POZNAN UNIVERSITY OF TECHNOLOGY ACADEMIC JOURNALS
No 70 Electrical Engineering 2012

Przemysław WŁODARSKI*

APPROXIMATED SELF-SIMILAR PROCESS FOR GENERATION OF ICT TRAFFIC

Self-similar nature of ICT traffic causes big impact on queuing performance needed in providing a certain level of Quality of Service. Therefore, synthesis of fractal process for simulation and analysis purposes should be fast and accurate. In this article it is proposed how to simplify distribution of interarrivals in the renewal process in order to implement the inverse distribution technique to generate random variables. Proposed method improves simulation performance by reducing computational complexity leaving the accuracy almost unchanged.

1. INTRODUCTION

Performance of ICT (Information and Communications Technologies) systems depends not only on the throughput but on network traffic structure which is quite complex [1], [5], [7], [16]. Since throughput involves higher costs and not always determines good performance, especially for high variable ICT traffic, one should focus more on statistical structure of the process of incoming and outgoing packets. Studies on network traffic modeling revealed the presence of phenomenon called self-similarity and long-range dependence (LRD) [2], [3], [10], [15]. [11], [14]. Self-similar nature of ICT traffic involves significant correlations across large time scales and causes big impact on queuing performance needed in achieving a certain level of Quality of Service [5], [6], [7], [8], [9]. Thus, traditional models such as Poisson or ARMA processes, that can capture only short-range dependence (SRD) cannot reflect real properties of the traffic. In this article we present the method of generating approximated self-similar random sequences with application to network traffic modeling. Our approach is based on a simplification of distribution function of inter-renewal times and makes it possible to directly apply an inverse distribution method instead of piecewise-linear approximation.

2. SELF-SIMILARITY

Basic properties of self-similarity can be shown on the basis of fractional Brownian motion (fBm) which is the theoretic exactly self-similar process with stationary increments [5], [15]. The main properties are:

* West Pomeranian University of Technology.

- invariance in distribution – statistical self-similarity

$$Y(at) = a^H Y(t), \quad t, a \in R, \; a > 0 \tag{1}$$

- has the stationary Gaussian increments with the density distribution ($\mu = 0$):

$$f_G(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right) \tag{2}$$

- has the covariance function defined as:

$$Cov(Y(s), Y(t)) = K_H(s,t) = \frac{\sigma^2}{2}\left(|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right) \quad t, s \in R \tag{3}$$

- for $\sigma > 0$ and $0.5 \le H < 1$

$$E\left(Y^2(t)\right) = \sigma^2 |t|^{2H} \tag{4}$$

- the increment sequence $X_k = Y(k+1) - Y(k)$, $k \in N$ is called fractional Gausian noise (fGn) and has the following properties:

$\{X_k\}$ is stationary

$E(X_k) = 0$

$E\left(X_k^2\right) = \sigma^2 = E\left(Y(1)^2\right)$

the autocovariance function of the process $\{X_k\}$ is given by

$$r(k) = \frac{\sigma^2}{2}\left[|k+1|^{2H} - 2|k|^{2H} + |k-1|^{2H}\right] = \frac{\sigma^2}{2}\delta^2\left(|k|^{2H}\right), \quad k \in N \tag{5}$$

where δ^2 denotes the second difference operator
for $k \ne 0$: $r(k) = 0$ if $H = 0.5$ and $r(k) > 0$ if $0.5 < H < 1$

The index H is called the *Hurst exponent* and is the measure of the degree of self-similarity ($0.5 \le H < 1$). When $H = 0.5$, the fBm is the usual Brownian motion with $Cov(Y(s), Y(t)) = K_{0.5}(s,t) = \min(s,t)$ and $r(k) = 0$ for $k \ge 1$.

