

On the Self-Similar Nature of ATM Network Traffic

Humam Elagha
Arab Academy for Banking and
Financial Sciences, Faculty of
Information Systems and
Technology, Amman, Jordan

humam@students.aabfs.org

Maher AlShafee
Royal Scientific Society,
Information Technology
Centre, Amman, Jordan

maher_s@rss.gov.jo

Abstract

Modeling multimedia traffic is an important issue in performance analysis and design of communication networks. With introduction of new applications, the characteristics of data traffic changes. In this paper, a measurement study of ATM Network traffic has been carried out and it is shown that the recorded data exhibit self-similar features. The conclusions are supported by a comprehensive analysis using one of the most popular statistical methods called Indices of Dispersion. Our results validate one of the most striking findings of the present teletraffic research: a broad range of packet network traffic has fractal-like behavior. We also investigate three popular synthetic self-similar traffic models and find out the most accurate one for the measured traffic.

Keywords: ATM, Modeling, Self-similar traffic, Indices of Dispersion, Hurst parameter, Performance Testing.

Introduction

The characterization of real traffic is a critical issue to the success of efficient traffic engineering in ATM networks. Research in this field has resulted in numerous models and techniques over the last decade (Bjorkman, Latour-Henner, Hasson, pers, and Miah, 1995; Stamoulis, Anagostu, and Georgantas, 1994). However, the developed traffic characterization methods have been, in general, rather complicated and demand intensive computation of several statistical parameters. There is a lack of simple and accurate methods that can be of practical use to network operators.

Extensive data studies indicate that traffic in high-speed communications networks has long-memory and heavy tailed (impulsive) characteristics. With the rising popularity of multimedia applications over networks, these properties of the traffic are only likely to become more dominant, posing unique new challenges to designers of network systems and protocols. Traditional

Material published as part of this publication, either on-line or in print, is copyrighted by the Informing Science Institute. Permission to make digital or paper copy of part or all of these works for personal or classroom use is granted without fee provided that the copies are not made or distributed for profit or commercial advantage AND that copies 1) bear this notice in full and 2) give the full citation on the first page. It is permissible to abstract these works so long as credit is given. To copy in all other cases or to republish or to post on a server or to redistribute to lists requires specific permission and payment of a fee. Contact Publisher@InformingScience.org to request redistribution permission.

teletraffic theory cannot capture these traffic characteristics. During the last few years, significant research results have been proposed on models that capture self-similarity of traffic. These models, however, are inadequate for predicting queuing performance, delays, and buffer dimensions since the implications of the combination of self-similarity and impulsiveness queuing performance can be dramatically differ-

ent from that of self-similarity alone. To the best of my knowledge there are no models that have been derived based on real traffic dynamics that also capture the data impulsiveness?

The modeling of self-similar traffic appeared as an emerging and challenging field of the present teletraffic research. It seems that there are different promising approaches to capture this complex fractal-like behavior. Norros (1993, 1995) used a Gaussian self-similar process known as the Fractional Brownian Motion. Willinger, Taquq, Sherman and Wilson (2004) applied the superposition of on/off sources with heavy-tailed on and off period. Erramilli (1994) and Singh and Erramilli (1999) studied different chaotic maps.

The rest of this paper is organized as follows. In section 2 we give information about the self-similar traffic characteristics and the used techniques for measuring the self-similarity level (Hurst-parameter H). In section 3 we present the analysis based on the real measurements taken from the Eastern Mediterranean University (EMU) ATM network and find out the self-similarity level. In section 4 we investigate the three promising self-similar modeling approaches to capture the observed properties, and we find out the most appropriate model for EMU ATM network traffic. Finally, section 5 summarizes the results of the paper and identifies areas for future research.

The Self-Similarity Phenomena and its Testing

A self-similar phenomenon represents a process displaying structural similarities across a wide range of scales of a specific dimension. In other words, the reference structure is repeating itself over a wide range of scales of diverse dimensions (geometrical, or statistical, or temporal). However, these properties do not hold indefinitely for real phenomena and, at some point, this structure breaks down.

