Cf. Proc. 5th International Conference on Computer Communications and Networks (IC3N), pp. 308-312, October 1996.

A Fast Method for Combining/Generating Synthetic Traffic Traces Exhibiting Short- and Long-Range Dependence

Gam D. Nguyen

Naval Research Laboratory, Washington DC 20375

Abstract

We consider a heuristic approach for combining short traffic traces of various degrees of self-similarity to produce a longer trace that has the following desirable properties. The mixed trace is capable of modeling computer networking traffic that exhibits both short- and long-range dependence in a unified model. Synthetic traces of self-similar characteristics can be generated quickly due to the simplicity of our method, which is faster than the traditional approach of generating/approximating traffic traces from nothing.

1 Introduction

Synthesizing traces that exhibit self-similarity is important because modern computer networking traffic is reported to possess self-similar (SS) characteristics (see [6] and its references). Several algorithms, which have complexity ranging from $O(n)$ to $O(n^2)$ for approximating/generating SS traces of length *n*, are reviewed in [6]. Traditional approaches are concentrated on generating SS traces from scratch; i.e., generating one data point after another until a specified length is reached. In contrast, our method produces a longer trace by combining previously generated shorter traces. In comparing with many existing methods for generating traces from scratch, our combining method is orders of magnitudes faster; hence it can be used to generate traces in real time. For example, we can cross-multiply 1000 points of a given SS trace (a short trace) to produce a new 1,000,000-point trace (a much longer trace) that exhibits some degree of self-similarity. Furthermore, the new method is capable of generating traces possessing both short- and long-range dependence at various time scales.

2 Long-Range Dependence and Self-Similarity

Let $\{X_n\}$ be a wide-sense stationary (WSS) process with autocorrelation $r(n)$ and power spectral density $g(f)$. Assume that $\sum_{n=1}^{\infty} r(n)$ diverges, then $\{X_n\}$ is long-range dependent (LRD) [otherwise, if $\sum_{n=1}^{\infty} r(n)$ converges, we define $\{X_n\}$ to be short-range dependent (SRD)]. Let $\{X_k^{(m)}\}$ be a WSS process formed by averaging the process $\{X_n\}$ in non-overlapping blocks of *m*, i.e.,

$$X_k^{(m)} = (X_{km-m+1} + \quad + X_{km})/m .$$

Then the LRD process $\{X_n\}$ is (*asymptotically*) *SS* with parameter $0 < \beta < 1$ if the following 3 (equivalent) properties hold [1]:

$$\text{var } X_k^{(m)} \quad m^{-\beta} \text{ as } m \to \infty , \tag{1}$$

$$r(i) \quad i^{-\beta} \text{ as } i \to \infty , \tag{2}$$

$$g(f) \quad f^{-(1-\beta)} \text{ as } f \to 0 . \tag{3}$$

One often uses the *rescaled range statistic* R/S to estimate β as follows. Let $\bar{X}$ and $S^2(n)$ be the sample mean and variance of $\{X_n\}$, respectively, and let $W_k = X_1 + \quad + X_k - k\bar{X}$, $k = 1, \quad , n$. The R/S statistic is defined by $R(n)/S(n)$, where $R(n) = \max(0, W_1, \quad , W_n) - \min(0, W_1, \quad , W_n)$. Then

$$E[R(n)/S(n)] \quad n^H \text{ as } n \to \infty , \tag{4}$$

where $H = 1 - \beta/2$ is the Hurst parameter. *H* is a measure for burstiness; i.e., the higher *H* (or lower β) the burstier the process [2]. The underlying WSS process is SRD when $H = 0.5$ and LRD otherwise. The following result shows that self-similarity is both persistent and dominating in heterogenous networking environment.

Theorem 1: The asymptotic behavior of a process formed by multiplexing two SS processes will be that of the burstier one.

