Comparison of Models for Self-Similar Network Traffic Generation

Igor Tomić, Nebojša Maletić
Faculty of Electrical Engineering
Banja Luka, Bosnia and Herzegovina
igor.tomic@etfbl.net, nebojsa.maletic@etfbl.net

***Abstract*—This paper gives an overview of five different models for self-similar network traffic generation (iterated chaotic maps, fractional Gaussian noise model, Pareto model and finite and infinite Markov chain model). The models are compared on the basis of Hurst parameter and mean value of generated sequences, and also on the basis of algorithm efficiency. R/S plot, Variance-Time plot and Periodogram method are used for the Hurst parameter estimation. According to simulation results, the model which gives sequences whose parameters are close enough to the given ones is fronted.**

Keywords-Hurst parameter; long-range dependence; self-similarity; traffic model

I. Introduction

In modern packet-oriented networks, telecommunication traffic is considered to be self-similar process. Unlike the Poisson process (used for classic telephone traffic modelling), which get smooth when averaged over large timescales, self-similar process retains its burstiness over a wide range of time scales [1].

There are several factors contributing to self-similarity of network traffic, such as link bandwidth [2], file size distribution, reliability and flow control mechanisms in the transport layer [3], VBR (Variable Bit Rate) video streaming [4] and others, so that the degree of traffic burstiness differs between network segments.

A number of papers have studied the impact of self-similarity on network performance in terms of packet loss rate, queueing delay or throughput [2], [5], [6]. Synthesized self-similar traffic is important in evaluating the performance of various switch architectures under realistic conditions. A good traffic model may lead to a better designing routers and network devices which handle long packet bursts. According to differences in the traffic statistic within different parts of the network, the generator is recquired to be flexible enough.

A lot of self-similar traffic generator models are developed, but none of them simulates all aspects of real traffic. This paper provides comparison of some of the most referred models.

II. Parameters of Self-Similar Processes

In this section, an overview of basic concepts and terms of self-similar network traffic is given.

As a measure of self-similarity the *Hurst* parameter (H) is used, where higher H ($H \in (1/2,1)$) implies higher self-similarity.

The packet traffic trace can be represented by binary sequence, where one stands for a packet and zero for an interpacket gap. If the sequence x is divided into non-overlapping adjacent blocks of size m and then blocks are averaged, m-th order aggregation of the sequence is given as

$$x_t^{(m)} = \frac{1}{m} \sum_{i=tm-m+1}^{tm} x_i \,. \qquad (1)$$

Two traffic traces with the aggregation scale m=1000 for different values of H are shown in Fig. 1. It can be noted that process with lower H becomes smooth when aggregated over a large scale.

A stochastic process is considered to be self-similar with parameter β, 0<β<1, if, for all m=1, 2, ..., the following applies:

$$Var\left(x^{(m)}\right) \sim \frac{Var(x)}{m^{\beta}} \,. \qquad (2)$$

The parameter β is related to H by H=1–β/2. The variance of the process with higher H decays more slowly with the increase of aggregation degree.

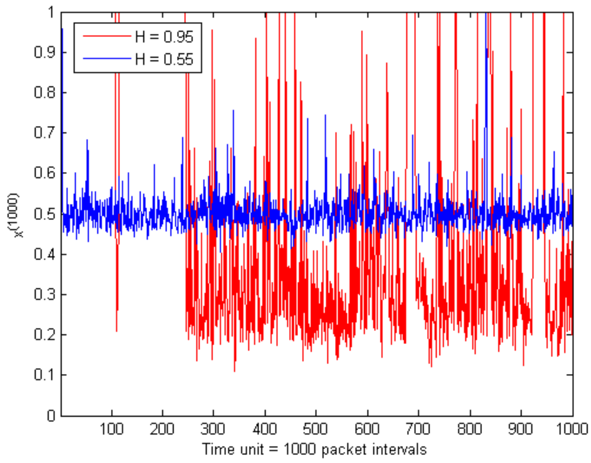

Fig. 1. Traces with the aggregation scale m=1000 for different values of H.

If distant time samples of the process are correlated, it is said to be LRD (*Long-Range Dependence*) process. The LRD process has a hyperbolically decaying autocovariance function:

$$C(\tau) \sim |\tau|^{-\beta}, \text{ as } |\tau| \to \infty, 0 < \beta < 1. \tag{3}$$

The important characteristic of the self-similar process is that packet train and interpacket gap duration have a heavy-tailed distribution. One of the most frequently used heavy tailed distributions is the Pareto with cumulative distribution function (CDF) given by:

$$F(x) = 1 - \left(\frac{\beta}{x}\right)^{\alpha}, (x \geq \beta, \alpha > 0). \tag{4}$$

Parameter α, 1<α<2, denotes the tail index.

