INSTITUTO DE COMPUTAÇÃO
UNIVERSIDADE ESTADUAL DE CAMPINAS

OBSOLETE

SEE TR-IC-03-021

OBSOLETO

VEJA TR-IC-03-021

An Envelope Process for Multifractal Traffic Modeling

Cesar Augusto Viana Melo
Nelson Luis Saldanha da Fonseca

Technical Report - IC-03-10 - Relatório Técnico

April - 2003 - Abril

OBSOLETE - SEE TR-IC-03-21

OBSOLETO - VEJA TR-IC-03-21

An Envelope Process for Multifractal Traffic Modeling

Cesar Augusto Viana Melo* Nelson Luis Saldanha da Fonseca †

22 de abril de 2003

Resumo

In this paper, a novel envelope process for multifractal traffic modeling is introduced. The envelope process is an upperbound for the amount of work arrived in a multifractal Brownian motion process. The time scale of interest of a queueing system fed by a multifractal stream is computed. Simulation experiments using both real and synthetic data show that the proposed model is accurate. Moreover, a new estimator for the Holder function is presented.

1 Introduction

Since the seminal work of Leland et al [1], several studies have shown that network traffic presents scale invariance, or "scaling", which is the absence of any specific time scale at which the "burstiness" of a traffic stream can be characterized. Instead, it is necessary to describe the traffic across different time scales. Self-similar or (mono) fractal processes have been used for modeling network traffic since then.

Scaling of fractal traffic is defined by a single constant value: the Hurst parameter, H. One of the most popular fractal processes for traffic modeling is the Fractal Brownian Motion process (fBm) due to its parsimonious representation of the modeled traffic. fBm is an accurate model when: i) the traffic results from the aggregation of several sources streams with low activity compared to the link bandwidth, which leads to a Gaussian representation of the marginal distribution of counts, ii) the impact of flow control is not relevant and iii) the time scale of interest is within the scaling region.

However, both Internet Protocol (IP) and Variable Bit Rate (VBR) video traffic present non-trivial scaling structure at small scales in addition to long memory [2][3][4]. At small scale, traffic is highly variable, more complex and follows less definitive scaling laws. For these traffics the marginal distribution of counts is clearly non-Gaussian, calling for a representation beyond second-order statistics. Moreover, the scaling exponent of the variance on time scale shorter than a typical (cut-off) one is smaller than an asymptotic exponent.

If on one hand, at the network core, traffic variability at small time scales is not as important as long time correlations due to traffic aggregation (additive property)[5]. On

*This work was partially sponsored by FAPESP (grant 00/09772-6)
†This work was partially sponsored by CNPq (grant 300064/95-0)

the other hand, at the network edge, where admission control is performed, small time scale plays a major role [4]. Such pattern can be modeled by multifractal processes which are able to capture both long memory and high variability at small scales.

Multifractal processes exhibits highly irregular patterns as a function of time. Local Holder regularity describes the local regularity of the sample path of a process. It is a measure of scaling and can be regarded as a generalization of the Hurst parameter [5].

Studies have pointed out that the multiscaling nature of IP traffic is highly influenced by the Transmission Control Protocol (TCP) congestion control mechanism rather than solely by network-related variability, such as the diversity of link capacity in the Internet [6]. Nonetheless, the multiplicative origin of IP traffic [7] remains to be fully explained. Understanding the interaction between TCP congestion control and IP multiscaling is undoubtfully relavant for predictive purpose. However, this work is concerned with open loop aspects of TCP/IP traffic for network dimensioning, which is of paramount importance to problems such as the dynamic resizing of MPLS LSP's and admission control to DiffServ domains.

The major contribution of this paper is a novel envelope process for multiscaling traffic. The envelope process is an upper bound for the accumulated amount of traffic arrived up to a certain time from a multifractal Brownian motion process (mBm)[8]. It is shown that although mBm is a steady state Gaussian process, the envelope process is a tight bound for the amount of traffic arrived from real network streams. One of the advantages of this envelope process is the parsimonious representation of traffic, which allows a simple computation of the solution of queueing systems fed by multifractal streams. An expression for the time at which a finite queueing system overflows is computed. Moreover, a robust and accurate estimator for the Holder function is introduced.