Proof: Let $\{Z_n\} = \{X_n + Y_n\}$ be the sum process, where $\{X_n\}$ and $\{Y_n\}$ are two component SS processes of parameters β_1 and β_2, respectively. We also assume that $\{X_n\}$ and $\{Y_n\}$ are jointly WSS; i.e., $E\{X_{n+k}Y_n\}$ and $E\{X_nY_{n+k}\}$ depend only on *k*. The WSS aggregated process $\{Z_i^{(m)}\}$ then has the variance

$$\text{var}Z_i^{(m)} = \text{var}X_i^{(m)} + \text{var}Y_i^{(m)} + 2a\,(\text{var}X_i^{(m)}\,\text{var}Y_i^{(m)})^{1/2} , \tag{5}$$

where *a* is the correlation coefficient of $X_i^{(m)}$ and $Y_i^{(m)}$, which is bounded (i.e., $|a| \le 1$). Since the component processes are asymptotically SS, when $m \to \infty$, (5) becomes

$$\text{var}Z_i^{(m)} \to a_1 m^{-\beta_1} + a_2 m^{-\beta_2} + a_3 a m^{-(\beta_1+\beta_2)/2} ,$$

where a_i are positive constants [see (1)]. When $m \to \infty$, $\text{var}Z_i^{(m)} \to a_4 m^{-\min(\beta_1, \beta_2)}$, for some positive a_4; therefore, $\{Z_n\}$ is asymptotically SS by (1). Q.E.D.

3 The Combining Algorithm

Our goal is to generate a synthetic trace of length *N*, where *N* = *TW* for some positive integers *T* and *W*. That is the entire trace is composed of *T* time windows; each window contains *W* samples. Furthermore, the required trace must behave like a SS process of Hurst parameter *H*; and it must look like another SS process in each window with Hurst parameter H_W. To this end, let $\{X_1, \quad , X_W\}$ be a SS trace of Hurst parameter H_W and $\{Y_1, \quad , Y_T\}$ be (another) SS trace of Hurst parameter *H*. Then the required trace

$\{Z_1, Z_2, \quad , Z_N\}$ is simply the cross-multiplication of all elements of $\{X_i\}$ and $\{Y_j\}$:

$$\begin{aligned} &Y_1X_1, Y_1X_2, \quad , Y_1X_W, \\ &Y_2X_1, Y_2X_2, \quad , Y_2X_W, \\ &\qquad\qquad\qquad\qquad\qquad\qquad (6) \\ &Y_TX_1, Y_TX_2, \quad , Y_TX_W. \end{aligned}$$

Therefore, we need only $T + W$ SS data points (which can be computed on-line or off-line) and TW multiplications to produce the longer trace of TW data points. One can think of the integrated trace $\{Z_i\}$ as a train of time windows; all windows share the same level of burstiness (i.e., the same Hurst parameter H_W). However, the load intensity varies from one window to the next according to a SS process of Hurst parameter H.

To see how the trace $\{Z_i\}$ specified by (6) meets most of our objectives, notice that in the *i*th window (of size *W*), the samples of $\{Z_i\}$ are $Y_iX_1, Y_iX_2, \quad , Y_iX_W$; therefore, $\{Z_i\}$ is a scaled version of $X_1, X_2, \quad , X_W$, which has Hurst parameter H_W.Thus, $\{Z_i\}$ also has Hurst parameter H_W in each time window as desired. On the other hand, it is shown in [4] that the averages of $\{Z_i\}$ in non-overlapping blocks of size jW are $Z_i^{(jW)} = Y_i^{(j)} X_1^{(W)}$, where $X_1^{(W)} = (X_1 + X_2 + \quad + X_W)/W$. Thus in large time scales, $\{Z_i\}$ retains some SS characteristics of $\{Y_i\}$. In other words, $\{Z_i\}$ is approximately asymptotically SS with Hurst parameter H. We consider only 2 Hurst parameters (H and H_W) for the sake of clear illustration in this paper; more parameters can be incorporated into the model at the cost of increasing complexity.

One often uses fractional Gaussian noises (FGNs) to generate the two component traces $\{X_i\}$ and $\{Y_i\}$, which take negative and non-integer values. In the next section we will transform the FGNs to have non-negative integer values as required in computer networking applications.

4 An Application

In this section we apply the algorithm (6) to generate a non-negative, integer-valued synthetic trace that possesses both SRD and LRD characteristics in different time scales. Additionally, the trace must satisfy a specified load constraint and a level of traffic denseness.

