TECHNICAL RESEARCH REPORT

Self-Similar Traffic Models

by Pradeep Ramakrishnan

CSHCN T.R. 99-5
(ISR T.R. 99-12)

CENTER FOR SATELLITE AND HYBRID COMMUNICATION NETWORKS

The Center for Satellite and Hybrid Communication Networks is a NASA-sponsored Commercial Space Center also supported by the Department of Defense (DOD), industry, the State of Maryland, the University of Maryland and the Institute for Systems Research. This document is a technical report in the CSHCN series originating at the University of Maryland.

Web site http://www.isr.umd.edu/CSHCN/

Supported by: NASA, Hughes Network Systems and, Lockheed Martin Corporation

SELF - SIMILAR TRAFFIC MODELS

Pradeep Ramakrishnan

TABLE OF CONTENTS

TABLE OF FIGURES

1.0 INTRODUCTION

With the advent of broadband communications which is characterized by a heterogeneous traffic mix (e.g. video conferencing applications, ftp, browsing the web....), commonly held assumptions of traditional traffic models have been put into question. Essentially the present type of traffic, is of a highly bursty nature, which is not captured by the traditional traffic models (e.g. Poisson Process). This has a major impact on the design of a network. New models that characterize this burstiness effect is required for the analysis, design, planning, engineering and congestion management of broad band networks [1].

Measurements using high-resolution traffic equipments of wide area network traffic, have confirmed this particular traffic phenomenon. The features shown by the traffic have been called “self-similar or fractal traffic”. Their important properties are stated below [1] :-

- Distributions of the actual traffic processes decay more slowly (“heavy tailed”, e.g. of such a distribution is the Pareto distribution) than exponential (“light tailed” e.g. a Poisson distribution) (see definition of heavy tail and light tailed distribution in the appendix)
- Correlations exhibit a hyperbolic (“long range dependence”) rather than an exponential (“short range dependence”) decay.

Traditional traffic models used in queueing analysis assume variations only in limited time scales while the long range dependent or self-similar processes have fluctuations over a wide range of time scales. This report tries to present various traffic models that represent these properties and the important parameters that need to be estimated which will hopefully enable in the design of an optimum network.

2.0 MODELING WIDE AREA ETHERNET TRAFFIC

In the paper by Murad Taqqu, Walter Willinger and Robert Sherman [7] mathematical proof is given as to how by aggregating simple renewal (ON-OFF) processes results in self-similar behavior. The aggregation of individual ON-OFF sources also allows for the explanation of observed self-similarity in wide area ethernet traffic.

In this particular case the traffic source is either transmitting packets at a constant rate R during the ON period or is idle in the OFF period. The time spent during the ON state (t_{on}) and during the OFF state (t_{off}) is i.i.d and is of a heavy tail distribution e.g. the Pareto distribution with finite mean and infinite variance (see Appendix A for properties of pareto distribution).

i.e. $P[X > x] \sim cx^{-\alpha}$ as $x \rightarrow \infty$, $1 < \alpha < 2$ and c is a finite positive constant.

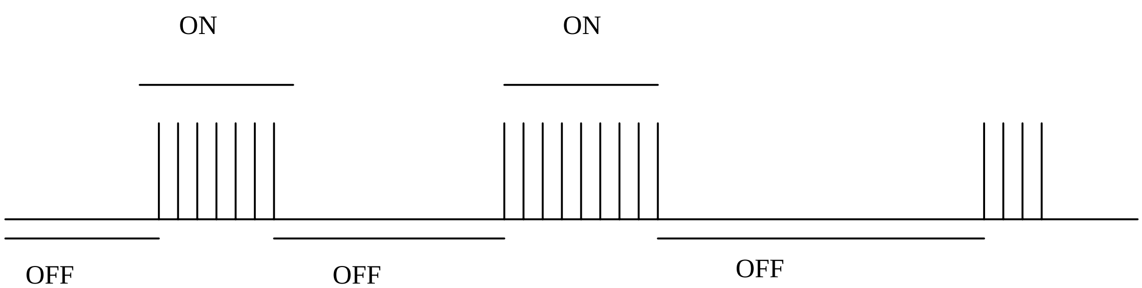


Figure 1

When a large number of these sources are aggregated it results in traffic having fractal (self-similar) characteristics and the model is called the **Fractional Gaussian Noise** Model.

Given below is a trace of FGN traffic, simulated in MATLAB [13] :-

The results in [7] also provide evidence that for a large number of sources, the self similarity property observed in wide area ethernet traffic doesn't depend on the underlying access schemes used (CSMA/CD, as is the case in Ethernet).