2 Related work

A. Erramilli, O. Narayan, A. Neidhardt and I Saniee [3][4] proposed that traffic should be modeled by random cascades at time scales smaller than a cutoff value and be represented as an fBm at larger scales. For IP traffic, the cutoff scale is of the order of one Round Trip Time (RTT), while for VBR video it is typically of the order of a frame duration. Erramili et. al. showed that much more accurate results can be obtained by using their model rather than using monofractal models.

Other models based on multiplicative cascade have been proposed. These models map a given sample into a binary multiscale tree [9]. Each node in the tree corresponds to the aggregation of the traffic mapped into its descendents. Thus, nodes at higher levels of the tree correspond to coarser time scale whereas nodes at lower levels correspond to finer time scales. The multipliers (weights) assigned to each descendent of a node can be set to represent a specific marginal distribution and scaling. In the wavelet-domain independent model (WIG)[10], multipliers are independent additive innovations and correspond to the Haar wavelet coefficient of the process represented by the binary multiscaling tree. As the depth of the tree goes to infinite, the marginal traffic distribution tends to a Gaussian. In the Multifractal Wavelet Model (MWM)[11], multipliers are multiplicative innovations, generating a log-normal marginal distribution, approximately. Both models require the setting of $2 + log_2 N$ parameters where N is the sample size. It has been shown that MWM captures more precisely the dynamics of real traces than does WIG. The major drawback of these models, however, are the number of parameters to be fitted. Moreover, they require the construction of multiscaling binary tree which is not suitable for on-line characterization. These aspects prevent the use of these models for real-time bandwidth management since the parameters of processes resulting from the aggregation of distinct traffic streams need to be computed on-line.

3 The Multifractional Brownian Motion Process

The local Holder regularity is related to scaling at small time scales since it expresses the regularity of the sample path of a process by comparing it to a power-law function[5]. The exponent of this power law, $h(t)$, is called Holder exponent and depends both on time and on the sample path. The Holder exponent is the largest value of h, $0 \leq h \leq 1$, such that

$$|X(t+\gamma) - X(t)| \leq k|\gamma|^h \quad \text{for } \gamma \to 0$$

For monofractal processes the Holder function (Hurst parameter) is a constant value whereas for multifractal processes the Holder function changes randomly with time. Let $H : (0, \infty) \to (0, 1)$ be a Holder function. The multifractional Brownian motion is a continuous Gaussian process with non-stationary increments defined on $(0, \infty)$ as:

$$\begin{aligned} W_{H(t)} &= \frac{1}{\Gamma(H(t)+1/2)} \\ &\quad \left\{ \int_{-\infty}^{0} [(t-s)^{H(t)-1/2} - (-s)^{H(t)-1/2}] dB(s) + \right. \\ &\quad \left. \int_{0}^{t} (t-s)^{H(t)-1/2} dB(s) \right\} \end{aligned}$$

where $B(s)$ is the Brownian motion.

The multifractal brownian motion process is a generalization of the fractal brownian motion process and exhibits the nice property that locally it is assintoticaly self-similar (lass), i.e.

$$\lim_{\rho \to 0^+} \left\{ \frac{W(t+\rho u) - W(t)}{\rho^{H(t)}} \right\}_{u \in R^+} = \{B_{H(t)}(u)\}_{u \in R^+}$$

where $W(.)$ is an mBm and $B_{H(t)}(u)$ is an fBm process with Hurst parameter H, given by $H(t)$.

4 An Novel Estimator for the Holder Function

Evaluating the Holder function value is crucial to the characterization of multifractal traffic. In this section, a novel estimator is presented. It is based on the parametric estimator introduced in [8], which assumes that the Holder function is continuous as well as that it is a constant value in the neighbourhoods of a point. Let N be the number of data samples of an mBm $W(.)$, $H\left(\frac{i}{N-1}\right)$ can be estimated as [8]

$$\hat{H}_{\frac{i}{N-1}} = \frac{-\log(\sqrt{\frac{\pi}{2}} S_{k,N}(i))}{\log(N-1)} \quad 1 \leq i \leq N-2,$$

where $S_{k,N}(i) = \frac{m}{N-1} \sum_{j \in [i-k/2, i+k/2]} \|W(j+1) - W(j)\|$ e $m = \frac{N}{k}$.