4.1 Model Parameters

Recall from Section 3 that we need the following 4 basic parameters to generate a longer trace from two shorter traces:

H — (global) asymptotic Hurst parameter. The generated trace looks asymptotically SS with Hurst parameter *H*.

H_W — (local) Hurst parameter as seen in each time window. The traffic in each time window behaves as a SS process of Hurst parameter H_W, generally $H_W \neq H$.

T — the number of time windows. The generated series consists of *T* time windows. Each window has the same degree of burstiness (i.e., same Hurst parameter H_W); and the traffic load will change from one window to the next according to a SS process of Hurst parameter *H*.

W — window size. Each time window has *W* data points.

To refine the model a bit further, we introduce 2 more parameters as follows:

L — window maximum traffic load. Cumulative traffic in each time window can not exceed *L*. However, instantaneous traffic values in finer time stamps can exceed the corresponding (translated) *L*. For example, the maximum ethernet load is 10 Mbps; however, in some rare 0.1-second interval, the load value can exceed 1 Mb (which is translated to the corresponding load that exceeds 10 Mbps).

r — an indicator of traffic denseness, $r \quad (0, 1)$. The generated trace has lower peaks and shallower valleys for larger *r* (i.e., the traffic looks denser at larger *r*).

Our simple model is meant to capture only important traffic characteristics such as the degree of SRD/LRD, the level of traffic denseness, and the extent of traffic burstiness. Traffic details at microscopic levels are not considered here because they are application-specific; however, one can tailor the model to fit the required applications. In other words, we model only intrinsic properties of the traffic; it is up to the users of the model to do the fine-tuning. The model accepts 6 input parameters and produces the specified numbers of output data points, which represent the number of network traffic units in a time unit (e.g., byte count per 0.01 second on a backbone network).

4.2 Procedures

One can use the following steps to generate the traffic trace specified in Section 4.1.

(1) Generate a FGN series $\{E_1, E_2, \quad , E_T\}$ of Hurst parameter *H* by some known technique such as Hosking Fractional Differencing. Notice that samples of FGN often assume negative and non-integer values.

(2) Translate $E_1, E_2, \quad , E_T$ into a positive series $\{F_i\}$; e.g. (cf. [6]),

$$F_i = 2^{\frac{(1-r)}{(r+e)} E_i}, \qquad (7)$$

where *r* is an indicator of traffic denseness, $r \quad (0, 1)$, and *e* is a positive number to make the denominator in the above exponent positive.

(3) Set $Y_i = (LF_i)/max\{F_1, \quad , F_T\}$ to ensure that $Y_i \leq L$. The traffic will be governed by the (load) parameter Y_i in each window, and the traffic load will change from one window to the next.

(4) Repeat steps (1) and (2), using the Hurst parameter H_W, to generate another positive series $G_1, G_2, \quad , G_W$. Let $X_i = G_i/(G_1 + G_2 + \quad + G_W)$. Using $\{X_i\}$ and $\{Y_i\}$ to generate the trace $\{Z_i\}$ as specified by (6), the total load in each time window of $\{Z_i\}$ satisfies $WZ_i^{(W)} = WY_iX_1^{(W)} = Y_i \leq L$ as desired. Then the integer parts of $\{Z_i\}$ constitute the required synthetic trace.

We assign the following values to the 6 input parameters to test the model: $H = 0.99$, $H_W = 0.50$, $r = 0.45$, $L = 10^6$, $T = W = 1000$. Therefore, the output trace has $N = TW = 10^6$ data points. We arbitrarily define each data point to be the number of traffic bytes per 0.01 second. Figure 1 shows the generated trace in four different time scales.

