STOCHASTIC REPRESENTATION OF TELEWORKER ACTIVITIES IN ETHERNET NETWORKS

Ján TURÁN, Juraj FUTÓ
Department of Electronics and Multimedia Communication, Faculty of Electrical Engineering and Informatics,
Technical University of Košice, Park Komenského 13, 040 21 Košice, Slovak Republic, tel. 055/6335692,
E-mail: jan.turan@tuke.sk, futo@napri.sk

SUMMARY

In this paper stochastic representation of teleworker activities in Ethernet networks are analyzed. The emphasis is on properties of self-similarity description and effective bandwidth estimation using the Hurst parameter method. Experimental results are presented for chat and desktop sharing trace.

Keywords: *teleworking, Hurst parameter, long-range dependence, quality of service, resource allocation*

1. INTRODUCTION

The main objective in telecommunications network engineering is to have as many happy users as possible. In other words, the network engineer has to resolve the trade-off between communication resource capacity and QoS requirements. Accurate modeling of the offered traffic load is the first step in optimizing resource allocation algorithms such that provision of services complies with the QoS constraints [1].

Multimedia communication is enabled by wide variety of services. These ones are characterized by heterogenity and that's why the description of these by simply model is a complicated problem [2,3].

Teleworking – the act of working outside of the conventional workplace and communicating with coworkers by means of information and communications technologies will be a standard work arrangement for majority of all corporate employees in near future. The main problem of teleworking workplace planning is modelling and dimensioning of communication resources which are used by teleworkes. The other question is to measure effectiveness of teleworking introducing [4,5].

In the paper we are proposing stochastic model of teleworker activities which is based on new traffic models with self-similar characteristics. Than the effective bandwidth estimation is analyzed using the Hurst parameter found by variance-time plot method. Experimental results are provided for chat and desktop sharing trace.

2. BRIEF REVIEW OF LRD PROCESS

A self-similar phenomenon displays structural similarities across a wide range of timescales [6]. Traffic that is bursty on many or all timescales can be described statistically using the notion of self-similarity. Self-similarity is the property what is typical for the fractals, which are objets whose appearances are unchanged regardless of the scale at which are viewed [7,8,10] .

In the case of stochastic objects like time series, self-similarity is used in the distributed sense: when viewed at varying timescales, the objects relational structure remain unchanged.

Statistical analysis of high-resolution traffic measurements from a wide range of working packet networks (i.e. Ethernet LANs [7,10], compressed video streams [8], wide area TCP/IP traffic [11] and World Wide Web (WWW) [12]) traffic load have convincingly shown of fractal or self-similar properties in both local area and wide area traffic traces.

The fact that network traffic is inherently fractal or long-range dependent (LRD) poses a problem for many teletraffic related traffic engineering problems, e.g. traffic measurements and performance, buffer sizing, admision control and congestion control.

2.1. Properties of Self-Similarity

Let $X=\{X_t\}=(X_1,X_2,....)$ be a covariance stationary or wide sense sattionary stochastic process; that is, a process with constant mean $\mu = E\left[X_t\right]$ and let $x_t = X_t - \mu$, finite variance $\sigma^2 = E\left[(X_t-\mu)^2\right]$ and an autocorrelation function

$$r(k)=\frac{E[x_t x_{t+k}]}{Ex_t^2} \quad \text{for} \quad (k=0,1,2,....) \tag{2.1.1}$$

that depends only on *k*. In particular, we assume that X has an autocorrelation function of form

$$r(k) \sim a_1 k^{-\beta}\frac{1}{2} \quad \text{as} \quad k \to \infty \tag{2.1.2}$$

where $0<\beta<1$ and $(a_1,a_2,......)$ denote finite positive constants.

For each $m=1,2,3,.....$, let $X^{(m)}=(X_k^{(m)}:k=1,2,3,....)$ denote a new time series obtained of size *m*. That is, for each $m=1,2,3,.....$, $X^{(m)}$ is given by

$$X_k^{(m)} = \frac{1}{m}(X_{km-m+1} + ... + X_{km}),\ k \geq 1 \tag{2.1.3}$$

Note that for each m, the aggregated time series $X^{(m)}$ defines a covariance stationary process; let $r^{(m)}$ denote the corresponding autocorrelation function.

The process X is called exactly second-order self-similar with self-similarity parameter

$$H = 1 - \frac{\beta}{2} \tag{2.1.4}$$

if corresponding aggregated processess $X^{(m)}$ have the same correlation structure as X i.e.,

$$r^{(m)}(k) = r(k) \text{ for } m = 1, 2, ...$$

and

$$k = 1, 2, 3, ... \tag{2.1.5}$$

In other words, X is exactly self-similar if the aggregated processes $X^{(m)}$ are indistinguishable from X at least with respect to their second order statistical properties.