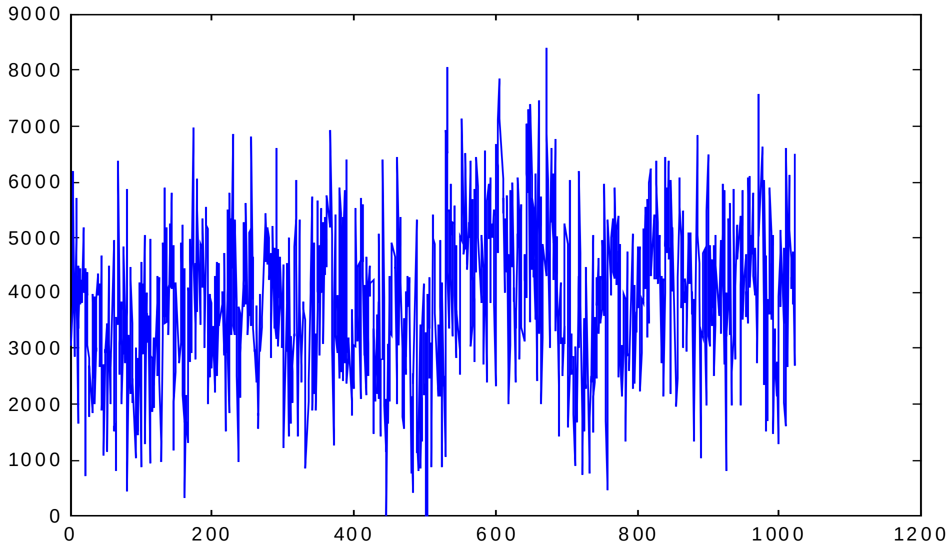


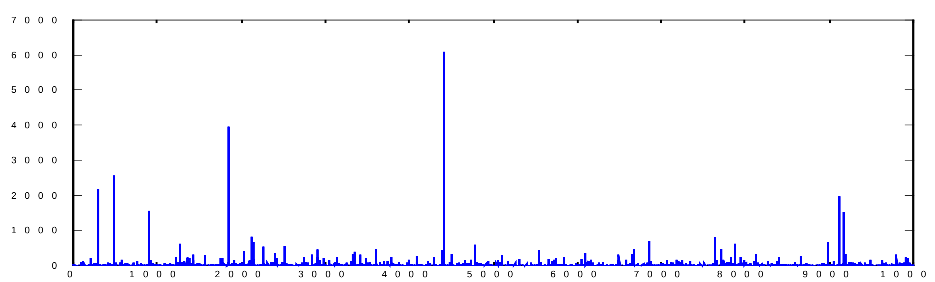
Figure 2

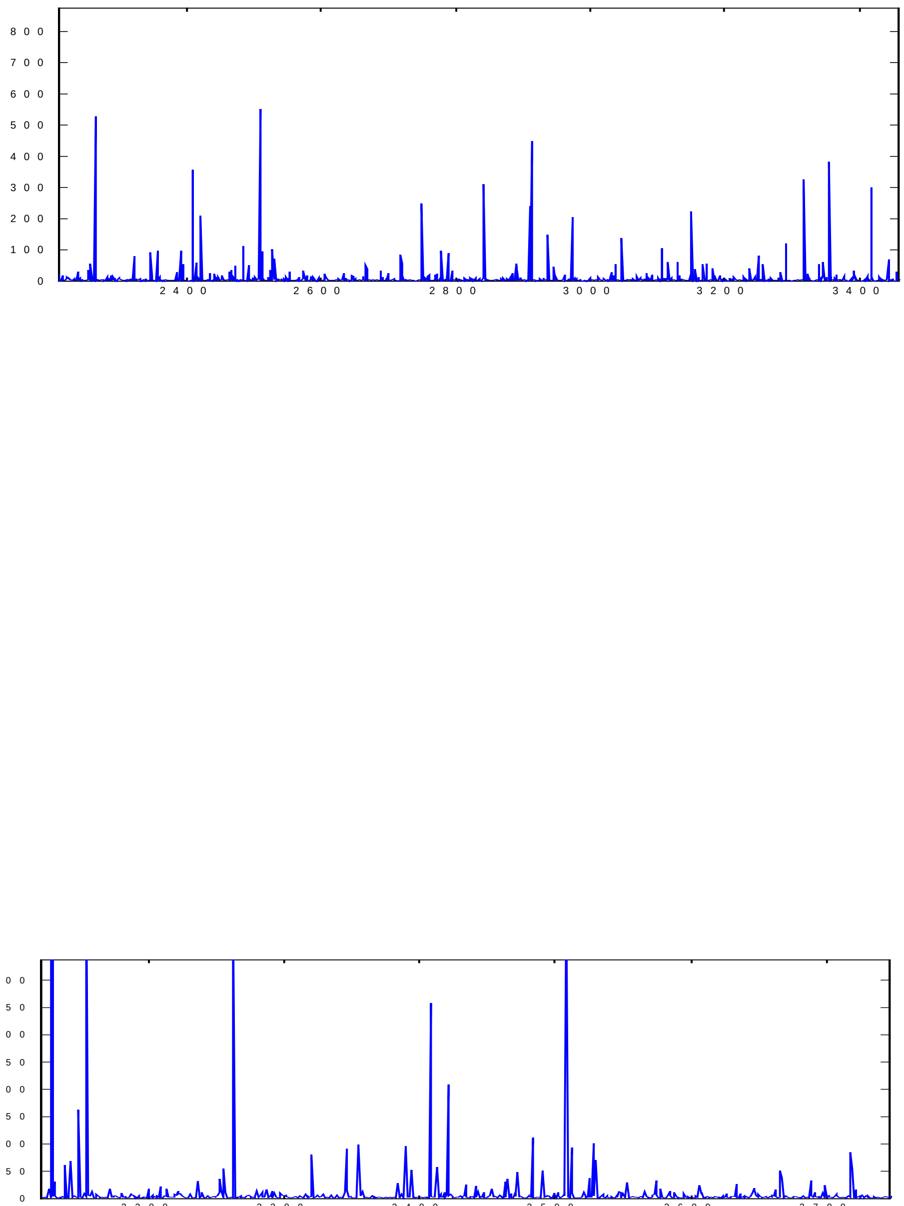
The Hurst Parameter 'H' represents the degree of self-similarity in the observed traffic. When the value of the Hurst parameter is between 0.5 and 1 the traffic is said to be self-similar (values of H closer to 1 indicate a high degree of self-similarity). The relation between H and α is given by $H = (3 - \alpha)/2$.

When trying to fit the synthetic generated data to the actual WAN traffic certain parameters play crucial roles. The first one is α which describes the intensity of self-similarity, once this is estimated H can be calculated using the above equation.

"The other important parameter is the number of sources (M), since the value of M is considered to be large it can chosen to be for e.g. 500, 1600 [7]". Other parameters to be considered are the rate at which packets are generated during the ON period and the lower cutoff of the pareto distribution.

As shown in [7], self-similar traffic generated by superimposing a number of ON-OFF sources, easily passes the visual test of actual Ethernet traffic. The above model has been implemented in OPNET4.0. A intiutive feel for self-similarity can be obtained by observing the following traces at various time-scales (the zoom level used was a factor of 2).





800
700
600
500
400
300
200
100
0
2400
2600
2800
3000
3200
3400
400
350
300
250
200
150
100
50
0
2200
2300
2400
2500
2600
2700

Figure 3

3.0 MODELING TELNET CONNECTIONS

Network connections such as telnet and TCP traffic can be modeled using the M/G/∞ queue. In the M/G/∞ queue, customers arrive according to a poisson process. The service time is obtained from a heavy tailed distribution with infinite variance. Empirical distributions of telnet packet inter-arrival times have shown that they are heavily tailed. Modeling telnet packet inter-arrival times with exponential distribution underestimates the burstiness of the traffic for a single connection as well as that of multiplexed traffic.

"The M/G/∞ model implies that multiplexing constant-rate connections that have poisson connection arrivals and a heavy tailed distribution for connection lifetimes would result in self-similar traffic [9]".

