TELECOMMUNICATIONS AND ELECTRONICS
ISSN 1407-8880 TELEKOMUNIKĀCIJAS UN ELEKTRONIKA 2008-8880

Statistical parameters estimation of the self-similar input traffic for the Measurement-based Admission Control

M. Kulikovs,
TET Department, ETF Faculty, RTU
Riga, Lomonosova str. 1, LV-1019, E-Mail: mihails.kulikovs@tet.rtu.lv

In recent years Admission control scheme based on traffic parameters measurements has become popular. The main reason is that such Input Traffic Management gives an opportunity to raise productivity of networks. The current research is focused on the self-similar input traffic generation and analysis for the statistic parameters estimation. The parameters include autocorrelation function, spectral density and the autocorrelation interval.

Keywords**: QoS, measurement-based admission control, correlation interval, autocorrelation function, spectral density function**

I. INTRODUCTION

Maximizing the utilization of the network while maintaining service guarantees (Quality of Service) to network users is one of the tasks of network management. Admission Control (AC) is a well known mechanism for QoS provision. AC procedure is employed to maintain a high utilization of network resources while preserving the guarantees made to existing flows.

There are two types of Service guaranties mechanisms, deterministic and statistical. The deterministic service guarantee management technique, Measurement-based Admission Control (MBAC), uses worst-case analytical bound. Statistical or measurement-based decision algorithms use the a priori traffic characteristics only for newly arrived flows while characteristics have to be measured for already existing in the system flows. The MBAC mechanism plays an important role in Quality of Service (QoS) guarantee. Therefore, in the present study we will focus on the main characteristics necessary for the correct traffic estimation in the measurements based admission control mechanism. They are the autocorrelation and spectral density functions.

II. MEASURING AND ESTIMATION

A. *Autocorrelation and Spectral density functions.*

Internet traffic exhibits intriguing temporal correlation properties, such as self-similarity and long memory (long-range dependence) on various time scales (K. Park, 1996),(K. Park, 1998). Contrary to the classical Poisson-model ($H = 0.5$) assumption, these properties emphasize long-range time correlation between packet arrivals.

Many research works show that Internet traffic has H parameters varying in the range $0.7 < H < 0.95$. Fig. 1 presents autocorrelation functions for Poisson traffic where $H = 0.5$, and self-similar traffic with H parameter equals to 0.75 and 0.95.

An interesting feature of self-similar processes is the fact that the autocorrelation function does not degenerate when $m \to \infty$. This feature is in contrast to stochastic processes where the autocorrelation function degenerates as $m \to \infty$. For the self-similar process the autocorrelation function with the dependence on Hurst parameter can be described as follows:

$$\rho(k) = \frac{1}{2}\left[(k+1)^{2-\beta} - 2k^{2-\beta} + (k-1)^{2-\beta}\right] \qquad (1)$$

and is depicted on Fig. 1. Such behavior of the autocorrelation function corresponds to the Fractal Brownian Motion (fBm) that is a classical example of the self-similar process.

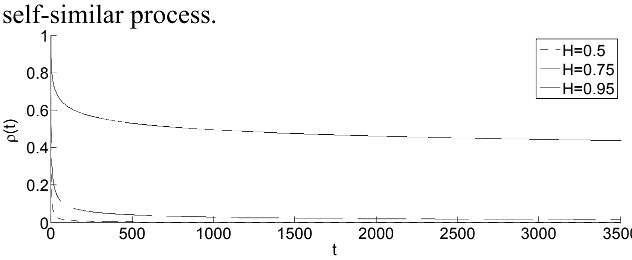

Fig. 1 Autocorrelation function of the self-similar process

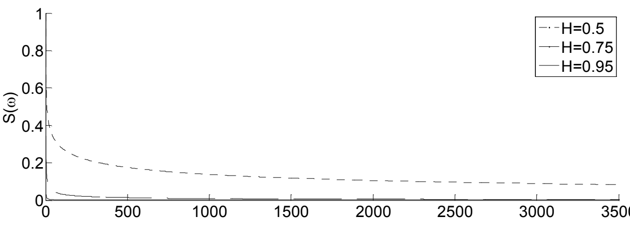

Fig. 2 Spectral density function of the self-similar process

Long-range dependent processes are characterized by the autocorrelation function which decays hyperbolically. This implies that the auto-correlation function is non-summable, unlike more conventional short-range dependent processes, which have auto-correlation functions that decay exponentially (Morin, Dec. 1995).