The optimum size of the neighbourhood of a point to produce accurate estimations was not estabilished in [8]. Extensive experiments with synthetic data were conducted in order to evaluate this optimum value. It was observed that the estimated value converges well to the true value for neighbourhoods of size $Ne = 2^{j-2}$ for a data set of $N = 2^j$ samples. Such neighbourhood leads to the waste of half of the data. To decrease such high overhead, the smoothing parametric estimator was defined. Firstly, the smoothing parametric estimator computes the Holder function values using small neighbourhoods and then applies the Savitzky-Goloy [12] smoothing filter to the computed values. The smoothing parametric estimator is, thus, given by:

Step 1: Let $Ne = \lfloor j/2 \rfloor - 1$ be the neighbourhood size where $N = 2^j$ and N is the size of the sample;

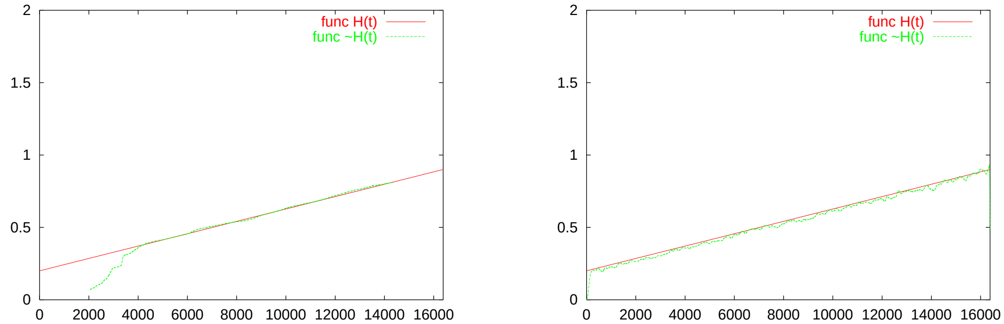
Step 2: Use the parametric estimator defined in [8] with the neighbourhood definde in Step 1 to compute the Holder function values of process $W(t)$;

Step 3: Apply the Savitzky-Goloy smoothing filter to the data set computed in Step 2.

Figures 1 to 2 compare the accuracy of the estimated values by the parametric estimator and by the smoothing parametric estimator for the following Holder functions:

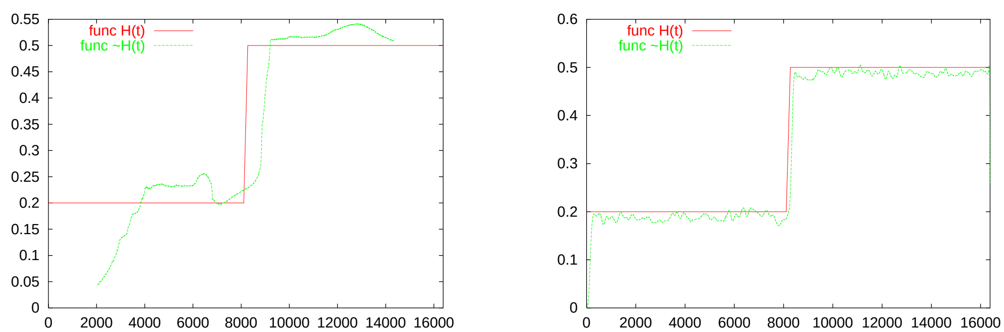
$$\begin{aligned} H(t) &= (0.7 * t + 0.2) \qquad t \in (0,1) \text{ (Figure 1)} \qquad (1) \\ H(t) &= \begin{cases} 0.2 & \text{if } t \in (0, 0.49] \\ 0.5 & \text{if } t \in [0.5, 1.0) \end{cases} \qquad \text{(Figure 2)} \end{aligned}$$

It is clear that the smoothing parametric estimator produces more accurate estimations than the parametric estimator. Note in Figure 1 that it takes 2000 data samples for the parametric estimator to stabilize and start producing reliable estimations, whereas the smoothing parametric estimator produces accurate results for the whole trace. Note also in Figure 2, that the parametric estimator takes the same amount of data to detect the change in the Holder function value and 3000 data to approximate the initial step value. Moreover, its estimations deviate significantly from the true values. Such deviations are not observed for the values produced by the smoothing parametric estimator which follow closely the Holder function values, even at discontinuities.



