PERGAMON

Nonlinear Analysis 47 (2001) 5765–5774

Nonlinear Analysis

www.elsevier.nl/locate/na

Multiplicative Multifractal Modeling of Long-Range-Dependent (LRD) Traffic in Computer Communications Networks

Jianbo Gao and Izhak Rubin

Electrical Engineering Department, University of California, Los Angeles Los Angeles, California, 90095, U.S.A. [1]

Abstract

Source traffic streams as well as aggregated traffic flows often exhibit long-range-dependent (LRD) properties. In this work, we model traffic streams using multiplicative multifractal processes. We develop two type of models, the multifractal point processes and multifractal counting processes. We demonstrate our model to effectively track the behavior exhibited by the system driven by the actual traffic processes. We also study the superposition of LRD flows. We prove that the superposition of a finite number of multiplicative multifractal traffic streams results asymptotically in another multifractal stream. Furthermore we demonstrate numerically that the superimposed process can be effectively modeled by an ideal multiplicative process.

Key words: Computer communications networks; Traffic modeling; Long-range-dependent traffic; Multiplicative multifractal processes

1 Introduction

Recent analysis of high–quality traffic measurements have revealed the prevalence of long-range-dependent (LRD) (or self-similar) features in traffic processes loading packet switching communications networks. With LRD traffic measured in many data networks, two related questions arise. One is how to parsimoniously model LRD traffic? The other is: what is the impact of LRD traffic on network performance?

[1] E-mails: {jbgao, rubin}@ee.ucla.edu

PII: S0362-546X(01)00713-1

Much work has been done along the first line, and many LRD traffic models have been proposed. More recently, the relevance of multifractal to network traffic modeling has also been studied. The first work was done by Taqqu et al. [10], who concluded that multifractal may not be needed when self-similar traffic models can be applied. Later, Feldmann et al. [1,7] concluded that the short-time variations of network traffic are of multifractal nature, and suggested a hybrid model: short-time multifractal model combined with long-time other LRD model. A more pure multifractal traffic model was suggested by Riedi et al. [9]. However, their model is not very straightforward, since it describes the wavelet coefficients, nor is the model parsimonious, since it requires $\log_2 n$ parameters, where n is the length of the traffic trace data.

In this paper, we describe two types of new multiplicative multifractal traffic models. Both types of models only contain one or two parameters, and are very easy to construct. We present a number of interesting properties of multiplicative processes in their role as models for traffic streams. To evaluate how well those processes can represent the measured network traffic, we consider a single server queueing system which is loaded, on one hand, by the measured processes, and, on the other hand, by properly parameterized multifractal processes. In comparing the system-size tail distributions, we demonstrate our model to effectively track the behavior exhibited by the system driven by the actual traffic processes.

We also study the superposition of LRD flows. We prove that the superposition of a finite number of multiplicative multifractal traffic streams results asymptotically in another multifractal stream. Furthermore we demonstrate numerically that the superimposed process can be effectively modeled by an ideal multiplicative process.

2 LRD traffic and its impact on network performance

Network traffic is often measured by collecting interarrival-time and packet-length statistics. For reference purposes, we refer to such a description as the customary model for network traffic. Aggregated traffic flows measured at a network node are presented as a stochastic counting process. The counting process is a more compact representation of a network traffic process. These two types of descriptions are schematically shown in Fig. 1. To conform the counting process description to the point process formulation, we can record each count $\overline{B}_i$ of the total number of bits inside the i-th slot at the beginning of the slot. We have shown that when the length of the time slot is on the order of the mean packet delay time or smaller, in terms of queueing performance, the two types of models are equivalent.