A covariance stationary process X is called asymptotically (second order) self-similar with self-similarity parameter H given by (2.1.4) if $r^{(m)}(k)$ agrees asymptotically i.e. for large m and large k with the correlation structure $r(k)$ of X given by (2.1.2).

A stationary stochastic process $x(t)$ is considered long-range dependent (LRD) if its autocovariance function decays as a power-law, slower than a negative exponential [7,8,13].

$$\gamma_x(k) \sim c\gamma |k|^{-(1-\alpha)} \tag{2.1.6}$$

Equivalently, LRD can also be defined in the frequency domain as the power-law divergence at the origin of the spectrum is

$$S_x(f) \sim \frac{c_f}{|f|^{\alpha}} \quad \text{when } |f| \to 0 \tag{2.1.7}$$

The LRD parameters are α and c_f. The scaling parameter α is related with the intensity of the LRD phenomenon (a qualitative measure) and is usually expressed as the Hurst parameter $H = \frac{(1+\alpha)}{2}$, while c_f has dimensions of variance and can be interpreted as a quantitative measure of LRD. Hurst parameter has received more attention, but c_f is not negligible, since it appears in the expression of loss probability when LRD traffic is fed into a queue [7], and also in the variance of the sample mean of an LRD process, determining the confidence intervals for its estimation.

2.2. Effective Bandwidth Definition

The notion of effective bandwidth provides a measure of resource requirements of a traffic stream with certain QoS constraint [5]. Statistical properties of the traffic stream have to be considered as well as system parameters (i.e., buffer size, service discipline) and the traffic mix. The terms equivalent bandwidth and equivalent capacity are often used as synonyms for effective bandwidth.

A mathematical framework for effective bandwidth has been defined based on the general expression

$$\alpha(s,t) = \frac{1}{st} \cdot \log E\left[\exp(s \cdot A_t)\right] \tag{2.2.1}$$

which depends on the space parameter s and the time parameter t [13].

When traffic is composed of flows from many independent sources, then, providing each source is heavy tailed [7,14], the resulting aggregation is self-similar [15]. Mathematical characterization of the effective bandwidth is provided in [16], required by a self-similar traffic source, as

$$f(H) = \frac{1}{(1-H)^{2(1-H)} H^{2H}} \tag{2.2.2}$$

where, m is the mean traffic arrival rate of the traffic stream, a is the variance coefficient of the traffic stream, H is the Hurst parameter in the range $0.5 \leq H \leq 1$, b is the buffer size and ε is the target loss rate for the traffic stream. The variance coefficient a is calculated as the variance to mean ratio of the traffic stream.

3. HURST PARAMETER ESTIMATION METHODS

In order to determine if a given series exhibits self-similarity, a method is needed to estimate H for a given series. Currently, there are several used approaches to doing that:

a./ Variance-Time Plot [7,17],

b./ Rescaled Range (R/S) [6,7,17],

c./ Periodogram [10],

d./ Whittle Estimator [12],

e./ Wavelet analysis [9].

Explanation of these methods may be found in the references [6,7,9,10,12,17,19,20]. In our experiment we have used Variance-Time Plot method [7,17].

Parameters :	Chat session	Desktop sharing
Number of packets	8775118	839177
Number of bytes	104513	66970690
Time in seconds	1809	3692