The auto correlation function r(k) for the arrival process is as follows:-

$$r(k) = \text{cov}\{X(t), X(t+k)\} = \rho\int_k^\infty (1 - F(x))dx$$

ρ = poisson arrival rate, *F* = distribution of the service time (heavy tailed in this case) and *X(t)* is the number of customers at time *t.*

Therefore applications such as FTP, TELNET and WWW can be modeled in such a way that the sessions arrive in a Poisson manner and the duration or size of each session has a heavy tailed distribution. This results in an asymptotically self-similar traffic.

i.e. the auto correlation traffic $r(k) \approx k^{-(2-2H)} L(k)$ as $k \rightarrow \infty$.

This has been implemented in OPNET4.0.

4.0 Hurst Parameter Estimation Methods

4.1 Whittle Estimator

[*Beran, 1994*] has suggested various methods for maximum likelihood estimation of different parameters. One of them the ***Whittle estimator*** (non graphical method that provides confidence intervals) is widely used, "it provides asymptotically consistent and normally distributed estimators of the unknown parameters for both Gaussian [*Fox and Taqqu, 1986*] and non Gaussian time series [*Giraitis and Surgailis, 1990*]. The reliability of the Whittle estimator was empirically tested by running Monte Carlo experiments by [*Taqqu and Teverovsky, 1997*] and [*Kokoszka and Taqqu, 1996*]". One important point to note about the Whittle estimation technique is that, it is assumed, the underlying process is actually self-similar. It gives an estimate of the Hurst parameter with a certain confidence. To determine whether the actual time series is self-similar or not, methods like, the R/S statistic, variance time plot have to be used. This has been explained in [3].

The spectral density (Fourier transform of Eq. 3) is (from [3])

$$f(\lambda) = C_H \left(2\sin\frac{\lambda}{2}\right)^2 \sum_{k=-\infty}^{\infty} \frac{1}{|\lambda + 2\pi k|^{2H+1}} \approx C_H |\lambda|^{1-2H} \text{ as } \lambda \to 0 \quad \textbf{- Eq. 4}$$

where C_H is a constant and is = $\left(\frac{2}{2H+1} + \frac{1}{H+2} - \frac{2}{H+1}\right)$

and λ is the frequency.

Given a data sample of size N, the estimation process essentially involves minimizing the following function.

$$Q(\theta) = \sum_{j=1}^{m} \frac{I(\lambda_j)}{f(\lambda_j, \theta)} \quad \text{- Eq. 5}$$

'θ' is the parameter H or α when dealing with fractional gaussian noise.

'm' is the integer part of (N-1)/2, λ_j are the Fourier frequencies ($\lambda_j = 2\pi j/N$), the Whittle estimator is the value of θ which minimizes Q.

$$I(\lambda) = \frac{1}{2\pi N} \left| \sum_{j=1}^{N} X_j e^{ij\lambda} \right|^2 \quad \textbf{- Eq. 6}$$

and f(λ) is the fourier spectral density of the model represented in Eq. 4.

Also the variance can be computed as follows:-

$$\sigma_H^2 = 4\pi\left[\int_{-\pi}^{\pi}\left(\frac{\partial \ \log f(\omega)}{\partial \ H}\right)^2 \partial\omega\right]^{-1} \textbf{ - Eq. 7}$$

This method also provides confidence intervals.

Another method suggested in [Mcleod and Hipel, 1978] is as follows:-

Given a set of observations $z_1, z_2, \ldots, z_N$ the log likelihood of μ, σ^2 and H in the FGN model is

$$\log\ L(\mu,\sigma^2,H) = -\frac{1}{2}\log\left|C_N(H)\right| - \left(2,\sigma^2\right)^{-1}S(\mu,H) - \left(\frac{N}{2}\right)\log\sigma^2$$

where $C_N(H)$ is the correlation matrix and is given by :-

$$C_N(H) = \left[\rho_{|i-j|}\right]$$

$$S(\mu,H) = (z-\mu 1)^T\left[C_N(H)\right]^{-1}(z-\mu 1)$$

where z_T equals $z_1, z_2, \ldots, z_N$ is a 1*N vector and 1_T equal 1,1, ..., 1 is 1*N vector.

For fixed H the MLE of μ and σ^2 are

$$\hat{\mu} = \left\{z^T\left[C_N(H)\right]^{-1}1\right\} \Big/ \left\{1^T\left[C_N(H)\right]^{-1}1\right\}$$

$$\hat{\sigma}^2 = N^{-1}S(\hat{\mu},H)$$

The maximized likelihood function of H is

$$\log L_{max}(H) = -\frac{1}{2}\log\left|C_N(H)\right| - N/2\ \log\left[S(\hat{\mu},H)\right]/N$$

$\log L_{max}(H)$ can then be maximized using the inverse quadratic interpolation search method to determine $\hat{H}$, the MLE of H. The variance of $\hat{H}$, is approximately

$$Var(\hat{H}) = -1 \Big/ \left.\frac{\partial^2 \log L_{max}(H)}{\partial H^2}\right|_{H=\hat{H}}$$

By using numerical differentiation the variance can be computed. One drawback of this approach is that it’s only feasible for small values of N(maximum suggested value of N = 200). The above algorithm is available in MATLAB at [13].

The trace for aggregated traffic was generated in OPNET 4.0 with parameters alpha = 1.2 i.e H = 0.9, 1000 data values were obtained, about 1000 data values were passed on as input to the Whittle estimator. Given below is the minimization graph obtained in MATLAB where H was obtained to be around 0.93.