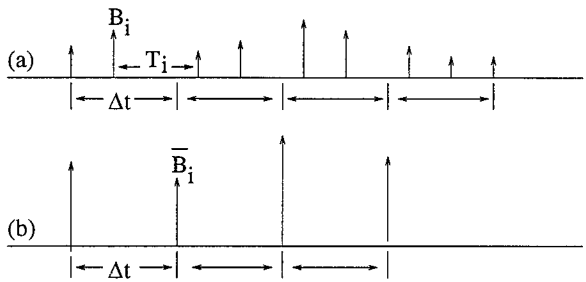

Fig. 1. (a) Point processes description: interarrival time series $\{T_i\}$ and packet length sequence $\{B_i\}$ (in bits). (b) Counting process description: choose a time slot of length Δt, record, at the start of each time slot, the count $\overline{B}_i$ of the total number of bits inside the slot.

One of the key concepts in describing network traffic is its burstiness. There are a number of (more or less) equivalent definitions for the burstiness of network traffic. We define it as follows:

Definition: A traffic process A is said to be more bursty than a traffic process B if a single server queueing system yields a longer system size tail distribution when the process A is used to drive the queueing system.

The burstiness of network traffic is often due to the long-range-dependent (LRD) features of the traffic. Intuitively, a burst group may consist of a random number of subsequent burst periods. The number of periods can be unbounded. This results in variation of traffic over all or many time scales. Furthermore, when traffic congestion occurs, the congestion tends to worsen. Formally, the LRD nature of network traffic is defined as follows.

Let $X = \{X_i : i = 0, 1, 2, \ldots\}$ be a covariance stationary stochastic process with mean μ, variance σ^2, and autocorrelation function $r(k), k \geq 0$. Assume $r(k)$ to be of the form

$$r(k) \sim k^{-\beta}, \quad as \quad k \to \infty \tag{1}$$

where $0 < \beta < 1$. $\sum_k r(k) = \infty$ is referred to as the LRD property.

For each $m = 1, 2, 3, \ldots$, let $X^{(m)} = \{X_i^{(m)} : i = 1, 2, 3, \ldots\}$ denote the new covariance stationary time series obtained by averaging the original series X over non-overlapping blocks of size m, i.e.,

$$X_i^{(m)} = (X_{im-m+1} + \cdots + X_{im})/m, \quad i \geq 1 \tag{2}$$

Self-similarity for X means that the process $X^{(m)}$ exhibits the same second order statistics as those characterizing the process X. In other words, effectively, no smoothing takes place for $X^{(m)}$ with large m values.

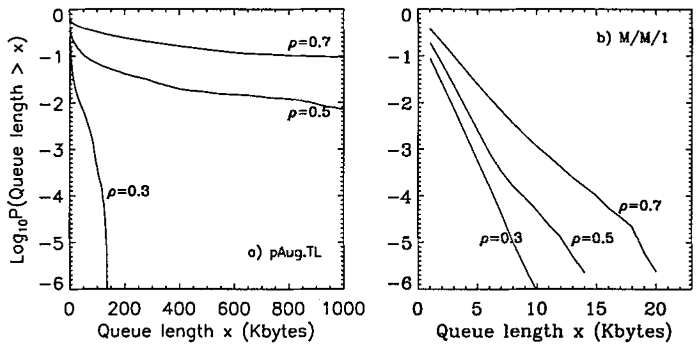

Fig. 2. Complementary queue length distributions for utilizations $\rho = 0.7$, 0.5, and 0.3. (a) FIFO single server queue with infinite buffer with Bellcore's LAN trace data pAug.TL as input traffic, (b) M/M/1 queueing system with mean interarrival time and mean packet length same as those for pAug.TL.

In characterizing the self-similarity of such processes, the most important parameter identified is the Hurst parameter H. It is equivalent to β, $1/2 < H = 1 - \beta/2 < 1$. The value of H measures the degree of persistence of the correlation: the larger the H value, the more persistent the correlation is.

LRD properties of measured network traffic were originally found from the aggregated traffic. We have found that the inter-arrival time series and packet length sequences also possess such properties. This finding opens up a new avenue to model network traffic, i.e., model the inter-arrival time series and packet length sequences instead of the counting process [3].

