A Multiplicative Multifractal Model for TCP Traffic *

J. Lévy Véhel and B. Sikdar
Projet Fractales, INRIA Rocquencourt
B. P. 105, 78153 Le Chesnay Cedex, France.
email: {jlv,sikdar}@bora.inria.fr

Abstract

Recent studies have shown that TCP traffic displays strong multifractal scaling. However, a physical explanation of why such a behavior occurs is still eluding. In this paper, we propose a cascade model that is based on the retransmission and congestion avoidance mechanisms of TCP. At the same time, it relates to the physical, tree-like organization of networks. This model allows to relate the most salient multifractal features with basic traffic parameters as the RTT and the loss probability. Numerical experiments confirm that such a parsimonious model is able to give a satisfactory explanation for a number of features pertaining to multifractality, including the range of scales where it is observed. We believe these results open the way to a more profound understanding of the small time scale properties of TCP traffic.

1. Introduction

Self-similarity of network traffic [8] has been investigated and its presence proved in traffic originating from various network environments and associated applications. More recent research, beginning with [13], has in addition shown the multifractal nature of network traffic [6, 9].

These two fractal aspects of traffic (self similarity and multifractality) are not contradictory. They may coexist, as the example of the binomial cascade shows [9]. While long range dependence studies the ”large” time scale properties, more recent work has focused on ”high” frequency aspects as investigated by multifractal analysis. In this paper, we investigate the underlying causes which result in the multifractal nature of TCP traffic. We develop a cascade structure based on TCP's retransmission and congestion control mechanism and use it to explain the multifractality in TCP traffic. Past research on the causes for multifractality have suggested the protocol hierarchy of IP data networks and the multiplicative nature of data passing through a number of queues in succession [5, 9]. In [5] simulations with *ns* were used to examine the effect of user or session behavior and network configuration on the scaling behavior of IP traffic. However, the impact of the adaptive nature of TCP and the traffic dynamics introduced by its loss recovery and congestion control mechanisms remained unexplained.

Through our cascade construction, we analytically show the impact of the loss rates experienced by a TCP flow, the round trip time (RTT) and the expected duration of the timeouts on the scaling behavior of traffic. Our analysis shows the high sensitivity of the shape and peak of the multifractal (Legendre) spectrum on these parameters. We also report on the results of multifractal analysis on TCP traces which show the dependence of the scaling behavior on the loss rates experienced by the flows. Thus we are able to make similar observations on TCP traces collected from the Internet and on our model.

The rest of the paper is organized as follows. In Section 2 we give a brief introduction to multifractal analysis. We present the cascade construction of the traffic generated by a TCP flow and its analysis in Section 3. In Section 4 we validate our model using tests on TCP traces from the Internet. Finally, Section 5 presents the concluding remarks.

2. Multifractal Analysis

Multifractal analysis deals with the description of the singularity structure of "signals" (which can be measures [1, 11], functions [7] or capacities [10]), both in a local and a global way. The local information is given by the *Hölder exponent* at each point, while the global information is captured through a characterization of the geometrical or statistical distribution of the occurring Hölder exponents, called *multifractal spectrum*. The way to measure local irregularity mainly depends on if one is dealing with functions or measures. For a nowhere differentiable signal X on $[0,1]$ one usually defines the *coarse Hölder exponents* through

$$\alpha_n^k = -\frac{1}{n}\log\left|X((k+1)2^{-n}) - X(k2^{-n})\right|$$

*This work was supported Project Epsilon, INRIA, France

where all logarithms are taken to the base 2 and where $\log 0 := -\infty$. For a fixed t in [0,1] let (k_n) be such that $I_n(t) := I_n^{k_n}$ contains t. The limiting exponent at t

$$\alpha(t) := \liminf_{n\to\infty} \alpha_n^{k_n} \quad (1)$$

is called the *local* Hölder exponent of X at t. For a measure μ[1] one sets:

$$\alpha_n^k = \frac{\log \mu(I_n^k)}{-n}$$

where $I_n^k = [k\,2^{-n}, (k+1)2^{-n}[, k = 0\ldots 2^n - 1$. Again, for any fixed t in [0,1], $\alpha(t)$ is defined as in Eqn. 1.

