

STUDY OF SELF-SIMILAR PROCESSES TRAFFIC BASED ON/OFF MODEL

The article examines the relationship of the various characteristics of the random process associated with the effect of self-similarity. A simulation of the transfer traffic. The results Hurst parameter estimation by different methods.

Keywords: random process, on/off model, traffic, self-similarity, network.

Statement of the problem.

In the study and the study of various properties and characteristics of information transmission systems poses the problem of constructing models of incoming traffic, the characteristics of which are as near to the characteristics of real-world data flows. Study the different types of network traffic over the last years prove that network traffic is self-similar (self-similar) or fractal (fractal) by their nature [1-5]. From this it follows that the methods used for modeling and calculation of network systems based on the use of пуассонівських flows not provide a complete and accurate picture of what is happening in the network.

Analysis of the last researches.

In [6], based on analysis of a large number of works on research of traffic in IP networks is the classification of traffic, and each traffic type the appropriate distribution law. The report States that the three most commonly used types of distributions (Pareto, Weibull and lognormal. Pareto distribution, as a distribution with heavy tail is used to simulate the intervals between requests to the web-resources, size of files uploaded, VoIP traffic. The Weibull distribution is used to model the processes of receipt of the Protocol blocks FTP. Lognormal, the early model of self-similar traffic is used to simulate the intervals of receipt of packets in the local computer networks, the time between calls to Call centers, the size of transmitted files. Special attention deserves the fact that the aggregation of flows from multiple sources, if at least one of them has the property of self-similarity, self-similarity properties will own and the resulting total flow. Association flows from sources generating traffic, the described process with infinite variance, leads to self-similar network traffic, which seeks to traffic, described by model fractal Brownian motion. Self-similarity is also saved when you merge threads and from homogeneous and heterogeneous sources of traffic. Widespread in the description of the package traffic received ON / OFF model.

The formulation of the objectives of article (problem).

According to the results of the analysis of the publications devoted to the study of self-similar flows, in the framework of the issues addressed in this article, interest is the value of the parameter Hirst traffic transmission remote work stations that transmit information flow, the data is packaged for distribution of Pareto.

Summary of the basic material.

As noted, one of the important moments in the development of models of the network traffic is the analysis of compliance of the behaviour of these models to experimental data, which indicates that корельованість values traffic in a wide time range or extended its dependence of the correlation

function. Analysis of different approaches and solutions to identify the conformity of the experimental data taken on the packet level with various internal switches in modern high-speed networks, the subject of a number of works [2 - 5]. The determining factor of the availability of the property for ON/OFF model is the so-called «heavy» distribution characterizes the fact that the probability of long ON and OFF the intervals of the order of T (long series of packages and міжсерійних intervals) may be significant $P(T > t) \sim t^{-\beta}$, $t \rightarrow \infty$, $0 < \beta < 1$. Interval ON identified with the passed a series of packages, and the range of OFF - the lack of a packet. These definitions without specifying the type of distribution function ON/OFF-intervals taken as a basis of construction of model of network traffic. That is, ON/OFF-models the traffic is considered as a combination of sources, which generate it. The size of ON-and OFF-periods is a random variable, which should have a finite expectation and infinite variance. Note that the ON/OFF frequently classified as дежурящий fractal recovery process is one of the varieties of the FPP. FPP - fractal point processes, which are very visual for modeling of self-similar traffic. The easiest dot process appears on the time axis nondecreasing step function, moments of growth which are random. There are many modifications of FPPs that enough economic and computationally efficient.

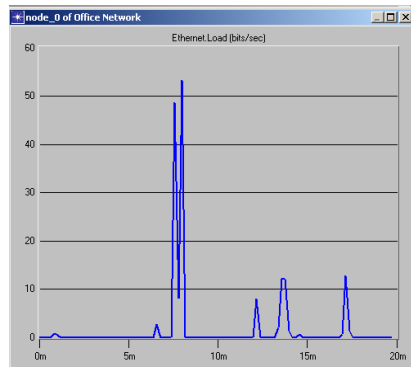
Consider further that the sequence of packets in ON the interval has regular stationary, and the randomness in the network traffic due to the statistical nature of the ON/OFF intervals. In the paper, we built a simulation module traffic in a computer network environment OPNET. Configure the appropriate settings for stations and graphics download the channels are presented in figures 1 and 2.

Dispersion of these intervals are large or even tend to infinity. Experimental data indicate a peculiar behavior of network traffic that does not fit into the framework of the behavior of known models queues (пуасонівських, Markov modulated etc). For the latest models корельованість events is on limited periods of time. With an extensive dependence is closely connected property of self-similarity, which determines the fractal nature of this function.

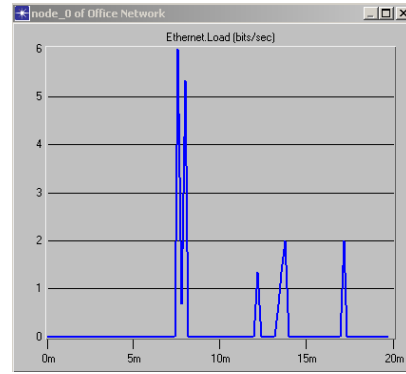
The model of the generalized traffic can be obtained from the aggregation (superposition) of a large number of independent identically distributed stationary point of the processes of recovery. Assuming strict alternation intervals, we obtain the sequence of these intervals the totality of the stream of points. Describing the behavior of these random intervals «heavy» distribution, we come to a consolidated network traffic, which has фрактальними properties.

