MODELING TELECOMMUNICATIONS TRAFFIC USING THE STOCHASTIC MULTIFRACTAL CASCADE PROCESS

Lyudmyla Kirichenko, Tamara Radivilova, Eias Kayali

Abstract: *In this work the simulation of realizations of telecommunications traffic, which has multifractal properties. The mathematical model of traffic is based on a stochastic binomial multiplicative cascade process with beta-distributed weighting coefficients. There was carry out computer simulation of model multifractal traffic advancing over the communication channel. The emergence of queuing in the infinite buffer size and number of losses with limited buffer size has been studied.*

Keywords*: self-similar stochastic process, Hurst exponent, multifractal stochastic processes, scaling exponent, wavelet transform modulus maxima method, telecommunications traffic, stochastic binomial multiplicative cascade.*

ACM Classification Keywords*: : G.3 Probability and statistics - Time series analysis , Stochastic processes, G.1 Numerical analysis, G.1.2 Approximation - Wavelets and fractals.*

Introduction

Experimental and numerical researches in recent decades indicate that traffic in many media networks has fractal properties. This traffic has a special structure, continuing on many scales – in effect, there is always presence of a number of extremely large bursts relative to a small average traffic. These bursts cause significant delay and packet loss, even when the total loading of all flows is more less than maximum allowable values. The reasons for this effect are features of distribution of files on servers, their sizes, the typical behavior of users, and to a large extent are due to changes in network resources and network topology. [Leland, 1994, Self-similar network traffic and performance evaluation, 2000, Stollings, 2002, Sheluhin, 2007, Loiseau, 2010, Шелухин, 2011].

Self-similar traffic properties led to a number of traffic models based on self-similar (monofractal) stochastic processes [Stollings, 2002, Sheluhin, 2007]. In the last decade, the multifractal properties of traffic have been intensively studied. Multifractal traffic is defined as an extension of similar traffic by taking into account the scaling properties of the statistical characteristics of the second and higher orders. The using multifractal stochastic processes for simulation telecommunications traffic is rather new, and the list of multifractal models is much shorter [Abry, 2002,Reidi, 2002, Veitch, 2005, Шелухин, 2011].

The purpose of this work is model development of telecommunications traffic, which has multifractal properties, based on a mathematical model of the stochastic binomial multiplicative cascade.

Characteristics of self-similar and multifractal random processes

Consider the basic concepts of self-similar and multifractal random processes [Feder, 1991, Calvet, 1997, Reidi, 2002, Kantelhardt, 2008]. Self-similarity of stochastic processes is to conserve the statistical characteristics of a change of the time scale. The stochastic process $X(t)$ is self-similar with the parameter H, if the processes $a^{-H}X(at)$ and $X(t)$ have same finite-dimensional laws of distributions:

$$\text{Law}\{a^{-H}X(at)\} = \text{Law}\{X(t)\}, \quad \forall a > 0,\ t > 0 \ . \tag{1}$$

The parameter H ($0 < H < 1$) is called the Hurst exponent and is a measure of self-similarity or a measure of long-range dependence of process. For values $0,5 < H < 1$ time series demonstrates persistent behaviour. In other words, if the time series increases (decreases) in a prior period of time, then this trend will be continued for the same time in future. The value $H = 0,5$ indicates the independence (the absence of any memory about the past) of values of time series. The interval $0 < H < 0,5$ corresponds to antipersistent time series: if a system demonstrates growth in a prior period of time, then it is likely to fall in the next period.

One can show by choosing in (1) $a = 1/t$, that for the self-similar process, the following equality is held:

$$\text{Law}\{X(t)\} = \text{Law}\left\{\left(\frac{1}{t}\right)^{-H} X(1)\right\} = \text{Law}\{t^H X(1)\}. \tag{2}$$

Using (2), the moments of the self-similar random process can be expressed as

$$\text{M}\left[|X(t)|^q\right] = \text{M}\left[|t^H X(1)|^q\right] = t^{qH}\,\text{M}\left[|X(1)|^q\right] = C(q)\cdot t^{qH}, \tag{3}$$

where the quantity $C(q) = \text{M}\left[|X(1)|^q\right]$.