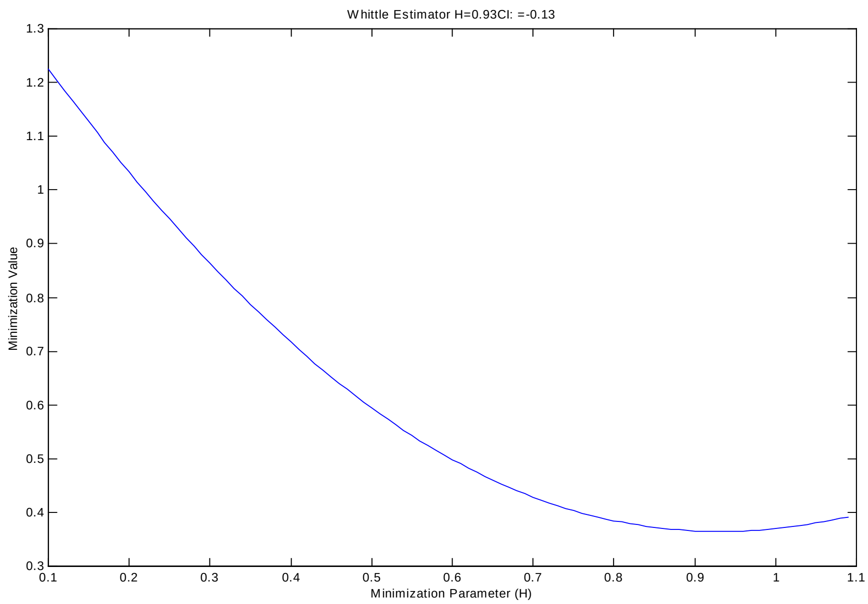


Figure 4

4.2 Local Whittle Estimator

The advantage of using this technique is that one doesn't have to make an assumption of the underlying spectrum for e.g whether it is a fractional gaussian noise model or not.
The only assumption one is making, is that the trace is long-range dependent which might be valid, as recent papers have shown that traffic is more bursty than previously assumed.
The above estimators have been known to give incorrect estimates for certain cases, but these are few and far between. One of the inputs to this function for yielding the estimate of the hurst parameter in the local whittle estimator is a parameter called the number of frequencies. The suggested value to be used is = 32.

4.3 Wavelet Method

Recent research by Feldmann, Gilbert, Willinger have shown that internet traffic over small time scales is multifractal and the effect of TCP leads to a multiplicative property, rather than the previously assumed additive property.

In these scenarios to capture the Hurst parameter at these fine time scales a technique called the wavelet method [by Abry and Veitch,1998] is an efficient(though biased) approach. The algorithm is available in MATLAB script form (at [13]). Following is a brief description from [13]:-

"In wavelet method differences in aggregated series is analysed compared to the variance plot method where aggregated series within a fixed interval was looked into.
Therefore if Y^j is the aggregated series then

$$Y_k^{j+1} = \left(Y_{2k}^j - Y_{2k-1}^j\right).\frac{1}{\sqrt{2}}, k = 1,2,....., N/2j \text{ and } j = 1,2,..$$

Since expectation of Y is zero. In the frequency domain, the variance is equivalent to the signal energy in a frequency band depending on j. E_j vs 2^j is plotted on a log-log scale.

Linearity is checked for all scales *j*. Hurst parameter is then calculated by calculating the slope of the line".

Comparisons are shown in the next two graph, the first one is log E_j vs $\log 2^j$ plot of a poisson process, the Hurst parameter estimated using the above method is approximately 0.45 indicating the traffic is not bursty while the second graph plots the same functions of an asymptotically self-similar process, the Hurst parameter obtained in this case is approximately = 0.65.

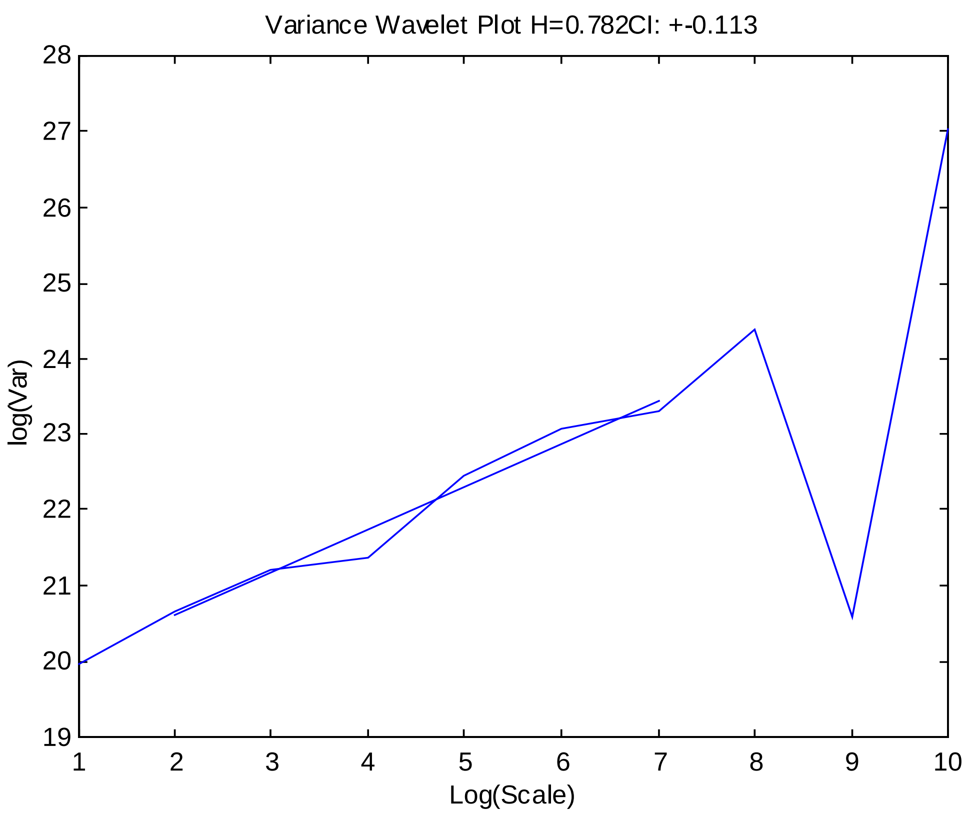


Figure 5

5.0 F-ARIMA (Fractional AutoRegressive Integrated Moving Average)

The F-ARIMA process can model both short range and long range dependence. Observations produced by the ARMA structure shows short range dependence, while the fractional differencing parameter 'd' decays hyperbolically hence showing long range dependence. The F-ARIMA(p,d,q) model can be defined as follows:-

$\phi(B)\nabla^{d}X_{t} = \theta(B)\varepsilon_{t}$ **- Eq. 8**

where the parameter 'd' is between 0 and 1/2 and $\nabla^{d} = (1-B)^{d}$ is the fractional difference operator and can be expressed using the binomial expansion as:

$$(1-B)^{d} = \sum_{k=0}^{\infty}\binom{d}{k}(-1)^{k}B^{k}$$

$$\binom{d}{k} = \frac{d!}{k!(d-k)!} = \frac{\Gamma(d+1)}{\Gamma(k+1)\Gamma(d-k+1)}$$

where Γ(x) represents the gamma function.