The correct traffic parameters estimation strictly depends on the time interval between measurements. Traffic estimator module of the MBAC mechanism can estimate the interval using the correlation interval of the data flow.

The autocorrelation function of the stochastic process can be evaluated in the following way:

$$\rho(\tau) = E[x(t)x(t+\tau)] \qquad (2)$$

The autocorrelation of an ergodic process is sometimes defined as equated

$$\rho(\tau) = \frac{1}{N}\sum_{i=0}^{N} x_i x_{\tau+i} \qquad (3)$$

and correlation interval:

$$\tau_k = \frac{1}{2}\int_{-\infty}^{\infty} |\rho(\tau)|\, d\tau = \int_{0}^{\infty} |\rho(\tau)|\, d\tau \qquad (4)$$

The τ_k value approximately shows the time interval where the correlation of a random process is. Based on the Fig. 1 and above mentioned correlation interval, we can argue that the random process with the high self-similarity degree has the higher correlation ("long memory") and, consequently, correlation interval is higher than the process with low self-similarity property.

III. TRAFFIC MODELLING AND ANALYSIS

A. Modelling

To evaluate the MBAC algorithm the data flows with different traffic characteristics, e.g. Hurst parameter and network utilization were generated and analyzed. Two models of the traffic generation were used and compared in the paper. One of the methods is a well-known model for self-similar traffic generation, e.g. traffic of the independent ON-OFF sources, firstly presented by Mandelbrot (Mandelbrot, 1969).

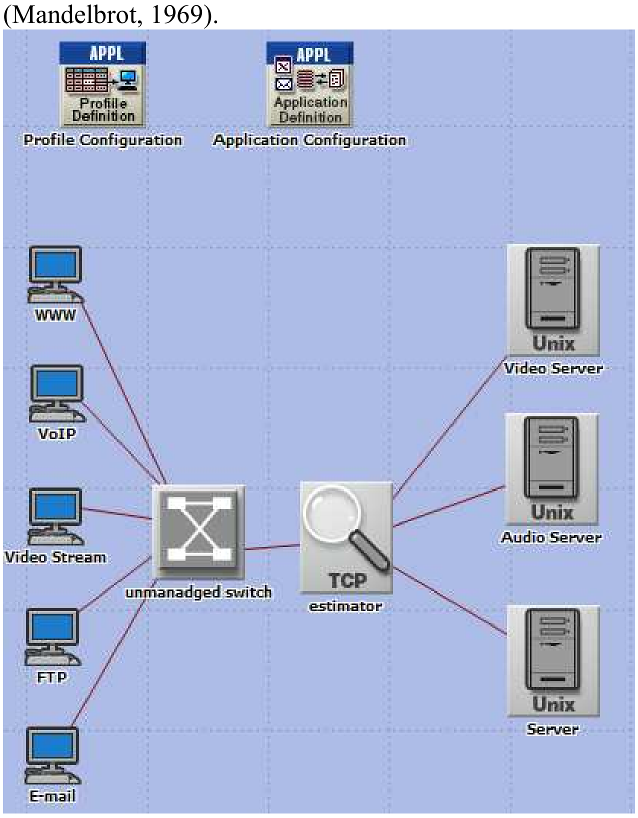

Fig. 3 MBAC Model in OPNET framework

This method is based on superposition of many (strictly alternating) independent and identically distributed (i.i.d.) ON-OFF sources. Here, by a strictly alternating ON-OFF sources we mean a model where ON- and OFF-periods strictly alternate and where ON-periods are independent and identically distributed (i.i.d.). The OFF-periods are i.i.d., and the ON- and OFF-period sequences are independent from one another. The ON- and OFF-periods do not need to have the same distribution. To simulate ON-OFF sources the OPNET simulation environment was used.

The other model is a simplified ON-OFF model where packet train consists only of one packet with the constant size, and the interarrival times are generated according to Pareto distribution.