For multifractal processes, a more general relationship is considered:

$$\text{Law}\{X(a\,t)\} = \text{Law}\{M(a)\cdot X(t)\},$$

where $M(a)$ - is the random function independent of $X(t)$. In the case of the self-similar process $M(a) = a^H$. The multifractal processes are often more flexible scaling relationship for the moment characteristics:

$$\text{M}\left[|X(t)|^q\right] = c(q)\cdot t^{\tau(q)+1}, \tag{4}$$

where $c(q)$ - is some deterministic function, $\tau(q)$ - the scaling exponent, generically nonlinear function. The value $\frac{\tau(q)+1}{q}$ at $q = 2$ is the degree of self-similarity H. For time series, which are responsible monofractal process, the scaling exponent $\tau(q)$ is linear.

The wavelet transform modulus maxima method

One of the most popular mathematical technique of multifractal analysis is the method of continuous wavelet transform modulus maxima (MMWT). It is based on wavelet analysis, which is called a «mathematical microscope» because of the ability to maintain a good resolution at different scales. Because wavelet functions are localized in time and frequency, the MMWT method is a powerful tool for the statistical description of nonstationary processes. [Muzy, 1993, Mallat, 1998, Kantelhardt, 2008]

The continuous wavelet transform of function $X(t)$ is defined as $W(a,b) = \frac{1}{\sqrt{a}}\int_{-\infty}^{\infty} X(t)\psi_{ab}(t)dt$, where $\psi_{ab}(t)$ - is the wavelet function with the parameters of scale a and translation b. The function $W(a,b)$ is called the wavelet spectrum and can be represented as the surface of the wavelet coefficients in three dimensions. The most important information is contained in the surfaces lines of local maxima $W(a,x)$, which are searched on each scale *a*. The MMWT method amounts to compute the following partition function:

$$Z(q,a) = \sum_{l \in L(a)} \left(\sup_{a' \le a} \left| W(a', x_l(a')) \right| \right)^q ,$$

where *L*(*a*) - is the set of all lines *l* modulus maxima of wavelet coefficients at scale *a*; $x_l(a)$ - is the location of the maximum on this scale. To calculate $Z(q,a)$, the absolute maximum value of the wavelet coefficients along each line is selected at scales smaller than the given value of the scale *a*. In this case the relationship is held:

$$Z(q,a) \approx a^{\tau(q)} ,$$

where $\tau(q)$ is the scaling exponent from (4), which is defined for each value of *q*, by calculating the slope of $\ln Z(q,a)$ from $\ln(a)$.

Stochastic multifractal cascade processes

The simplest model of a multifractal process with the desired properties is a deterministic binomial multiplicative cascade [Feder, 1991, Calvet, 1997, Reidi, 2002,]. In its construction, the initial unit interval is divided into two equal intervals, which are assigned weights p_1 and $p_2 = 1 - p_1$, respectively. Then the same procedure is repeated with each of the intervals. As a result, the second step has 4 intervals with weighting coefficients p_1^2, $p_1 p_2$, $p_2 p_1$ and p_2^2. If the number of steps $n \to \infty$ and $p_1 \ne p_2$, we arrive at a limit measure, which is a heterogeneous fractal set. Fig. 1 (a) shows the time series of values a binomial cascade for values $p_1 = 0.6$ (top) and $p_1 = 0.8$ (bottom). The number of iterations $n = 10$, i.e. the length of the realization equals 2^{10}. It is obvious that with increasing the weighting coefficient p_1 the heterogeneity increases.

In the deterministic case the scaling exponent $\tau(q)$ of the binomial process depends only on the weighting coefficient p_1: $\tau(q) = \dfrac{-\ln(p_1^q + p_2^q)}{\ln 2}$. Fig. 1 (b) shows the theoretical scaling exponents $\tau(q)$ for the values $p_1 = \{0.6, 0.7, 0.8, 0.9\}$.