(a) Parametric (b) Smoothing parametric

Figura 1: $H()$ e $\hat{H}()$ for linear function



(a) Parametric (b) Smoothing parametric

Figura 2: $H()$ and $\hat{H}()$ for step function

5 An Envelope Process for the Multifractal Brownian Motion Process

To solve a queueing system fed by an input process, it is necessary to know the amount of work that arrived to the system. Envelope processes are upper bounds to such amount, and are, thus, less complex to characterize than the exact input process. Envelope processes can be either deterministic or probabilistic. In deterministic envelopes, the amount of work arrived never surpasses the envelope value whereas in probabilistic envelopes it may surpass with a certain pre-defined probability. Consequently, probabilistic envelope processes are tighter bounds than deterministic envelopes. Dimensioning based on deterministic envelope processes may lead to waste of bandwidth, since the provision of bandwidth needs to take into account the maximum amount of work arrived at any time. When probabilistic envelopes are used, there is no need to consider spikes of work up to a certain amount defined by the probability of violation. However, loss of packets may occur.

An upper bound for the accumulated amount of work arrived can be computed as the mean amount of work plus an upper bound to the accumulated increments. An upper bound for mBm increments can be expressed as upper bounds for the local fBm increments, given that in the neighbourhood of time t, an mBm can be approximated by an fBm with Hurst parameter $H(t)$. It is known that [13][14]:

$$Z_H(m) \leq \kappa H t^{H-1} \tag{2}$$

As the size of local infinitesimal neighbourhood of t goes to zero, the envelope process, $\hat{A}(t)$, of an mBm with mean $\bar{a}$, standard deviation σ and Holder function $H(.)$ is given by:

$$\hat{A}(t) = \int_0^t \bar{a} + \kappa\sigma H(x) x^{H(x)-1} dx \tag{3}$$

which is called mBm enveloope process.

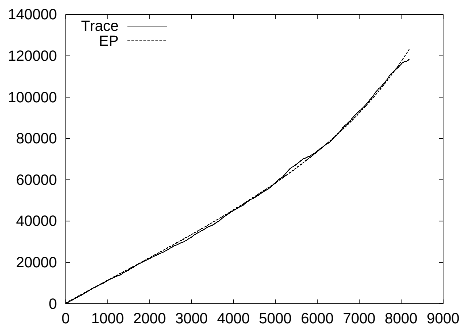
This envelope reduces to the fBm envelope derived in [13] when $H(.)$ is a constant, i.e.,

$$\hat{A}(t) = at + \kappa\sigma t^H$$

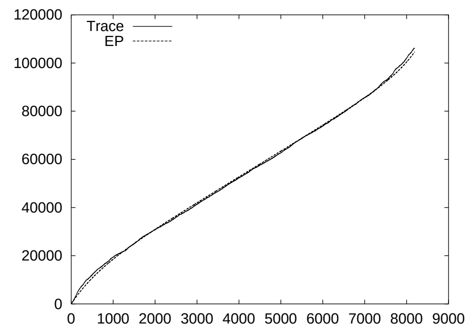
Extensive simulation experiments using both synthetic traffic and real network traffic were conducted in order to assess the accuracy of the proposed envelope. The mBm generator defined in [8] was utilized to generate samples of mBm of up to 10^6 data sample. Different Holder functions were experimented. In Figure 3, results are shown for the following Holder functions:

$$\begin{aligned} H(t) &= 1.9 * (t-0.5) * (t-0.5) + 0.51 \quad t \in (0,1); \\ H(t) &= 0.5 + t/2.0 \quad t \in (0,1). \end{aligned} \tag{4}$$

It can be observed that the mBm envelope process is a tight bound to the modeled process regardless of its Holder function. Violations of the established bound do exist but are within the expected value given by the pre-defined violation probability value.