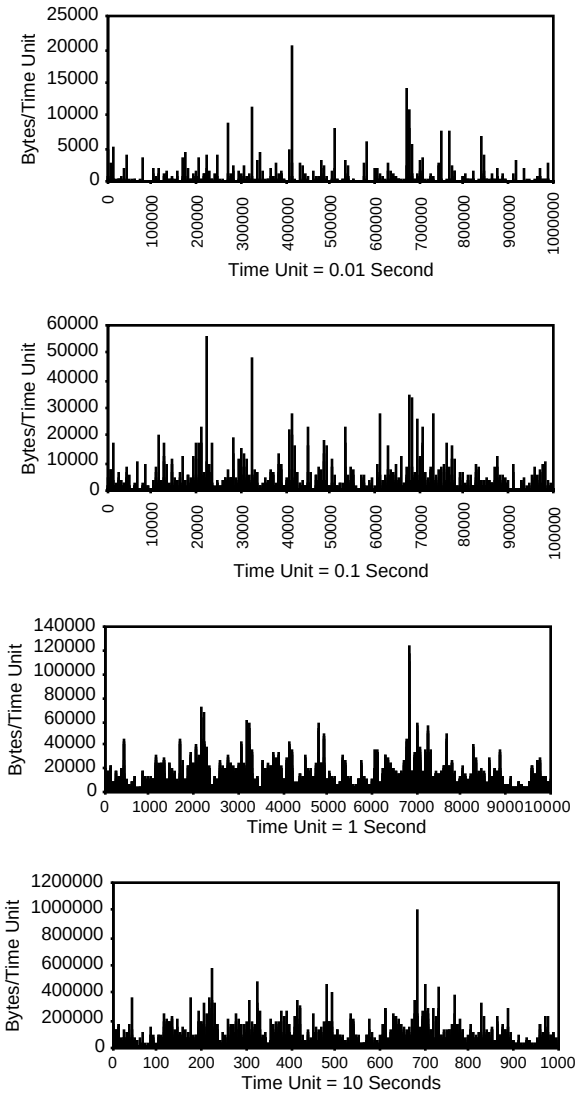


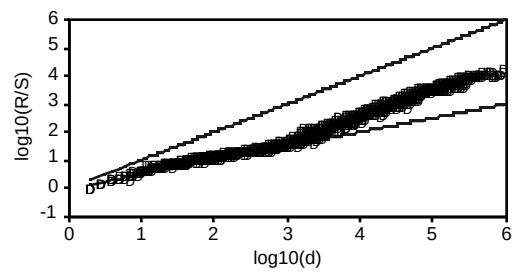
Fig. 1 Synthetic Trace at 4 Time Scales: 0.01 - 10 Seconds ($H = 0.99$, $H_W = 0.50$, $r = 0.45$, $L = 10^6$, $T = W = 10^3$). The traffic shows burstiness in all 4 time scales.

4.3 Output Analysis

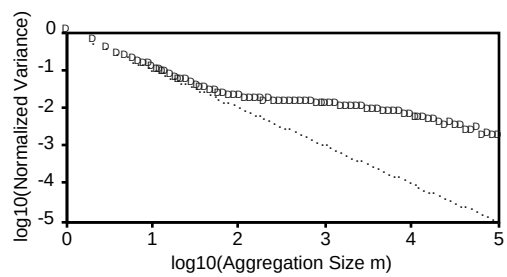
Several graphical techniques exist for testing if the generated trace agrees, to some extent, with SS properties (1) — (4). In this subsection we use two of the most popular techniques, namely the variance-time and *R/S* plots, which are graphical demonstrations of (1) and (4), respectively.

In each *R/S* plot, the straight line of slope 0.5 corresponds to an uncorrelated (i.e., low bursty) process of Hurst parameter $H = 0.5$; the line of slope 1 corresponds to a process of extremely high degree of burstiness ($H = 1$). In each variance-time plot, the line of slope –1 corresponds to an uncorrelated process ($\beta = 1$), whereas the line of slope 0 corresponds to an extremely bursty process ($\beta = 0$). In computing the *R/S* statistics, the *n* samples are segmented into time bins of equal size *d*. Therefore, more *R/S* statistics are produced for smaller values of *d*. Figure 2 shows that the generated trace indeed is SRD (i.e., $H_W \approx 0.50$) in small time scales and is LRD (i.e., $H > 0.9$) in larger time scales.

