# Wavelet analysis of packet loss in video transfers over UDP

Fei Xue, Velibor Markovski, and Ljiljana Trajkovic*
School of Engineering Science
Simon Fraser University
Vancouver, British Columbia, Canada
*http://www.ensc.sfu.ca/research/cnl*

**Abstract** *Understanding packet loss patterns in IP networks is important for achieving the desired quality of service in multimedia transfers. In this paper, we study the loss patterns in video transfers over UDP in a congested packet network. We use wavelet analysis to investigate the behavior of packet loss on various time-scales. We show that time scales are essential for understanding loss behavior and that packet loss exhibits long-range dependence over the coarser time scales.*

*Keywords:* Packet loss, user datagram protocol, video transfer, long range dependence, wavelets, time scales.

## 1 Introduction

High quality multimedia applications in IP (Internet Protocol) networks have stringent requirements in terms of packet loss, delay, and delay jitter. In IP networks, loss characteristics impact the quality of the established video and/or voice connections, and affect the throughput of bulk data transfers. These loss characteristics cannot be captured by specifying only the loss probability, but require a detailed study of loss patterns [3].

In our study we use *ns* simulator [12] and genuine traffic traces to explore the loss behavior of video connections over User Datagram Protocol (UDP). We use wavelet analysis [4] to study the packet loss behavior on various time-scales. It proved to be a useful tool for analyzing the scale-dependent properties of packet loss via the coefficients of wavelet decomposition.

In Section 2, we first describe the simulation scenario, the topology, and the choice of the traffic sources. Section 3 presents the quantitative simulation results of packet loss bursts. In Section 4, we perform the wavelet analysis of packet loss and discuss our findings. We conclude with Section 5.

## 2 Simulation scenarios

We consider a network consisting of $n$ sources that generate video traffic and feed a common router buffer that is connected to a traffic sink. By selecting the number of video sources $n$ in the input scripts for the *ns* simulator, we generate automatically the network topology. In our simulation scenarios, the link speed between the sources and the router is 10 Mbps, and the link speed between the router and the sink is 44.736 Mbps. The UDP protocol is used to transfer the packets. Packet sizes are fixed to 200 Bytes. The router uses a FIFO buffer. The maximum buffer size $B$ varies from $B = 128$ to $B = 1{,}024$ packets (in steps of 128 packets), which correspond to maximum queueing delay of 4.58 to 36.62 ms, respectively. These values for delays are reasonable for video communications where applications can sustain latencies between 5 and 300 ms. Delays below 10 ms

*This research was supported by the NSERC Grant 216844-99 and the BC Advanced Systems Institute Fellowship.

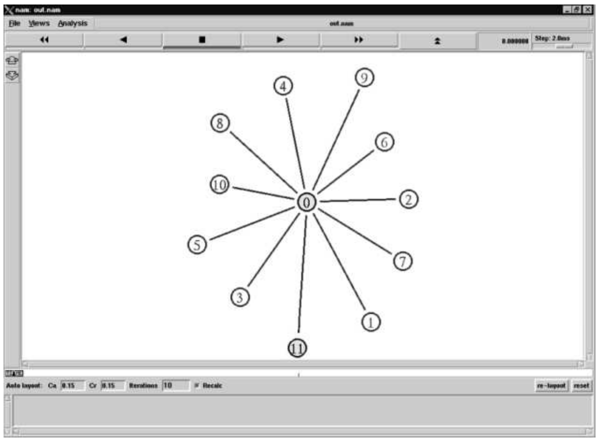

Figure 1: Screen dump from network animator *nam* for topology with $n$ = 10 sources. The topology is generated automatically. Node 0 is the router, nodes 1 to 10 correspond to the traffic sources, and node 11 is the traffic sink. The sources generate UDP traffic according to a supplied trace.

are within the range of latencies specified for high-end interactive video [7]. We use genuine MPEG-1 encoded *Star Wars* video trace [8] to drive our simulations. The trace can be obtained in both ASCII and in *ns* format [13].

Fig. 1 represents a screen dump from network animator *nam* [11], a companion software tool to *ns* [12], used for visualization of the topology layout, packet level animation, and data inspection. In this example, a network with $n$ = 10 sources, one router, and a sink is shown. The number of sources in the topology shown in Fig. 1 is only for illustration purposes. In our simulations, we experiment with larger values of $n$, thus covering utilization levels from 50% to 90%.