Multifractal analysis consists in giving a compact representation of the local singularity structure $\alpha(t)$ of a signal. In that view, two approaches arise naturally: a *geometrical* description, or a *statistical* one. The former leads to the Hausdorff spectrum f_h which we will not consider here, while the latter leads to the large deviation spectrum f_g. A third spectrum is also usually considered, the Legendre spectrum, the interest of which is to provide a simple way to compute f_g when the conditions of the Gärtner-Ellis theorem [1, 9] are met. In this paper we use the Legendre spectrum as it is easier to compute. Define for any real number q

$$S_n(q) = \sum_{k=0}^{2^n-1} \mu(I_n^k)^q \quad (2)$$

with the convention $0^q := 0$ for all q. Next, define the 'structure function'

$$\tau(q) = \liminf_{n\to\infty} \frac{\log S_n(q)}{-n}$$

and the so-called *Legendre spectrum*

$$f_l(\alpha) = \tau^*(\alpha) := \inf_q(\alpha q - \tau(q)).$$

3. Cascade Construction of TCP Flows

For traffic models in the multifractal frame, two directions have been explored. The first one uses a generalization of fractional Brownian motion known as multifractional Brownian motion [2]. The second class of models relies on multinomial cascades [6, 13] which are very appealing from a multifractal point of view. However, to be plausible, such a model must be supported by physical evidence that real traffic behaves as a cascading process. To our knowledge, such a justification is still lacking. In this section we take a few steps in this direction.

[1]For our purpose, a measure will be a function $\mu : \mathcal{B} \to [0,1]$ (where $\mathcal{B}$ is the set of Borel subsets of [0,1]) which is σ-additive, i.e. $\mu(\cup_{n\geq 1} A_n) = \sum_{n\geq 1} \mu(A_i)$ for all countable collections $(A_n)_{n\geq 1}$ of elements of $\mathcal{B}$ with $A_i \cap A_j = \emptyset \; \forall\, i \neq j$.

Developing a cascade model amounts to modeling the distribution of the traffic load on the different nodes of the network with an equivalent cascade. We thus have the following intuitive interpretation for our multiplicative process: at a given fixed time, the whole network N has a load L. Dividing the network into P sub-networks $N_1^1, N_2^1, \ldots N_P^1$ we assume that N_1^1 supports the load $M_1^1 L$, $\ldots N_P^1$ supports $M_P^1 L$. Now each sub-network N_l^1 can itself be divided into sub-sub-networks $N_{l,1}^2, \ldots N_{l,P}^2$ which support $M_{l,1}^2 M_l^1 L, \ldots M_{l,p}^2 M_l^1 L$, and so on.

A plausible explanation for the occurrence of a multiplicative process is then that the global network is indeed composed of several such layers, and that the topology distributes itself along these layers in a random way but with approximately the same law at each stage. Reports in [3] and [4] have shown that the Internet topology can be closely modeled using graph theoretical approaches and that there are indeed power laws which can be used to describe the various graph properties. This adds credibility to our proposed explanation.

3.1. A Cascade Construction

In general, when looking for multifractal scaling, it is always is important to identify which range of scale is physically relevant. For TCP traffic, we now identify three distinct regions in scale. At the lowest scale we consider periods lower than the RTT where the scaling is controlled by by variations queueing delays, ACK compression etc. At the highest level we have times ranging from the order of a few seconds and above. This region corresponds to the "steady-state" behavior of TCP flows. Finally, the middle time scale corresponds to the times between the lower and upper region where packet dynamics are dependent on the variations in window size in each RTT.