$\phi(B) = \left(1-\phi_{1}B-\ldots-\phi_{p}B^{p}\right)$ and $\theta(B) = \left(1-\theta_{1}B-\ldots-\theta_{q}B^{q}\right)$

The parameter d represents the long-range dependence of the process while p and q models the short-range behavior. The relation between d and Hurst parameter(H) is as follows:-

H = d + 1/2

The spectral density of F-ARIMA is comparatively simpler than the fractional Gaussian noise model and is given by [3]:

$$f(\lambda) = \frac{1}{2\pi}\left(2\sin\frac{\lambda}{2}\right)^{-2d} \approx \frac{1}{2\pi}|\lambda|^{-2d} \, as \lambda \rightarrow 0.$$

F-ARIMA models are useful in modeling sequences that do not vary much between successive observations e.g. VBR traffic.

5.1 Estimation Techniques

Initial estimation techniques involved a two step approach [5]. The first step involved estimating the fractional differencing parameter i.e. '$\hat{d}$' using for example the R/S statistic [Mcleod and Hipel (1978)], by estimating d, the Hurst parameter can be obtained (H = d+1/2). The second step involves transforming the observed series using the estimated differencing parameter into a series that follows the ARMA(*p*,*q*) model. But the problem is, in reality we only have a finite sample of X_t (in Eq. 8) and by definition $\nabla^d = (1-B)^d$ is of an infinite realization. Two different procedures (one in the time domain and the other in frequency domain) have been suggested in [5]. The time domain procedure is given below [5] :-

say $u_t = (1-B)^d X_t$ then

using the Binomial theorem

$$u_t = \sum_{j=0}^{\infty} \frac{\Gamma(-d+j)}{\Gamma(-d)\Gamma(j+1)} X_{t-j}$$

u_t is approximated by using the estimated value of *d* in the above expression and setting $z_{t-j} = 0$ for *t-j* outside of the sample.

The problem with the two step approach is that the transformed series in the second step does not have the ARMA model of the u_t series, hence the parameters estimated will not be correct [5].

To solve these problems encountered in the two step approach a one step approach where all the parameters are estimated is suggested in Li and Mcleod (1986) and the other in Fox and Taqqu (1986). But as noted in [5] and [6] Li and Mcleod's method is computationally expensive and is of order n^3 while method which Fox and Taqqu uses, gives a poor estimate if the spectrum of the ARIMA series contains peaks near zero, which exists for positive values of 'd'.

In [6] the exact likelihood function based on n observations $Y_n = (y_1,....,y_n)^T$ from a Gaussian ARFIMA (or ARMA) process is given $\{y_t\}$ by

$$f(Y_n;\Psi) = (2\pi\sigma^2)^{-n/2} |\Omega|^{-1/2} \exp\left[\frac{-(Y_n - \mu 1_n)^T \Omega^{-1}(Y_n - \mu 1_n)}{2\sigma^2}\right]$$

where $\Psi=(\Phi^T,\Theta^T,d,\mu,\sigma^2)^T$ of vector dimension (p+q+3) for ARFIMA processes $\Phi = (\phi_1,...,\phi_p)^T$, $\Theta = (\theta_1,...,\theta_q)^T$, 1_n is the vector of 1's and $\sigma^2\Omega$ is the covariance matrix of Y_n with elements γ_k^y, the autocovariances of $\{y_t\}$ of lag k.

Also, [6] presents an exact likelihood function for estimating the partial regression coefficients from the ARFIMA(p,d,q) process, by a transformation where computation are done from a simpler ARFIMA(0,d,0) process. The complexity of the computation is of order n^2.

6.0 SYNTHESIZING SELF-SIMILAR TRAFFIC

Many methods have been proposed to generate self-similar traffic which approximates the real time data trace. For example, in [7] self-similar traffic is synthesized by superimposing various individual ON-OFF sources. The packet trains produced by each individual ON-OFF source was synthesized by an individual processor in a parallel computing environment.

To produce a synthetic trace of length 100,000, required a massively parallel computer with more than 16000 processors and the time required was the order of a few minutes.

Fast, efficient and accurate schemes are needed to synthesize this type of traffic. In [8] a technique to synthesize approximate self-similar traffic using the fast fourier transform technique is presented. It is as follows [8]: -

First the assumption is made that the power spectrum of the time series corresponds to a Fractional Gaussian Noise Model. The second step is to create a series of complex numbers (z_i) corresponding to the FGN power spectrum. Thirdly inverse discrete fourier transform technique is applied to obtain the time series equivalent.

The difficulty behind this approach is to accurately compute z_i which corresponds the FGN power spectrum. In terms of speed [8], a sample path of 32,768 points took about 11 seconds on a SPARC station IPX and 262,144 sample points took less than 2 minutes. This method compared to other schemes such as the Random Mid Point Displacement method was twice as fast. Also in [8], tests were carried for various values of the Hurst parameter (H) to see if the samples produced, matched what is expected for FGN. Using the Whittles estimating technique, the data generated using the fast fourier transform method was consistent with FGN for the desired value of H. In comparison to the random mid point displacement method (RMD), RMD suffers from biases at certain values of the Hurst parameter which is not observed in the FFT technique.

The point to note is that, there is speed vs. accuracy tradeoff when synthesizing self-similar traffic. Aggregation of many ON-OFF sources is more suitable in a parallel computing environment and the generated traffic is close to the real time ethernet traffic trace. The FFT and RMD technique is fast but results only in an approximate self-similar traffic.

7.0 OPNET SIMULATION

Given below is a simulation conducted in OPNET. The dominant application over the internet that generates majority of the traffic is surely the WWW; users browsing, downloading files and doing other interactive stuff can be thought of as a typical scenario. A typical (although simplified) topology would be as shown in the Figure 6. Users browsing and downloading files from different servers located in different sites around the world. The packets travelling to and fro are switched or routed via a gateway to the respective hosts and servers. The gateway could be thought of as a simplistic view of an ISP.