To better appreciate the impact of LRD traffic on the performance of a network, we consider a FIFO (first in, first out) single server queueing system with a measured traffic trace as the input traffic, and compare its queueing behavior to that of a M/M/1 queueing system (with mean interarrival time and mean packet length the same as those for the measured traffic). Fig. 2 shows an example of such a comparison (by using a LAN traffic pAug.TL from Telcordia). Note the huge difference in the x axis range for these two queueing systems. Thus a Poisson process underestimates buffer size or packet loss probability by several orders of magnitude.

It is argued [8] that a larger H value corresponds to a more bursty traffic. However, we have observed that a burstier traffic even may at times be associated with a smaller value for H. Hence, the H parameter is not a consistent measure of the burstiness of network traffic. This observation has an important implication to the VBR video traffic modeling. That is, video traffic can possess LRD property, on one hand, and be not too bursty (so that effectively modeled by Markovian models), on the other hand. We show below that certain key parameter(s) of multifractal traffic models can serve as both simple and consistent indicators of the burstiness of the traffic.

3 Multiplicative multifractal traffic models

3.1 Construction of multiplicative multifractals

Consider a unit interval. Associate it with a unit mass. Divide the unit interval into two (say, left and right) segments of equal length. Also partition the mass into two fractions, r and $1-r$, and assign them to the left and right segments, respectively. The parameter r is in general a random variable, governed by a probability density function $P(r)$, $0 \leq r \leq 1$. The fraction r is called the multiplier, and $P(r)$ is called the multiplier function. Each new subinterval and its associated weight (or mass) are further divided into two parts following the same rule. Note the scale (i.e., the interval length) associated with stage i is 2^{-i}. We assume that $P(r)$ is symmetric about $r = 1/2$, and has successive moments $\mu_1, \mu_2, \cdots$. The weights at the stage N, $\{w_n, n = 1, ..., 2^N\}$, can be expressed as $w_n = u_1 u_2 \cdots u_N$, where u_l, $l = 1, ..., N$, are independent identically distributed random variables having multiplier function $P(r)$. The multifractality of the multiplicative process refers to the fact that $M_q(\epsilon) = E(\sum_{n=1}^{2^N}(w_n(N))^q) \sim \epsilon^{\tau(q)}$, with $\epsilon = 2^{-N}$, $\tau(q) = -\ln(2\mu_q)/\ln 2$ [5].

We have proven a number of interesting properties about the distribution and correlation structures of the weights in stage N [5]. Specifically, we have proven that when $N \gg 1$, the weights at stage N have log-normal distribution. This feature can be used to check if a measured traffic trace data is consistent with a multiplicative multifractal process model. We have also proven that the Hurst parameter for a multiplicative process is given by

$$1/2 \leq H = -\frac{1}{2}\log_2 \mu_2 \leq 1 \tag{3}$$

Hence, multiplicative processes possess LRD property.

3.2 Two types of multiplicative multifractal traffic models

We have developed two types of multifractal traffic models: multiplicative multifractal point process model and multiplicative multifractal counting process model. In the former, we model the inter-arrival time series and packet length sequences separately using two multiplicative multifractals. This type of model is more faithful to the measured LAN and WAN traffic, because we have found that the inter-arrival time series and packet length sequences are multiplicative multifractals at certain time scale ranges. In the latter, we model the counting process by a multiplicative multifractal. This type of model is quite faithful to the measured VBR video traffic. It is simpler than the former, and is especially useful and convenient for the study of superposition of LRD traffic streams. Each model contains only one or two parameters.