In this paper, we concentrate exclusively on the highest scale. For this region, we propose a cascade construction for the number of packets transmitted by TCP in a given unit of time. In this paper we use TCP Reno as our model for TCP as it is the most widely implemented flavor of TCP. We assume that the reader is familiar with the basic concepts of TCP like the congestion window $cwnd$, slow start, delayed acknowledgments etc and is referred to [14] for further details.

We assume that the receiver sends one ACK for every b packets it receives. In the following, we assume an infinite TCP flow to make the illustrations clearer. Now, TCP responds to loss indication with either a fast retransmit or a timeout. Thus the flow can be seen as a sequence of linear and silent periods corresponding to the congestion avoidance and timeout periods respectively. We break up each flow in pieces which are marked by two successive timeout periods. The two successive timeouts are separated by a sequence of loss indications each of which is recovered using a fast retransmit. We take the expected duration of the time between two successive timeouts as our basic unit of time and the multiplicative nature of cascade comes into effect for larger time scales. This expected time is directly proportional to the packet loss probability and the RTT.

We denote the packet loss probability by p. The expected duration of a timeout is denoted by $E[TO]$ and from [12]

$$E[TO] = T_O \frac{1 + p + 2p^2 + 4p^3 + 8p^4 + 16p^5 + 32p^6}{1 - p} \quad (3)$$

where T_O is the period of time a sender waits before retransmitting an unacknowledged packet. Also, the probability that an arbitrary loss indication leads to a timeout is denoted by Q, the expected duration of a congestion avoidance phase by $E[A]$ and the expected number of packets transmitted in the timeout and the congestion avoidance periods by $E[R]$ and $E[Y]$ respectively (see Fig. 1). Finally, we denote by $E[W_u]$ the expected value of the $cwnd$ without the effects of window limitation. From [12]

$$E[W_u] = \frac{2+b}{3b} + \sqrt{\frac{8(1-p)}{3bp} + \left(\frac{2+b}{3b}\right)^2} \quad (4)$$

Accounting for the effect of window limitation, the expected window size $E[W]$ is now equal to $\min(W_{max}, E[W_u])$. Following the derivation in [12]

$$Q = \min\left(1, \frac{(1 + (1-p)^3(1-(1-p)^{E[W]-3}))}{(1-(1-p)^3)^{-1}(1-(1-p)^{E[W]})}\right) \quad (5)$$

$$E[A] = \begin{cases} \left(\frac{b}{2}E[W_u] + 1\right) RTT & \text{if } E[W_u] < W_{max} \\ \left(\frac{b}{8}W_{max} + \frac{1-p}{pW_{max}} + 2\right) RTT & \text{otherwise} \end{cases} \quad (6)$$

$$E[Y] = \begin{cases} \frac{1-p}{p} + E[W] & \text{if } E[W_u] < W_{max} \\ \frac{1-p}{p} + W_{max} & \text{otherwise} \end{cases} \quad (7)$$

$$E[R] = \frac{1}{1-p} \quad (8)$$

The expected throughput of the TCP connection as a function of the loss probability p can be calculated as

$$B(p) = \frac{E[X]E[Y] + E[R]}{E[X]E[A] + E[T_O]} \quad (9)$$

where X is a geometric random variable with parameter Q and denotes the number of congestion avoidance phases between two successive timeout periods.

