A Locally Stationary Semi-Markovian Representation for Ethernet LAN Traffic Data

S. Vaton and E. Moulines
Ecole Nationale Supérieure des Télécommunications
46 rue Barrault, 75634 Paris Cedex 13, France.

Abstract

In the past few years a large number of teletraffic measurements have been extensively studied by many contributors. Most of the authors agree on saying that the traffic measured on today's broadband networks is long-range dependent. Oddly enough these contributors do not question the stationarity of the traffic on these long time scales when the hypothesis of stationarity is essential to speak of long-range dependence. Concurrently it has been demonstrated that some kind of non stationarities in a short-range dependent process can lead, if they are not detected, to the untrue conclusion of long-range dependence.

We prove on the basis of different tests of stationarity that the hypothesis of short range dependence and the hypothesis of stationarity are contradictory on long time-scales. Contrary to many authors who decide in favor of the long-range dependence we propose to model the measured traffic as a locally stationary and markovian process. We exhibit a new markovian model and we show how one can track the varying parameters of this model by means of a recursive maximum likelihood algorithm.

We then generate a non stationary markovian traffic whose varying parameters are matching the parameters of the measured traffic. We verify that the use of a classical visual index of long-range dependence brings to the same conclusion of long-range dependence for the non stationary and markovian model than for the measured traffic.

Keywords

Long-range dependence, long memory, non-stationarity, tests, Hidden Markov Model.

1 INTRODUCTION

In the past few years a large number of contributions have been devoted to the statistical study of different traffics measured on today's broadband networks.

Broadband Communications P. Kühn & R. Ulrich (Eds.)

There is now a consensus to say that this traffic is long-range dependent. These contributions are very significant. They are indeed at variance with markovian models of traffic such as the Poisson process or the Markov Modulated Poisson Process for which queuing results have been established.

It has what is more been established that some quality measures such as the overflow probability or the average packet delay are strongly underestimated by markovian models if the traffic is long-range dependent. The only queuing results that can be established for long-range dependent processes are overestimations of the overflow probability of a queue fed by a long-range dependent process. These results appeal to the difficult theory of great deviations.

We wonder if the traffic might not be a stationary and long-range dependent process but a non stationary and markovian process. It is known ([1],[2],[3]) that deterministic jumps or trends in the mean of a time series without long-range dependence can mislead to the conclusion of long-range dependence if one relies on visual indexes of long-range dependence such as the variance time plot. In the above-mentioned contributions the authors investigate hours of traffic and oddly enough they do not test the stationarity of the traffic on these time-scales.

To support our intuition we investigate a traffic stream that is commonly studied in the litterature (LBL-PKT3). This trace was originally investigated by Paxson and Floyd [4] who conclude that it exhibits a high degree of burstiness that can not be explained by markovian models and who discuss how this burstiness might mesh with self-similar models of traffic.

The rest of the paper is organised as follows. In Section 2 we expose different tests of stationarity for mixing processes and their application to the LBL-PKT3 stream. We conclude that the hypotheses of stationarity and of mixing are incompatible on long time scales which proves that stationary markovian models are inadequate on long time-scales. In Section 3 we propose a new model, the Shifted Exponential Hidden Markov Model (SEHMM) and we briefly recall how one can track the parameters of this model. We then simulate a non stationary SEHMM. We show that for this synthetic non stationary and markovian model the variance time plot index leads to the same conclusions as those obtained by Paxson and Floyd for the real traffic. These findings question the consensus according to which the traffic measured on modern broadband networks is long-range dependent.

2 TESTS OF STATIONARITY

2.1 Theory of the tests

(a) General Framework

Basically the tests of stationarity that we propose rely on the comparison of different empirical statistics calculated on two neighbour segments of finite length of the stream. The hypothesis of stationarity is rejected if the empirical statistics for the two neighbour segments are significantly different.

Denote by $\{X_t\}$ the sequence of the inter arrival times (IAT) from which a set of finite length $\{X_t\}_{1\leq t\leq T}$ is observed. Suppose that one aims at testing if this finite observation is strict sense stationary. Denote by τ_1 the presumed change point. For the sake of homogeneity we also define $\tau_0 = 0$ and $\tau_2 = T$. We do as if $\{X_t\}_{1\leq t\leq \tau_1}$ and $\{X_t\}_{\tau_1+1\leq t\leq T}$ were two realizations of finite length $T_i = \tau_i - \tau_{i-1}$ of two processes $\{X_t^1\}$ and $\{X_t^2\}$.