III. Self-Similar Traffic Generation

In this section, five models for self-similar traffic generation, frequently found in literature, are presented.

A. *Iterated Chaotic Maps*

The iterated chaotic map (ICM) model is described by traffic load $d \in (0, 1)$ and parameters m_1, $m_2 \in (1.5, 2)$, which determine the slope of curve plotted in Fig. 2.

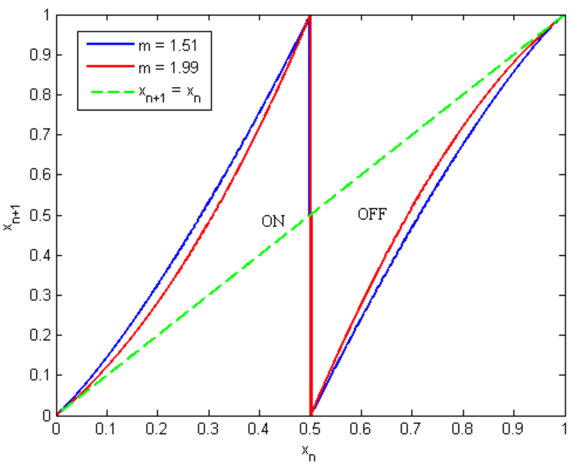

Fig. 2. Graph of an iterated chaotic map for *d*=0.5.

Based on initial value $x_0 \in (0, 1)$, the next one is given by:

$$x_{n+1} = \begin{cases} x_n + \dfrac{1-d}{d^{m_1}} x_n^{m_1}, & 0 < x_n < d \\ x_n - \dfrac{d}{(1-d)^{m_2}} (1-x_n)^{m_2}, & d < x_n < 1 \end{cases}. \tag{5}$$

The sequence x takes values between 0 and 1. The binary time series $\{y_n: n \in N\}$ is generated by applying the following rule: $y_i = 1$ if $x_i < d$ and $y_i = 0$, otherwise. If $m_1 = m_2 = m$, the relation with H is given by $H = (3m-4)/(2m-2)$. Thus, parameters d and H describe the model.

When close to the mean value d, short sequences of zeros and ones are generated. Long sequences are generated when x_n is found near points zero or one, by which traffic burstiness is obtained. A detailed explanation about the model can be found in [7].

B. *Pareto Traffic Generation*

The starting model hypothesis is that the time of packet train and interpacket gaps have a Pareto distribution (Fig. 3).

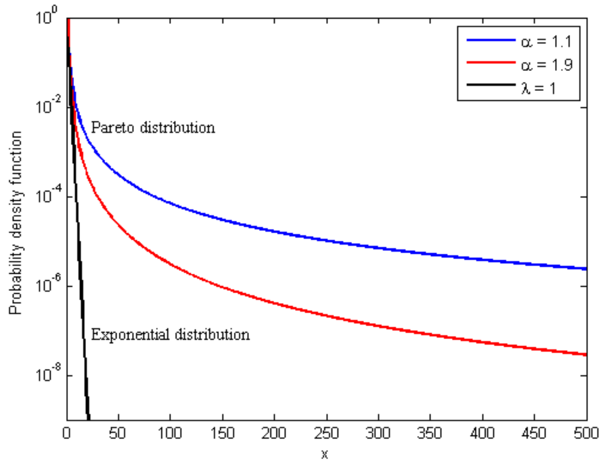

Fig. 3. Pareto vs exponential density distribution function.

Two random variables are required to generate binary sequences. One variable stands for packet sequence duration, and the other for interpacket gap duration measured in time intervals equal to packet duration. As $x \geq \beta$ (eq. 4), parameter β presents the shortest length of the packet train. Tail index, α (1<α<2), defines how fast CDF decays. As α →1, the rate of decay is low, indicating that appearance of longer bursts is more probable. Poisson distribution (used for classic voice traffic modelling) decays much faster (exponential) compared to Pareto distribution.

The parameter H depends on both parameter value, α and β. Larger β, as well as α closer to unity, causes longer bursts. Anyway, it was showed that the effect of β is not so influential on the degree of self-similarity as α is. The relation with H is $H=(3-\alpha)/2$. Details of the algorithm can be found in [6].

C. *FGN Traffic Generation*

The Fractional Gaussian noise (FGN) traffic generator produces the streams of variable-length packets with self-similar statistics, unlike Pareto generator, which gives fixed length packets [6].