Consider a time series $X = \{X_n, n \in Z^+\}$ and define another time series (m -aggregated) $X^{(m)} = \{X_n^{(m)}, n \in Z^+\}$ by averaging the original time series over non-overlapping, adjacent blocks of size m :

$$X_n^{(m)} = \frac{1}{m} \sum_{i=nm-(m-1)}^{nm} X_i$$

$X^{(1)}$ represents in this case the highest resolution that is possible for the process. Lower resolution evolutions of the process $X^{(m)}$ can be obtained by m -aggregating the X_n process, like for instance

$$X_n^{(4)} = \frac{X_{4n-3} + X_{4n-2} + X_{4n-1} + X_{4n}}{4}$$

There are two classes of self-similar processes, namely *exactly self-similar* processes and *asymptotically self-similar* processes.

A process X is said to be *exactly self-similar* with parameter β ($0 < \beta < 1$) if, for $m \in Z^+$ the following conditions are fulfilled

- Variance

$$Var[X^{(m)}] = \frac{Var[X]}{m^\beta}$$

- Autocorrelation

$$R(k, X^{(m)}) = R(k, X)$$

The parameter β is related to the so-called Hurst parameter (H) by the following formula

$$\beta = 2(1 - H)$$

It is observed that for *stationary ergodic* processes, $\beta = 1$ and the variance decays to zero. On the contrary, for *exactly self-similar* processes the variance decays more slowly.

Another class of self-similar processes is the so-called *asymptotically self-similar* processes. A process X is said to be *asymptotically self-similar* if, for k large enough

- Variance

$$\text{Var}[X^{(m)}] = \frac{\text{Var}[X]}{m^\beta}$$

- Autocorrelation

$$R(k, X^{(m)}) \rightarrow R(k, X)$$

as $m \rightarrow \infty$

It is observed that, for both classes of self-similar processes, the variance of $X^{(m)}$ decreases more slowly than $\frac{1}{m}$ as $m \rightarrow \infty$. This is to compare with the case of stochastic processes where the

variance decreases proportional to $\frac{1}{m}$ and approaches zero $m \rightarrow \infty$ (consistent with white noise, i.e., uniform power spectrum).

The most striking feature of self-similar processes is however that the autocorrelation function does not degenerate when $m \rightarrow \infty$. This is in contrast to stochastic processes, where the autocorrelation function degenerates as $m \rightarrow \infty$.

Properties

- **Long Range Dependence:**

A self-similar process with parameter $0.5 < H < 1$ is long-range dependence. By definition the process X is called long-range dependent if

$\sum_{k=0}^{\infty} \text{Cov}(k) = \infty$ is non-summable. Otherwise it is called short range dependence

- **Slowly Decaying Variance**

The variance of the sample mean decreases much more slowly than the reciprocal of the sample size in the case of a self-similar process.

$$\text{Var}[X^{(m)}] \sim m^{-\beta}$$

for sufficiently large m .

• **Fractal Dimensions**

The fractal dimension d of an object is defined as

$$d = \frac{\ln N}{\ln \frac{1}{\eta}}$$

where N is the number of self-similar objects/pieces (versions of itself) to cover a given d -dimensional object, and η is the linear scale size of the covering objects.

Tests

The level of self-similarity in a time series is indicated by the Hurst parameter H . This H parameter has a value range of $0.5 \leq H \leq 1.0$ and strong self-similarity means that H has larger values (close to 1.0). Several methods are available today to estimate self-similarity in a time series (Stallings, 1998; Abry and Veitch, 2005; Beran, 1994).

1. Indices of Dispersion : A commonly used measure for capturing the variability of traffic over different time scales is provided by the *index of dispersion for counts* (IDC), Cox and Lewis (1966),

$$IDC(L) = \frac{\text{var}\left\{\sum_{j=1}^L X_j\right\}}{E\left\{\sum_{j=1}^L X_j\right\}}$$

Self-similar processes produce a monotonically increasing IDC. Plotting $\log IDC(L)$ against $\log L$ results in an asymptotic straight line with slope $2H - 1$ (Leland, Taquu, Willinger and Wilson, 2003).