Attribute	Value
name	node_0
model	ethernet_station
Traffic Generation Parameters	(...)
Start Time (seconds)	constant (5.0)
ON State Time (seconds)	exponential (10.0)
OFF State Time (seconds)	exponential (90.0)
Packet Generation Arguments	(...)
Interarrival Time (seconds)	exponential (1)
Packet Size (bytes)	pareto (0.1, 0.8)
Segmentation Size (bytes)	No Segmentation
Stop Time (seconds)	Never

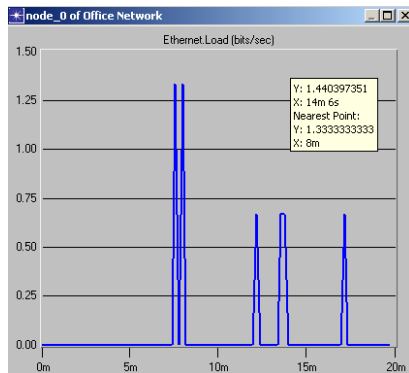
Fig. 1 - Parameters of traffic



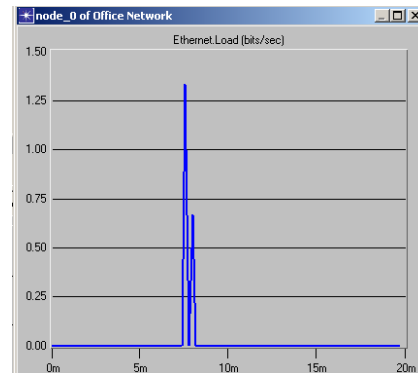
a) pareto (0.1,0.8)



б) pareto (0.1,1.2)



в) pareto (0.1,1.6)



г) pareto (0.1,1.8)

Fig. 2 - Graphics download channel

Figure 4 shows a graph of the autocorrelation function of the investigated time series. The calculation was performed using the classical formula

$$r(k) = \frac{1}{N-k} \frac{\sum_i (X_i - \bar{X})(X_{i+k} - \bar{X})}{\sigma^2(X)} \quad (1)$$

\bar{X} – the sample mean of a number o X ;

$\sigma^2(X)$ – the sample variance of the number X ;

$k \in \mathbb{Z}^+ = \{0, 1, 2, \dots\}$.

Exploring the auto-correlation function of the energy spectrum (figure 3) you can also note the presence of harmonic component and small fall ACF this type of traffic.

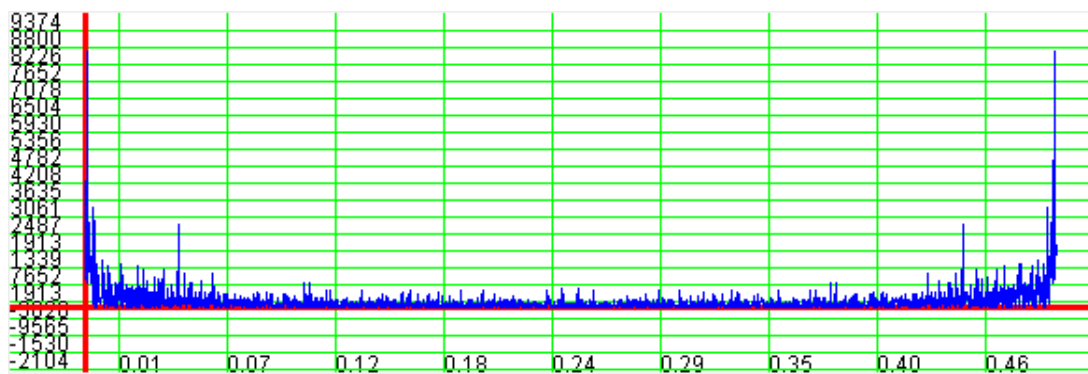


Fig.3 - Range power network

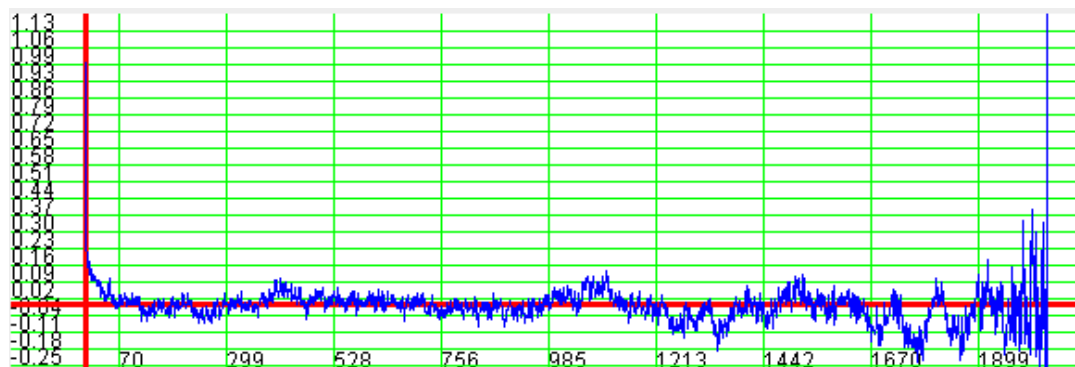


Рис.4 – Автокореляційна функція

The determined value of the coefficient of the Hearst family methods:

- 1) Method of the aggregate differences (Dispersion method). Logarithmic sample variance compared with the level of aggregation should be a straight line with a slope greater than -1. In this case, $H=1+B/2$.
- 2) Method of normalized range (R/S plot method). Logarithmic sample R/S statistics in comparison with the number of points of the aggregated series is a straight line with the slope.
- 3) Method of absolute moments (Absolute Moments). Aggregated series $X(m)$ defines the use of different block sizes m . Logarithmic aggregation in comparison with a level of absolute first moment of the aggregated series of X must be a straight line with a slope of $H-1$.
- 4) the Method of difference residues (Variance of Residuals). Logarithmic aggregation in comparison with the average variance of residuals should be a straight line with a slope of $H/2$.
- 5) the Method of calculating using Periodogram. This method plots of the logarithm of the spectral density of the time series in comparison with the logarithm of the frequency. Tilt provides an estimate of H . Періодограма is given by the formula

$$I(\nu) = \frac{1}{2\pi N} \left| \sum_{j=1}^N X(je^{i\nu}) \right|^2 \quad (2)$$

where ν – frequency,

N – the length of the time series,

- 6) Method of calculating energy bursts (Abry-VeichEstimator) for assessment of H use the energy of the series, at different scales.
- 7) Method of calculation of the minimal values (Whittle Estimator, based on the minimization of the likelihood function, which applies to the period of time series gives an estimate of H and conducts dependence on confidence interval.

Each of these methods has its precision and accuracy of calculations. The results obtained are summarized in table 1.

able 1 - the results Obtained by evaluating Hurst parameter

Option period ON/OFF	Aggregate assessment	R/S assessment	Periodogram assessment	Whittle assessment	Average value H
Pareto(10,0.8)	0.647	0.773	0.636	0.628	0.67100
Pareto(10,1.2)	0.640	0.699	0.576	0.608	0.63075
Pareto(10,1.6)	0.637	0.669	0.518	0.546	0.59250
Pareto(10,1.8)	0.598	0.671	0.625	0.559	0.61325

Hurst exponent value for all cases examined more stable 0.5, which demonstrated the effect of self-similarity in the study implementation. Hurst Exponent (H) is a measure of the sustainability of the statistical phenomenon or the extent of the duration of long-term dependency process. The closer the value is to 1, the higher the degree of sustainability of long-term dependence. At $0 \leq H \leq 0.5$, the time series is трендонестійким (антиперсистентним). It is more volatile than random number, because it consists of frequent reverses the decline is the rise. At $0.5 \leq H \leq 1$ number is трендостійким. Trend changes can be predicted. The simulation results showed that the method of energy bursts are the most effective for the analysis of processes in modern telecommunication networks. Самоподібні flow properties of weakly depend on the characteristics of the period of ON-OFF, as well as the location of the alpha values to 1 and more of it, the traffic does not have more properties of self-similarity, which is in contradiction with the theoretical data.

Conclusions and perspectives of further research.

Each of the methods has its precision and accuracy of the calculation. From these results we can conclude that for real traffic, the average value of the Hurst parameter H more than 0.5. It says about the manifestation of the effect of self-similarity in the investigated the implementation of computer network, since the index Hirst, which is a measure of the sustainability of statistical phenomena or the extent of the duration of a long-term process. The closer the value H is to 1, the greater the level of resistance. This enables us to predict the behavior of traffic.

1. Norros, I. A storage model with self-similar input / I. Norros // *Queueing System*. – 1994. – V. 8. 2. Taqqu M.S., Proof of a Fundamental Result in Self-Similar Traffic Modeling / Taqqu M.S., Willinger W., Sherman R. // *SIGCOMM Comput. Commun. Rev.* – 1997. – Vol. 27, Issue 2. – P. 5 - 23. 3. Шелухин О.И., Тенякишев А.М., Осин А.В. Фрактальные процессы в телекоммуникациях. Монография / Под ред. О.И. Шелухина. - М.: Радиотехника, 2003.- 480 с. 4. Большаков, И.А. Прикладная теория случайных потоков / И.А. Большаков, В.С. Ракошиц. – М. : Сов. радио, 1978. 5. Городецкий А.Я., Заборовский В.С. Информатика: фрактальные процессы в компьютерных сетях.-СПб.:Изд-во СПбГТУ, 2000. 102 с. Симонина О.А. Яновский Г.Г. Характеристики трафика в сетях IP // *Труды учебных заведений связи*. — 2004. — № 177. — С. 8—14.