The model is based on features of LRD process in the frequency domain. The first step in synthesis of self-similar sequence is the estimation of FGN power spectrum with given Hurst parameter (Fig. 4). The Inverse Fourier Transform (IFFT) is used to obtain the time trace which looks like noise. It is necessary to modify time trace to obtain the desired variance and mean value. Every sample then represents a flow in number of packets per time unit.

To obtain a sequence of packets of variable lengths, it is necessary to know pakcet distribution in the observed network segment. More details on the model can be found in [6], [8], [9].

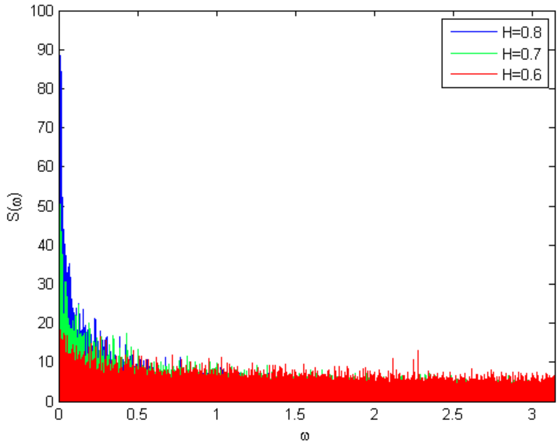

Fig. 4. The power spectrum of FGN process.

This method requires that the size of sequence which needs to be generated to be defined.

D. Markov Chain Based Models

An infinite Markov chain (IMC) can be used to generate a time series exhibiting LRD (see Fig. 5). If the chain is in the state i, $i \neq 0$, it moves to state $i-1$ in next time instant generating one in the output sequence. Only in the state $i=0$ (zero state), zero is generated in the output sequence. Any state, i, can be reached from the zero state with the transition probability f_i. In brief, if the chain changes its state from 0 into n ($n \neq 0$), then the sequence of n ones will be generates. The sequence of zeros will be generated if the chain remains in the zero state.

In practical implementation, IMC does not have an infinitive number of states. It is specified to an maximum state $N_{\max}$. For finite Markov chain (FMC) model, the number of states, N, is fixed. It is also the maximum packet burst size. Since burstiness is the condition for the self-similarity, the sequence with large H cannot be obtained with low N.

In infinite Markov chain realization, the number of states is determined dynamically. The initial number of states, N, a relatively small number, is specified. If the probability f_i, generated in the zero state, does not correspond to any state in range from 0 to N, then states from N+1 to 2N are generated and the state corresponding to given probability is looked for. The procedure is repeated until the maximum state, N_{max}, is reached. Therefore, this chain is also constrained by N_{max}, but N_{max} could be very large number.

Details about these models can be found in [7],[10].

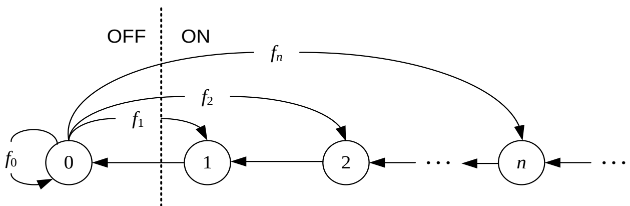

Fig. 5. An infinite Markov chain.

IV. Hurst Parameter Estimation Techniques

The properties of self-similar processes (see Section II) lead to the different methods to estimate H.

The R/S statistic is a method that estimates the parameter H as a slope of the straight line in plot of log(R/S) versus log(T), where R is a measure of the range of the process, S is the standard deviation of the sample and T is the time interval.

The Variance-Time plot (VTP) method estimates the Hurst parameter on the basis of the variance of the aggregated time series. Considering eq. (2), a plot of log[Var($x^{(m)}$)] versus log(m) will yield a straight line with slope of $-\beta$. The Hurst parameter is related to β by $H=1-\beta/2$.

The Periodogram plot is a frequency domain technique, unlike the previous two methods which are the time domain estimators. By plotting log(S_n) versus log(ω), where S_n denotes the power spectrum density of n-length time-sequence block, and ω is the frequency, the points of the periodogram scattered around a negative slope are obtained. An estimate of the Hurst parameter is given by $H=(1-\gamma)/2$, where γ is the slope.

More about these techniques can be found in [10].

V. Simulation Results

For traffic generator models presented in the Section III, H and traffic load values of generated sequences are estimated by simulation analysis.