3. PROPERTIES OF SOME RENEWAL PROCESS

Renewal process which is a generalization of Poisson process is broadly used for describing randomness in many systems where events occur at some points of time. The variant of this process called spatial renewal process which has some interesting properties was introduced by Jelenković, Lazar and Semret [4] and

developed in [14] by Taralp Davetsikiotis and Lambadaris. In order to generate single process, we should consider two different random sequences:
- inter-renewal time sequence $\{T_i\}$, $i = 0,1,\ldots$ with distribution function $F_T(t)$ which determines times among "jumps" of the second random sequence
- independent random sequence $\{X_i\}$, $i = 0,1,\ldots$ with distribution function $F_X(x)$

The process takes at $t_i = \sum_{m=0}^{i} T_m$ values from $\{X_i\}$ for time intervals $\{T_i\}$: $Y(t = t_i) = X_i$. Distribution function $F_T(t)$ is associated with autocovariance function $\rho(t)$ in the following way:

$$1 - \rho(t) = \frac{1}{\mu}\int_0^t (1 - F_T(u))du \tag{6}$$

$$\text{where } \mu = \int_0^\infty (1 - F_T(u))du \text{ and } t \geq 0$$

Let $\rho(t,H)$ stands for an autocovariance function for self-similar process which depends on Hurst exponent H (eq. 5). After differentiation of (6), for a discrete form we get:

$$F_T(n,H) = 1 - \mu[\rho(n,H) - \rho(n+1,H)] \tag{7}$$

The constant μ can be obtained from the condition $F_T(0,H) = 0$:

$$\mu = \frac{1}{\sigma^2 - \rho(1,H)} = \frac{1}{\sigma^2\left(2 - 2^{2H-1}\right)} \tag{8}$$

Thus the distribution function of interarrivals has the following form:

$$F_T(t,H) = \begin{cases} 0 \text{ for } 0 \leq t < 1 \\ 1 + \mu \frac{d}{dt}\rho(t,H) = 1 + \mu\sigma^2\Big[H(t-1)^{2H-1} + \\ \qquad - 2Ht^{2H-1} + H(t+1)^{2H-1}\Big] \text{ for } t \geq 1 \end{cases} \tag{9}$$

Authors of [14] propose to approximate equation (9) by piecewise linear function. It is not easy task because of shape of the distribution (fig. **2**) – one needs to use many of segments of different length in order to obtain good approximation, especially near the bend. Furthermore, generated random sequence is very sensitive to the number and location of the points for segments.

4. PROPOSED APPROXIMATION

Instead of piecewise linear approximation of the distribution function $F_T(t,H)$, one proposes a simplification that makes it possible to directly apply an inverse distribution technique in order to generate random variable. Expanding autocovariance function using Taylor series for $t \geq 1$ and taking into consideration terms up to second derivation for $f(x) = x^{2H}$, we obtain:

$$\hat{\rho}(t,H) = \sigma^2 t^{2H-2}(2H-1)H \tag{10}$$

Based on the result in (10), the inter-renewal time distribution has the following form:

$$\hat{F}_T(t,H) = \begin{cases} 0 \text{ for } 0 \leq t < 1 \\ 1 + \dfrac{(2H-2)(2H-1)H}{2-2^{2H-1}} t^{2H-3} \text{ for } t \geq 1 \end{cases} \tag{11}$$

Now, the inverse distribution method can be directly applied in order to generate the random sequence $\{T_i\}$:

$$T_i = \left(\frac{-R \cdot \left(2 - 2^{2H-1}\right)}{(2H-2)(2H-1)H} \right)^{\frac{1}{2H-3}}, \quad i = 0,1,\ldots \tag{12}$$

where R is an independent random variable uniformly distributed between 0 and 1.