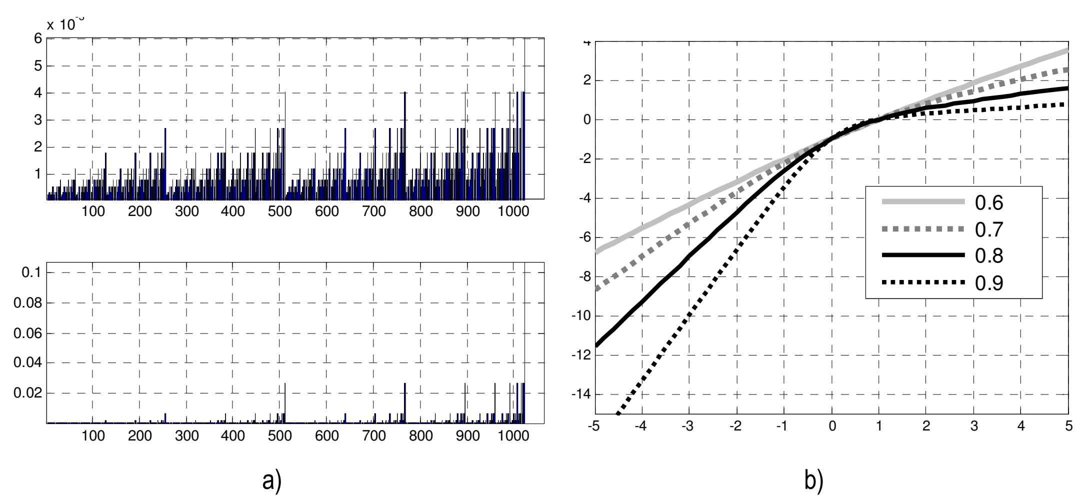


Fig.1. Realizations of the cascade (a) and the scaling exponent $\tau(q)$ for different values of p_1 (b)

Realizations of a deterministic cascade are completely determined by the value p_1, that is unacceptable for the simulation of random processes. When constructing the stochastic cascade, weighting coefficients are independent values of a random variable W [Reidi, 2002, Calvet, 1997, Шелухин, 2011]. A random variable W is chosen so that at each iteration the expectation of the sum of weights amounts 1. If you select the random variable defined on an interval $[0,1]$, then the sum of the coefficients at each iteration will be equal to 1. In this case, the first two intervals would be assigned weights w_1 and $1-w_1$, respectively. In the second step two new independent random values w_2 and w_3 are added. We obtain four intervals with weights $w_1 w_2$, $w_1(1-w_2)$, $(1-w_1)w_3$ and $(1-w_1)(1-w_3)$. When $n \to \infty$ we come to the limit measure, which is a heterogeneous fractal set.

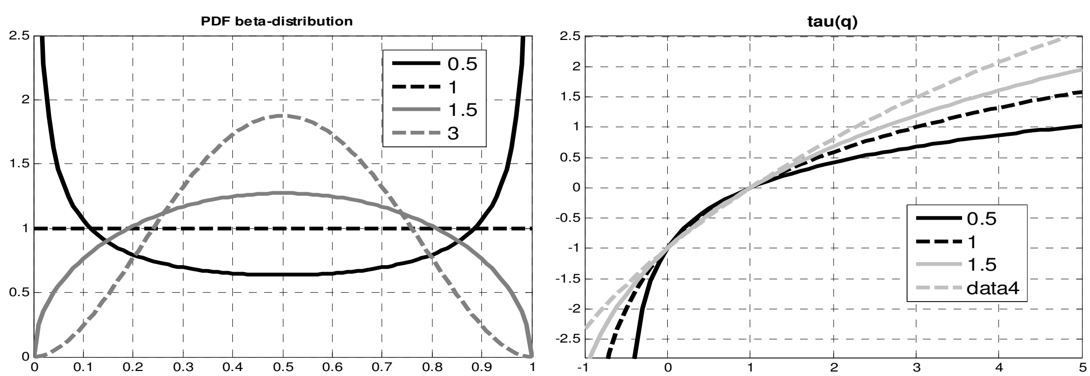
In this work we proposed a random variable that generates the weights, using a random variable having a beta distribution. The probability density function of the beta distribution with parameters $a>0$, $b>0$ is:

$$p(x) = \begin{cases} \frac{1}{B(a,b)}(1-x)^{b-1}, x \in [0,1], \\ 0, \qquad\qquad x \notin [0,1] \end{cases},$$

where $B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}dx$ - is the beta function. For the beta distribution with the equal parameter values $a=b$, for which the function of the density distribution is symmetric, you can analytically determine the scaling exponent $\tau(q)$ [Calvet, 1997, Reidi, 2002]:

$$\tau(\text{q}) = -\log_2 \frac{\text{Beta}(\alpha+\text{q},\alpha)}{\text{Beta}(\alpha,\alpha)} -1, \tag{5}$$

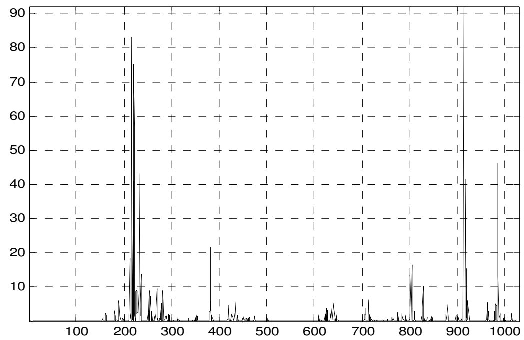
Fig. 2 (a) shows the different types of probability density function of the symmetric beta distribution for values $a=\{0.5,1,1.5,3\}$. For parameters $a=b=1$ we obtain a random variable having a uniform distribution on the interval $[0,1]$. Fig. 2 (b) shows plots of the scaling exponents $\tau(q)$ for the corresponding values of the parameter a of the symmetric beta distribution.



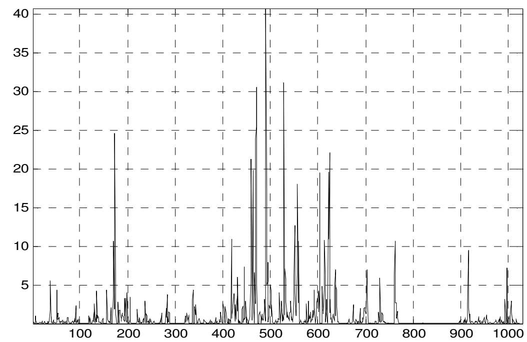
a) b)

Fig.2. Density distribution and the scaling exponent $\tau(q)$ for different values of the parameter a

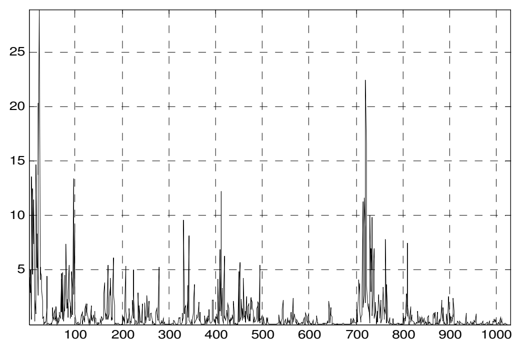
It is obvious that with increasing values of the parameter a is a weakening of the multifractal properties of the time series. Fig. 3 shows the corresponding realizations of binomial cascades.