All generated data sequence were evaluated for the Hurst parameters and compared with the defined ones. Additionally, the data were investigated for the autocorrelation function and spectral density function. The traffic of the simplified model was generated in Matlab framework, while for the second traffic model the OPNET simulation framework was used. The OPNET simulation tool is used for the MBAC model evaluation.

B. Data analysis

The data analysis was done for four sets of data: two different methods of the traffic generation (simple and ON-OFF methods) and two different forms of Pareto distributions (2-parameters and Generalized Pareto distribution).

Parameters for the traffic generation were estimated in the following way. Firstly we generated and analyzed Poisson data flows for the different network utilization values $(0.5, 0.7, 0.9)$. For the scenarios the server capacity of 1000packets/sec has been chosen.

Further, we started to model the simple traffic with Pareto distributed interarrival time. The data traffic was generated for three different Hurst parameters: $(0.5, 0.75, 0.95))$ and corresponding parameters for distribution were estimated in the way that mean values of packets would be similar to mean values in Poisson data flow.

After that the parameters for ON-OFF sources were estimated for the similar assumption of network utilization and Hurst parameter. The difference in parameter estimation is OFF time periods, when traffic generation is in the idle state. During the ON- state ("packet train") packets are generated according to Exponential distribution and to estimate parameters of the packets two parameters should be taken into account: number of ON-OFF sources and ON- OFF- time periods relation. For our experiments we assumed that OFF period is four time longer than ON period ("packet train"). And the number of ON-OFF sources was ten identical sources.

On the Fig. 4 and Fig. 5 the autocorrelation function of the 2-parameters Pareto distribution generated data is depicted. Just after a brief look at the pictures it can be realized that autocorrelation function for generated data coordinally differ from the theoretical background. There is aurocorrelation function of two types of traffic on the Fig.

4: the autocorrelation function of the Piosson data flow and the autocorrelation function of the simplified ON-OFF model with Hurst parameters $H = 0.5; H = 0.75$ and $H = 0.95$. As it was mentioned earlier autocorrelation function of the Pisson data flow and self-similar traffic with $H = 0.5$ have the same charecteristics. From Fig. 4 we can conclude that the statistic characteristics of the Poisson data flow and self-similar data flow with $H = 0.5$ and $H = 0.75$ have the same statistic characteristics, e.g. autocorrelation function. The second conclusion is that correlation of the process with „high degree" of the self-similarity is less than correlation of the process with „low degree". Both of these statements are wrong and and makes us to look for the explanation of this phenomena.

Autocorrelation functions of the data set generated based on the Generalized Pareto distribution are presented on the Fig. 6 and Fig. 7. Comparing that functions with theoretical background we can say that it has correct statistical charecteristics of the traffic with self-similar behavior.

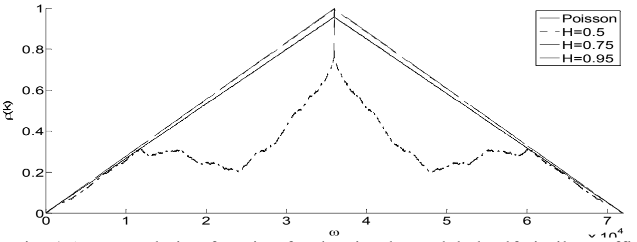

Fig. 4 Autocorrelation function for the simple modeled self-similar traffic based on 2-parameters Pareto distribution and $\rho = 0.75$

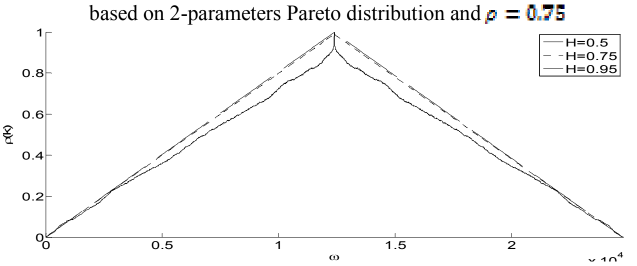

Fig. 5 Autocorrelation function for the ON-OFF modeled self-similar traffic based on 2-parameters Parto distribution and $\rho = 0.75$