In what follows we test the stationarity of the IAT process in the sense of (i) the mean of the process (ii) the sampled cumulative distribution function $\mathbb{E}(g(X_t))$ where $g(x) = (\mathbb{I}_{\Delta_1}(x), \cdots, \mathbb{I}_{\Delta_N}(x))'$, $(\Delta_i)_{1\leq i\leq N}$ being a partition of $\mathbb{R}^+$ and (iii) the first covariance coefficients $\mathbb{E}((X_t^2, X_tX_{t+1}, \cdots, X_tX_{t+N-1})')$. We introduce a new time series $\{Z_t\}$ that is defined as (i) $Z_t = X_t$ (ii) $Z_t = g(X_t)$ or (iii) $Z_t = (X_t^2, X_tX_{t+1}, \cdots, X_tX_{t+N-1})'$ depending of the non stationarities that we want to detect.

(b) Central Limit Theorems

Different Assumptions are needed to establish a Central Limit Theorem for the vector of the empirical statistics.

Assumption 1 *$\{X_t\}$ is strict sense stationary.*

Assumption 2 $\frac{T_i}{T} \overset{T\to\infty}{\to} c_i > 0$.

Assumption 3 *$\{X_t\}$ is α-mixing with an α-mixing coefficient that verifies $\sum_{n=0}^{+\infty} \alpha_n^{\delta/(2+\delta)} < +\infty$ and with $\mathbb{E}(|X_t|^{2+\delta}) < +\infty$.*

Assumption 4 *$q_t \overset{t\to\infty}{\to} +\infty$ and $q_t = o(t)$.*

Let us recall that the α-mixing coefficient of the process $\{X_t\}$ is defined as $\alpha_n = \sup_{A,B} |\mathbb{P}(A\cap B) - \mathbb{P}(A)\mathbb{P}(B)|$ the supremum being taken on all sets A in $\mathcal{M}_{-\infty}^t$ and B in $\mathcal{M}_{t+n}^{+\infty}$ where $\mathcal{M}_a^b = \sigma(X_t, a\leq t\leq b)$.

The Assumption 3 is verified by many usual processes and in particular by a large class of Markov processes. It is in particular verified by any ARMA process if the density of the innovation is strictly positive on $\mathbb{R}$ [5] and by any finite state irreducible Hidden Markov Chain. For a survey about mixing processes and about the Central Limit Theorem for such processes we refer the reader to [6] and the references therein.

Theorem 1 *Assume (A1-A2-A3). Then it holds that*

$$\sqrt{T}\Big(\frac{1}{T_i}\sum_{t=\tau_{i-1}+1}^{\tau_i} Z_t - \mathbb{E}(Z_t)\Big) \sim \mathcal{AN}(0, c_i^{-1}\Gamma_0), \quad 1\leq i\leq 2$$

where $\Gamma_0 = \sum_{\tau=-\infty}^{+\infty} \gamma_Z(\tau)$ *with* $\gamma_Z(\tau) = \mathbb{E}(Z_t Z'_{t+\tau}) - \mathbb{E}(Z_t)\mathbb{E}(Z'_t)$

Note that Γ_0 is equal to the spectral density matrix of $\{Z_t\}$ at zero frequency. This remark permits the construction of a consistent estimator

$$\hat{\Gamma}_0 = \frac{1}{T}\sum_{0}^{+m_T} w(k)\Re((\sum_{1}^{T}(Y_t - \hat{\rho})e^{j\frac{k+1}{T}t})'(\sum_{1}^{T}(Y_t - \hat{\rho})e^{j\frac{k+1}{T}t}))$$

where $m_T = \sqrt{T}$ and $w(k) = \mathbb{I}_{k=0} + \frac{2}{2m_T+1}\mathbb{I}_{1\le k\le m_T}$.

Denote by $\hat{Z}^i_T = \frac{1}{T_i}\sum_{t=\tau_{i-1}+1}^{\tau_i} Z^i_t$ the empirical statistics for the segment of index i and denote by $\hat{Z}_T = ((\hat{Z}^1_T)'(\hat{Z}^2_T)')'$ the vector of the empirical statistics for the two segments.

Theorem 2 *Assume (A1-A2-A3). Then it holds that*

$$\sqrt{T}(\hat{Z}_T - (11)'\otimes\mathbb{E}(Z_t)) \sim \mathcal{AN}(0,\Gamma) \quad \text{where} \quad \Gamma = \begin{pmatrix} c_1^{-1}\Gamma_0 & 0 \\ 0 & c_2^{-1}\Gamma_0 \end{pmatrix}$$

and where $A\otimes B$ *denotes the Kronecker product of A and B.*

To demonstrate the Theorem 2 we mimic the approach of Epps in [7]. We define a new estimator $\tilde{Z}^i_T$ where the first q_T terms are removed

$$\tilde{Z}^i_T = \frac{1}{T_i}\sum_{\tau_{i-1}+q_T+1}^{\tau_i} Z_t$$

The basic idea consists in proving that $\sqrt{T}\hat{Z}^i_T$ and $\sqrt{T}\tilde{Z}^i_T$ converge in distribution to the same normal distribution and in proving that $\sqrt{T}\tilde{Z}^1_T$ and $\sqrt{T}\tilde{Z}^2_T$ are asymptotically independent, in the sense that

$$T(\mathbb{E}(\exp(i\tilde{Z}^1_T u^H + i\tilde{Z}^2_T v^H)) - \mathbb{E}(\exp(i\tilde{Z}^1_T u^H))\mathbb{E}(\exp(i\tilde{Z}^2_T v^H))) \overset{T\to\infty}{\to} 0$$