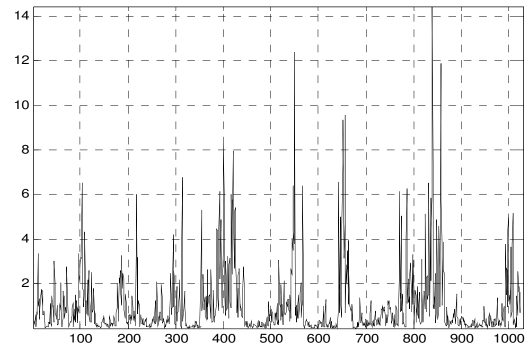
a) $a = 0.5$



b) $a = 1$



c) $a = 1.5$



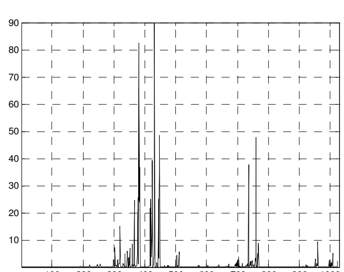
d) $a = 3$

Fig.3. Realizations of the binomial cascade for different values of a

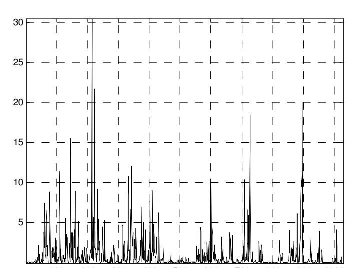
In this case of a symmetric beta distribution, the multifractal properties of the cascade are completely determined by the parameter a. The Hurst exponent H, in consideration of (5), in this case is

$$H = \frac{\tau(2)+1}{2} = -\log_2 \frac{\mathrm{Beta}(\alpha+q,\alpha)}{2\ \mathrm{Beta}(\alpha,\alpha)} .$$

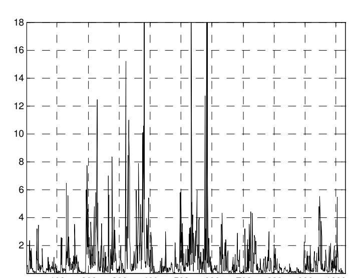
In this work the properties of multifractal cascades generated by beta distributions with different values of the parameters a and b have been investigated. Numerical relationship between the values of the parameter H and the various scaling exponents $\tau(q)$ were obtained. In this case, you can choose not only defined scaling exponent, but also defined Hurst exponent, which determines the degree of long-range dependence. Fig. 4 shows the realizations of cascade processes with the parameter $H = 0.8$ (top) and different multifractal properties $\tau(q)$ (middle), which are determined by the density of beta-distribution (bottom) of different values a and b.



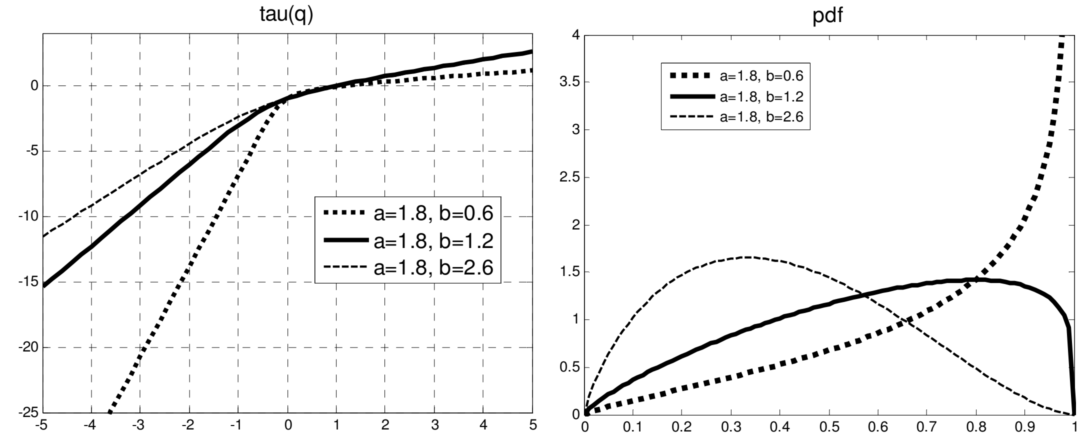
a) $a = 1.8,\ b = 0.6$



b) $a = 1.8,\ b = 1.2$



c) $a = 1.8,\ b = 2.6$



d) e)

