the case of a structural modeling (section 3), the validation of these assumptions has been performed at the level of individual source/destination pairs in the network environment of Local Area Networks (LANs) where the network activity is generated by source/destination *ON* (at a constant packet transmission rate)/*OFF* behaviour; at the infinite source level, in the network environment of Wide Area Networks (WANs, for example the Internet), where the network activity is characterised by an arrival process of applications (*e.g.* file transfer FTP, information retrieval HTTP *i.e.* the World Wide Web, electronic mail), and corresponding connection requests, transmitting packets at a constant rate (*i.e.* in a fluid fashion) during their lifetime or connection length. In the first case the validation that the traffic is consistent with the assumed *ON-OFF* behaviour has been performed by Willinger et al. (1998) using textured plots; the validation that the traffic is consistent with the heavy-tailed distribution of the *ON* and/or *OFF* periods has been performed by Willinger et al. (1998) using *QQ*-plots and Hill plots (Hill, 1975). In the second case the validation that the arrival process is homogeneous (over an hour) Poisson has been performed using tests of exponential and independent nature of interarrival times (Paxson et al., 1994); empirical evidence that the traffic is consistent with the heavy-tailedness of the application lifetime has been performed by Willinger et al. (1998) using *QQ*-plots and Hill plots. In both cases the observed heavy tailedness of the activity period (*ON* period or application lifetime) may be related to the heavy tailedness of the file sizes or Web documents residing in network file systems, that is preserved under packet based transmission and mapped into heavy tailed activity periods. The observed heavy tailedness of the inactivity period (*OFF* period) seems to be a result of user thinking time.

As far as a non structural modeling approach is concerned, it is necessary to work directly on the aggregate cumulative arrival process $A_M(t)$ to validate the assumption that the corresponding increment process $\{X_k\}$ is long range dependence. Parametric and semiparametric estimates of the parameters (H, σ^2) by the maximum likelihood approach have been suggested. Among the semiparametric ones, the “aggregated Whittle” (Taqqu et al., 1996) and the “local Whittle” (Robinson, 1995) approximate maximum likelihood estimators. The theoretical background supporting many of these estimation methods is based on convergence results for the averaged process $\{X_j^{(m)}; j \geq 1\}$, defined in formula (2), showing that when the averaging level m is large it approaches FGN if the original process $\{X_k\}$ is Gaussian with long range dependence (Samorodnitsky et al., 1994). It is essentially the same conclusion as in the structural modeling approach. As a consequence, these semiparametric methods test long range dependence on the averaged arrival process using

the spectral density function of the FGN. Due to the equivalence between asymptotic second order self similarity and long range dependence (section 2), the estimation of H in the averaged process is equivalent to the estimation of H in the original one.

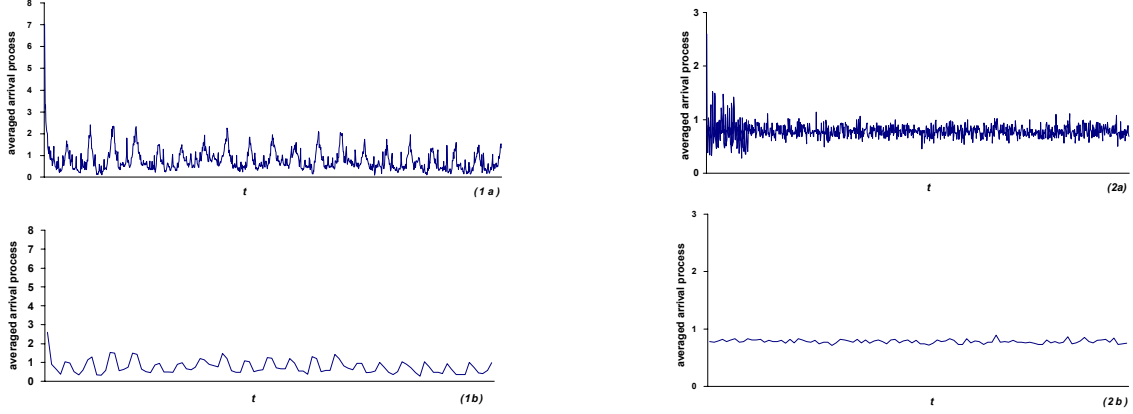


Figure 3: *Averaged arrival process of cells: real (1a, 1b), Poisson simulated (2a, 2b)*