It results from the Davydov Theorem [8] that

$$|\mathbb{E}(\exp(i(\tilde{Z}^1_T u^H + \tilde{Z}^2_T v^H))) - \mathbb{E}(\exp(i\tilde{Z}^1_T)u^H)\mathbb{E}(\exp(i\tilde{Z}^2_T)v^H)| \le \alpha(q_T) \overset{T\to+\infty}{\to} 0$$

and though $\tilde{Z}^1_T$ and $\tilde{Z}^2_T$ are asymptotically independent.

Denote by $D^i_T = \sqrt{T}(\hat{Z}^i_T - \tilde{Z}^i_T) = \frac{\sqrt{T}}{T_i}\sum_{t=\tau_{i-1}+1}^{\tau_{i-1}+q_T} Z_t$ the difference between $\sqrt{T}(\hat{Z}^i_T$ and $\sqrt{T}\tilde{Z}^i_T$. It results from (A2-A4) that the covariance matrix of D^i_T tends to zero as T tends to infinity. It then results from the Theorem 1 and from the Slubtski Theorem [9] that $\sqrt{T}(\tilde{Z}_T - (11)'\otimes\mathbb{E}(Z_T)) \sim \mathcal{AN}(0,\Gamma)$ which concludes the proof of Theorem 2.

(c) Tests

- **Stationarity of the mean and of the marginal distribution**

As stated above the test of stationarity consists in comparing $\hat{Z}_T^1$ and $\hat{Z}_T^2$. For the mean and for the marginal distribution of the process we consider the difference between $\hat{Z}_T^1$ and $\hat{Z}_T^2$ on the two neighbour segments $\hat{Z}_T^1 - \hat{Z}_T^2 = U\hat{Z}_T$, where $U = (1 - 1)$ in the test of stationarity of the mean of $\{X_t\}$ and $U = (1 - 1) \otimes I_N$ in the test of stationarity of the marginal distribution of $\{X_t\}$.

Theorem 3 *Assume (A1-A2-A3). Then it holds that*

$$\sqrt{T}U\hat{Z}_T \sim \mathcal{AN}(0, U\Gamma_0 U')$$

Theorem 4 *Assume (A1-A2-A3). Then it holds that*

$$T\hat{Z}_T(\Gamma^{-1/2})^H \Gamma^{-1/2}\hat{Z}_T \sim \chi^2(N)$$

where $\Gamma^{\frac{1}{2}}$ *denotes the square root of* Γ, $\Gamma = \Gamma^{\frac{1}{2}}(\Gamma^{\frac{1}{2}})'$.

- **Stationarity of the first correlations**

In [10] Mauchly introduces the sphericity statistics to test whether two gaussian random vectors have the same covariance matrix. Drouiche and Mokkadem ([11],[12],[13],[14]) generalize this test to the case of time series to test whether two processes have proportional spectra. Vaton [15] proposes to use this measure of spectral similarity to test whether a process is second order stationary. We briefly recall the principles of the test of second order stationarity proposed by Vaton [15].

For any positive sequence $\rho = (\rho_0, \rho_1, \cdots, \rho_{N-1})'$ denote by $T_N(\rho)$ the Toeplitz matrix $T_N(\rho) = \sum_{\tau=0}^{N-1} \rho_\tau M_\tau$ where M_τ is the matrix whose entry (i,j) is equal to $M_\tau(i,j) = \delta_\tau(|i-j|)$. Denote by μ and ν two positive sequences and define

$$S(\mu,\nu) = \frac{(det(T_N(\mu)T_N^{-1}(\nu)))^{1/N}}{\frac{1}{N}Tr(T_N(\mu)T_N^{-1}(\nu))}$$

It results from the arithmetico-geometric inequality that $S(\mu,\nu) \leq 1$ with equality when $\mu = \alpha\nu$ where α is a proportionality constant.

Our idea is to derive the asymptotic distribution of the ratio $S(\hat{Z}_T^1, \hat{Z}_T^2)$ normalized by a factor that depends on the length T of the observation and to reject the hypothesis of stationarity if the obtained value is lower than a prescribed threshold determined by the false alarm probability.

Note that $S(\hat{Z}_T^1, \hat{Z}_T^2)$ is a deterministic function of $\hat{Z}_T$. This permits the derivation of an asymptotic result for $S(\hat{Z}_T^1, \hat{Z}_T^2)$. The demonstration of this result is based on a Taylor development of S at point $(\mathbb{E}(Z_t), \mathbb{E}(Z_t))$. As S is maximum at point $(\mathbb{E}(Z_t), \mathbb{E}(Z_t))$ a second order Taylor development is needed.