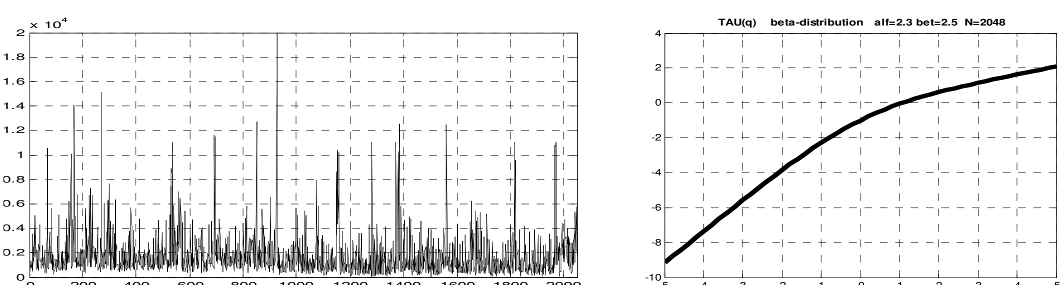
Fig. 4. Realization of the cascade for different values of a and b (a-c)

the corresponding scaling exponents (d) and the densities of beta distributions (e)

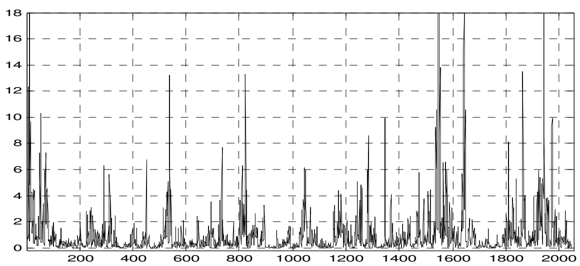
Construction of the model realizations TCP-traffic

The proposed in this work multifractal traffic model has three basic parameters $(I, H, \tau(q))$, where I - is the intensity (average) of the traffic, H - is the Hurst exponent, which determines the degree of long-range dependence, $\tau(q)$ - is the scaling exponent, which determines the heterogeneity (bursts) of realization. To construct a model realization, it is necessary to estimate the parameters of telecommunication traffic and choose the appropriate beta distribution law, which generates a weighting coefficients of multifractal cascade.

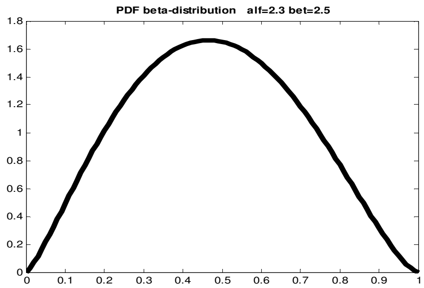
In this work we have investigated traffic realizations of different protocols, which showed their apparent multifractal properties. Fig. 5 (a) shows a sample realization of the TCP-protocol traffic. Also calculated using the MMWT method scaling exponent $\tau(q)$ is shown in fig. 5 (b). For this realization, the estimated value of the Hurst exponent equal to 0.83. Realizations of cascades with same multifractal properties can be derived from the beta distribution with parameters $a = 2.3$, $b = 2.5$, the density of which is shown in fig. 5 (d). One of the realizations of the cascade model of the given multifractal properties is shown in fig. 5 (c).



a) b)



c)



d)

Fig.5. Realization of traffic (a), sample scaling exponent $\tau(q)$,

the density of the beta distribution $a = 2.3,\ b = 2.5$ (d), model realization (c)

The paper presents the simulation of a channel loading and the emergence of queuing in the buffer for the realizations of network traffic. Taking into account the digital nature of modern high-speed communication networks, the communication system can be viewed as a queuing system (QS) of the form, $G/D/C/B/d$, where G means that the input traffic has an arbitrary distribution; D – a deterministic service time equal to one; C – the number of servers equal to the channel bandwidth; B means an infinite buffer size and d - discipline operating the system [Stollings, 2002].