# 3 Packet loss and loss bursts

We used the simulation scenario described in Section 2 to conduct a series of experiments in which we varied the number of sources $n$ and the maximum buffer size $B$. We used *ns* simulator to record lost (dropped) packets and the corresponding instances of loss. Our analysis of the nature of the packet loss indicated that packet loss is bursty. A packet loss burst begins

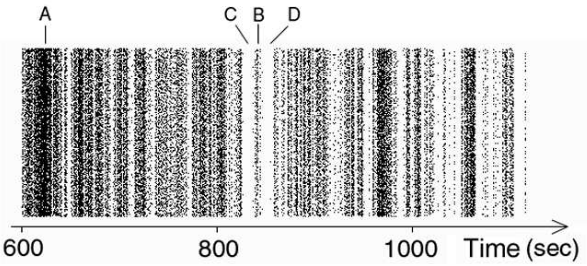


Figure 2: Textured dot strip plot of packet loss instances at the common buffer from a simulation run with $n$ = 80 sources and buffer size $B$ = 512 packets.

with a lost packet if the previous packet was successfully received. For example, if packets with sequence numbers 1, 4, and 6 have been successfully received and if packets 2, 3, and 5 have been lost, then the first loss burst begins with packet number 2 and ends with packet number 3, while the second loss burst begins and ends with packet number 5.

To illustrate our simulation results we use textured dot strip plots [14]. The plot depicting the loss pattern of a simulation run with 80 sources and buffer size of 512 packets is shown in Fig. 2. Each point represents the time instance when the loss occurred. The points are randomly placed along the vertical axis within the height of the plot for better visibility of the bands with more or less losses. We can identify periods of high congestion between 615 and 625 seconds (band A), lower congestion between 836 and 846 seconds (band B), and long periods with no loss, as for example between 826 and 836 seconds (band C), or between 846 and 854 seconds (band D).

Table 1 presents the contribution of lost packets from loss bursts of various lengths in the 10-second time bands A and B. Our results indicate that longer loss bursts have much larger share of the overall loss during the periods of high congestion (band A in our example).

We also found that consecutive packet losses represent the majority of the overall packet loss. For example, Fig. 3 illustrates the contribution of loss bursts of various lengths to the overall number of loss bursts. Fig. 4 de-

Table 1: Contribution of lost packets (in percents) from loss bursts of length $i$ in band A (higher congestion) and band B (lower congestion), shown in Fig. 2. The number of sources $n = 80$ and the buffer size $B = 512$ packets.

| *i (pkt)* | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Band A | 25.98 | 28.75 | 19.77 | 12.59 | 6.56 |
| Band B | 43.84 | 28.99 | 16.30 | 7.25 | 3.62 |

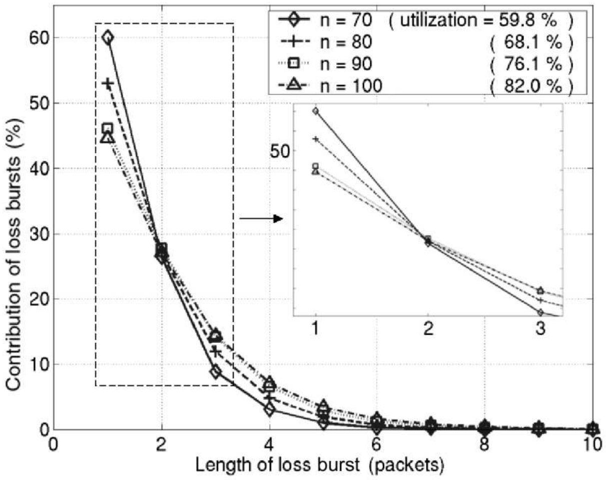


Figure 3: Contribution of loss bursts of various lengths to the overall number of loss bursts. The number of sources is $n$ and the buffer size is $B = 512$ packets.

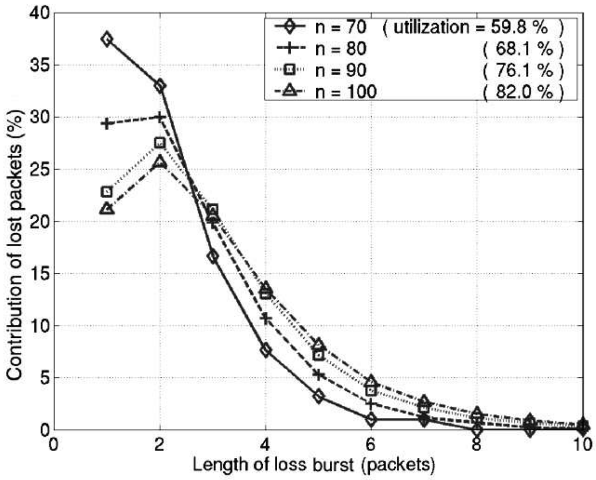


Figure 4: Contribution of lost packets from loss bursts of various lengths to the overall number of lost packets. The number of sources is $n$ and the buffer size is $B = 512$ packets.