We now propose a dyadic cascade. The interval I on which the cascade develops is made of a large number of sequences of congestion avoidance and timeout periods. By convention, we take $I = [0, 1]$. We denote the number of sequences in I by 2^m. At the i^{th} level, each subinterval

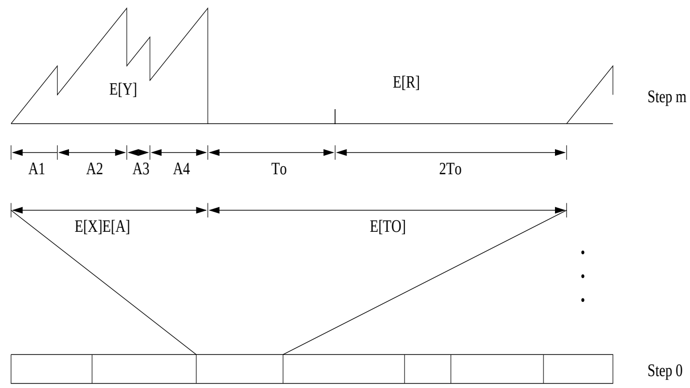


Figure 1. The timeout and congestion avoidance phases in a TCP flow and their relation to the cascade construction.

is composed of 2^{m-i} such sequences. The cascade has a mass $2^m(E[Q]E[Y] + E[R])$. In a dyadic cascade, each subinterval of the previous stage is divided into two to form the next stage. Denote the k^{th} subinterval at stage p by

$$I_p^k = [a_p^k, b_p^k] \quad (10)$$

Then the endpoints of the subintervals resulting from the splitting of this interval in the next stage, I_{j+1}^{2k} and I_{j+1}^{2k+1} are given by

$$\begin{aligned} I_{p+1}^{2k} &= [a_p^k, a_p^k + r_p^k(b_p^k - a_p^k)] \\ I_{p+1}^{2k+1} &= [a_p^k + r_p^k(b_p^k - a_p^k), b_p^k] \end{aligned} \quad (11)$$

where r_p^k is a random number in $(0, 1)$ drawn according to a law described below.

$$r_p^k = \frac{X_1 E[A] + E[TO]}{(X_1 + X_2)E[A] + 2E[TO]} \quad (12)$$

where X_1 and X_2 are i.i.d. geometric random variables with parameter Q. For mathematical convenience, we in fact draw X_1 and X_2 from a truncated geometric distribution i.e. $P\{X = k\} = (1-Q)/(1-Q^{N+1}) \cdot Q^{k-1}$ for $k = 1, \cdots, N$, where N is some large number. This truncation has little practical impact. Similarly, the mass of the k^{th} subinterval at stage p, C_p^k, is split into two parts in the $p+1^{\text{th}}$ stage. The subinterval I_{p+1}^{2k} gets a mass $m_p^k C_p^k$ while the subinterval I_{p+1}^{2k+1} gets the mass $(1 - m_p^k)C_p^k$. The random variable m_p^k takes values in (0,1) and is given by

$$m_p^k = \frac{X_1 E[Y] + E[R]}{(X_1 + X_2)E[Y] + 2E[R]} \quad (13)$$

where X_1 and X_2 are the same geometric random variables used for the time process.

3.2. Computation of the multifractal spectrum

Our cascade model has the following features:

- At each step, the ratio between the size of the intervals I_p^k and I_{p+1}^{2k} is r_p^k, and between I_p^k and I_{p+1}^{2k+1} is r_p^{k+1}, with $r_p^k + r_p^{k+1} = 1$.
- The ratios between the masses in the same intervals are m_p^k and m_p^{k+1}, with again $m_p^k + m_p^{k+1} = 1$.
- For all k, p we have: $0 < r_{\min} \leq r_p^k < 1, 0 < m_{\min} \leq m_p^k < 1$ with probability one.
- r_p^k and m_p^k have the same distribution for each p.
- For all k, p the interiors of I_p^k and I_k^{p+1} do not overlap.