(a) Linear



(b) Quadratic

Figura 3: The mBm envelope process evaluation for synthetic traces

Traces of real network traffic used here are the same ones used in previous investigations by others [4][11]. Figure 4 shows the mBm envelope process for the trace [15] used in [4]. These traces were collected at DEC network and at LBL, and have approximately $2.4 * 10^7$ packets. mBm envelope processes were derived for all these traces and it was verified that the mBm envelope processes are tigh bounds to all traces used, as can be seen, in Figure 4, for a trace of $5.6 * 10^6$ packtes collected at DEC network.

Traces gathered at the public Internet access point at Auckland University [16], utilized in [11], were also used here. Data were collected with a precision of the order of microseconds. Several time serie can be obtained from these traces, e.g. packet interarrivel time, connection time duration and packet traffic. Figure 5 shows the mBm envelope process caracterization for seven hours time serie of the TCP/IP packets (data set Auckland-II). It can be observed that the mBm envelope process is also a reasonably accurate bound for this data set.

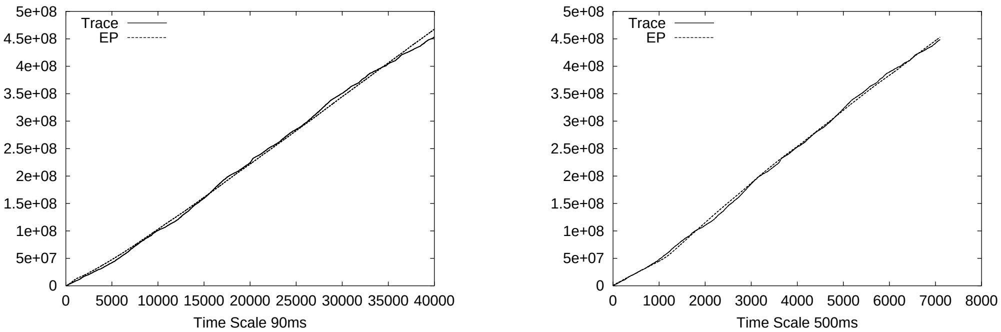
6 Time Scale of Interest

This section characterizes the time at which a queue reaches its maximum occupancy, in a probabilistic sense. The queue size at this time provides a simple delay bound [13]. Consider a continuous-time queueing system, with deterministic service given by C. The cumulative arrival process is represented by $A_{H(t)}(t)(A_{H(t)}(0) = 0)$. Let $\hat{A}_{H(t)}(t)$, continuous and differentiable, be the probabilistic envelope process of $A_{H(t)}(t)$, such that $P(A_{H(t)}(t) > \hat{A}_{H(t)}(t)) \leq \epsilon$.

During a busy period, which starts at time 0, the number of cells in the system at time t is given by $q(t)$. Thus, $q(t) = A_{H(t)}(t) - ct \geq 0$.

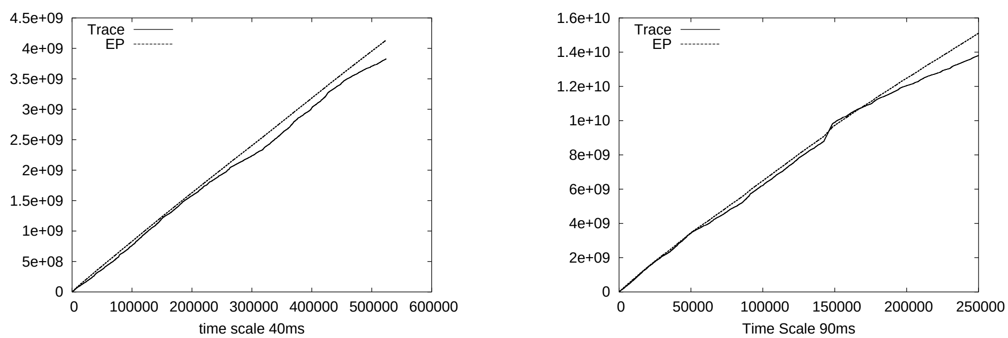
By defining $\hat{q}(t)$ as

$$\hat{q}(t) = \hat{A}_{H(t)}(t) - Ct \geq 0, \tag{5}$$



(a) (b)