Theorem 5 *Assume (A1-A2-A3). Then it holds that*

$$2TS^*(\hat{Z}_T) \xrightarrow[(d)]{T\to\infty} Z^H \nabla^2 S((11)' \otimes \mathbb{E}(Z_t))Z \quad \text{with} \quad Z \sim \mathcal{N}(0,\Gamma)$$

where $\nabla^2 S^*((11)' \otimes \mathbb{E}(Z_t))$ *denotes the Hessian of* S *at point* $(11)' \otimes \mathbb{E}(Z_t)$.

The only technical points that are needed to establish the expression of $\nabla^2 S$ are the second order differential of the determinant and of the inverse of any matrix M. These expressions can be obtained by differential calculus :

$$\log|M+\Delta M| = \log|M| + Tr(M^{-1}\Delta M) - Tr(M^{-1}\Delta M M^{-1}\Delta M) + o(||\Delta M||^2)$$
$$(M+\Delta M)^{-1} = M^{-1} - M^{-1}\Delta M M^{-1} + 2M^{-1}\Delta M M^{-1}\Delta M M^{-1} + o(||\Delta M||^2)$$

- **Thresholds**
  The Theorems 4 and 5 permit to reject the set of Assumptions (A1-A3) with a false alarm probability of α. If the obtained statistics is superior to the $(1-\alpha)$ quantile of the asymptotic distribution one concludes that (A1-A3) is wrong which means that A1 and A3 are mutually exclusive. Note that in the Theorem 5 the asymptotic distribution is a quadratic form in a multidimensional Gaussian random variable and that the prescribed threshold is obtained by Monte-Carlo simulation.

2.2 Results

The simulations are replicated for thirteen time-scales ranging from six seconds to one hour and thirty minutes and for ten pairs of neighbour segments for each time-scale.

On the Figures 1 and 2 we plot $T\hat{Z}_T^H(\Gamma^{-1/2})^H\Gamma^{-1/2}\hat{Z}_T$ for all the pairs of neighbour segments. The 90% and 99% quantiles of $\chi^2(N)$ are represented in dotted lines. (A1-A3) is rejected when $T\hat{Z}_T^H(\Gamma^{-1/2})^H\Gamma^{-1/2}\hat{Z}_T$ is superior to the $(1-\alpha)$ quantile of $\chi^2(N)$.

On the Figure 3 we plot the cumulative distribution function $P(X \leq 2TS(\hat{Z}_T^1, \hat{Z}_T^2))$ for the asymptotic distribution $X = Z^H \nabla^2 S \mid_{(\mathbb{E}(Z_t),\mathbb{E}(Z_t))} Z$ where $Z \sim \mathcal{N}(0,\Gamma)$. The 90% and 99% fractiles for the distribution of X are represented in dotted lines. ((A1-A3) is rejected if $P(X \leq 2TS(\hat{Z}_T^1, \hat{Z}_T^2))$ is superior to $(1-\alpha)$.

The conclusions of our simulations is that (A1-A3) is wrong for most pairs of neighbour segments for long time-scales. Classical models such as the stationary Poisson process or the stationary Markov Modulated Poisson Process are consequently not adapted to the traffic that we investigate on these long time-scales.

Note that the Assumption A3 is wrong for long-range dependent processes such as the Fractional Gaussian Noise or the fractionally integrated autoregressive moving average process. Consequently the tests developed do not permit to reject A1 for long-range dependent processes. The difficulty to decide between long-range dependence and non stationarities has already been discussed by Duffield *et al.*in [2].

It is thus difficult to decide if the evidences of auto-similarity mentioned by many authors result from a real auto-similarity of the traffic or from some non-stationarities that might have mislead to the conclusion of auto-similarity ([1],[2],[3]) or from the coexistence of both phenomena.

3 A NON STATIONARY AND SEMI-MARKOVIAN MODEL

3.1 The Shifted Exponential Hidden Markov Model

As mentioned in Section 2 it is difficult to decide between a real auto-similarity of the traffic and some non stationarities. Our intuition is that the hypothesis of local stationarity is as plausible as the hypothesis of long-range dependence. Contrary to many authors who suggest modeling the BISDN traffic as a long-range dependent process we propose to model the measured traffic as a locally stationary and markovian process.

One way of modeling time series that are suspected to be locally stationary consists in using a parametric model whose parameters are jumping from time to time or are drifting with time at a rate that is sufficiently fast for the non stationarities to be perceptible and sufficient slow for parameters tracking to be possible.

We propose to model the observed process as a locally stationary Hidden Markov Chain with conditional laws that are shifted exponential. We call this model the Shifted Exponential Hidden Markov Model (SEHMM).