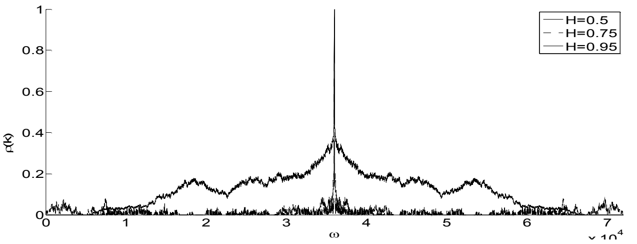

Fig. 6 Autocorrelation function for the simple modeled self-similar traffic based on 3-parameters

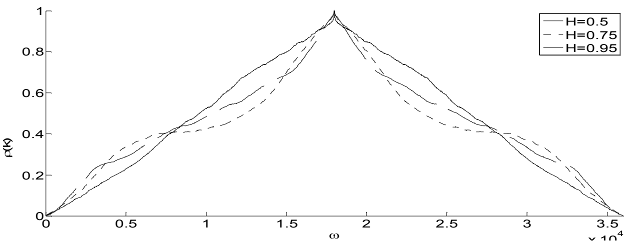

Fig. 7 Autocorrelation function for the ON-OFF modeled self-similar traffic based on 3-parameters

Using Fourier transform on autocorrelation function spectral density function could be obtained. After comparison of autocorrelation and spectral density functions of generated data with theoretical background the next conclusions are done. Data generated using the Generalized Pareto distribution provide better fitted statistical characteristics of the theoretical and then generated with well known 2-parameters Pareto distribution.

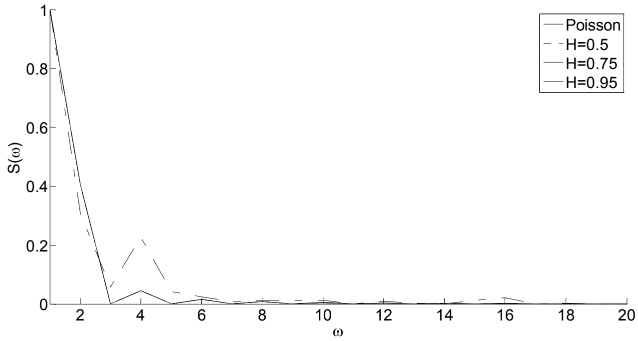

Fig. 8 Spectral density function for the simple modeled self-similar traffic based on 2-parameters Parto distribution

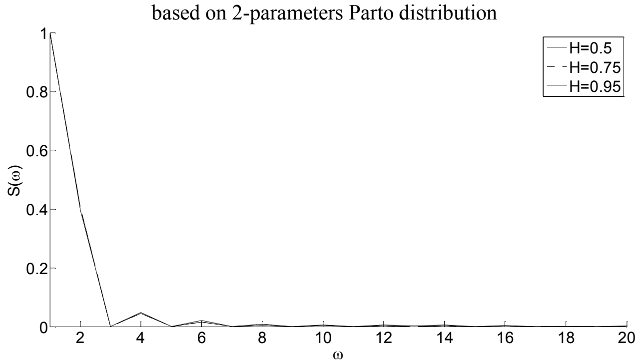

Fig. 9 Spectral density function for the ON-OFF modeled self-similar traffic based on 2-parameters Parto distribution

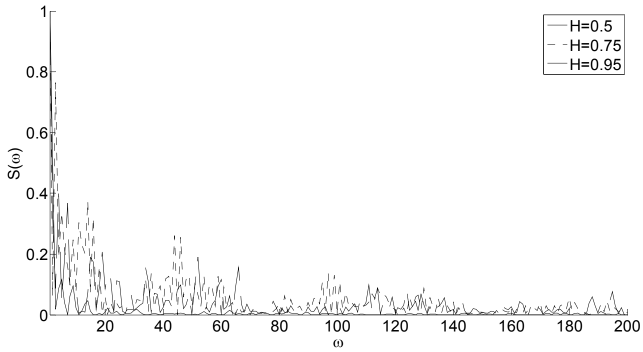

Fig. 10 Spectral density function for the simple modeled self-similar traffic based on 3-parameters Parto distribution