Figura 4: The mBm envelope process evaluation for real network traffic (The Digital data set)



(a) (b)

Figura 5: The mBm envelope process evaluation for real network traffic (Auckland-II data set) we can see that $P(q(t) > \hat{q}(t)) = P(A_{H(t)}(t) > A_{H(t)}(t)) \leq \epsilon$

The maximum delay in a FIFO queue is given by the maximum number of cells in the queue during the busy period, which can be defined as

$$q_{max} = max(\hat{q}(t)) \qquad t \geq 0$$

Therefore,

$$P(q(t) > q_{max}) \leq P(q(t) > \hat{q}(t)) \leq \epsilon$$

$$P(q(t) > q_{max}) \approx \epsilon.$$

The queue length at time t, $q(t)$, will thus only exceed the maximum queue length q_{max} with probability ϵ. In other words, only when the arrival process exceeds the envelope process, will the maximum number of cells in the system exceed the estimated value. Intuitively, by bounding the behaviour of the arrival process, it is possible to transform the problem of obtaining a probabilistic bound for the stochastic system defined by $q(t) = A_{H(t)}(t) - Ct$, into an easier problem of finding the maximum of a deterministic system, described by $\hat{q}(t) = \hat{A}_{H(t)}(t) - Ct$.

The mBm process is inserted into Equation 5 giving:

$$\hat{q}(t) = \hat{A}_{H(t)}(t) - Ct = \int_0^t \bar{a} + \kappa\sigma H(x) x^{H(x)-1} dx - Ct \tag{6}$$

In order to compute q_{max}, it is necessary to find $t^\star$ such that

$$\frac{d\hat{q}(t)}{dt} = 0$$

or equivalently,

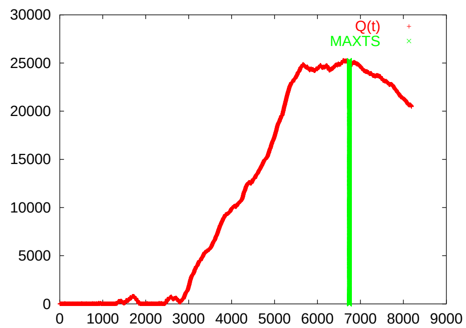
$$\frac{d\hat{A}(t)}{dt} = C \tag{7}$$

Hence, $t^\star$ is given by

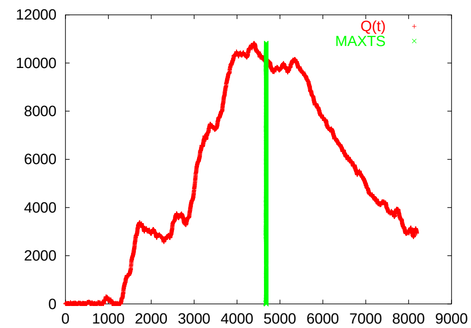
$$t^\star = \left[\frac{\kappa\sigma H(t^\star)}{(C - \bar{a})} \right]^{\frac{1}{1-H(t^\star)}} \tag{8}$$

The time-scale of interest is defined by the time until a queue size reaches its peak, denoted by $t^\star$. This is denominated the Maximum Time-Scale (MaxTS), and it defines the point in time where the unfinished work in the queue achieves its maximum in a probabilistic sense.

To evaluate the precision of the expression for the time scale of interest, simulation experiments where conducted. A queue with constant service rate was fed by an mBm and the queue length was recorded. The estimated time scale of interest was computed and compared to the time at which the queue length reaches its maximum in the simulation



(a) H() a quadratic function



(b) H() a cubic function

Figura 6: The MaxTS for synthetic mbm processes

experiments. Figure 6 shows results for mBms with Holder functions defined in Equation 4. The computed value is the same found in the simulation experiments for the linear and for the quadratic functions, whereas it closely approximates the one found in the simulation experiments for the cubic function. Note that, such deviation is within the known error margin established by the violation probability value.