Denote by $\{O_t\}$ the successive inter-arrival times and denote by $\{Q_t\}$ a finite state Markov Chain whose transition matrix is denoted by P and whose initial distribution is denoted by π. The parameters of the distribution of O_t conditionally to $Q_t = i$ are the shift s_i and the intensity λ_i of the shifted exponential distribution. Denote by $\mathcal{F}_t = \sigma(o_t, o_{t-1}, \cdots)$ and by $\mathcal{G}_t = \sigma(q_t, q_{t-1}, \cdots)$ the filtrations associated to the processes $\{O_t\}$ and $\{Q_t\}$. Then

$$\begin{aligned} & P(Q_t = i \mid \mathcal{F}_{t-1}, \mathcal{G}_{t-1}) = P(Q_t = i \mid q_{t-1}) \\ \forall A \in \mathcal{B}, \quad & P(O_t \in A \mid Q_t = i, \mathcal{F}_{t-1}, \mathcal{G}_{t-1}) = P(O_t \in A \mid Q_t = i) \\ & = \int_A \mathbb{I}_{[s_i, +\infty[}(u) \lambda_i exp(-\lambda_i (u - s_i)) du \end{aligned}$$

The SEHMM is a new model proposed by Vaton *et al.*[16] from the analysis of the traffic measured by Paxson and Floyd [4]. This model has many attractive features, among which the existence of simple on-line and off-line algorithms of estimation and control for such models.

One should remark that the SEHMM is a generalization of the model of Kofman *et al.*. According to Kofman *et al.*who analyzed the traffic measured by Jain and Routhier [17] the distribution of the interarrival times is a mixture of exponentials. The SEHMM is also close to the Markov Modulated Poisson

Process (MMPP). In both cases the model is semi-markovian and the marginal distribution of is a mixture of exponential distributions.

The vindication of this new model as well as off-line and on-line procedures of estimation of the parameters of this model are detailed in [16]. Note that the estimation of the shifts s_i is particularly involved. This estimation can not be performed in a maximum likelihood sense since the likelihood has many discontinuity points. The Cramer-Rao variance lower bound that justifies the use of the maximum likelihood estimator is not even defined; the conditions under which this bound is derived are indeed not fulfilled. Vaton and Chonavel [18] propose an algorithm of estimation of several shifts in the case of incomplete data. This algorithm is based on a Fourier transform of the marginal distribution of the process; it exploits the fact that the shift of a distribution is equivalent to a modulation by a complex exponential function of the Fourier transform of this distribution. In what follows the shifts s_i are supposed to be constant and known. This point has been verified on the traffic that we investigate in this contribution.

3.2 A recursive estimation procedure

In our context it is of major interest to derive a recursive algorithm of estimation of the model that we propose since we wish to cope with varying parameters. The recursive estimation of the parameters of a HMM has been studied by several authors (see Elliott [19] for a review). In this contibution we briefly recall the original procedure developed by Mevel [20].

Denote by $\theta(t)$ the value of the parameters at time t. Contrary to the procedure developed by Elliott [19] the procedure developed by Mevel is not based on the EM paradigm but it exploits directly the particular structure of the log-likelihood derivative. Denote by

$$b_i(o_t;\theta(t)) = \mathbb{I}_{[s_i(t),+\infty[}(o_t)\lambda_i(t)exp(-\lambda_i(t)(o_t - s_i(t)))$$

the probability density function of O_t conditionally to $Q_t = i$ and to $\Theta = \theta(t)$ and denote by α_t the one step ahead prediction filter at time t

$$\alpha_t(i) = P(Q_{t+1} = i \mid o_t, o_{t-1}, o_{t-2}).$$

Because of the semi-Markov property the one step ahead prediction filter, its gradient $g_t = \nabla\alpha_t$ and its Hessian $h_t = \nabla^2\alpha_t$ can be computed recursively

$$\alpha_{t+1} = F(\alpha_t, o_{t+1}) \quad \text{and} \quad g_{t+1} = G(\alpha_t, g_t, o_{t+1}) \quad \text{and} \quad h_{t+1} = H(\alpha_t, g_t, h_t, o_{t+1}) \tag{1}$$

The computation of the log-likelihood of $\{o_1, \cdots, o_t\}$ and of its gradient is

$$\log p(o_{1:t};\theta) = \sum_{\tau=0}^{t-1} \log p(o_\tau \mid o_{1:\tau-1}) = \sum_{\tau=0}^{t-1} \log(\sum_i \alpha_\tau(i) b_i(o_{\tau+1})) \tag{2}$$

which permits a stochastic approximation procedure

$$\theta(t+1) = \theta(t) + \gamma_t \frac{\partial \log p(o_t \mid o_{1:t-1}\theta_t)}{\partial \theta} \tag{3}$$

where γ_t is a sequence of step-size (in the tracking context, we set $\gamma_t = \gamma$). The idea is to exploit the recursive formulae 1 to compute recursively the term of excitation in this stochastic gradient algorithm. Mevel [20] demonstrates that the stationary points of this algorithms are the extrema of the Kullback information and he demonstrates the asymptotic normality of the estimator under the hypothesis of stationarity.