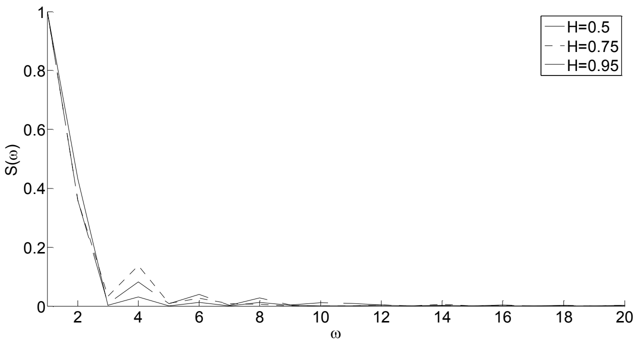

Fig. 11 Spectral density function for the ON-OFF self-similar traffic based on 3-parameters Parto distribution

2-parameters Pareto distribution is the most popular distribution for self-similar traffic generation is. Our previous investigation shows that this distribution does not provide correct statistical characteristic of the generated traffic. The discovery of this phenomenon makes us to get into the root of the Pareto distribution.

C. *Non-ergodic process*

After serious theoretical explorations, experiments with traffic generation and traffic analysis we have found the explanation of this phenomena. Data traffic where interarrival time of the packets has Pareto distribution is not *ergodic.*

Taking into account non-ergodic property of the self-similar traffic the autocorrelation function has to be used as follows:

$$\rho(\tau) = E[(x(t) - \mu)(x(t + \tau) - \mu)] \quad (5)$$

An example of the autocorrelation function of the data set obtained by treatment of 2-parameter Pareto distribution is presented in Fig. 12 and Fig. 13. Comparison of the behavior of the centered autocorrelation function of the modeled data and theoretical background shows that behaviors are similar. The Fig. 16 and Fig. 17**Error! Reference source not found.** shows spectral density of the self-similar process obtained with Fourier transform of the centered autocorrelation function. The graph is similar to the theoretical behavior.

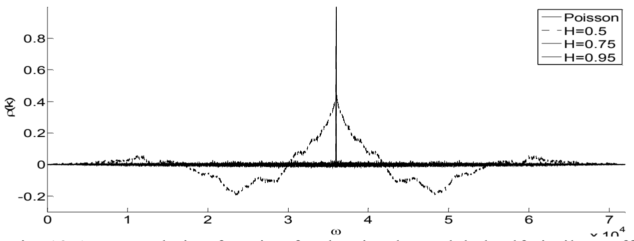

Fig. 12 Autocorrelation function for the simple modeled self-similar traffic based on 2-parameters Pareto distribution and $\rho = 0.75$

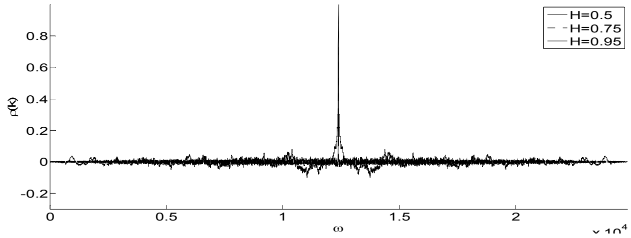

Fig. 13 Autocorrelation function for the ON-OFF modeled self-similar traffic based on 2-parameters Parto distribution and $\rho = 0.75$

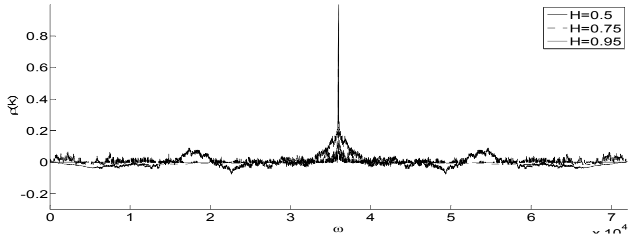

Fig. 14 Autocorrelation function for the simple modeled self-similar traffic based on 3-parameters

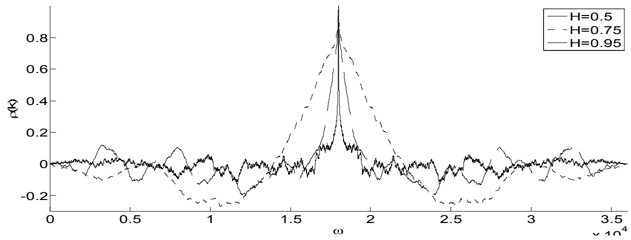

Fig. 15 Autocorrelation function for the ON-OFF modeled self-similar traffic based on 3-parameters