For ICM model, parameters m_1 and m_2 are equal. The shortest packet burst length for Pareto generator is $\beta=1$. For FMC model, different chain lengths are specified for various H values in order to attain the appropriate burstiness [10]. Chain sizes are $N=2^{13}$, 2^{14}, 2^{15}, 2^{16}, 2^{17}, for H=0.55, 0.65, 0.75, 0.85 and 0.95, respectively. For IMC model, the initial chain length is N=256, and maximum one is $N_{\max}$=65636 for H<0.8 and $N=2^{10}$, $N_{\max}=2^{20}$ for H>0.8.

The R/S plot, Variance-Time plot and Periodogram plot are used to estimate H from simulated sequences.

Each model, except the FGN, provides the binary sequence of ten million bits. The parameter H is estimated for the sequence aggregation of degree m=100. The FGN model generates a sequence of 100 000 points. The mean value H_{avg} and standard deviation σ_H are calculated for the Hurst parameter value estimated for 50 sequences. Traffic load is equal to ρ=0.5. The results are shown in Table I.

Based on the results, it is noticed that R/S method provides small standard deviation with lower H values when compared to other two methods. For higher H values, R/S method provides much higher standard deviation.

For H=0.55 and H=0.65, the FGN model generates series with H close enough to specified one. For higher values of H, the FMC model provides better results than FGN in terms of estimated H. Though, it should be bared in mind that the FGN has the lowest standard deviation for all specified H values. For ICM and Pareto models, the difference between specified and estimated H is large. Also, the variance of H is larger compared to other models.

TABLE I. ESTIMATION OF THE HURST PARAMETER.

Generator model	H	R/S $H_{avg} \pm \sigma_H$	Variance-Time $H_{avg} \pm \sigma_H$	Periodogram $H_{avg} \pm \sigma_H$
Iterated chaotic map	0.55	0.570 ± 0.025	0.567 ± 0.040	0.579 ± 0.038
	0.65	0.611 ± 0.037	0.634 ± 0.072	0.642 ± 0.043
	0.75	0.684 ± 0.114	0.729 ± 0.077	0.727 ± 0.033
	0.85	0.751 ± 0.117	0.805 ± 0.046	0.810 ± 0.024
	0.95	0.893 ± 0.192	0.890 ± 0.037	0.900 ± 0.017
Fractional Gaussian noise	0.55	0.561 ± 0.016	0.548 ± 0.010	0.538 ± 0.011
	0.65	0.655 ± 0.019	0.646 ± 0.011	0.629 ± 0.012
	0.75	0.743 ± 0.018	0.737 ± 0.010	0.723 ± 0.012
	0.85	0.824 ± 0.017	0.821 ± 0.013	0.811 ± 0.011
	0.95	0.895 ± 0.020	0.889 ± 0.011	0.899 ± 0.011
Pareto distribution	0.55	0.575 ± 0.019	0.561 ± 0.064	0.569 ± 0.049
	0.65	0.613 ± 0.044	0.626 ± 0.069	0.652 ± 0.053
	0.75	0.677 ± 0.060	0.712 ± 0.082	0.731 ± 0.036
	0.85	0.768 ± 0.089	0.809 ± 0.059	0.814 ± 0.029
	0.95	0.903 ± 0.252	0.894 ± 0.033	0.906 ± 0.017
Finite Markov chain	0.55	0.586 ± 0.023	0.577 ± 0.029	0.598 ± 0.037
	0.65	0.634 ± 0.043	0.644 ± 0.045	0.678 ± 0.051
	0.75	0.724 ± 0.058	0.753 ± 0.037	0.775 ± 0.036
	0.85	0.847 ± 0.081	0.836 ± 0.030	0.824 ± 0.025
	0.95	0.946 ± 0.068	0.922 ± 0.008	0.925 ± 0.018
Infinite Markov chain	0.55	0.569 ± 0.017	0.547 ± 0.032	0.559 ± 0.039
	0.65	0.616 ± 0.037	0.626 ± 0.064	0.652 ± 0.056
	0.75	0.664 ± 0.055	0.687 ± 0.060	0.712 ± 0.037
	0.85	0.750 ± 0.085	0.785 ± 0.048	0.797 ± 0.027
	0.95	0.808 ± 0.099	0.847 ± 0.030	0.872 ± 0.021

Results for IMC model are much lower than expected for H>0.7. None of the observed models can reach very high Hurst parameter values as H=0.95.

Performances of network devices are usually analysed under high traffic loads conditions. Also, higher values of H parameter are of interest. Fig. 6. plots traffic load ρ for three H parameter values. Resultant load is calculated as average value of loads estimated for 50 sequences.