3.3 A visual index of long-range dependence

A time series $\{X_t\}$ is long-range dependent if its autocovariance function $r(j)$ is $r(j) \sim Cj^{2H-2}$ as $j \to +\infty$ where $1/2 < H < 1$. It is well known [21] that if $\{X_t^{(m)}\}$ denotes the aggregated series

$$X_t^{(m)} = \frac{1}{m} \sum_{k=(t-1)m+1}^{tm} X_k$$

the sample variance $\mathrm{var}X^{(m)}$ of the aggregated series is $\mathrm{var}X^{(m)} \sim \sigma_0^2 m^{2H-2}$ as $m \to +\infty$.

This permits the construction of a visual index of long-range dependence. One plots $\log \mathrm{var}X^{(m)}$ versus $\log m$ for various aggregation levels. If the series is long-range dependent the graphic fits a straight line with a slope $-1 < \beta = 2H - 2 < 0$. The slope of this straight line provides an estimate of the Hurst parameter H.

This visual index of autosimilarity is the main evidence of many contributors to sustain that the traffic measured on modern broadband networks is long-range dependent.

As we suspect that some non stationarities might have mislead to the conclusion of long-range dependence we mean to exhibit the same visual index for a non stationary SEHMM that we simulate. We compare our conclusions with the conclusions of these authors. The parameters of the non stationary SEHMM are matching the varying parameters of the real traffic that we estimate with the stochastic gradient algorithm exposed above.

The logarithm of the sample variance of the aggregated process is plotted versus the logarithm of the aggregation level on Figure 4 for both the real traffic and the traffic that we simulate. The full line is a reference that corresponds to a time series that is not correlated. The graphic resembles a straight line in the case of the locally stationary SEHMM as well as in the case of the real traffic. The estimates of H deduced from the slope of the straight lines are close if one takes into account the bad quality of this estimate.

4 CONCLUSION

In this contribution we have developed different tests of stationarity for mixing processes. Thanks to these tests we have established by intensive simulation that any stationary and semi markovian model is inadequate on long time scales for the traffic measured on modern broadband networks. The hypothesis of stationarity and the hypothesis of mixing that these models postulate are indeed incompatible on these time scales.

We suspect that the apparent auto-similarity revealed by many contributors results at least partly from some non stationarities of the traffic. We have proved by simulation that a locally stationary and markovian model exhibits the same evidence of autosimilarity as the real process. Our findings question the consensus that has become established around the long range dependence of the traffic measured on modern broadband networks. They lead the way to some new realistic and tractable models.

Acknowledgments
This work was supported by France Telecom research center under contract number PE95-7633. This work was accomplished while S. Vaton was an invited student at the Ecole Nationale Supérieure des Télécommunications de Bretagne.

REFERENCES

[1] V. Klemes, "The hurst phenomenon: a puzzle?," *Water Resour. Res.*, vol. 10, pp. 675–678, 1974.

[2] N. Duffield, G. Lewis, N. O'Connel, and F. Toomey, "Statistical issues raised by the bellcore data," Preprint.

[3] V. Teverovski and M. Taqqu, "Testing for long-range dependence in the presence of shifting means or a slowly declining trends, using a variance type estimator," *J. of Time Series Analysis*, vol. 18, no. 3, pp. 279–304, 1997.

[4] V. Paxson and S. Floyd, "Wide-area traffic: The failure of poisson modeling," in *SIGCOMM'94*, London, England, aug 1994, pp. 257–268.

[5] A. Mokkadem, "Entropie des processus linéaires," *Probability and mathematical statistics*, vol. 11, no. 1, pp. 79, 1990.

[6] P. Doukhan, *Mixing: properties and examples*, Lecture Notes in Statistics. Springer-Verlag, 1994.

[7] T.W. Epps, "Testing that a gaussian process is stationary," *The Annals of Statistics*, vol. 16, no. 4, pp. 1667, dec 1988.

[8] P. Billingsley, *Probability and measure*, Wiley series in Probability and Mathematical Statistics. Wiley, 1979.

[9] P.J. Brockwell and R.A. Davies, *Time Series: Theory and Methods; Second Edition*, Springer Series in Statistics. Springer-Verlag, 1991.

[10] A. Mauchly, "Significance test for sphericity of a normal n-variate distribution," *Ann. Math. Stat.*, vol. 11, pp. 204–209, 1940.

[11] A. Mokkadem and K. Drouiche, "A new test for time series," Prepublications d'Orsay.

[12] A. Mokkadem, "On some new tests for time series," *CRAS Paris, Series I*, vol. 318, pp. 755–758, 1994.

[13] K. Drouiche, "Measuring randomness," Preprint. To appear in the Journal of the Royal Society B.

[14] K. Drouiche, "A new test for whiteness," Preprint. To appear in the IEEE Trans. on Signal Processing.