2. Rescaled Range Analysis

For a given set of observations $X = \{X_n, n \in Z^+\}$ with sample mean $\bar{X}(n)$, sample variance $S^2(n)$, and range $R(n)$, the rescaled adjusted range or R/S statistic is given by

$$\frac{R(n)}{S(n)} = \frac{\max(0, \Delta_1, \Delta_2, \dots, \Delta_n) - \min(0, \Delta_1, \Delta_2, \dots, \Delta_n)}{S(n)}$$

Where

$$\Delta_k = \sum_{i=1}^k X_i - k \bar{X}$$

for $k= 1, 2, \dots, n$.

It is also observed that for many natural phenomena

$$E\left[\frac{R(n)}{S(n)}\right] \approx cn^H$$

as $n \rightarrow \infty$ and c being a finite positive constant independent of n . Taking the logarithm of both sides, we have

$$\left\{ E \left[\frac{R(n)}{S(n)} \right] \right\} \approx H \log(n) + \log(c)$$

Thus, one can estimate H by plotting $\log \left\{ E \left[\frac{R(n)}{S(n)} \right] \right\}$ versus $\log(n)$ on a graph, and least-squares fitting a straight line, with slope H , through the resulting points.

EMU's ATM Traffic Measurements and Analysis

The EMU Network

The measurements were made on the Eastern Mediterranean University ATM network. University's ATM Network is completely built with FORE SYSTEMS, Inc.'s devices. Some of the devices in this ATM network can be listed as an ATM backbone switch, ATM workgroup switches and special devices for video and voice over ATM.

Measuring Tool and Configuration

The measurements were made in the EMU Computer Center using the ATM Sniffer Pro. ATM Sniffer Pro is a full bandwidth, Asynchronous Transfer Mode analysis and management system. It is designed to be used by network managers, technicians, and development engineers involved with all aspects of ATM technology. ATM Sniffer Pro allows the user to monitor, analyze, and manage the physical interface to each ATM device, as well as the traffic passing through each port.

There are two analysis devices which can be used with the ATM Sniffer Pro, the ATM Book and UTP-5 Pod. UTP-5 Pod presents the data as ATM frame size in multiples of 48 bytes (the same size of the payload of ATM cells), but the ATM book presents the data as ATM cells. Since the ATM book is not available, we used the UTP-5 Pod.

We can connect the UTP-5 pod to the network in either of the following two methods:

1. User-to-Switch. In this method, the pod is installed between an ATM switch and a workstation with ATM adaptor card.
2. Switch-to-Switch. In this method, the pod is installed between two switches in the ATM network.

The Measured Data

In both of the above configurations the measured data was the number of cells received in one-second intervals.

In the first scenario, the ATM Sniffer Pro is connected to the network as user-to-Switch Configuration. Two data files were captured, EMUN1, EMUN3; each file contains about 620,000 cells in a 60 hours capture duration.

In the second measurement scenario, the ATM Sniffer Pro is connected to the network as switch-to-Switch Configuration. Two data files were captured, EMUN2, EMUN4. Even though, all files contains approximately the same number of cells, the capture duration was about 30 minutes to 6 hours according to the network load, which is very small compared with the first scenario. A summary description of these data sets is given in Table 1.

Table 1: Summary of Data Sets

File name	Number of cells	Length (hour)	Configuration
EMUN1	620,000	62 hour	User-to-switch
EMUN2	656,128	6.5 hour	Switch-to switch
EMUN3	620,000	59.5 hour	User-to-switch
EMUN4	688,882	½ hour	Switch-to switch

Analysis

To test the self-similarity parameter H , we performed the IDC plots analysis. Figure 1, shows the IDC curve corresponding to the trace EMUN1. The sequence of cell counts in every one-second interval was analyzed. The IDC curve for the EMUN1 file increases monotonically throughout a time span that covers 3-4 orders of magnitude and shows an asymptotic slope that is strictly different from the horizontal line resulting in an estimate H of 0.84. This behavior is in stark contrast to conventional traffic models such as Poisson processes, where the IDC is constant ($H = 0.5$). The same analysis is made for the second data set (EMUN2); the analyzed process was the sequence of cells in each second interval as in the case of EMUN1.