Tab. 1 The Traces Parameters

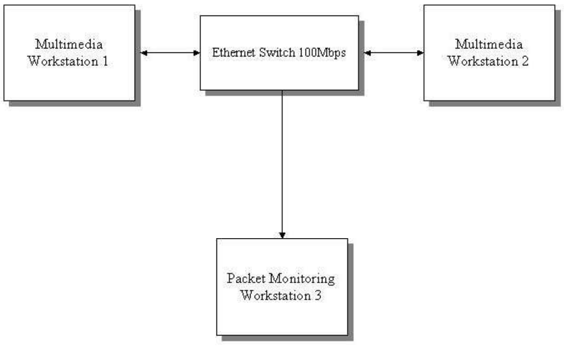

Fig. 1 Experiment arragement

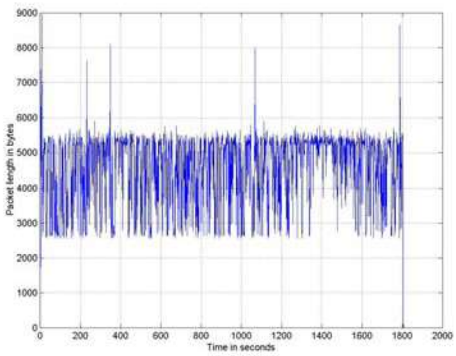

Fig. 2 Chat trace

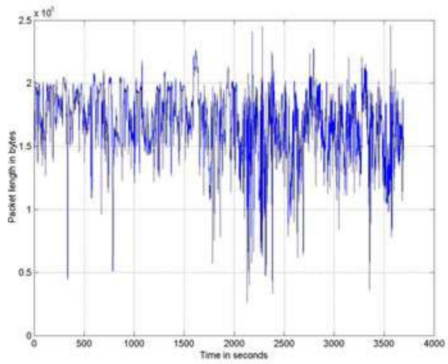

Fig. 3 Desktop sharing trace

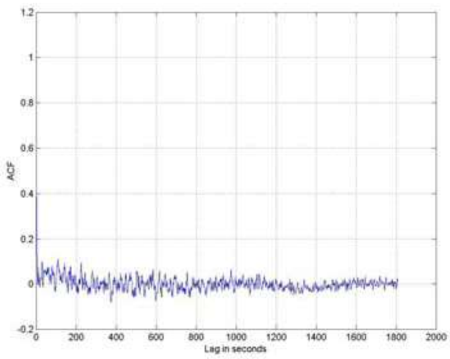

Fig. 4 Chat trace autocorrelation

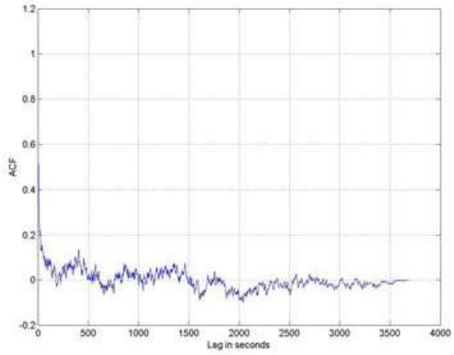

Fig. 5 Desktop sharing trace autocorrelation

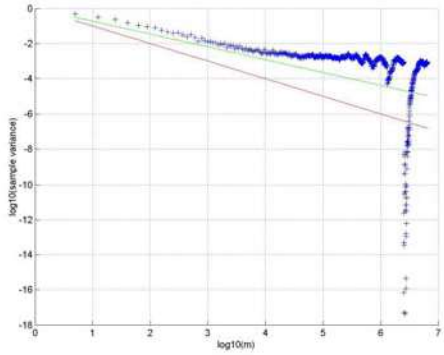

Fig. 6 Variance-Time Plot method for H=0.6353

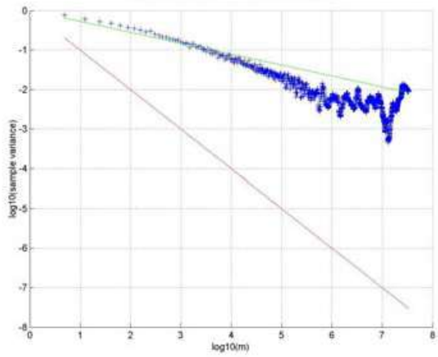

Fig. 7 Variance-Time Plot method for H=0.86224

4. EXPERIMENTAL RESULTS

The experiment is based by measurement of packet arrivals and lengths by software based packet catching program WinPcap [18]. The traffic measurements was collected on experimental network which is represented by LAN Ethernet network created by three PC workstations and network switch displayed on Fig. 1. Our goal was to collect packet traces generated by Windows NetMeeting 3 sessions. This application is often used for remote user management tasks and it is useful mainly for remote user and system administrator communication. For teleworkers it can viewed as main remote management tool.

First trace is generated by half hour chat session, second one is generated by desktop sharing. The traces parameters are described in Table 1 and on Fig. 2 and Fig. 3 are displayed the both traces. Autocorrelation are displayed on Fig. 4, Fig. 5.

Variance-time plot is aplicated for both traces and the results are displayed on Fig. 6 and Fig. 7.

5. CONCLUSIONS

The Hurst parameter is used to measure the user teleworker activity in Ethernet network. Two different traces are used for showing the relation between user activities and Hurst parameter. We can see that such simply application as chat is generating packet arrivals which cannot be modelled by simply model such as Poisson process. Desktop sharing is worse in comparison with chat, because number and volume of transmitted bytes is higher. Netmeeting connection can be an impact for communication network. Network administrator can use effective bandwidth computation for such sessions planning and dimensioning. From the experiments results that the Hurst parameter estimation is playing key role at stochastic process specification. In future it will be important to clasify typical network applications in relation to Hurst parameter. In such way it will be possible to prevent the network congestion.