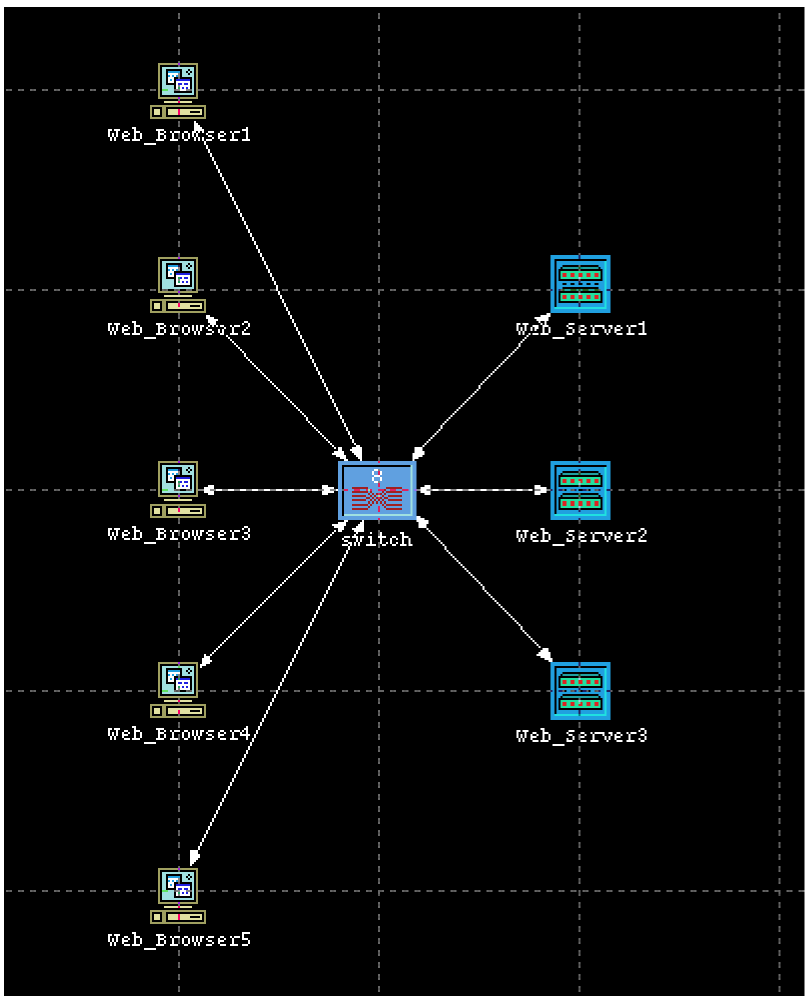


Figure 6

Comparison of various parameters like offered load, the web server throughput etc… when an exponential and a heavy tail distribution is specified are illustrated. The analysis of these results is presented below.

7.1 Http Obect Size

Self-Similar nature of WWW traffic has been attributed to the presence of heavy tailed file-sizes at the server end. This feature has been incorporated into the OPNET models by specifying the transmitted response size of an application when a user request for a file to be of heavy tailed.

In [16] analysis of actual packet traces suggests that the response sizes (above 1KB) are well modeled using Pareto distribution where the parameter 'α' is in the range from α = 1.04 and α = 1.14. These were the values used for the simulation.

Show below is the object size in bytes for a heavy tail and an exponential distribution. Clear distinction can be observed when the object size is of a heavy distribution (Figure 7) where the traffic is more bursty, compared to the exponential case (Figure 8) where it is less so.

A WWW page usually consists of various objects for example images, audio files, text …etc. The size of these objects represents the page size. By setting the object sizes to be of a heavy tailed distribution, translates to a page which is of the same distribution. Various other attributes relating to the page size such as number of pages per file, typical user charactersitics such as low, medium or heavy intensity browsing can be specified as attributes in the simulation.