We can then make use of results in [1], which yield the multifractal spectrum of the cascade. In our case where in particular r_1^1 and r_1^2 (resp. m_1^1 and m_1^2) are identically distributed, these take the following form: for any real q, there exists a unique $\beta(q)$ such that $E(m^q)E(r^{\beta(q)}) = \frac{1}{2}$ where m (resp. r) is distributed as m_1^1 (resp. r_1^1). Let $\alpha^{\min} = ess\inf \frac{\log m}{\log r}, \alpha_{\max} = ess\sup \frac{\log m}{\log r}$. It is easy to see that, in our case,

$$\alpha_{\min} = \frac{E[A] + E[TO]}{E[A]N\log N}, \alpha_{\max} = \frac{E[Y]N\log N}{E[Y] + E[R]}.$$

Then, almost surely

$$\begin{aligned} f_g(\alpha) &= -\infty \text{ if } \alpha \notin [\alpha_{\min}, \alpha_{\max}] \\ f_g(\alpha) &= q\alpha(q) + \beta(q) \text{ otherwise,} \\ \text{where } \alpha(q) &= -\beta'(q) \end{aligned}$$

Note also that $f_l(\alpha) = q\alpha(q) + \beta(q)$ is equal to $f_g(\alpha)$ inside $[\alpha_{\min}, \alpha_{\max}]$, but assumes non trivial negative values outside this interval. The usual interpretation of the negative value is that an α with $f_l(\alpha) < 0$ does not occur in any given trace with probability one, but might be observed in one of a set of $\exp(-n\log(2)f_l(\alpha))$ independent traces [11]. Finally, the mode α_0 of $f_g(\alpha)$, which corresponds to the almost sure behavior in any given trace is given by the following formula, which readily follows from the law of large numbers:

$$\alpha_0 = \frac{E[\log(m)]}{E[\log(r)]}$$

An easy computation yields (for $j, k = 1, \ldots, N$)

$$P(m = x) = 0 \text{ if } x \neq m_{j,k} := \frac{a_j + b}{a(j+k) + 2b}$$

$$P(m = m_{jk}) = P(X_1 = j)P(X_2 = k)$$

So that

$$\alpha_0 = \frac{\sum_{j,k=1}^N \log(m_{j,k})Q^{j+k-2}}{\sum_{j,k=1}^N \log(r_{j,k})Q^{j+k-2}}$$

with $r_{j,k} = \frac{E[A]j + E[TO]}{E[A](j+k) + 2E[TO]}$. It is interesting to investigate how $\alpha_{\min}, \alpha_0$ and $\alpha_{\max}$, which are the three most important parameters in f_g, vary with our parameters $E[A], E[TO], E[Y], E[R], Q$ and N. It is easy to see that α_0 will be an increasing or decreasing function Q depending on whether

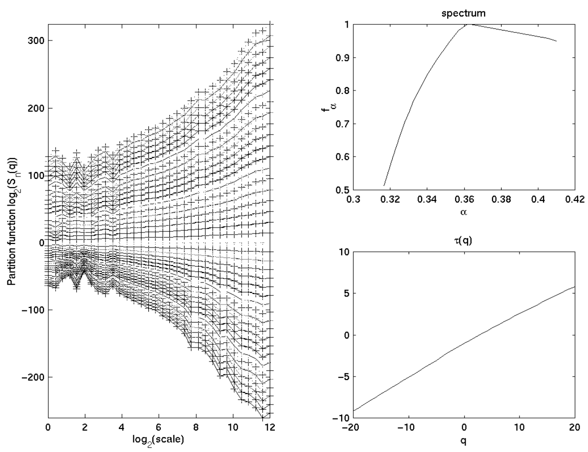
$$\sum_{j,k,l,m=1}^N \log(m_{j,k})\log(r_{j,k})(Q)^{j+k+l+m-4}(j+k-l-m)$$

is positive or negative. Note that while $\alpha_{\min}$, which correspond to the highest degree of burstiness only depends on $E[A], E[TO]$ and N, $\alpha_{\max}$ depends on $E[Y], E[R], N$.