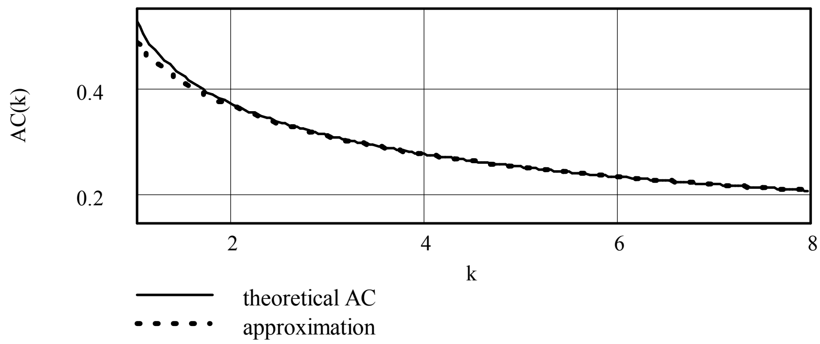

Fig. 1. Theoretical (eq. 5) and approximated (eq. 10) autocovariance

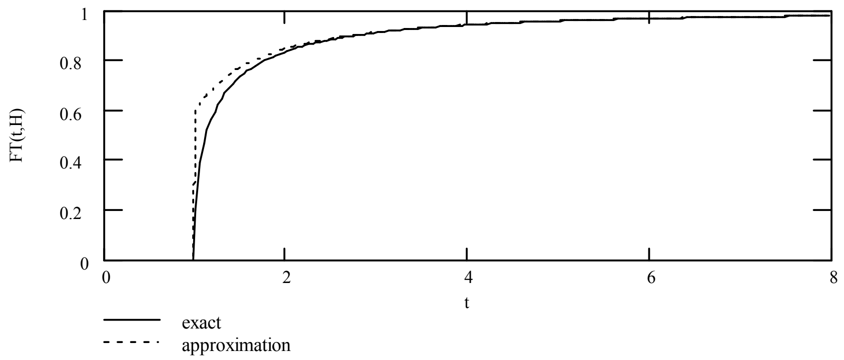

Fig. 2. Exact (eq. 9) and approximated (eq. 11) distribution of inter-renewal times for H=0.8

Compatibility of the autocovariance for self-similar sequences with the autocovariance for inter-renewal times is preserved only if $\{X_i\}$ have normal distribution (fig. 4). For other distributions there is a big discrepancy from the desired autocovariance. Moreover, in order to get a better approximation of Gauss distribution, authors of [14] propose to aggregate single renewal processes (≈ 10), but it involves long running times as well as worse approximation of autocovariance function. The value of Hurst exponent for this case is smaller than expected (0.8) – the mean value for 10 aggregated time series (each consists of 10 single sequences) is $\widetilde{H} = 0.708$, and for 10 single sequences, the mean value is $\widetilde{H} = 0.801$. In figure 4, there is the mean empirical autocovariance function for described method based on proposed approximation.

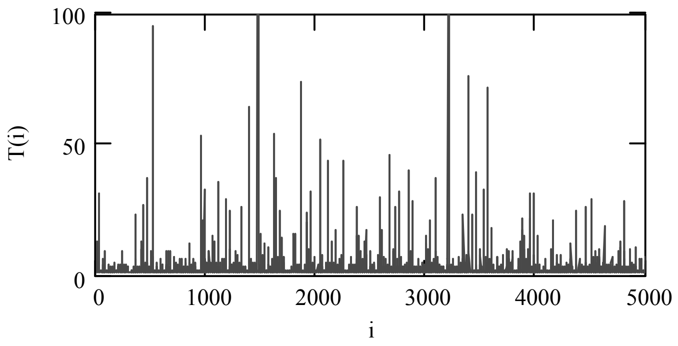

Fig. 3. An example of inter-renewal times samples generated using eq. 12

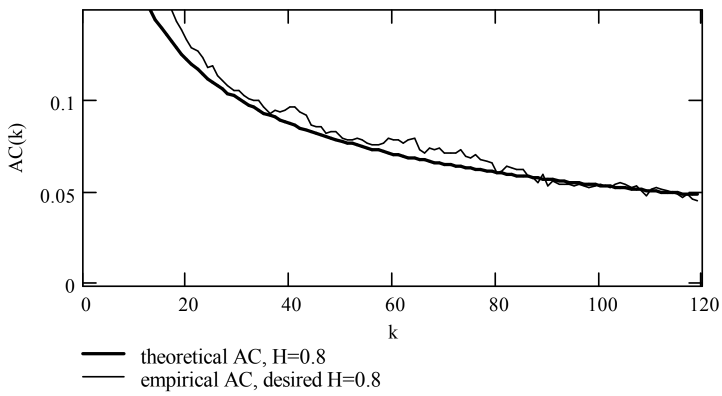

Fig. 4. An empirical mean autocovariance (AC) for 10 random sequences of length 16384, H=0.8, normal distribution N(0,1)