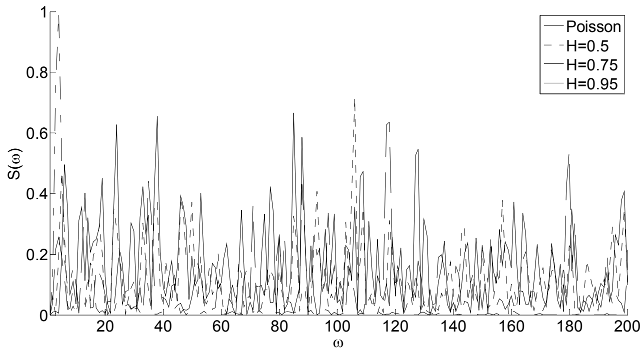

Fig. 16 Spectral density function for the simple modeled self-similar traffic based on 2-parameters Parto distribution

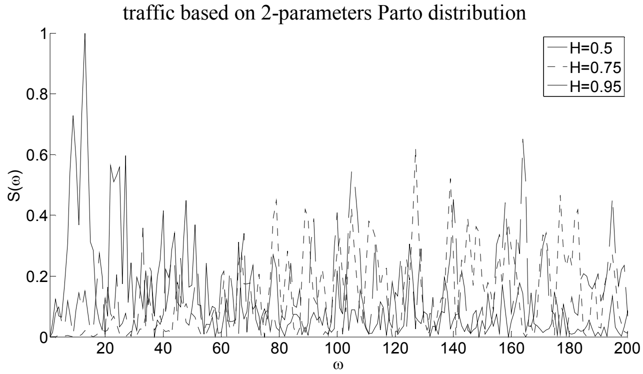

Fig. 17 Spectral density function for the ON-OFF modeled self-similar traffic based on 2-parameters Parto distribution

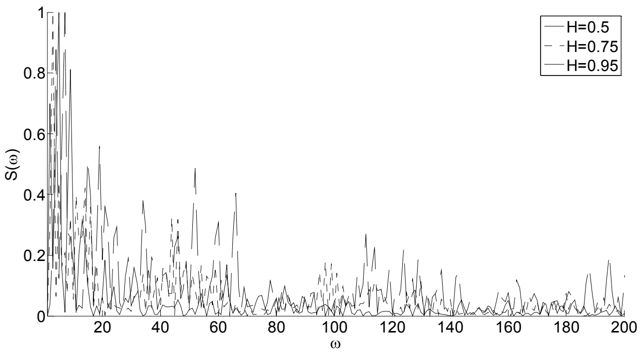

Fig. 18 Spectral density function for the simple modeled self-similar traffic based on 3-parameters Parto distribution

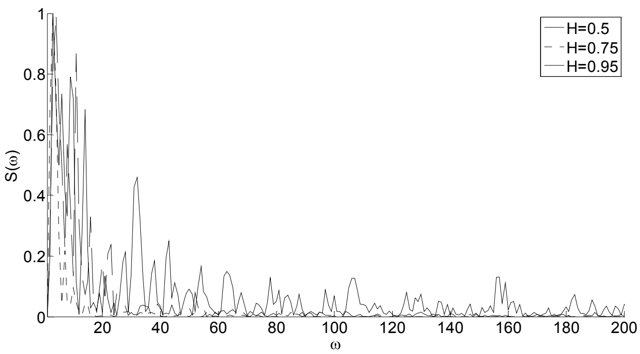

Fig. 19 Spectral density function for the ON-OFF self-similar traffic based on 3-parameters Parto distribution

D. Correlation interval

Based on the autocorrelation function the important property of the traffic can be estimated.
Correlation interval can be estimated as follows:

$$\tau_k = \frac{1}{\rho_0}\sum_{i=0}^{k}\rho_i \qquad (6)$$

For example, the correlation intervals for data set generated with an assumption that utilization of the network $\rho = 0.5$ (in our scenarios it means that mean interarrival rate is $\lambda = 500\ packet/se$) are:

$$\tau_{k_{Poisson}} = 50.82sec$$
$$\tau_{k_{H=0.5}} = 52.04sec$$
$$\tau_{k_{H=0.75}} = 55.92sec$$
$$\tau_{k_{H=0.95}} = 117.01sec$$