picts the contribution of lost packets from loss bursts of various lengths to the overall number of lost packets. For the case with buffer size of 512 packets and 100 sources, we can see from Fig. 4 that single packet losses amount to only 21% of the total number of lost packets, while the majority of packet losses are from loss bursts with more consecutively lost packets.

# 4 Wavelet analysis of loss

Recent studies have shown that video traffic possesses complex correlation structure [2, 8] and exhibits long-range dependence (LRD) [5]. Wavelet analysis proved to be a useful tool to analyze the LRD properties of network traffic [1]. We investigate here the packet loss using wavelet analysis in order to understand the behavior of packet loss on various time-scales and to detect the presence of LRD in loss bursts.

## 4.1 Long-range dependence

Let $X(t), t = 0, 1, ...,$ be a *covariance stationary* (sometimes called *wide-sense stationary*) stochastic process with mean $m$, variance $\sigma^2$, and autocorrelation function $r(k), k \geq 0$. $X$ is said to exhibit *long-range dependence* if

$$r(k) \sim c_r k^{-(2-2H)}, \qquad k \to \infty, \qquad (1)$$

where $c_r > 0$ and $0.5 < H < 1$. The key element in this description is the Hurst parameter $H$ that measures the degree of long-range dependence. For short-range dependent processes $H = 0.5$. From Eq. (1) we see that long-range dependence is characterized by an autocorrelation function that decays hyperbolically as the lag increases. In the frequency domain, an equivalent property of an LRD process is that the spectral density $f(\lambda)$ of $X(t)$ takes the following form for a range of frequencies $\lambda$ close to 0:

$$f(\lambda) \sim c_f |\lambda|^{1-2H}, \qquad \lambda \to 0, \qquad (2)$$

where $c_f > 0$ and $0.5 < H < 1$.

Recently, long-range dependence has been found in network traffic measured from a wide range of high-speed networks such as Ethernet local area networks, wide area networks, and WWW [16]. It has also been found that LRD is an inherent feature of VBR video traffic, independent of scene and codec algorithms. For example, the genuine *Star Wars* video trace, which is used to drive our simulation in this paper, exhibits LRD with Hurst parameter $H \approx 0.8$ [2].

## 4.2 Discrete wavelet transform

The wavelet transform is the representation of a signal *X(t)* by means of its inner products with a set of basic wavelet functions $\psi_{a,t}(u) = \psi_0(\frac{u-t}{a})/\sqrt{a}$, called the wavelets, that are scaled and shifted versions of an adequately chosen mother wavelet $\psi_0$. *X(t)* is then transformed into a scale-time wavelet domain $(a,t)$, $a \in R^+$ and $t \in R$. The wavelet transform presents a method for simultaneously observing a time series at a full range of different scales $a$, while retaining the time dimension of the original data. Multi-resolution analysis (MRA) theory shows that no information is lost if we sample the continuous wavelet coefficients at a sparse set of points in the scale-time plane known as the dyadic grid. This grid, defined by $(a,t) = (2^j, 2^j k)$, $j,k \in N$, leads to the discrete wavelet transform. Detailed background on wavelets and MRA is given in [4].

An efficient implementation of the discrete wavelet transform is given by the filter-bank-based pyramidal algorithm. In this algorithm, the mother wavelet $\psi_0$ is represented in terms of the so-called scaling function $\phi_0$:

$$\psi_0 = \sqrt{2} \sum_n (-1)^n h(1-n) \phi_0(2t-n). \qquad (3)$$

$\phi_0$ is defined by a two-scale difference equation, known as the dilation equation:

$$\phi_0 = \sqrt{2} \sum_n h(n) \phi_0(2t-n). \qquad (4)$$

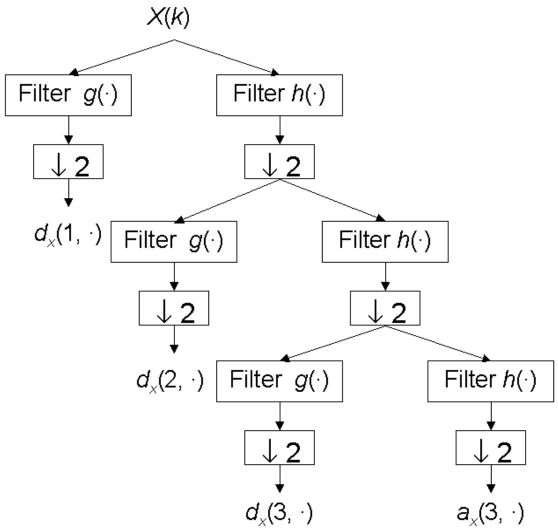


Figure 5: The filter-bank-based pyramidal algorithm. At each level $j$ in the recursive structure, $d_x(j,k)$ and $a_x(j,x)$ are the outputs of the filters *g(n)* and *f(n)*, respectively. They occur at half the sampling rate of the input $a_x(j-1,x)$.

The discrete-time filters $h(n)$ and $g(n) = (-1)^n h(1-n)$ have the low-pass and high-pass characteristics, respectively. Fig. 5 depicts the filter-downsample procedure of this transform. Given a time series *X(k)* of length $N$ and the initial approximation coefficients $a_x(0,k) = X(k)$, we can compute the so-called detail coefficients $d_x(j,k)$ at scale $2^j$ $(1 \leq j \leq log_2 N)$ by using the convolving approximation coefficients with filters $h(n)$ and $g(n)$:

$$\begin{aligned} d_x(j,k) &= g * a_x(j-1,k) \\ a_x(j,k) &= h * a_x(j-1,k). \end{aligned} \qquad (5)$$

## 4.3 Wavelets and the LRD detection

A wavelet-based Hurst parameter estimator, suggested in [1], is based on a spectrum estimator designed by performing a time average of $|d_x(j,k)|^2$ at a given scale:

$$\Gamma_x(2^{-j}\nu_0) = \frac{1}{n_j} \sum_k |d_x(j,k)|^2, \qquad (6)$$

where $n_j$ is the number of wavelet coefficients at octave $j$. The linear relationship between $log_2(\Gamma)$ and level $j$ for a range of scales [$j_1$, $j_2$] indicates the presence of an LRD behavior. We can find an estimator $\hat{H}$ for the Hurst parameter $H$ by performing a linear regression of $log_2(\Gamma)$ on scale level $j$ in the range [$j_1$, $j_2$]:

$$log_2(\Gamma_x(2^{-j}\nu_0)) = (2\hat{H} - 1)j + \hat{c}. \qquad (7)$$

We have evaluated the performance of this wavelet-based estimator for variety of genuine and synthetically generated traffic traces. Rather than focusing on the quantitative aspects of the LRD estimation emphasized in [1], we suggest here a more qualitative use of the wavelet-based analysis. We use the $log_2(\Gamma)$ vs. $j$ plots to examine the scaling behavior of data. From the plots we observed the linear relationship between $log_2(\Gamma)$ and $j$ over several scales that enabled us to identify the scale range over which the LRD is present. The plots also enabled us to determine the break-point scale beyond which we can accurately detect the LRD. Hence, the technique we propose may be used to detect both the presence and the location of LRD.

### 4.4 Main result: packet loss scaling behavior

In order to study the loss behavior on various time scales, we consider samples of the loss rate processes obtained from *ns* simulations. Each sample represents the number of loss packets during 1 ms. We present the results of our wavelet analysis for various buffer sizes and source numbers in Figs. 6—8. We observe from Figs. 6—8 that the plots are similar and, hence, that the overall scaling property of packet loss is robust. Furthermore, the packet loss behavior varies over the time scales and the loss property changes over different time scales.

The importance of time scales on the network traffic analysis and connection admission control has already been reported [6, 15]. For example, it has been shown in [6] that the measured Internet traffic is LRD (or asymptotically

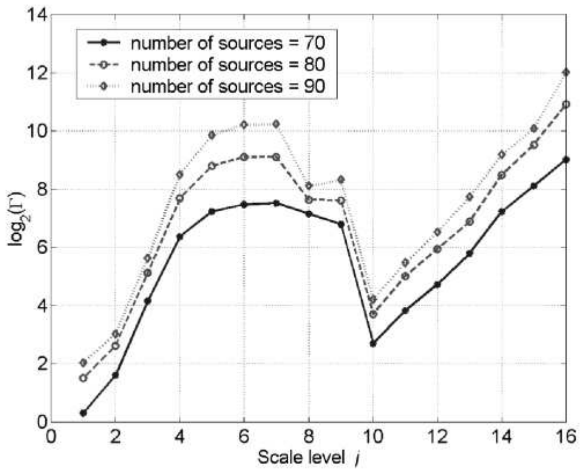


Figure 6: The $log_2(\Gamma)$ vs. $j$ plot of packet loss process for buffer size $B = 128$ packets.

self-similar) over coarser time scales, but that it exhibits multifractal behavior over finer time scales. Our results presented here show that time scales also play an important role for a comprehensive understanding of the behavior of packet loss.