Specifically, we have considered three types of multiplier distribution functions, namely, double exponential with parameter α_e,

$$P(r) \sim e^{-\alpha_e |r-1/2|} \tag{4}$$

Gaussian with parameter α_g,

$$P(r) \sim e^{-\alpha_g (r-1/2)^2} \tag{5}$$

and a function being of the form:

$$P(r) = \begin{cases} q + p\delta(r-1/2) & 1/2 - d \leq r \leq 1/2 + d \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

where $0 \leq d \leq 1/2$. Note that the three parameters d, p, and q are related by the equation, $p + 2qd = 1$. Hence, the function contains two independent parameters. We shall choose p and d as the two basic parameters. Note that parameter p indicates the mean of the counting process, while parameter d describes the variation of the traffic around the mean function.

The parameters α_e, α_g, and p and d are simple and effective indicators of the burstiness of the modeled traffic. When the double exponential or Gaussian multiplier distributions are used, then the burstiness decreases when α_e (or α_g) is increased. When the third distribution (Eq. 6) is used, the burstiness of the modeled traffic increases with d when p is fixed; and decreases with p when d is fixed. The monotonic dependence of the burstiness of the traffic on the parameters allows us to develop systematic procedures to choose the optimal parameter values for modeling measured traffic.

The above burstiness behavior is most easily understood if one notices that when α_e (or α_g) $\rightarrow \infty$, or $(p, d) = (1, 0)$, then $P(r) = \delta(r - 1/2)$. Hence, all the weights are identical. They constitute a non-bursty (or deterministic) traffic.

3.3 *Evaluation of the multifractal traffic models*

To evaluate how well the multifractal traffic models describe a measured traffic process, we consider a single server FIFO queueing system with an infinite buffer. We drive the queueing system, on one hand, by the measured traffic,

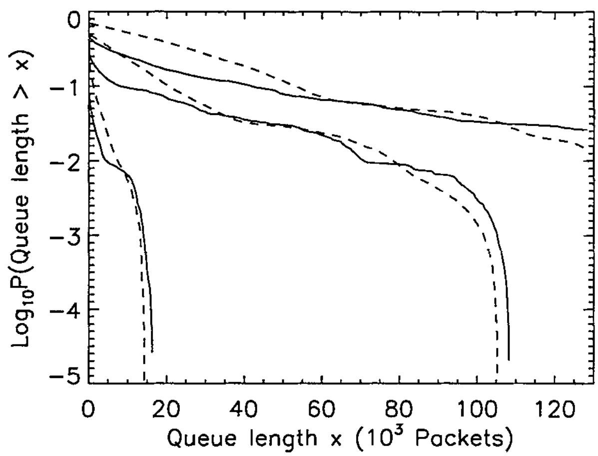

Fig. 3. System-size tail probabilities obtained when WAN traffic Sample-B (solid curves) and its corresponding multifractal counting traffic process (dashed curves) are used to drive the queueing system. Three curves, from top to bottom, correspond to $\rho = 0.7$, 0.5, and 0.3, respectively. The results for the multifractal point process model are similar.

and, on the other hand, by the modeled traffic. We then compare the system size tail distributions of these queueing systems. If those distributions simultaneously match for different loading conditions, we then say that the multiplicative multifractal traffic processes effectively describe the measured traffic.

We have considered a number of measured traffic processes, including LAN, WAN, WWW, and VBR video traffic. Both types of multifractal traffic models have produced similarly excellent fit to the system size tail distributions. An example is given by Fig. 3, showing how well a counting multiplicative multifractal traffic model describes a measured WAN traffic process. We observe that the queueing system driven by the multifractal traffic produces essentially the same tail distributions of that driven by the measured traffic, under different (light, medium, and high) loading conditions. For more examples, we refer to our earlier papers [2,3,6].