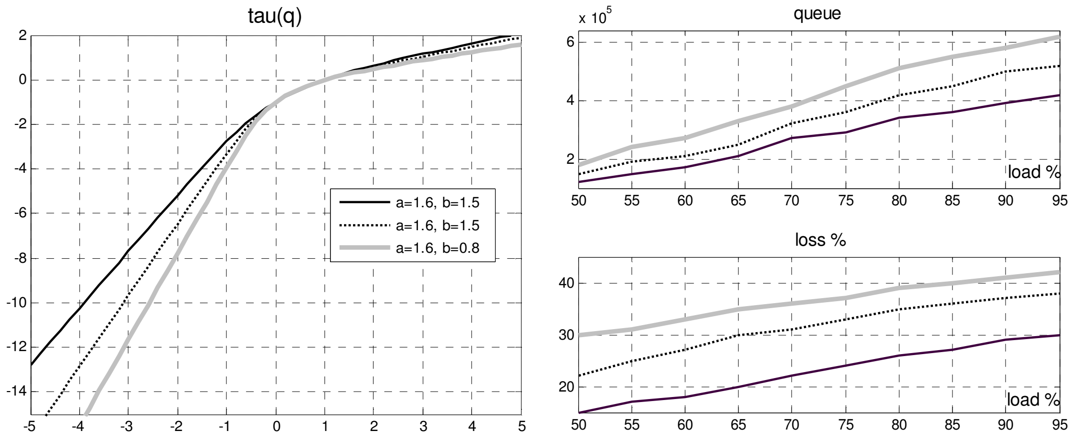
The input buffer receives traffic $Y = (Y_1, Y_2, Y_3, \ldots)$, where Y_t denotes the number of packets that arrive at a time moment t. It is assumed that in QS in every moment t the discipline decides which of the following alternatives should be applied to the package in the system: 1) to initiate the transmission (service) of the package at the time moment t; 2) to store the package in a buffer until the moment $t+1$; 3) to reset (to lose) the package at the time moment t.

In each window t (a window is a time interval $[t, t+1]$), the channel can transmit no more than C packages, which are taken either from the buffer, or from Y_t new packages. The package of the buffer which is passed to the window t, leaves the channel and the system itself at a time moment $t+1$.

The emergence of queuing in the infinite buffer size and number of losses with limited buffer size for self-similar input traffic has been studied sufficiently in many works. [Self-similar network traffic and performance evaluation, 2000, Stollings, 2002, Sheluhin, 2007]. They focused on the determination of the dependency of the queue lengths and the number of losses on the value of Hurst exponent. In this work, a numerical analysis of influence of multifractal properties traffic on the queue size and number of losses was performed.

During the simulation the average system loading was varied from 50 to 95 percent. We investigated the dependence the average queue size in an infinite buffer and a number of losses by limited buffer size on the channel loading for model realizations with the same Hurst exponent and varying degrees of nonlinearity of the scaling exponent $\tau(q)$.

Fig. 6 (a) shows the scaling exponents $\tau(q)$ for model realizations obtained using the beta distribution. The value of parameter Hurst $H = 0.77$. Figure 6 (a) shows the average size of the queue in an infinite buffer and the number of losses in the buffer size $B = 256000$ for realizations with an average traffic intensity $I = 1000$.



a) b)

Fig. 6. Scaling exponent $\tau(q)$ for the model realizations (a)

the average queue length in buffer (top) and the percentage of losses (bottom)

Conclusion

In this work were investigated the properties of stochastic multiplicative cascade processes with randomly weighted beta distribution functions. The proposed mathematical model of the traffic, which has three basic parameters: the average intensity of traffic, the Hurst exponent and scaling exponent. It is shown that traffic models obtained using the stochastic multiplicative cascades allow to flexible present the multifractal properties of actual telecommunication traffic. The offered model network traffic allows by using simulation to adjust the network parameters at the design stage or during its operation.

Acknowledgements

The paper is published with financial support by the project ITHEA XXI of the Institute of Information Theories and Applications FOI ITHEA Bulgaria www.ithea.org, and the Association of Developers and Users of Intelligent Systems ADUIS Ukraine www.aduis.com.ua.