The linear relationship between $log_2(\Gamma)$ and $j$ is evident for the coarser scales beyond the break-point 10. As we discussed earlier, this effect indicates that the time scale granularity for possible detection of LRD is close to 1,024 ($2^{10}$) ms.

Once the time scale break-point is observed, we can obtain new packet loss rate processes with a numerically more suitable time scale. For example, we created a loss rate process in which each sample represents the number of loss packets over 1,000 ms. We then apply well known Hurst parameter estimators [9], such as variance-time and R/S analysis [10], to test the presence of LRD and to estimate Hurst parameters. Estimated Hurst parameters for various scenarios are given in Tables 2—4.

As shown in Tables 2—4, all estimated Hurst parameters are in the range [0.5, 1]. This indicates packet loss process is LRD over time scales coarser than the break-point scale 10, which corresponds to time scales of the magnitude of approximately 1 second (1,024 ms) or

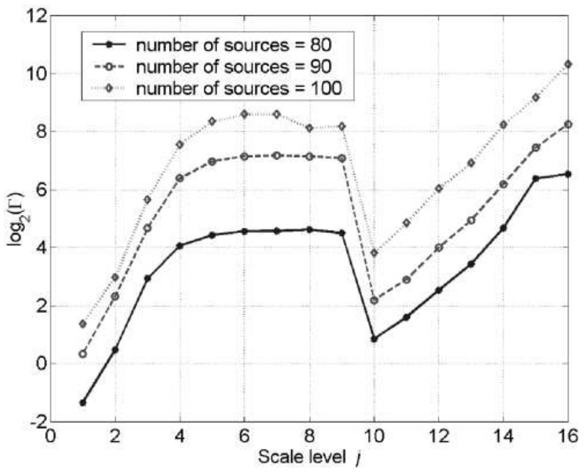


Figure 7: The $log_2(\Gamma)$ vs. $j$ plot of packet loss process for buffer size $B = 512$ packets.

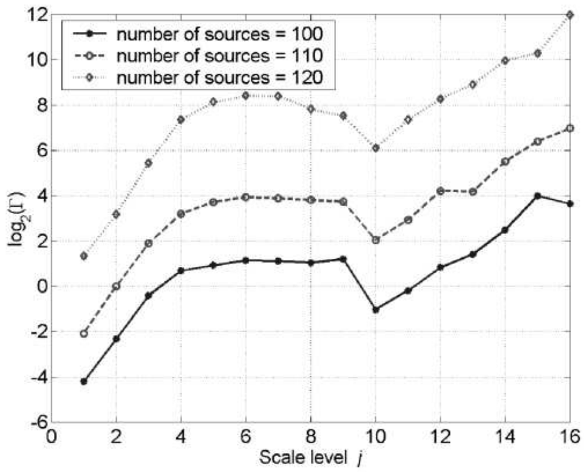


Figure 8: The $log_2(\Gamma)$ vs. $j$ plot of packet loss process for buffer size $B = 1{,}024$ packets.

Table 2: Estimated Hurst parameters of packet loss process with 1,000 ms granularity, $n$ sources, and buffer size $B = 128$.

| $n$ | Variance-time | R/S |
|---|---|---|
| 70 | 0.8273 | 0.8326 |
| 80 | 0.6667 | 0.8663 |
| 90 | 0.6647 | 0.7971 |

Table 3: Estimated Hurst parameters of packet loss process with 1,000 ms granularity, $n$ sources, and buffer size $B = 512$.

| $n$ | Variance-time | R/S |
|---|---|---|
| 80 | 0.7829 | 0.8662 |
| 90 | 0.6758 | 0.9007 |
| 100 | 0.7898 | 0.8563 |

larger. Hence, our findings confirm the importance of time scale on the quantitative estimation of loss burst.

# 5 Concluding remarks

In this paper, we studied packet loss in a video transfer over UDP with *ns* simulations driven by genuine video traffic. Our simulation results indicated that packet loss is bursty and that consecutive packet losses significantly contribute to the overall packet loss.