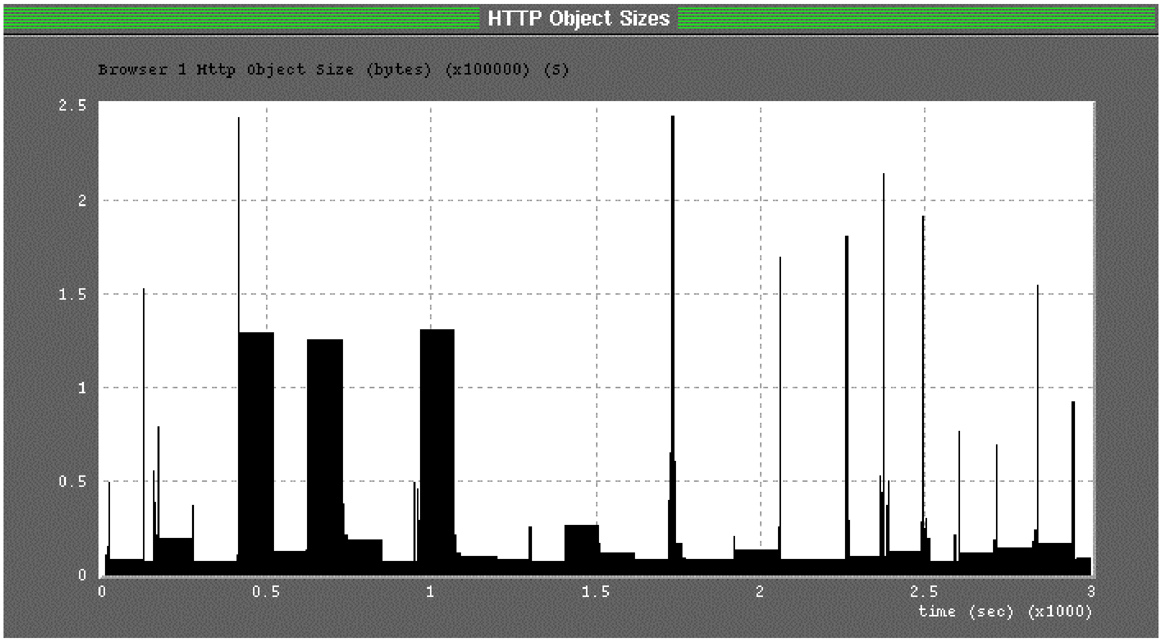


Figure 7

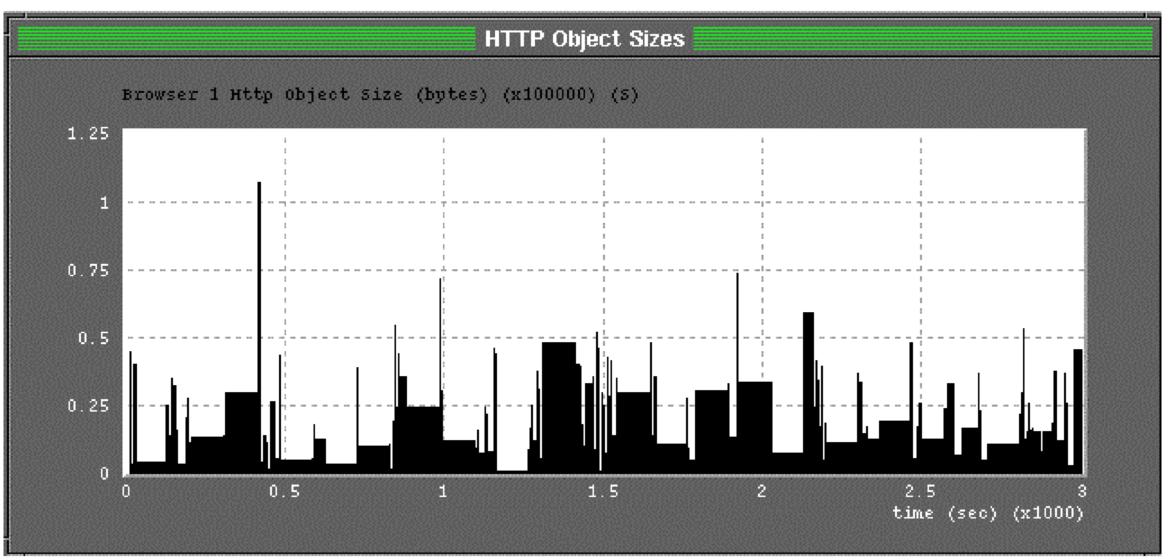


Figure 8

7.2 Server Throughput

The Server throughput (in terms of packets) is much lower when files sizes are of a heavy tailed distribution (Figure 9) compared to that of an exponential case (Figure 10). This could be possibly explained by the fact that the queues in the gateway are getting congested and therefore the server has to back off (probes should be created to analyse the more interesting efects in the gateway e.g delay, variation in queue depth etc.)