These results correspond to theoretical background stating that growing of self-similarity "degree" is associated with traffic being more correlated and statistical parameters stay constant for longer time.
The correlation interval is a very important parameter. Knowing correlation interval of the data flow it is possible to decrease system overhead related to traffic measurements. As presented in Fig. 20 - Fig. 23 and was estimated by **Error! Reference source not found.** the process with higher self-similar degree has bigger correlation interval. It means that knowing statistical charecteristics of the traffic in time T, they would be constant to time $T' = T + \tau_k$.

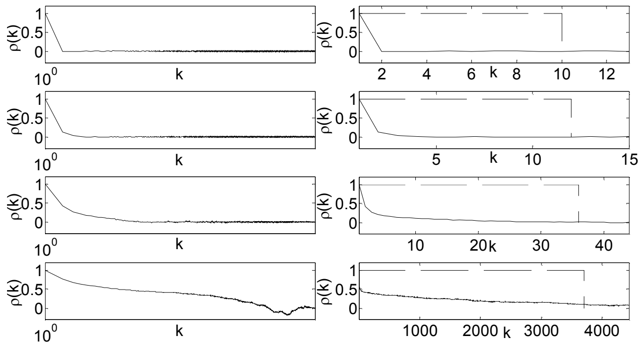

Fig. 20 Autocorrelation interval for the simple modeled self-similar traffic based on 2-parameters Pareto distribution and $\rho = 0.75$

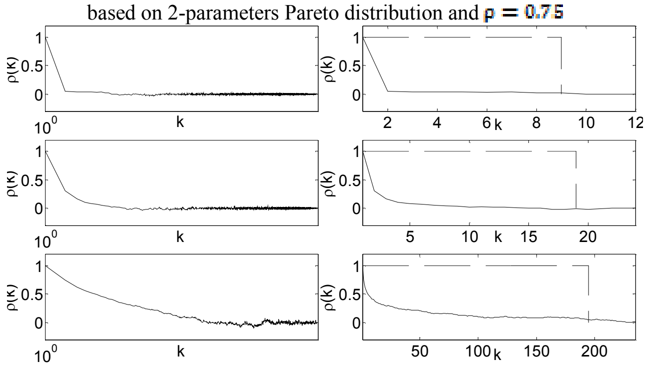

Fig. 21 Autocorrelation interval for the ON-OFF modeled self-similar traffic based on 2-parameters Parto distribution and $\rho = 0.75$

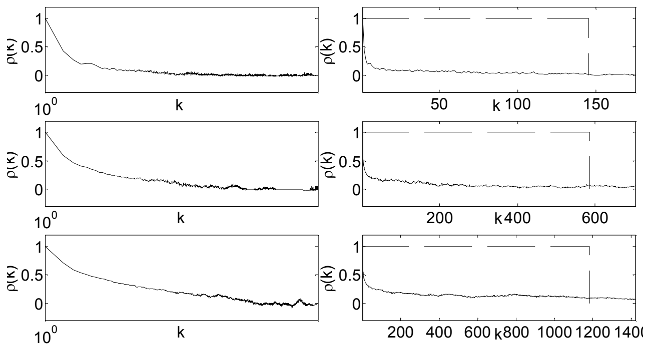

Fig. 22 Autocorrelation interval for the simple modeled self-similar traffic based on 3-parameters

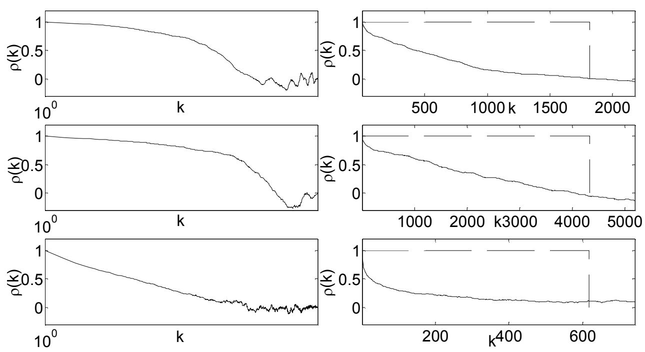

Fig. 23 Autocorrelation interval for the ON-OFF modeled self-similar traffic based on 3-parameters

IV. CONCLUSIONS

After serious theoretical explorations, experiments with traffic generation and traffic analysis we have found the explanation of this phenomena. Data traffic where interarrival time of the packets has Pareto distribution is not ergodic. The centered data should be used for autocorrelation and spectral density function estimation.