7 Conclusions

Scaling analysis of IP and video traffic have pointed out their multifractal nature. Models based on multiscaling have been proposed in the literature. These models, however, need the knowledge of the whole stream beforehand. Moreover, the number of parameters to be fitted depends on the sample size.

In this paper, a novel probabilistic envelope process for multifractal traffic modeling was introduced. The envelope process is an upperbound to the amount of work arrived from a multifractal Brownian motion process. Extensive simulation experiments using both synthetic and real network traffic show that the proposed model is a tigh bound to the modeled traffic. Moreover, expressions for the time scale of interest and a new estimator for the Holder function were presented.

The mBm envelope process is instrumental for dynamic bandwidth management. Statistical multiplexing of multifractal streams is currently under investigation as well as its use in admission control to DiffServ domains.

Referências

[1] W. Leland, M. Taqqu, W. Willinger, and D. Wilson. On the self-similar nature of ethernet traffic (extended version). *IEEE/ACM Trans. Networking*, pages 1–15, February 1994.

[2] R. Riedi and J. Levy-Vehel. Tcp traffic is multifractal: A numerical study. Technical Report 3129, INRIA Rocquencourt, March 1997.

[3] A.Erramilli, O.Narayan, A.Neidhart, and I.Saniee. Performance impacts of multi-scaling in area TCP/IP traffic. In *INFOCOM 2000*, pages 352–359, April 2000.

[4] A.Erramilli, O.Narayan, A.Neidhart, and I.Saniee. Multi-scaling models of TCP/IP and sub-frame VBR video traffic. *Journal of Communications and Networks*, 3(4):383–395, December 2001.

[5] P. Abry, R. Baraniuk, P. Flandrin, R. Riedi, and D. Veitch. The multiscale nature of network traffic: Discovery, analysis and modelling. *IEEE Signal Processing Mag.*, 19(3):28–46, 2002.

[6] Anja Feldmann, Anna C. Gilbert, Polly Huang, and Walter Willinger. Dynamics of ip traffic: a study of the role of variability and the impact of control. In *Proceedings of the conference on Applications, technologies, architectures, and protocols for computer communication*, pages 301–313. ACM Press, 1999.

[7] R. H. Riedi and Walter Willinger. *Self-Similar Network Traffic and Performance Evaluation*, chapter Toward and Improved Understanding of Network Traffic Dynamics. John Wiley & Sons, Inc., 2000.

[8] R. Peltier and J.L.Vehel. Multifractional brownian motion: definition and preliminary results. Technical report, INRIA, 1995.

[9] R. H. Riedi, M. S. Crouse, V. J. Ribeiro, and R. G. Baraniuk. A multifractal wavelet model with application to network traffic. *IEEE Trans. Inform. Theory*, 45(4):992–1018, 1999.

[10] Sheng Ma and Chuanyi Ji. Modeling heterogeneous network traffic in wavelet domain. *IEEE/ACM Trans. Networking*, 9(5):634–649, 2001.

[11] Vinay J. Ribeiro, Rudolf H. Riedi, Matthew S. Crouse, and Richard G. Baraniuk. Multiscale queuing analysis of long-range-dependent network traffic. In *INFOCOM (2)*, pages 1026–1035, 2000.

[12] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery. *Numerical Recipes in C : The Art of Scientific Computing.* Cambridge University Press, January 1993.

[13] Gilberto Mayor. *Performance Modeling and Network Management for Self-Similar Traffic.* PhD thesis, University of Southern California, 1997.

[14] N.L.Fonseca, G.S. Mayor, and C.A.V Neto. On the equivalent bandwidth of self-similar source. *ACM Transactions on Modeling and Computer Simulation*, 10(2):104–124, April 2000.

[15] http://ita.ee.lbl.gov/html/contrib/DEC-PKT.html.

[16] J. Micheel, I. Graham, and N. Brownlee. The auckland data set: an access link observed. In *14th ITC Specialist Seminar*, Barcelona, April 2000.