To see the effect of the denseness-indicator parameter r, we increase the value of r from 0.45 to 0.85 while keeping all other parameters the same; and then we plot the newly generated trace in Figs. 3 and 4. Comparing Figs. 1 and 2 with Figs. 3 and 4 confirms that higher r means denser traffic profile. Comparing Fig. 2 with Fig. 4 suggests that the estimated Hurst parameter seems to be higher at higher *r* (which is also a measure of traffic load); this is related to the observation in [2] that the degree of self-similarity tends to increase with the load level (on ethernets). Notice that the variance-time and the *R/S* plots of our synthetic traffic resemble those of the backbone traffic of a distributed interactive simulation reported in [3]; our variance-time plots are also comparable to those of some TCP traffic traces studied in [5].



(a) *R/S* Plot



(b) Variance-Time Plot

Fig. 2 Graphical Statistics for the Generated Trace ($H = 0.99$, $H_W = 0.50$, $r = 0.45$, $L = 10^6$, $T = W = 1000$).

Remark 1: By letting $X_i = Y_i$ when $W = T$ and $H = H_W$ in algorithm (6), we can generate a trace of length T^2, which is approximately SS with Hurst parameter *H*. Therefore, generating a trace of length *n* requires only *n* cross-multiplications and a trace of length $\sqrt{n}$ (since $n = \sqrt{n}\sqrt{n}$). As an example, let $H = H_W = 0.99$, $T = W = 1000$; and then use the procedures in Section 4.2 with $X_i = Y_i$. Thus the mixed trace has $N = TW = 10^6$ data points and is supposed to

approximate a SS trace of Hurst parameter 0.99 (see Fig. 5). This combining method is much faster than the traditional start-from-scratch approach. The resultant *R/S*, variance-time, power spectral density (periodogram estimate), and autocorrelation plots are shown in Fig. 6, which show that our mixing algorithm produces the longer trace that inherits SS characteristics from the shorter trace. Then we estimate the Hurst parameter by linear fitting on all time scales: the estimate from the *R/S* plot is ~ 0.933; whereas that from the variance-time plot is ~ 0.905. Thus the averaged estimate for the Hurst parameter is H ~ 0.92, which is 93% of the target value (0.99).

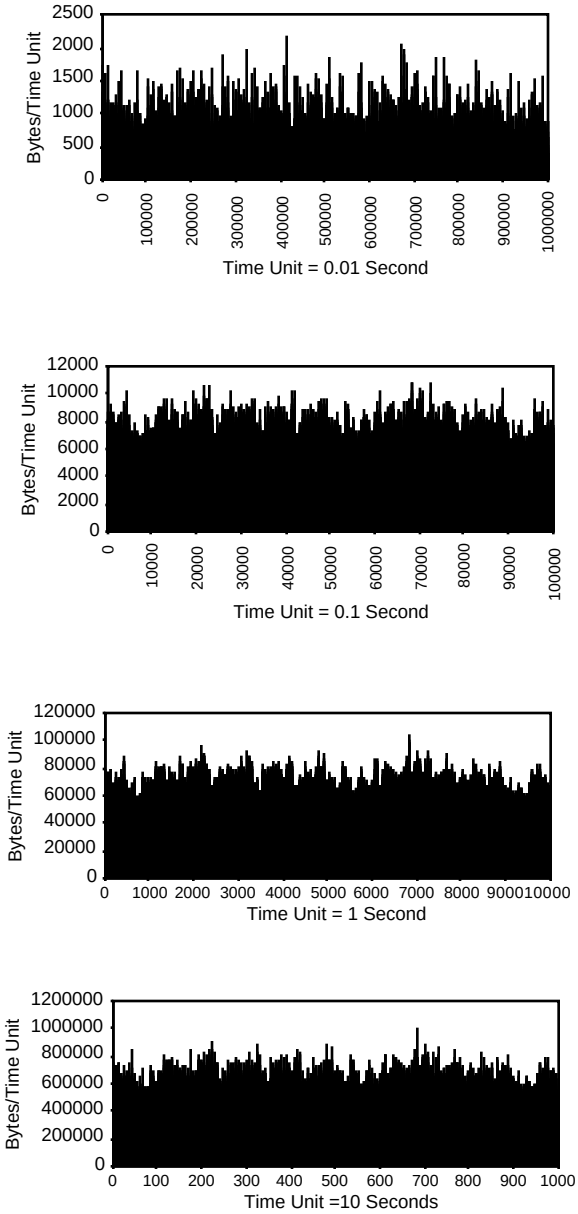


Fig. 3 Synthetic Trace at 4 Time Scales: 0.01 - 10 Seconds (H = 0.99, H_W = 0.50, r = 0.85, $L = 10^6$, $T = W$ = 1000). The traffic shows burstiness in all 4 time scales.