4 Superposition of multiplicative multifractal counting traffic streams

Data networks carry traffic from a multiplicity of sources. Superposition of LRD sources will result in LRD aggregated traffic [11]. When lower level LRD aggregated traffic flows are fed into a higher level backbone network, further superposition occurs. It is very desirable that a simple procedure can be applied to model the LRD traffic flows at different locations and levels across a network. In this section, we prove that the superposition of a finite number of multiplicative multifractal counting processes results asymptotically in another multifractal process, and demonstrate numerically that the superimposed process can be effectively modeled by an ideal multiplicative process. This property allows us to model LRD traffic streams at different locations and levels across computer communications networks. In particular, when there are a number of independent users each generating a LRD source traffic modeled by a multiplicative multifractal, in so far as the aggregated traffic is concerned, one needs only to simulate one multiplicative multifractal for the aggregated traffic instead of simulating a bunch of multifractals for all the users.

Consider the superposition of (an arbitrary) k independent multiplicative multifractal traffic streams. Let these multifractal traffic streams be denoted as MF_1, ..., MF_k. Their multiplier distributions are $P^{(1)}(r), ..., P^{(k)}(r)$. These distributions are assumed to be symmetric about 1/2, and have successive moments $\mu_q^{(i)}$, $i = 1, ..., k$, $q = 1, 2, \cdots$. The superimposed traffic stream is denoted by $MF^{(sk)}$, $MF^{(sk)} = \sum_{i=1}^{k} \lambda_i \cdot MF_i$, with $0 < \lambda_1, ..., \lambda_k < 1$, $\sum_{i=1}^{k} \lambda_i = 1$. A weight $w^{(sk)}(N)$ of $MF^{(sk)}$ at the stage N can then be expressed as $w^{(sk)}(N) = \sum_{i=1}^{k} \lambda_i w^{(i)}(N) = \sum_{i=1}^{k} \lambda_i u_1^{(i)} \cdots u_N^{(i)}$, where $w^{(i)}(N)$ is a weight of MF_i at stage N, and $u_j^{(i)}$, $j = 1, ..., N$ are i.i.d random variables governed by pdf $P^{(i)}(r)$, for $i = 1, ..., k$. Then we have the following interesting theorem [4]:

Theorem 1: $MF^{(sk)}$ is (asymptotically in N) a multifractal.

Furthermore, we have shown [4] that the multiplier distribution of $MF^{(sk)}$ are (asymptotically in N) stage-independent. This implies that the multiplier distribution of $MF^{(sk)}$ are stage-independent for sufficiently large stage number N, hence, for most practical situations, $MF^{(sk)}$ is asymptotically a multiplicative multifractal. This is indeed the case, as shown by the following numerical example.

Consider an example involving the superposition of two multiplicative multifractal traffic streams, MF_1 and MF_2, with the largest stage number being 18. Let $\lambda_1 = \lambda_2 = 1/2$, and the multiplier distributions for MF_1 and MF_2 be governed by Eq. 5 with $\alpha_g = 50$ and 100 for MF_1 and MF_2, respectively. Fig. 4 shows that $MF^{(s2)}$ can be amazingly well modeled by an ideal multiplicative multifractal, with its multiplier distribution being also Gaussian with $\alpha_g = 67$.

We have also examined the burstiness of the superimposed traffic streams by measuring the tail distributions they induce when applied to a single server queueing system. By examining a wide range of LRD traffic cases, we have demonstrated that the superimposed process is less bursty than the most bursty traffic component. We also note that when a Poisson process is superimposed with a bursty multiplicative multifractal traffic (as a nonnegligible

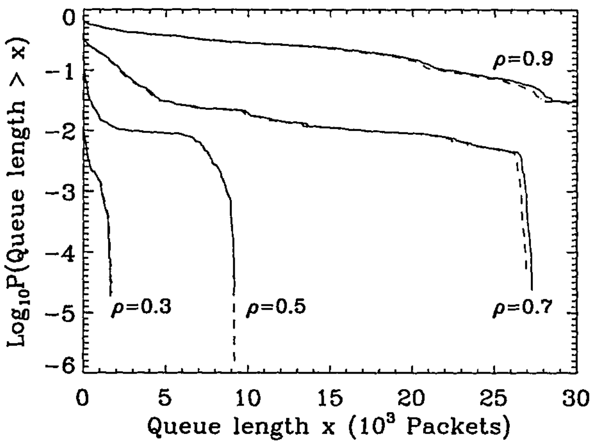

Fig. 4. System-size tail distributions obtained when $MF^{(s2)}$ (dashed lines) and an ideal multiplicative multifractal (solid lines) are used to drive identical single server queueing systems. The utilization levels are indicated in the figure.