On the Fig. 20 - Fig. 23 the calculations of the correlation interval are presented. On the left side there is a correlation function, while on the right the correlation interval is depicted with scaled autucorrelation function.

Fig. 20 - Fig. 23 present correlation intervals for the different way of generated traffic and the different value of the Hurst parameter. On the figurē, the correlation interval for Poisson data flow and flows with Hurst parameter where H=0.5, H=0.75, H=0.95, are presented.

Our investigation shows that correlation interval depends not only on the H parameter but on the traffic intensity as well. The correlation interval increases with the increase of H parameter, and it decreases with the traffic intensity increase.

REFERENCES

1. Adas A. and Mukherjee, A On resource management and QoS guarantees for long range dependent traffic [Conference] // IEEE INFOCOM'95. - 1995. - pp. 779-787.
2. Addie R., Zukerman, M., and Neame, T Fractal traffic: measurements, modelling and performance evaluation [Conference] // IEEE INFOCOM '95. - 1995. - pp. 977-984.
3. B.B. Mandelbrot Long-Run Linearity, Locally Gaussian Processes, H-Spectra and Infinite Variances [Journal]. - [s.l.] : International Economic Review, Vol.10, pp. 82-113, 1969.
4. Crovella M. E. & Bestravos, A. Self-Similarity in World Wide Web Traffic: Evidence and Possible Causes [Conference] // IEEE Trans. Networking, 5(6). - 1997.
5. Crovella M. E. Self-Similarity in WWW Traffic: Evidence and Possible Causes [Journal] // IEEE Trans. Networking, vol. 5, no. 6, Dec. 1997,. - [s.l.] : ," IEEE Trans. Networking, vol. 5, no. 6, Dec. 1997, pp. 835–45., 1997. - pp. pp. 835–45..
6. K. Park G. Kim, and M. Crovella On the Relation Between File Sizes,Transport Protocols, and Self-Similar Network Traffic [Conference] // IEEE Network Protocols. - 1996. - pp. 171-180.
7. K. Park G. Kim, and M. Crovella. On the Relation Between File Sizes,Transport Protocols, and Self-Similar Network Traffic [Conference]. - [s.l.] : IEEE Network Protocols. pp. 171-180., 1996.
8. K. Park T.Tuan Congestion Control for Self-Similar Network Traffic [Book]. - 1998.
9. K. Park T.Tuan Congestion Control for Self-Similar Network Traffic [Report]. - 1998.
10. Leland W. E On The Self-Similar Nature of Ethernet Traffic [Article] // IEEE Trans. Networking, . - [s.l.] : IEEE Trans. Networking, vol. 2, no. 1, Feb. 1994, pp. 1–15., Feb. 1994. - vol. 2. - pp. 1–15..
11. Likhanov N. and Tsybakov, B Analysis of an ATM buffer with self-similar ("fractal") input traffic [Conference] // IEEE INFOCOM '95. - 1995. - pp. 985-992.
12. Likhanov N. Bound of the buffer occupancy probability with self-similar input traffic [Book Section] // SELF-SIMILAR NETWORK TRAFFIC AND PERFOMANCE EVALUATION / book auth. Park. K Walter Willinger. - 2000.
13. Moore Andrew W. An Implementation-Based Comparison of Measurement-Based Admission Control Algorithms [Report]. - 2002.
14. Morin P. R. The Impact of Self-Similarity on Network Performance Analysis [Report]. - [s.l.] : Ph.D. dissertation, Carleton Univ., Dec. 1995.
15. Norros I. A storage model with self-similar input. [Journal] // Queueing Systems,16. - 1994. - pp. 387-396.
16. Park K. AFEC: an adaptive forward errorcorrection protocol and its analysis [Report]. - 1997.