[15] S. Vaton, "A new test of stationarity and its application to teletraffic data," in *ICASSP'98*, Seattle, USA, may 1998.

[16] S. Vaton, E. Moulines, H. Korezlioglu, and D. Kofman, "Statistical identification of lan traffic data," in *ATM'97*, Bradford, England, july 1997, pp. 15/1–10.

[17] R.Jain and S.A. Routhier, "Packet trains: and a new model for computer network traffic," *IEEE-JSAC*, vol. 4, no. 6, pp. 986–995, sep 1986.

[18] S. Vaton and T. Chonavel, "A fourier transform estimation of shifts in the case of incomplete data," in *SSAP'98*, Portland, USA, sep. 1998.

[19] R.J. Elliott, L. Aggoun, and J.B. Moore, *Hidden Markov Models: Estimation and Control*, Applications of Mathematics. Springer-Verlag, 1995.

[20] L. Mevel, *Statistique Asymptotique pour les Modèles de Markov Cachés*, Ph.D. thesis, Université de Rennes 1, 1997.

[21] J. Beran, *Statistics for long memory processes*, Chapmann and Hall, 1994.

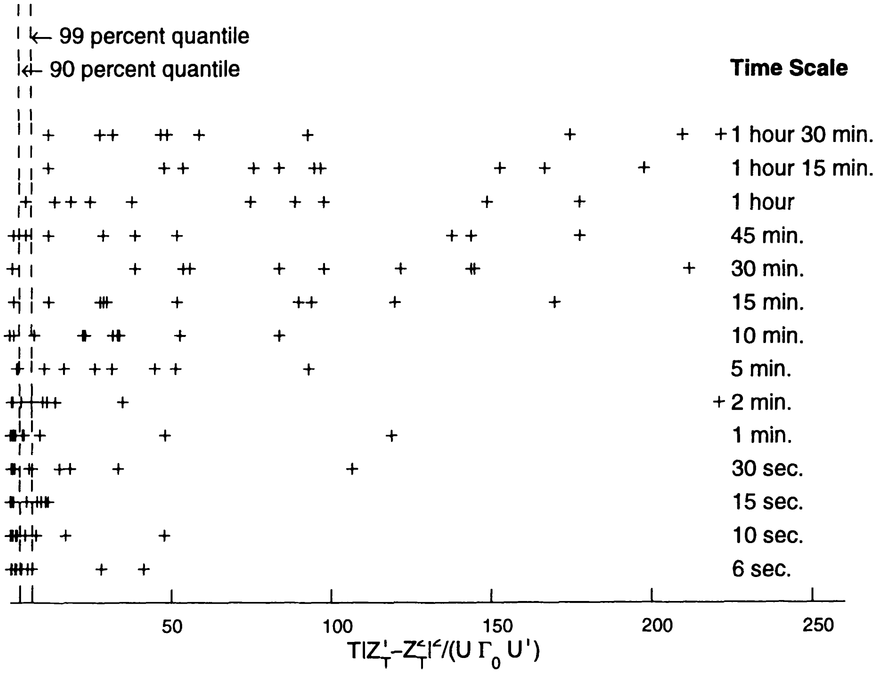

Figure 1 Stationarity of the mean

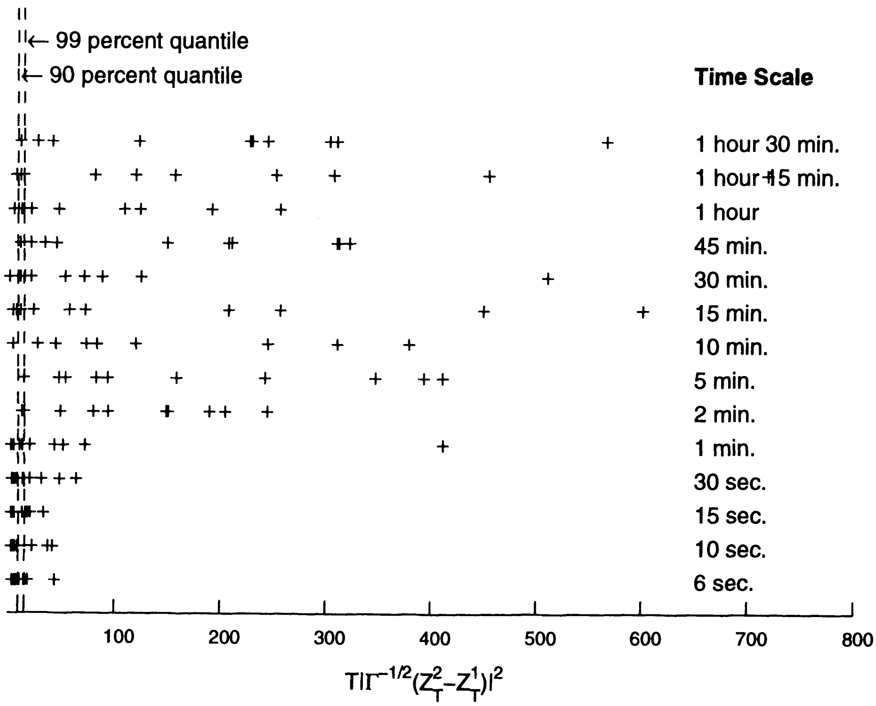

Figure 2 Stationarity of the marginal distribution

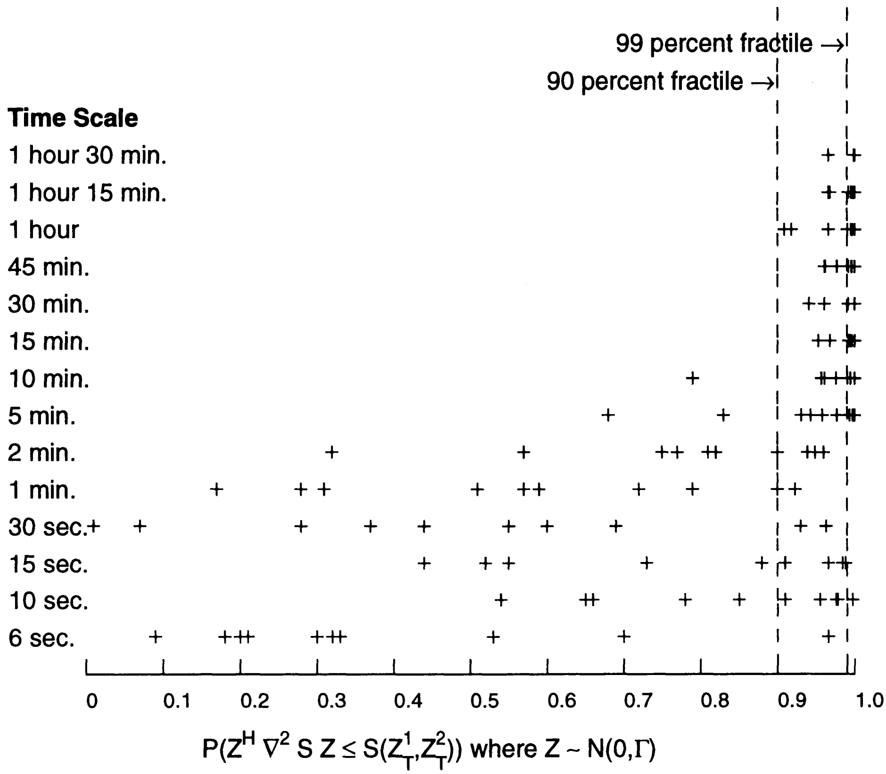

Figure 3 Stationarity of the first five correlations

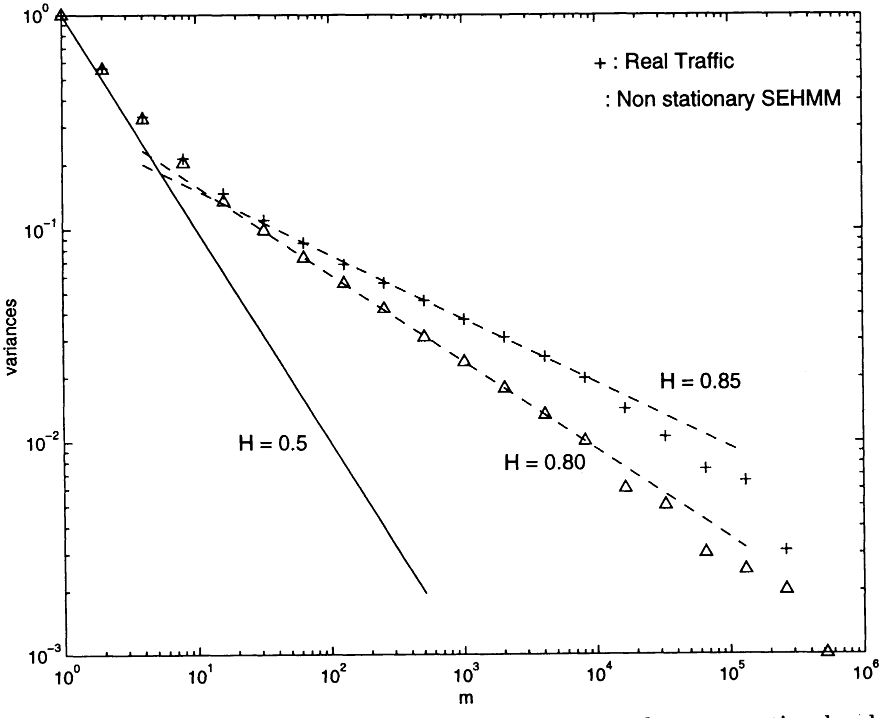

Figure 4 Variance of the aggregated process versus the aggregation level