We can conclude, that for the EMUN measurements the estimated value of H is about 0.85 for all the data sets and the measured ATM traffic definitely exhibit the features of self-similarity. To estimate the self-similarity parameter H more precisely, more refined statistical methods with confidence intervals for H and longer measurements with more data are needed.

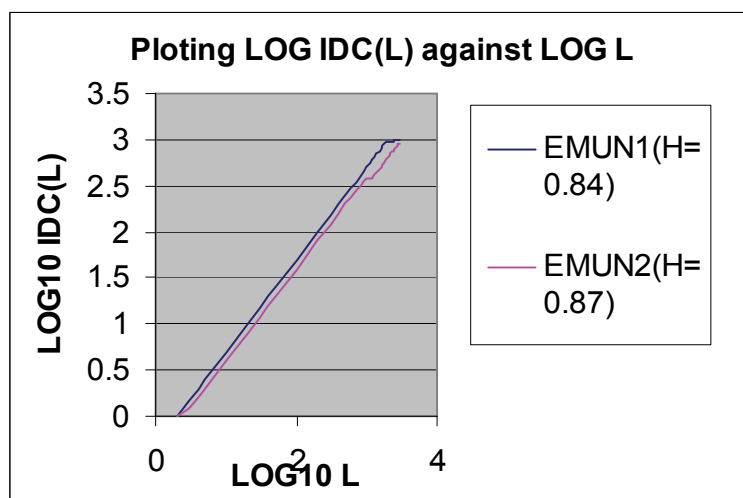


Figure 1: IDC Plots for EMUN1, EMUN2

Modeling of Self-Similar Traffic

There are three different promising self-similar traffic generators based on different modeling approaches:

1. Fractional Brownian traffic

This model was introduced by Norros (1993). The *fractional Brownian traffic* is a process of the form $At = mt + \text{sqrt}(am) Zt$, where At is the amount of traffic in $[0, t)$, and Zt is a normalized fractional Brownian motion, i.e., a strictly self-similar process. The process has three parameters: the mean input rate m , the variance coefficient a and the Hurst-parameter H of Zt .

2. Superposed on/off sources

The model was introduced by Mandelbrot (1994). The basic idea of the model is the construction of self-similar processes based on aggregating many simple on/off processes with heavy-tailed on and off periods.

3. Chaotic maps

Erramiili and Singh (1994, 1999) proposed *chaotic maps* for fractal traffic modeling. The underlying idea is based on a nonlinear map that describes the evolution of a state variable over discrete time.

Performance Testing

The self-similar traffic traces of the models were generated by Monlar and Vidacs (2005). They calculated the value of the Hurst-parameter of synthetic traffic traces as given in Table 2.

Table 2: Hurst-parameter for synthetic traffic models

Traffic Trace	Hurst-parameter value
Fractional Brownian traffic	0.71
Superposed on/off sources	0.73
Chaotic maps	0.80

The investigated three different models are all capable to capture the self-similar properties of ATM traffic but the chaotic map model was found to be the closest model for EMU network traffic.

Conclusion

In this paper, we presented ATM traffic measurements, analysis and modeling results focusing on capturing the self-similarity phenomenon. On the basis of the considerations presented in our paper, the following conclusions can be drawn:

For the EMUN measurements, the calculated value of H is about 0.85 for all the data sets and the measured ATM traffic definitely exhibit the features of self-similarity. To estimate the self-similarity parameter H more precisely, more refined statistical methods with confidence intervals for H and longer measurements with more data are needed. The interpretation of the estimated Hurst parameter is problematic in practice.

The most promising modeling approach in our case is the chaotic map model. The synthetic traffic trace generated by the model has nearly the same statistical properties and self-similar features than the measured traffic trace.

Some important areas that are impacted by self-similarity are data analysis, statistical inference, mathematical modeling, queueing and performance analysis. These are questions that are still under investigation and research, and more efforts must be done in the future to answer them.