component), the Poisson component can be effectively replaced by a deterministic process in deriving a superimposed process that provides the same queueing tail features as those exhibited by the original process.

5 Conclusions

We have introduced two types of multiplicative multifractal traffic models, and shown that both are easy to construct, and can very well describe different types of measured network traffic flows. By considering a single server queueing system that is loaded, on one hand, by the measured LAN, WAN, WWW, and VBR video traffic processes, and, on the other hand, by the corresponding properly parameterized multifractal processes, we demonstrate our model to effectively track the behavior exhibited by the system driven by the actual traffic processes.

To shed light on why multiplicative multifractal processes can faithfully model the measured aggregated traffic across local and metropolitan areas, we have also studied the superposition of a finite number of multiplicative multifractal traffic streams. We have proven that the superimposed process is asymptotically a multifractal process, and shown numerically that the superimposed process can be effectively described by an ideal multiplicative process. Furthermore, when a Poisson process is superimposed with a bursty multiplicative multifractal traffic (as a non-negligible component), the Poisson component can be effectively replaced by a deterministic process in deriving a superimposed process that provides the same queueing tail features as those exhibited by the original process.

Acknowledgment

This work was supported by UC MICRO/SBC Pacific Bell research grant 98-131 and by ARO grant DAAGIJ-98-1-0338.

References

[1] A. Feldmann, A.C. Gilbert, and W. Willinger, 1998: Data networks as cascades: investigating the multifractal nature of Internet WAN traffic. *ACM SIGCOMM'98 Conference*. Vancouver, BC, Canada.

[2] J.B. Gao and I. Rubin, 1999: Multiplicative Multifractal Modeling of Long-Range-Dependent Traffic. *Proceedings ICC'99*, Vancouver, Canada.

[3] J.B. Gao and I. Rubin, 1999: Multifractal modeling of counting processes of Long-Range Dependent network traffic. *Proceedings SCS Advanced Simulation Technologies Conference*,San Diego, CA.

[4] J.B. Gao and I. Rubin, 2000: Superposition of Multiplicative Multifractal Traffic Streams. *Proceedings ICC'2000*, New Orleans, Louisiana.

[5] J.B. Gao and I. Rubin, 2000: Statistical Properties of Multiplicative Multifractal Processes in Modeling Telecommunications Traffic Streams. *Electronics Lett.* **36**, 101-102.

[6] J.B. Gao and I. Rubin, 2000: Multifractal analysis and modeling of VBR video traffic. *Electronics Lett.* **36**, 278-279.

[7] A.C. Gilbert, W. Willinger, and A. Feldmann, 1999: Scaling analysis of conservative cascades, with applications to network traffic. *IEEE Tran. Info. Theory* **45**, 971-991.

[8] W.E. Leland, M.S. Taqqu, W. Willinger, and D.V. Wilson, 1994: On the self-similar nature of Ethernet traffic (extended version). *IEEE/ACM Trans. on Networking*, **2**, 1–15.

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

[10] M.S. Taqqu, V. Teverovsky, and W. Willinger, Is network traffic self-similar or multifractal? *Fractals* **5** 63-73.

[11] W. Willinger, M.S. Taqqu, M.S. Sherman, and D.V. Wilson, 1997: Self-similarity through high-variability: Statistical analysis of ethernet LAN traffic at the source level. *IEEE/ACM Trans. on Networking*, **5** 71–86.