5 Discussions

We form a single longer trace by multiplying two shorter SS traces; the combined trace is shown by examples to inherit some SS characteristics from the shorter traces. So far the two initial traces are free of any restrictions besides the SS assumption. Suppose that the two initial SS traces have the same Hurst parameter $H = 1 - \beta/2$; naturally one would ask the following questions (thanks to a reviewer for good questions and comments). What are sufficient conditions that the two shorter traces must satisfy to yield a combined trace that has high quality? Can the algorithm be used recursively; i.e., by starting with two very short traces and then by repeating the procedure to generate a very long trace? These questions are dealt with in [4], where we show under mild conditions that var $Z_i^{(jW)} \quad (jW)^{-\beta}$, which resembles (1).

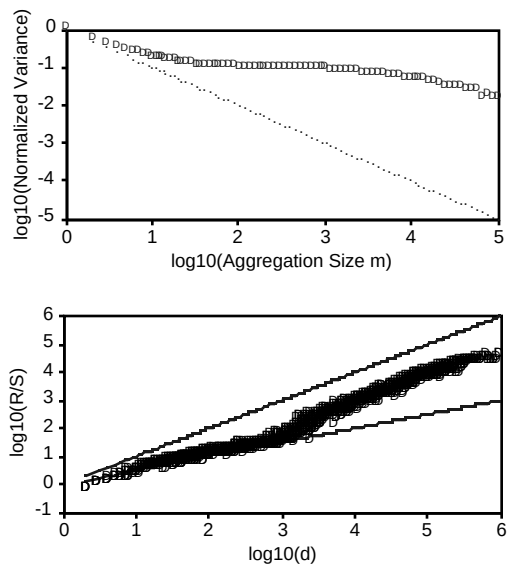


Fig. 4 *R/S* Plot and Variance-Time Plot for the Generated Trace (H = 0.99, H_W = 0.50, r = 0.85, $L = 10^6$, $T = W$ = 1000).

References

[1] D. Cox, Long-Range Dependence: A Review, in *Statistics*: *An Appraisal*, H. A. David and H. T. David (Eds), The Iowa State University Press, Ames, Iowa, pp. 55-74, 1984.

[2] W. Leland, M. Taqqu, W. Willinger, and D. Wilson, On the Self-Similar Nature of Ethernet Traffic, *IEEE/ACM Transactions on Networking*, **2**(1), pp. 1-15, February 1994.

[3] G. Nguyen and S. Batsell, Statistical Characteristics of DIS Traffic, *Proc. IEEE MILCOM*, pp. 637-641, October 1996.

[4] G. Nguyen, Mixing of Self-Similar Series, to appear.

[5] V. Paxson and S. Floyd, Wide-Area Traffic: The Failure of Poisson Modeling, *Proc. ACMSigcomm*, pp. 257-268, 1994.

[6] V. Paxson, Fast Approximation of Self-Similar Network Traffic, *Technical Report LBL*-36750, Lawrence Berkeley Laboratory, April 1995.