ACKNOWLEDGEMETS

The paper is partially supported by grant COST 276, COST 292 and VEGA grant No.1/0381/03.

REFERENCES

[1] The ATM Forum, Technical Committee: Traffic Management Specification, Version 4.1, AF-TM-0121.000, March 1999

[2] Rao, K.R. - Bojsovic, Z. - Milanovic, D.: Multimedia Communication Systems. Prentice Hall, New York, USA, 2002.

[3] Rao, K.R. - Bojsovic, Z. - Milanovic, D.: Introduction to Multimedia Communication: Application, Middleware and Networking. J.Wiley and Sons, New York, USA, 2004.

[4] Mandal, K.M.: Multimedia Signals and Systems. Kluwer Acad. Publishers, New York, USA, 2003.

[5] Clao, H.J. - Guo, X.: Quality of Service Control in High-Speed, Networks. J. Wiley and Sons, New York, USA, 2002.

[6] Arlt, J. - Arltová, M.: Finanční časové rady, Grada Publishing, Praha, 2003.

[7] Leland, W.E. - Taqqu, M.S. - Willinger W. - Wilson D.V.: On the Self-Similar Nature of Ethernet Traffic, www.argreenhouse.com/papers/wel/sigcomm93.pdf

[8] Rose, O.: Estimation of the Hurst Parameter of Long-Range Dependent Time Series, long-memory.martinsewell.com/rs-hurst/Rose96.pdf

[9] Abry, P. - Veitch, D.: Wavelet Analysis of Long-Range-Dependent Traffic, IEEE Transaction Information Theory,1998, 365-369.

[10] Leland, W. - Taqqu, M. - Willinger, W. - Wilson, D.: On the Self-Similar Nature of Ethernet Traffic (Extended Version), IEEE/ACM Transactions on Networking, No.2, 1994, 1-15.

[11] Paxson, V. - Floyd, S.: Wide Area Traffic: The Failure of Poisson Modelling, In Proc. ACM SIGCOMM'94, 1994, 257-268.

[12] Crovella, M.E. - Bestavros, A.: Explaining World Wide Web Traffic Self Similarity, October 1995, http://www.cs.bu.edu/faculty/crovella/paper-archive/self-sim/paper.html

[13] Norros, I.: A storage model with self-similar input, Queuing Systems, Vol.16, 1994, 387-396.

[14] Norros, I.: On the use of fractional Brownian motion in the theory of connectionless networks, http://dl.comsoc.org/cocoon/cpmsoc/servlets/GetPublication;jsessionid=EB03CCB61645CD65561625F3748D5054?id=1227032

[15] Tsybakov, B. - Georganas N.D.: On Self-Similar Traffic in ATM Queues: Definitions, Overflow Probability Bound and Cell Delay Distribution, IEEE/ACM Transactions on Networking, 5:397-409

[16] Patel, A.A. - Williamson, C. L.: Effective Bandwidth of self-similar Traffic Sources: Theoretical and Simulation Results, http://citeseer.ist.psu.edu/patel97effective.html

[17] Leland, W.E. - Taqqu, M.S. - Willinger W. - Wilson D.V.: On the Self-Similar Nature of Ethernet Traffic, In Proc. ACM SIGCOMM '93, 1993 183-193

[18] Degionanni, L. - Baldi, M. - Risso, F.- Varenni, G.: Profiling and Optimization of Software-Based Network-Analysis Applications

[19] Karagiannis, T. - Faloutsos, M. - Riedi, R.H.: Long-Range Dependence: Now you see it, now you don't, http://cmc.rice.edu/docs/docs/Kar2002Nov5Long-Rang.pdf

[20] Ghaderi, M.: On the Relevance of Self Similarity in Network Traffic Prediction, http://www.cs.uwaterloo.ca/cs-archive/CS-2003/28/TR-CS-2003-28.pdf

BIOGRAPHY

Ján Turán (Prof., Ing., RNDr., DrSc.) was born in Šahy, Slovakia. He received Ing. (MSc.) degree in physical engineering with honors from the CTU Prague, Czech Republic, in 1974, and RNDr. (MSc.) degree in experimental physics with honours from the UK Prague, Czech Republic, in 1980. He received a CSc. (PhD.) and DrSc. Degrees in radioelectronics from the TU Košice, in 1983, and 1992, respectively. Since March 1979, he has been at the TU Košice as Professor for electronics and information technology. His research interest include dgital signal processing and fiber optics, communication and sensing.

Juraj Futó (Ing.) was born in Košice, Slovakia. He received his Ing. (MSc.) degree in 1984 from the TU Košice. His research interest include digital signal processing and multimedia communication.