References

- Abry, P. & Veitch, D. (2005). Wavelet analysis of long range dependent traffic. *IEEE Transactions on Information Theory*, 44(1).
- Beran, J. (1994). *Statistics for long-memory processes*. Chapman and Hall.
- Bjorkman, N., Latour-Henner, A., Hasson, U., Pers, O., & Miah, A. (1995). Practical ATM resource dimensioning based on real-time traffic measurements and analysis. *GLOBECOM*, 1, 399-403.
- Cox, D.R & Lewis, P.A.W. (1966). *The statistical analysis of series of events*. Menthuen.
- Erramilli, A. (1994). Chaotic maps as models of packet traffic. *Proceedings of ITC-14*.
- Leland, W., Taquq, M., Willinger, W., & Wilson, D. (2003). On the self-similar nature of Ethernet traffic. *IEEE/ACM Transactions on Networking*, 2(1).
- Mandelbrot, B.B. (1994). Long-run linearity, locally Gaussian processes, H-spectra and infinite variances. *International Economic Review*, 10, 54- 60.
- Monlar, S. & Vidacs, A. (2005). On modeling and shaping self-similar ATM traffic. TR. High speed Networks Laboratory, Department of Telecommunications and Telematics, Technical University of Budapest, Budapest, Hungary.
- Norros, I. (1993). Studies on a model for connectionless traffic, based on fractional Brownian motion. *Conference on Applied Probability in Engineering, Computer and Communication Sciences*, Paris, France, June.
- Norros, I. (1995). On the use of fractional Brownian motion in the theory of connectionless networks. *IEEE JSAC*, 13(6), 953-962, August.
- Norros, I. (2004). A storage model with self-similar input. *Queueing Systems Theory and Applications*, 16, 387-396.
- Singh, R.P. & Erramilli, A. (1999). Application of deterministic chaotic maps to model packet traffic in broadband networks. *Proceedings of 7th ITC Specialists Seminar*, pp. 8.1.1- 8.1.3.
- Stallings, W. (1998). *High-speed networks TCP/IP and ATM design principles*. Prentice-Hall.
- Stamoulis, G.D., Anagostu, M.E., & Georgantas, A.D. (1994). Traffic source models for ATM networks: A survey. *Computer Communications*, 17(6), 428-438.
- Willinger, W., Taquq, M.S., Sherman, R. & Wilson, D.V. (2004). Self-similarity through high variability: Statistical analysis of Ethernet LAN traffic at the source level. *IEEE/ACM Transactions on Networking*, 5(1).

Biographies



Humam Elagha is currently a PhD student at Faculty of Information Systems and Technology at Arab Academy for Banking and Financial Sciences, Jordan. His main research interests include IT Governance, Strategic Alignment, IT Value Assessment and High Speed Networks Measurements and Analysis. He holds BSc and MSc in Computer Engineering from Eastern Mediterranean University, Cyprus. He worked as a full-time lecturer at Faculty of Information Technology at Islamic University of Gaza, Palestine. He teaches Database Management Systems, Management Information Systems, Operating systems, Software Engineering, Systems Analysis and Design, Logic Design, Computer Networks, Systems Simulation and Automata Theory Courses. He is a Certified ATM Network Manager. His work appears in number International Journals and Conferences. He is a member of ACM and IEEE.



Maher Alshafee is currently Programmer and Training Coordinator at the Information Technology Center / Royal Scientific Society, Jordan. His research interests include Information Systems Management, Technology Change Management, Human Resources Management, Software Engineering, Network Analysis and Quality Assurance. He holds BSc in Computer Studies and Information Technology from Eastern Mediterranean University, Cyprus. He worked at UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION “UNESCO” as Technical Support & Quality Assurance (QA) Supervisor for the International Computer Driving License (ICDL) Programme, executed by the UNESCO Cairo Office, the ICDL Program in Jordan. He is a Training Coordinator Professional. His work appears in number International Journals and Conferences. He is a member of ACM and IEEE.