Bibliography

[Abry, 2002] P. Abry. The multiscale nature of network traffic: discovery analysis and modeling / P. Abry, R. Baraniuk, P. Flandrin // IEEE Signal Processing Magazine. – 2002. – № 4 (2). – P. 5–18.

[Calvet, 1997] L. Calvet. Large Deviationsand the Distribution of Price Changes / L. Calvet, A. Fisher, B.B. Mandelbrot // Cowles Foundation Discussion Paper – 1997. –N.1165. –P. 1-30.

[Feder, 1988] J. Feder. Fractals. J. Feder. Plenum, New York, 1988.

[Kantelhardt, 2008] J.W. Kantelhardt. Fractal and Multifractal Time Series. – 2008 [Электронный ресурс]: http://arxiv.org/abs/0804.0747

[Leland, 1994] W. E. Leland. On the self-similarnature of ethernet traffic / W. E. Leland, M. S. Taqqu, W. Willinger, D. V. Wilson] // IEEE/ACM Transactions of Networking. – 1994. – № 2(1). – P. 1-15.

[Loiseau ,2010] P. Loiseau. Investigating self-similarity and heavy-tailed distributions on a large scale experimental facility Networking / P. Loiseau, P. Goncalves, G. Dewaele, P. Borgnat, P. Abry, P. Vicat-Blanc Primet // IEEE/ACM Transactions on.- 2010.-V. 18. – № 4. –P.1261 – 1274.

[Mallat, 1998] S. Mallat. A wavelet tour of signal processing. S. Mallat. Academic Press, San Diego, London, Boston, N.Y., Sydney, Tokyo, Toronto, 1998.

[Muzy, 1993] J.F. Muzy. Multifractal formalism for fractal signals: the structure-function approach versus the wavelet-transform modulus-maxima method / Muzy J.F., Bacry E., Arneodo A. // Phys. Rev. E. -1993. -V.47. -P.875−884.

[Reidi, 2002] R.H.Riedi. Multifractal processes, in Doukhan P., Oppenheim G., Taqqu M.S. (Eds.), Long Range Dependence: Theory and Applications: Birkhuser. -2002. -P. 625–715,

[Self-similar network traffic and performance evaluation, 2000] Self-similar network traffic and performance evaluation. Edited by K. Park, W. Willinger.- John Wiley & Sons, Inc. : 2000,-558 p.

[Sheluhin, 2007] O. I. Sheluhin. Similar processes in telecommunications / O. I. Sheluhin, S. M. Smolskiy, A. V. Osin – John Wiley & Sons Ltd, England: 2007. – 337 p.

[Stollings, 2002] W. Stollings. High-speed networks and Internets. Performance and quality of service. W. Stollings. New Jersey, 2002.

[Veitch, 2005] D. Veitch. Multifractality in TCP/IP traffic: the case against / D. Veitch, N. Hohn, P. Abry // Computer Networks -2005. – № 48(3). – P. 293-313.

[Шелухин, 2011] О. И. Шелухин. Мультифракталы. Инфокоммуникационные приложения / О. И. Шелухин . М.: Горячая Линия - Телеком, 2011. – 578 с.

Authors' Information

Lyudmyla Kirichenko *– Ph. D., Associate professor, Kharkiv National University of Radioelectronics; 14 Lenin Ave., 61166 Kharkiv, Ukraine; e-mail: ludmila.kirichenko@gmail.com.*

Major Fields of Scientific Research: Time series analysis, Stochastic self-similar and multifractal processes, Wavelets , Chaotic systems

Tamara Radivilova *– Ph. D., Associate professor, Kharkiv National University of Radioelectronics; 14 Lenin Ave., 61166 Kharkiv, Ukraine; e-mail: tomachka_7@yahoo.com.*

Major Fields of Scientific Research: Wavelets and fractals, Computer systems and networks, NGN and MPLS technologies

Eias Kayali *– postgraduate, Kharkiv National University of Radioelectronics; 14 Lenin Ave., 61166 Kharkiv, Ukraine.*

Major Fields of Scientific Research: Computer systems and networks