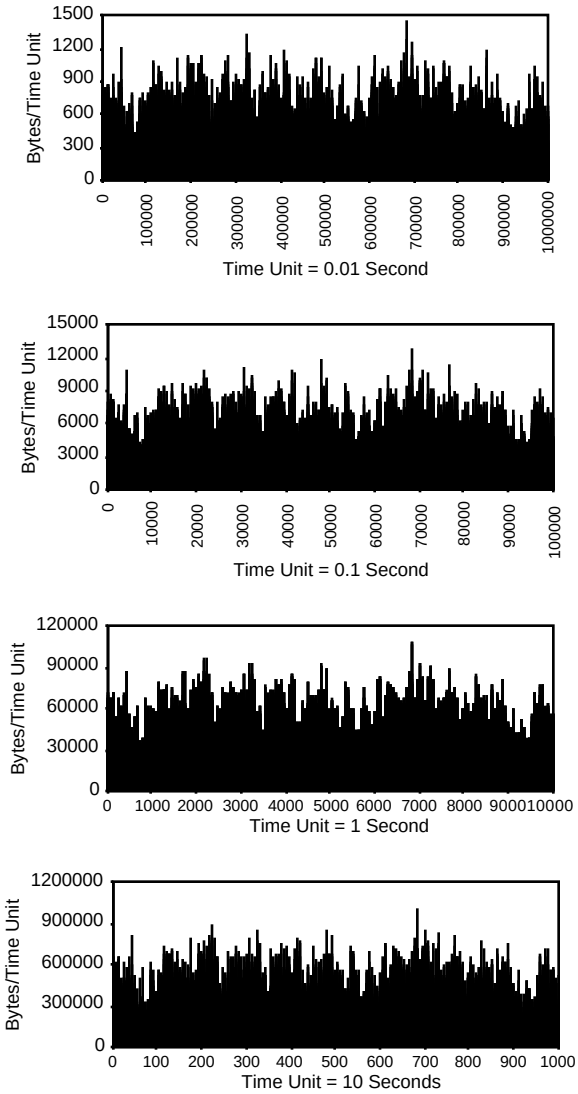


Fig. 5 Synthetic Trace at 4 Time Scales: 0.01 - 10 Seconds ($H = H_W = 0.99$, $r = 0.85$, $L = 10^6$, $T = W = 1000$, $X_i = Y_i$). The traffic shows burstiness in all 4 time scales.

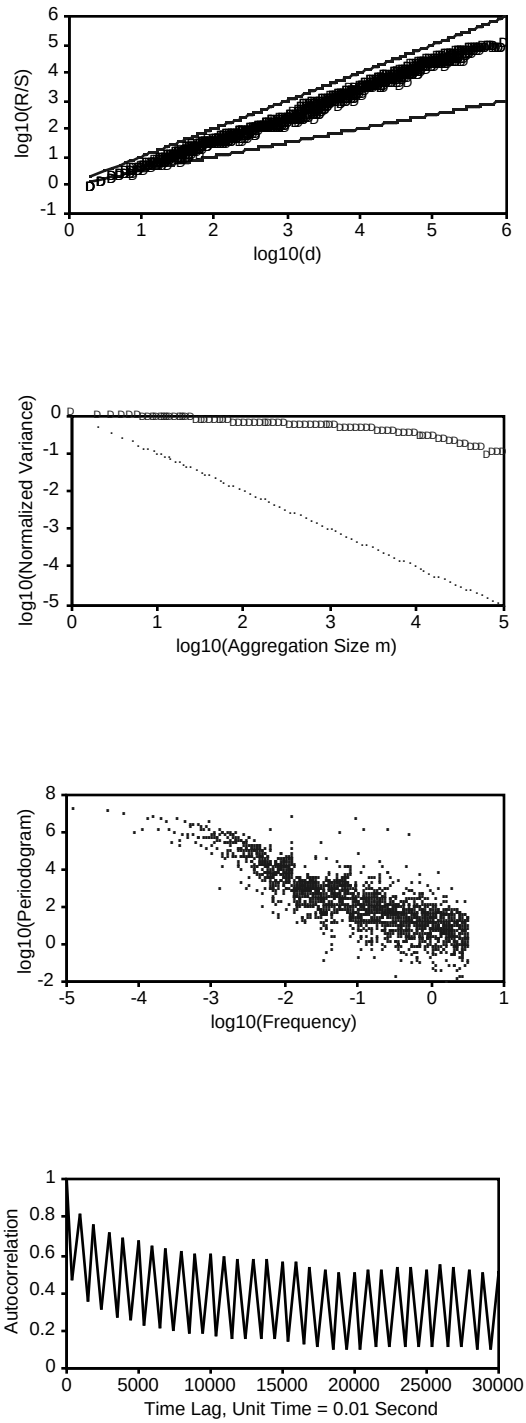


Fig. 6 Graphical Statistics for the Synthetic Trace ($T = W = 1000$, $r = 0.85$, $H = H_W = 0.99$, $X_i = Y_i$).