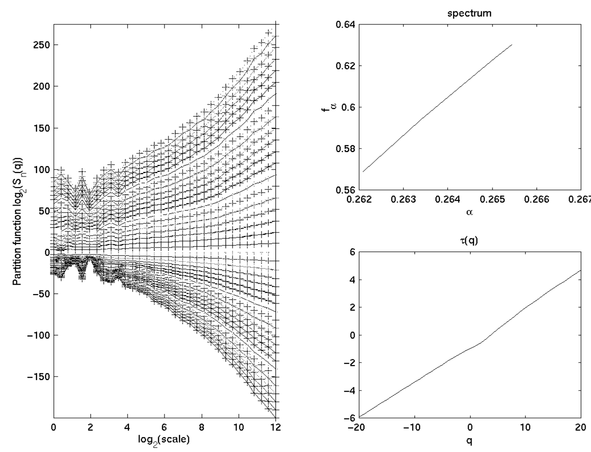
4. Numerical Results

We now carry out analysis for presence of multifractality in traces of TCP flows. Each trace shows multifractal behavior and as predicted by our model, the scaling exponents have a strong dependence on the loss rates, the RTTs and the average timeout durations. The traces were collected for TCP transfers originating from Troy, NY to Columbus OH, Los Angeles CA, Boston MA, and Pisa Italy. Due to space restrictions, we show results for only Ohio and Italy. The results for the others are similar. Each trace is 33 minutes long and was collected using `tcpdump` and have microsecond resolution.

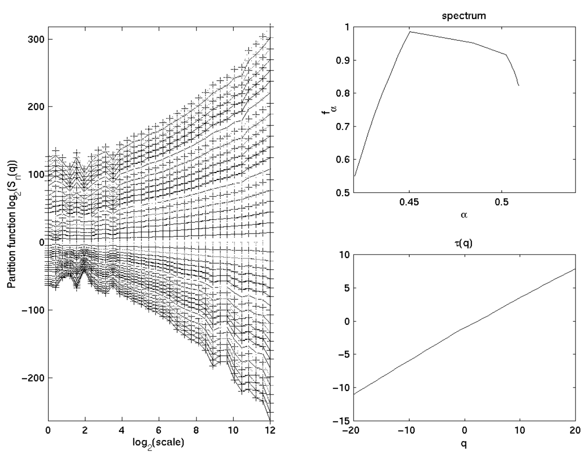
In Fig. 2 we show the scaling behavior of the traces for Ohio corresponding to loss rates of 0.002, 0.010 and 0.135 and for the Italy traces with loss probabilities of 0.000, 0.006 and 0.099. We see that the partition function becomes almost linear after a scale of roughly 5. Also the scaling exponents depend on the loss rates and higher loss rates lead to greater multifractal behavior. In Fig. 3 we plot the theoretical Legendre spectrum from our model with parameters estimated from the traces. The results on the traces qualitatively match the ones on the cascade model. The fact that the empirical spectra do not reach the x-axis is due to finite size effects i.e., the number of available scales is limited and has no physical significance. Also, the figure for the traces corresponding to the Italy has only two curves as for a loss probability of 0, the traffic is no longer multifractal. This is validated by the top right plot of Fig. 2 which shows the absence of any linear regions in the partition function. As a consequence, the computed multifractal spectrum is not relevant. To sum up, for both the estimated and theoretical spectra, we have: (1) As p increases, the width of the spectrum decreases, and (2) The general appearance, with a fast increase for $\alpha < \alpha_0$ and a slower decrease for $\alpha > \alpha_0$ is the same. Also, we get in both cases more asymmetry for low loss probability. Finally, the the lower cutoff scale where we have multifractal scaling is around 2^5 which is roughly equal to the basic unit interval of our model.