In Fig. 3 the averaged arrival process has been derived from the increments (arrivals in one cell time) of a real arrival process of ATM cells (superposition of 3 sources) with mean arrival rate 0.8, and compared with the averaged Poisson arrival process with same mean arrival rate ($m=100$ 1a, 2a, $m=1000$ 1b, 2b): the first one exhibits a correlation structure unchanged under averaging, giving evidence of asymptotic second order self similarity; the second one converges to an uncorrelated process.

New statistical techniques in this area include the estimation of (H, σ^2) using wavelets (Daubechies, 1992). Due to their ability to localize a given time series both in time and scale (or frequency), wavelets provide a robust efficient technique for detecting and quantifying scaling behaviour (Abry et al., 1998). A comparison on real data in ATM networks among “local Whittle”, “aggregated Whittle” and wavelet based estimators is in De Giovanni (2000), De Giovanni et al. (2002b).

Once the parameters describing the Gaussian process are estimated, plug-in estimates of the performance parameters are obtained and their statistical inference properties evaluated (Conti et al., 2001; De Giovanni et al., 2002a; where ATM networks are considered).

6. Conclusion: toward self-similar non Gaussian models

A weakpoint of the described models is the assumption that the traffic over a transmission link is made up of a large number of independent sources, where each source’s traffic cannot fluctuate too widely. A constant (unitary) packet rate has been assumed for all the *ON* periods in the *ON/OFF* source model and for the lifetime of the connection in the infinite source Poisson model. Unfortunately, the feedback congestion control implemented by the Internet Protocol has an impact on the interaction among sources competing for the same transmission link, allowing traffic rates of individual sources to vary depending on

network conditions. As a consequence, processes that combine heavy tails either in time or in space (*i.e.* the distributions of the duration as well as of the rates at which individual sources emit packets are heavy tailed with infinite variance) are required.

As far as the *ON/OFF* model is concerned, recent results (Levy et al., 1997) show that, in contrast to FGN, the superposition process of independent sources where in addition to the duration of the *ON/OFF* periods also the rate at which the source emits packets during the *ON* period is heavy tailed with infinite variance, is a non Gaussian, self similar, dependent stationary increment process with heavy tailed (stable with infinite variance) marginals. This implies that there is a higher probability than in the Gaussian case that the traffic differs greatly from the average (a phenomenon also known as “intermittency”).

As far as the infinite source Poisson model is considered, the relaxing of the assumption of a constant rate during the lifetime of a connection has been studied in Kurtz (1996). The corresponding impact on performance evaluation is under research (Park et al., 2000).

References

- Abry P., Veitch D. (1998) Wavelet Analysis of Long Range dependent Traffic. *IEEE Trans. Info. Theory*, 44 (1), 2-15.
- Addie R.G. (1999) On weak convergence of long range dependent traffic processes. *Journal of Statistical Planning and Inference*, 80, 155-171.
- Beneš V.E. (1962) *General Stochastic Processes in the Theory of Queues*. Addison-Wesley.
- Beran J. (1994) *Statistics for Long-Memory Processes*. Chapman and Hall.
- Conti P.L., De Giovanni L., (2001) Queueing models and statistical analysis for ATM based networks. Accepted for publication in *SANKHYA B*.
- Cox D.R. (1984) Long Range Dependence: a review. *Statistics: an appraisal*. The Iowa State University Press, Ames, Iowa, 55-74.
- Cox D.R., V. Isham (1980) *Point Processes*. Chapman and Hall.
- Dahlhaus, R. (1989) Efficient parameter estimation for self similar processes. *The Annals of Statistics*, 17, 1749-1766.
- Daubechies I. (1992) *Ten Lectures on Wavelets*. SIAM.
- De Giovanni L. (2000) Tecniche Statistiche per l'individuazione di dipendenze a memoria lunga: applicazione a dati reali in telecomunicazioni (in Italian). Proceedings of the *XL Conference of the Italian Statistical Society (SIS)*, 141-144.
- De Giovanni L., Conti P.L. (2002)a A non parametric estimate of performance in queueing models with long-range correlation, with applications to telecommunications. To appear: *METRON*.
- De Giovanni L., Naldi M. (2000) Identification of long range dependence in telecommunication networks. *Italian Journal of Applied Statistics*, 12, 1, 31-58.
- De Giovanni L., Naldi M. (2002)b. *A wavelet based estimator of long range dependence in high-speed data networks*. Submitted for publication.