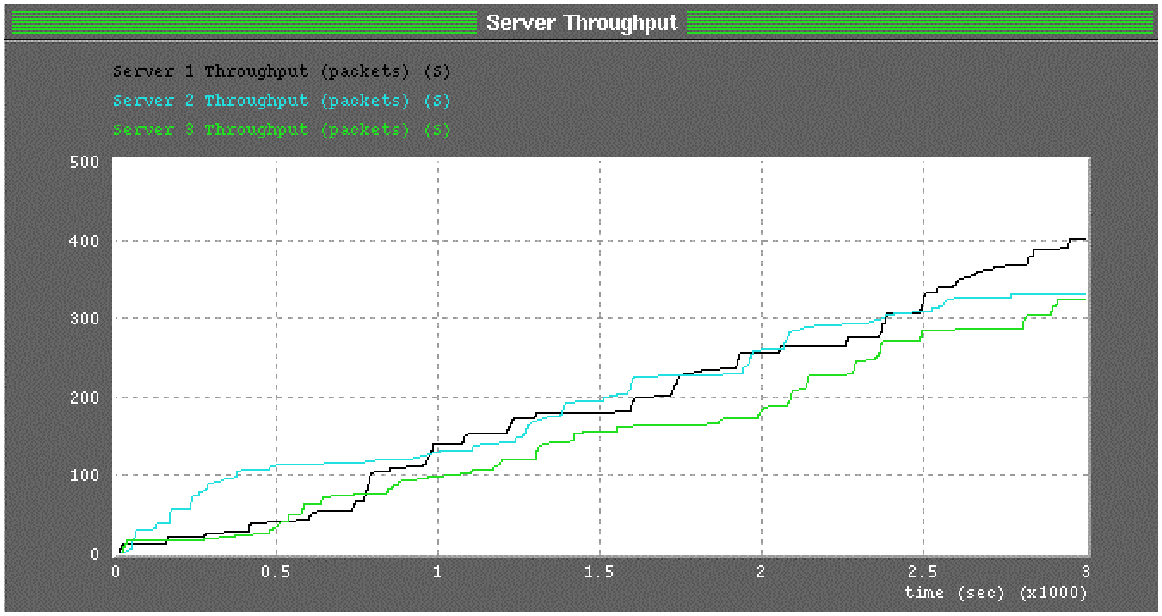


Figure 9

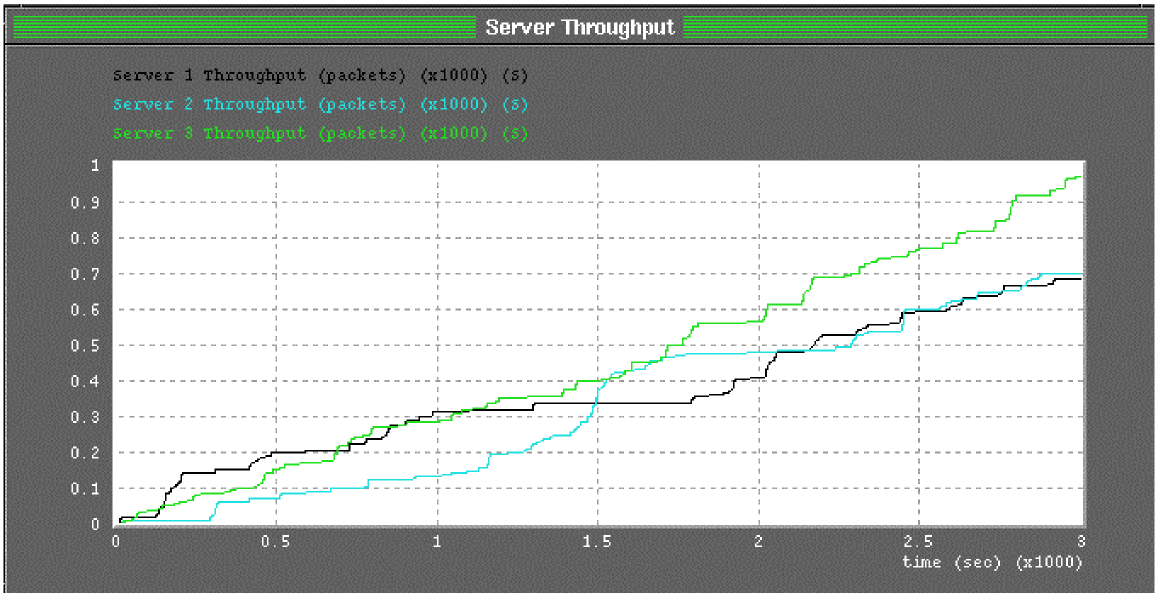


Figure 10

8.0 FUTURE WORK

Various researchers have claimed that high speed network traffic including traces obtained from a host to host TCP connection is actually "multi-fractal" rather than just being long range dependent. Long range dependency only captures the low frequency content of traffic(i.e. at large time scales) while multi fractal is about analyzing the high frequency content i.e. the property is observed when studying traffic over small time scales. The effect of TCP flow control and such actually exaggerates the multiplicative effect. This would be an interesting area to look into. Answers to questions like what parameters other than the Hurst parameter would be appropriate in capturing this effect and how this effect queue sizes and so on would be interesting.

In [15] a parameter called the Holder exponent is estimated using wavelet techniques, this parameter captures the burstiness at both large and small time scales. Also from a given multifractality can be detected by plotting a log – log plot of mean and block size 'm' for different values of 'q'

$$\mu^{(m)}(q) = E\left|X^{m}\right|^{q} = E\left|\frac{1}{m}\sum_{i=1}^{m} X(i)\right|^{q}, m = 1,2......$$

If the plot reveals linearity in different values of q then we can assume that the multifractal approach is suitable. Further explanation is provided in [15].

Suggested methods for the construction of multifractals is by using cascades. The Hurst parameter can be estimated using discrete wavelet transform method suggested by Abry and Veitch in 1998.

9.0 REFERENCES

[1] Fractal Queuing models *by Ashok Erramilli, Onuttom Narayan and Walter Willinger*

[2] Stochastic modeling of traffic processes *by David L. Jagerman, Benjamin Melamed and Walter Willinger*

[3] Estimators for long range dependence: an empirical study *by Murad S. Taqqu, Vadim Teverovsky and Walter Willinger in "Fractals", Vol 3, No. 4 (1995) 785-788*

[4] Statistics for Long-Memory Processes *by Jan Beran*

[5] Maximum Likelihood estimation of stationary univariate fractionally integrated time series models *by Fallaw Sowell in Journal of Econometrics 53 (1992) 165 -188*

[6] Exact Likelihood Function Forms for an ARFIMA Process *by J.S.Pai and N. Ravishanker*

[7] Proof of a Fundamental Result in Self-Similar Traffic modeling *by Murad Taqqu, Walter Willinger and Robert Sherman*

[8] Fast Approximation of Self-Similar Network Traffic *by Vern Paxson*

[9] Wide-Area Traffic: The Failure of poisson Modeling *by Vern Paxson and Sally Floyd*

[10] Modeling of Long-Range Dependent Traffic in ATM Networks Using Heterogeneous On-Off *Sources by Chih-Hsien Chou and Evaggelos Geraniotis*

[11] Queue Response to input Correlation Functions: Continuos Spectral *Analysis by S.Q.Li and C.L.Hwang*

[13] http://buggy.itm.hk-r.se/ttn410/models/

[14] Wavelet Analysis of Long-Range Dependent Traffic *by Abry and Veitech*

[15] The Multifractal Property of Bursty Traffic and its Parameter Estimation Based on Wavelets *by Huimin Chen, Hong Cai and Yanda Li*

[16] An Emperical Model of HTTP Network Traffic by Bruce A. Mah

9.0 APPENDIX

APPENDIX A

p.d.f of light tailed distribution

P[X>x] = $L_1(x)e^{-x}$ as $x \rightarrow$ì where L_1 is slowly varying at infinity

c.d.f

$$F(x) = P[X \le x] = 1 - e^{-\lambda x}$$

taking the inverse $x = \dfrac{-\ln(1-y)}{\lambda}$

p.d.f of heavy tailed distribution

P[X>x] = $L_2(x)e^{-\alpha}$ as $x \rightarrow$ì

c.d.f

$$F(x) = P[X \le x] = 1 - \left(\frac{\alpha}{x}\right)^{\beta}$$

taking inverse

$$f(x) = \beta\alpha^{\beta} x^{-\beta-1}$$

for $\beta \le 2$ the distribution has infinite variance, for $\beta \le 1$ distribution has infinite mean.

mean = $\dfrac{\alpha\beta}{1-\beta}$

APPENDIX B

Fractional Gaussian Noise

Fractional Gaussian Noise is an increment of fractional Brownian motion $\{B_H(t), t \geq 0\}$ with mean 0 and variance $EB_H^2(t) = t^{2H}$ and covariance [3]: -

$$EB_H(t_2)B_H(t_1) = \frac{1}{2}\sigma^2\left[t_2^{2H} - (t_2 - t_1)^{2H} + t_1^{2H}\right] \text{ - } \mathbf{Eq.\ 1}$$

The covariance of increments in two non-overlapping blocks is given by:-

$$E\left[B_H(t_4) - B_H(t_3), B_H(t_2) - B_H(t_1)\right] = E\left[B_H(t_4), B_H(t_2)\right] - E\left[B_H(t_4), B_H(t_1)\right] - E\left[B_H(t_3), B_H(t_2)\right] + E\left[B_H(t_3), B_H(t_1)\right]$$

$$= \frac{1}{2}\sigma^2\left[\left(t_4 - t_1\right)^{2H} - (t_3 - t_1)^{2H} + \left(t_3 - t_2\right)^{2H} - \left(t_4 - t_2\right)^{2H}\right] \text{ - } \mathbf{Eq.\ 2}$$

In the discrete situation substituting the increments t_1, t_2, t_3 and t_4 by n, n+1, n+k & n+k+1 respectively in Eq. 2 and dividing by σ^2 we have

$$\rho_k = \frac{1}{2}\left[(k+1)^{2H} - 2k^{2H} + (k-1)^{2H}\right] \text{- } \mathbf{Eq.\ 3}$$

This increment sequence is called **Fractional Gaussian Noise**