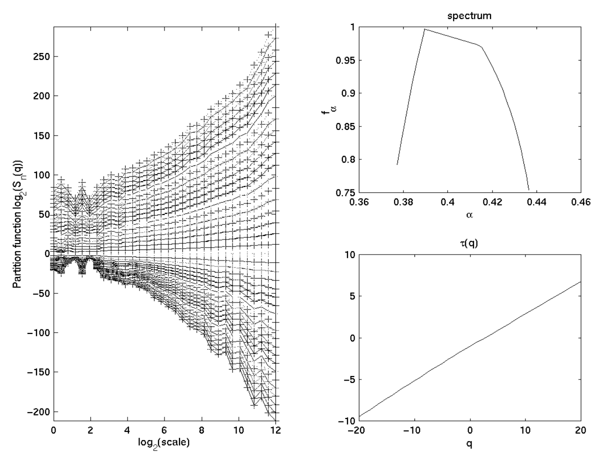
$p = 0.002$



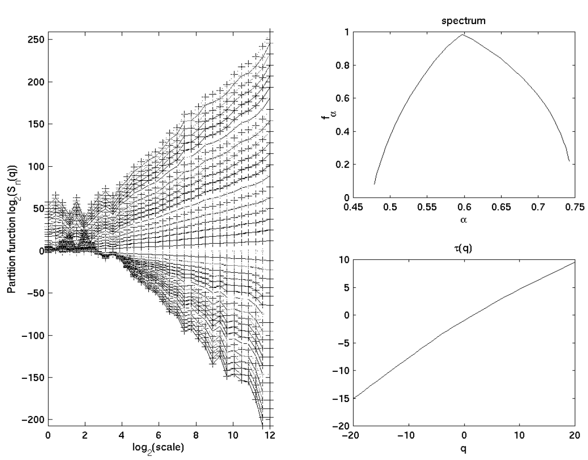
$p = 0.000$



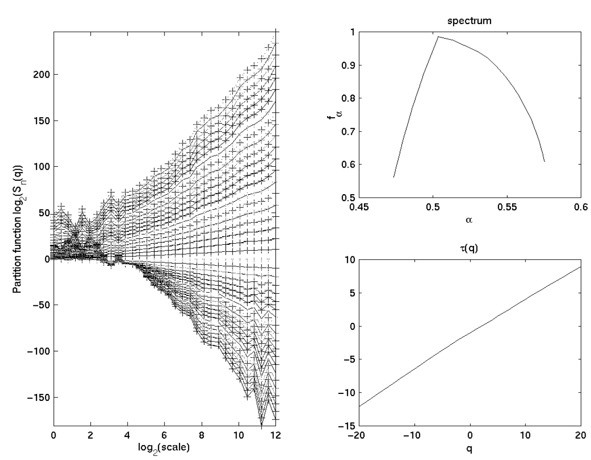
$p = 0.010$



$p = 0.006$



$p = 0.135$



$p = 0.099$

Figure 2. Partition function, Legendre spectrum and $\tau(q)$ for the traces corresponding to Ohio (left) and Italy (right).

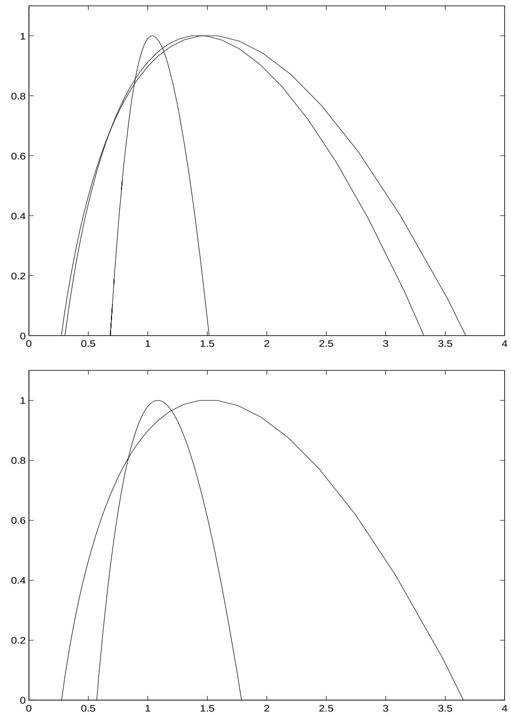


Figure 3. Legendre spectrum computed from our model for the parameters of the Ohio (top) and Italy (bottom) traces.