- Duffield, N.G., O'Connell N. (1995) Large deviations and overflow probabilities for the general singleserver queue, with applications. *Mathematical Proceedings of the Cambridge Philosophical Society*, 118, 363-374.
- Hill B. (1975). A Simple General Approach to Inference about the Tail of a Distribution. *The Annals of Statistics*, 3, 1163-1174.
- Husler J., Piterbarg V. (1998) Extremes of a certain class of Gaussian Processes. *International Conference 'Asymptotic Methods in Probability and Mathematical Statistics'*, S. Peterburg.
- Kobayashi, H. (1983) *Discrete Time Queueing Systems*. In G. Louchard, G. Latouche (Eds): *Probability Theory and Computer Science*. Academic Press, New York.
- Kurtz T. (1996) Limit Theorems for Workload Input Models. In Kelly, Zachary, Ziedins Eds., *Stochastic Networks: Theory and Applications*, Oxford Univ. Press, 339-366.
- Levy J.B., Taqqu M. (1997) Renewal Reward processes with Heavy tailed interarrival times and heavy tailed rewards. Preprint.
- Mandelbrot B.B., Wallis J.R. (1968) Noah, Joseph and operational hydrology. *Water Resources Research*, 4, 909-918.
- Montgomery, M., de Veciana G. (1996) On the relevance on time scales in performance oriented traffic characterization. Proceedings of *IEEE INFOCOM 96*, 513-520.
- Norros, I. (1994) A storage model with self similar input. *Queueing Systems*, 16, 387-396.
- Park K., Willinger W. (2000) *Selfsimilar network traffic and performance evaluation*. Wiley.
- Paxson V., Floyd S. (1994) Widearea traffic: The Failure of Poisson Modeling. Proceedings of the *ACM/SIGCOMM '94*, 257-268.
- Resnick T., Mikosch T., Rootzen H., Stegeman A. (2002) Is network traffic approximated by Stable Levy Motion or Fractional Brownian Motion?. To appear: *Annals of Applied Probability*.
- Robinson P.M. (1995) Gaussian Semiparametric estimation of long range dependence. *The Annals of Statistics*, 23, 5, 1630-1661.
- Samorodnitsky G., Taqqu M. (1994) *Stable Non-Gaussian Random Processes, Stochastic Models with Infinite Variance*. Chapman and Hall.
- Taqqu M., Teverovsky V. (1996) Semiparametric graphical estimation techniques for long-memory data'. *Lecture Notes in Statistics*, 115, 420-432.
- Taqqu M., Willinger W., Sherman R. (1997) Proof of a fundamental result in self similar traffic modeling. *Computer Communications Review*, 27, 5-23.
- Tsybakov B., Georganas N.D. (1997) On self-similar traffic in ATM queues. *IEEE Trans. on Networking*, 5, 3, 397-409.
- Willinger, W., Taqqu M., Leland W., Wilson D. (1995) Self similarity in high-speed traffic: analysis and modelling of Ethernet traffic measurement. *Statistical Science*, 10, 1, 67-85.
- Willinger W., Paxson V., Taqqu M., Wilson D. (1998) Self similarity and heavy tails: structural modeling of network traffic. In *A Practical Guide to heavy tails: Statistical Techniques and Applications Computer Communications*, Adler, Feldman, Taqqu (Ed.).