FGN and FMC models provide traffic load which is close to the desired one. For other models, deviations from the specified value are much higher, especially for large traffic load and H parameter value.The IMC model is invalid for some combinations of ρ and H [10]. Values for these combinations are labeled in Fig. 6. with "x".

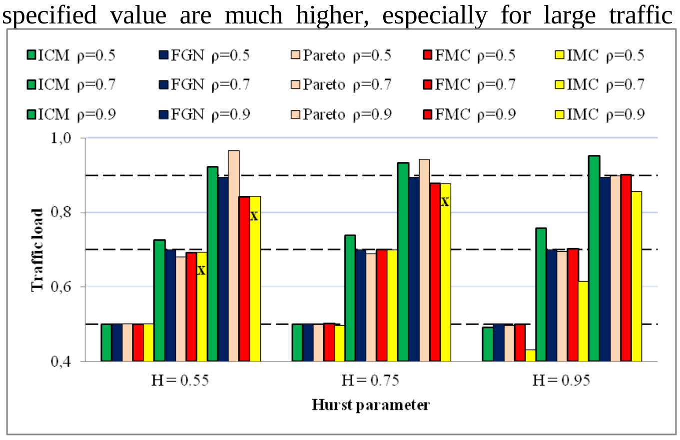

Fig. 6. Traffic load for different values of Hurst parameter.

All models are implemented in Matlab. To generate the binary sequence 10 million bits long, it took in average 0.3 seconds for FMC, 0.87 seconds for FGN, 3.01 seconds for ICM, 3.92 seconds for Pareto model and 4.45 seconds for IMC. The execution time of algorithms depends on values of parameters H and ρ, and other parameters, such as size of a Markov chain. Given values are determined on 750 runs with various combinations of H i ρ as in Table I.

VI. CONCLUSION

Considering all results, among analysed self-similar traffic generator models, according to performances referring to Hurst parameter and traffic load of generated series, and also to algorithm efficiency, FGN model and the one with finite Markov chain are emphasized. The advantage of FGN model is lower variance of the Hurst parameter and traffic load which is closer to expected. The difference between desired and obtained H is smaller when FGN is used with smaller H values, although finite Markov chain has proven to be better for higher values of H parameter. The advantage of Markov model is its efficiency and its capability to attain higher Hurst parameter values compared to FGN model. In applications where traffic generation in online manner is required, it is sometimes simpler to use Markov model because for FGN model the total length of sequence must be specified in advance.

REFERENCES

[1] V. Paxson, S. Floyd, "Wide Area Traffic: The Failure of Poisson Modeling", *IEEE/ACM Trans. Networking*, vol. 3, no 3, pp. 226-244, 1995.

[2] K. Park, G. Kim and M. Crovella, "On the effect of Traffic selfsimilarity on network performance", in *Proc. Of SPIE International Conf. on Performance and Control of Network System*, Nov. 1997., pp. 293-310.

[3] K Park, G Kim, M Crovella,"On the relationship between file sizes, transport protocols, and self-similar network traffic", *International Conference on Network Protocols*, Oct. 1996.,pp. 171-180.

[4] Garrett, M.W., Willinger, W., "Analysis, modeling and generation of self-similar VBR video traffic", In: Proc. ACM SIGCOMM 1994, London, UK, pp. 269–280

[5] A. Adas and A. Mukherjee. "On resource management and QoS guarantees for long range dependent traffic", In Proc. IEEE INFOCOM 1995, pages 779–787,

[6] S. Taebi Harandi, "Self-Similar Traffic Modeling and Clos-like Packet Switch Architectures," M.Sc. thesis, Electrical Engineering (Optical Networking) University of Ottawa, Ottawa, ON, Canada, (2005)

[7] R. G. Clegg and M. Dodson, "A Markov Chain Based Method for Generating Long-Range Dependence," Phys. Rev. E 72, 026118 DoI: 10.1103/PhysRevE.72.026118, (2005)

[8] S. Ledesma and D. Liu, "Synthesis of Fractional Gaussian Noise Using Linear Approximation for Generating Self-Similar Network Traffic," ACM Computer Communication Review, Volume 30, Number 2, pp 4-17, (2000)

[9] W. Paxon, "Fast, Approximate Synthesis of Fractional Gaussian Noise for Generating Self-Similar Network Traffic," Comp Comm. Rev. 27,5-18, (1997)

[10] R. G. Clegg, "The Statistics of Dynamic Networks," Ph.D. thesis, Dept. of Math., Uni. of York, York, (2004)