5. CONCLUSIONS

The method based on spatial renewal processes can be used to generate random sequences with normal distribution and any Hurst exponent from the range $0.5 \le H < 1$. The problem is to generate samples according to distribution function proposed in [14]. Instead of piecewise linear approximation it is proposed to substitute the autocovariance function of self-similar process by its approximation, which leads to a significant simplification of the method and stability of results. Furthermore, this simplification involves the possibility of using direct inverse distribution method (eq. 12) for generating random inter-renewal times, that makes the method simple and easy to implement. It may be used for example in network traffic generators as well as for analysis purposes of improving quality of service performance in ICT networks.

REFERENCES

[1] Elbiaze, H., et al., ”A new structure-preserving method of sampling for predicting self-similar traffic”. Telecommunication Systems 43 265-277, 2010.

[2] Field, A.J., Harder, U., Harrison, P.G. *Measurement and Modelling of Self-similar Traffic in Computer Networks*. IEE Proc-Commun., 151, 4 (Aug 2004), 355-386.

[3] He, G., Gao, Y., Hou, J., Park., K. *A case for exploiting self-similarity of network traffic in TCP congestion control*, Computer Networks, 45, 6 (Aug 2004), 743-766.

[4] Jelenkovic, P. R., Lazar, A. A., Semret, M., *Multiple time scales and subexponentiality in MPEG video streams*. Proc. of IFIP-IEEE, Conference on Broadband Communications, 1996, 64-75.

[5] Leland, W., Taqqu, M., Willinger, W., Wilson, D., *On the self-similar nature of Ethernet traffic (extended version)*. IEEE/ACM Transactions on Networking, 2, 1 (Feb 1994) 1-15.

[6] Lelarge, M., Liu, Z., Xia, C.H. *Asymptotic Tail Distribution of End-to-End Delay in Networks of Queues with Self-Similar Cross Traffic*. IEEE INFOCOM 2004.

[7] Lin G. C., Suda, T., *On the impact of long-range dependent traffic in dimensioning ATM network buffer*. Proc. Of IEEE INFOCOM, 1998.

[8] Medhi, J., *Stochastic models in queueing theory*. Academic Press, 2003.

[9] Norros, I., *A storage model with self-similar input*. Queueing Systems, 16, 1994, 387-396.

[10] Norros, I., *On the use of fractional Brownian motion in the theory of connectionless networks*. IEEE Journal on Selected Areas in Communications, 13 (6), 1995.

[11] Parulekar, M., Makowski A. M., *Tail probabilities for a multiplexer with self-similar traffic*. Proc. IEEE INFOCOM, 1996.

[12] Park, K., Kim G,. T., Crovella, M. E., *On the relationship between file sizes, transport protocols and self-similar network traffic*. Technical report, Boston University, Computer Science Department, 1996.

[13] Paxson, V. Fast, *Approximate synthesis of fractional Gaussian noise for generating self-similar network traffic*. ACM SIGCOMM Computer Communication Review 27, 5 (Oct 1997), 5-18.

[14] Taralp, T., Devetsikiotis, M., Lambadaris, I. *Efficient fractional Gaussian noise generation using the spatial renewal process*. Proceedings of the IEEE International Conference on Communications, Atlanta, GA, vol.3 (Jun 1998), 1456-1460.

[15] Taqqu, M., Willinger, W., Scherman, R. *Proof of a Fundamental Result in Self-Similar Traffic Modeling*. ACM SIGCOMM Computer Communication Review, 27, 2 (Apr 1997), 5-23.

[16] Yunhua, R. *Evaluation and estimation of second-order self-similar network traffic*. Computer Communications 27 (2004) 898-904.