5. Conclusions

In this paper, we have shown that it is possible to propose a valid cascade model for TCP traffic. This model is based on the retransmission and congestion control algorithms of TCP and allows a parsimonious explanation of the multifractal properties observed in real traces in a physically relevant way. We also identified three different regions in scale in TCP traffic which arise from various factors like the nature of TCP's response to loss indications, to random delays introduced in the paths of individual packets and phenomena like ACK compression.

Our cascade model for TCP traffic shows the important relationship between the scaling behavior and the loss rates experienced by a TCP flow. The model also allows the direct interpretation of the higher and lower cutoff points of the multifractal spectrum in terms of TCP related parameters like the evolution of the congestion window, duration of timeouts etc. From the analysis of our cascade model, it can be concluded that higher loss rates will lead to a more multifractal behavior. Intuitively, this can be explained by considering the fact that with increasing loss rates, the incidence of timeouts increases which in turn increases the burstiness of the traffic. The model also illustrates the dependence of the scaling regions on the physical aspects of the network like the round trip time.

References

[1] M. Arbeiter and N. Patzschke, "Self-Similar Random Multifractals," *Math. Nachr.* vol. 181, pp. 5-42, 1996.

[2] A. Ayache and J. Lévy Véhel, " The generalized multifractional Brownian motion," To appear in SISP.

[3] K. Calvert, M. Doar and E. W. Zegura, "Modeling Internet topology," *IEEE Communications Magazine,* vol. 35, no. 6, pp. 160-163, June 1997.

[4] M. Faloutsos, P. Faloutsos, C. Faloutsos, "On power-law relationships of the Internet topology," *Proceeding of ACM SIGCOMM,* Boston, MA, 1999.

[5] A. Feldmann, P. Huang, A. C. Gilbert and W. Willinger, "Dynamics of IP traffic: A study of the role of variability and the impact of control," *Proceeding of ACM SIGCOMM,* Boston, MA, August 1999.

[6] A. C. Gilbert, W. Willinger and A. Feldmann, "Scaling Analysis of Conservative Cascades, with Applications to Network Traffic," *IEEE Trans. on Information Theory,* vol. 45, no. 3, pp. 971-992, April, 1999.

[7] S. Jaffard, "Multifractal formalism for functions," *CRAS*, vol. 317 pp. 745–750, 1993.

[8] W. E. Leland, M. S. Taqqu, W. Willinger and D. V. Wilson, "On the self-similar nature of Ethernet traffic (Extended Version)," *IEEE/ACM Trans. on Networking,* vol. 2, no. 1, pp. 1-15, Feb 1994.

[9] J. Lévy Véhel and R. Riedi, "Fractional Brownian motion and data traffic modeling: The other end of the spectrum," *Fractals in Engineering,* Eds. J. Lévy Véhel, E. Lutton and C. Tricot, Springer Verlag, 1997.

[10] J. Lévy Véhel and R. Vojak, "Multifractal analysis of Choquet capacities: Preliminary Results," *Advances in Applied Math.*, vol. 20, no. 1, pp. 1-43, January 1998.

[11] B. B. Mandelbrot, "Negative Fractal Dimensions and Multifractals," *Physica A*, vol. 163, pp. 306-315, 1990.

[12] J. Padhye, V. Firoiu, D. Towsley and J. Kurose, "Modeling TCP Reno performance: A simple model and its empirical validation," *IEEE/ACM Trans. on Networking,* vol. 8, no. 2, pp. 133-145, April 2000.

[13] R. H. Riedi and J. Levy-Vehel, "TCP Traffic is multifractal: A numerical study," Preprint, 1997.

[14] W. R. Stevens, "TCP/IP illustrated volume 1," Addison Wesley, 1994.