We performed wavelet analysis on packet loss over various time scales. Wavelet-based analysis proved helpful not only for estimating Hurst parameter, as observed in the past, but also for finding the break-point level over which LRD detection can be performed.

We also observed that time scales are important for estimating quantitative behavior of packet loss. For time scales coarser than the break-point level, we applied variance-time and R/S analysis to obtain Hurst parameters of packet loss processes. Our results indicated the existence of LRD in packet loss processes on coarser time scales. Our findings are robust and hold for various simulation scenarios. We are currently investigating the packet loss behavior on finer time scales.

We conjecture that one possible cause for the LRD phenomena in packet loss may be that the source traffic itself (*Star Wars* trace) has been proven to be LRD [8]. It is an open question how to establish the relationship between the LRD behavior of the source traffic and the pa-

Table 4: Estimated Hurst parameters of packet loss process with 1,000 ms granularity, $n$ sources, and buffer size $B = 1{,}024$.

| $n$ | Variance-time | R/S |
|---|---|---|
| 100 | 0.6716 | 0.9681 |
| 110 | 0.6927 | 0.8913 |
| 120 | 0.6030 | 0.8671 |

rameters used to measure network performance (packet loss, delay, and delay jitter).

## References


[1] P. Abry and D. Veitch, "Wavelet analysis of long-range-dependent traffic," *IEEE Trans. Information Theory*, vol. 44, no. 1, pp. 2–15, 1998.

[2] J. Beran, R. Sherman, M. S. Taqqu, and W. Willinger, "Long-range dependence in variable-bit-rate video traffic," *IEEE Trans. Commun.*, vol. 43, no. 2/3/4, pp. 1566–1579, 1995.

[3] M. S. Borella, D. Swider, S. Uludag, and G. B. Brewster, "Internet packet loss: measurement and implications for end-to-end QoS," *Proc. Int. Conf. Parallel Processing*, Minneapolis, MN, pp. 3–15, Aug. 1998.

[4] C. K. Chui, *An Introduction to Wavelets.* Boston: Academic Press, 1992.

[5] D. R. Cox, "Long-range dependence: A review," in *Statistics: An Appraisal*, H. A. David and H. T. David, eds. Ames, IA: The Iowa State Univ. Press, 1984, pp. 55–74.

[6] A. Feldmann, A. C. Gilbert, and W. Willinger, "Data networks as cascades: Investigating the multifractal nature of Internet WAN traffic," *Proc. ACM SIGCOMM'98*, pp. 25–38, 1998.

[7] M. Fester, "Performance issues for high-end video over ATM," Aug. 1995: http://www.cisco.com/warp/public/cc/sol/mkt/ent/atm/vidat_wp.htm .

[8] M. Garrett and W. Willinger, "Analysis, modeling and generation of self-similar VBR video traffic," *Proc. ACM SIGCOMM'94*, London, U.K., Aug. 1994, pp. 269–280.

[9] W. Leland, M. Taqqu, W. Willinger, and D. Wilson, "On the self-similar nature of Ethernet traffic (extended version)," *IEEE/ACM Trans. Networking*, vol. 2, pp. 1–15, 1994.

[10] B. B. Mandelbrot and M. S. Taqqu, "Robust R/S Analysis of Long Run Serial Correlation," *Proc. 42nd Session Int. Statistical Inst.*, Manila, December 1979, pp. 1–37.

[11] Nam network animator, 1999, http://www-mash.cs.berkeley.edu/nam .

[12] ns-2 network simulator, 1999, http://www-mash.cs.berkeley.edu/ns .

[13] *Star Wars* traces: in ASCII format *ftp://ftp.telcordia.com/pub/vbr.video.trace*, and in *ns* format: *http://www.research.att.com/~breslau/vint/trace.html* .

[14] D. S. Swayne, D. Cook, and A. Buja, "XGobi: Interactive dynamic data visualization in the X window system," *J. Computational and Graphical Statistics*, vol. 7, no. 1, pp. 113–130, Mar. 1998.

[15] Lj. Trajkovic and A. Neidhardt, "Effect of traffic knowledge on the efficiency of admission-control policies," *ACM Comput. Commun. Rev.*, vol. 29, no. 1, pp. 5–34, Jan. 1999.

[16] W. Willinger, M. S. Taqqu, R. Sherman and D. V. Wilson, "Self-similarity through high-variability: statistical analysis of Ethernet LAN traffic at the source level," *ACM Comput. Commun. Rev.*, vol. 25, no. 4, pp. 100–113, Oct. 1995.