ABOUT METHODS FOR ANALYSIS OF SELF-SIMILARITY OF NETWORK TRAFFIC

Liudvikas Kaklauskas
Šiauliai State College
Lithuania

Annotation
The article research networks traffic self-similarity analysis methods. Applications of Hurst statistics for calculation of Hurst coefficient, frequency/wave features estimators methods - Periodograms, Whittle, Abby-Weich have been analyzed. Suitability of employed methods for analysis was tested by the way of computer-based simulation. The analyzed methods has been tested by applying Fractan programme's R/S statistics, Selfis programme by applying time analysis and frequency/wave features' estimation methods. R. Weron's (2004) algorithm of generation of random standard stable values has been used for forming self-similar network traffic time series, where stability index α=1.8 (H=0.56). The results obtained with Fractan and Selfis show that Hurst coefficient changes from 0.53 to 0.70, the stability index changes from 0.53 to 1.89.
Key words: *self-similarity, Hurst coefficient, α-stable distribution, traffic burstiness*.

Introduction

The self-similarity phenomenon is explained by a character of network service usage which is attributed with burstiness. In fact, data is inherently "bursty" in that it occurs in short bursts of communications followed by long periods of silence. Indeed, one can characterize data communication users who wish network resources to send their data as follows: users don't warn you exactly when they will demand access; one cannot predict how much they will demand, most of the time users do not need access to network; when users ask for it, they want immediate access (Kleinrock, 2002). Such situation is often met in distance learning networks when a learner receives tasks and sends one's answers only at the same time.

Empirical research of computer network packet traffic shows that it is attributed with self-similarity (Erramilli, Narayan, Willinger, 1996, Петров, 2004, Петров, 2003, Park, Willinger, 2000, Leland, Taqqu, Willinger, Wilson, 1994). After estimating the latter feature, it is possible to adequately prognosticate the change of traffic and to apply the prognosis results in increase of network throughput and improvement or its QoS quality of service, while regulating packet latency, fluctuation restriction and packet loss transportation on data and physical OSI layers (He, Gao, Hou, Park, 2004).

In contemporary university studies, computer networks are widely used; they usually undergo unprognosticated overloads. For effective network control, monitoring of network nodes is necessary to be carried out in order to prognosticate network node's loads and overloads. Researches have proved that classical Markov's models which are widely used in estimation of classical telephone network indexes are not suitable for modelling of contemporary computer network parameters. Network parameters' estimations obtained by a classical way are not exact and lead to unreasoned prognoses (Kaj, 2002). On the base of empirical research of 10 Mbps local area network Ethernet network flow carried out at Bellcore laboratory (120 people using network services, New Jersey State) by A. Erramilli, O. Narayan and W. Willinger in 1989, it was found out that Ethernet network flow characterisations bear some fractal features and are attributed with self-similarity with long-range dependence (Erramilli, Narayan, Willinger, 1996). I. Kaj (2002) in the monograph suggests a number of methods for statistical analysis of features of contemporary communication network flows by applying opportunities of contemporary mathematical modelling. V. V. Petrov (2003) analysis network flow as a fractal process attributed with a second-row statistical self-similarity characterised by a fractal measure. Lazaros K. Gallos, Chaoming Song, Herna´n A. Makse (2007) relate fractality with self-similarity of complex networks. For modelling and description of network processes methods of non-linear analysis of self-similar processes are applied, while estimating heavy-tails and regarding asymmetry, leptokurtosis and a short-long-range memory effect which are peculiar to distributions of network flows.

The parameter of self-similarity of flows is an important index of network's operation; and there are no worked out means for its dynamic monitoring in contemporary computer networks. Self-similarity of network flow impact the following: network Quality of Service (QoS), regulation of network flow, bandwidth, decreases loss of packets, decreases delay. Such flows are attributed with heavy tails, asymmetry and leptokurtosis. QoS refers to the capability of a network to provide better service to selected network traffic over various technologies. These technologies allow you to measure bandwidth, detect changing network conditions (such as congestion or availability of bandwidth), and prioritize or throttle traffic. For investigation of peculiarities of self-similar flows, a big number of methods are proposed. Research of self-similarity of a network is a complex task for solution of which methods and technologies are being constantly improved. Design of network self-similarity analysers is an incompletely solved scientific problem, especially in analysis of a network flow in real-time mode. Technologies of self-similarity research is a multiply task encompassing both measurement hardware and software recording and accumulating information on the flow; also, it includes analysers of accumulated data and selection and estimation of theoretical models of a self-similar flow. Familiar computer programmes for network flow research operate with data files prepared in advance; for their management, a graphic interface without team management means, *Fractan*, *Selfis*, etc. is used. Data measurement and accumulation in a computer network must be carried out in real-time mode and not interfere computer's work. Analysis means also must be attributed with pace ensuring dynamic flow analysis. Thus, methods that are reliable but computer time consuming ones are not suitable for analysis as a maximum method; that is why it is best to apply robust analysis methods.

This work aim is by using freely distributed programmes to estimate parameters data flow of network node by applying the algorithm of simulation of time series for validation of results.

1. Methods for analysis of self-similarity

One of the most popular methods for calculation of self-similarity is application of *Hurst* statistics for estimation of *Hurst* coefficient. *Hurst* statistics are applied for time series x_t^{Θ} not satisfying normal distribution (Beran, 1998). According to G. Samorodnitsky (2006), stochastic process $Y(t), t \geq 0$ is self-similar if it is possible to find such H which would satisfy the equation for all *c>0*:

$$(Y(ct), t \geq 0) \overset{d}{=} (c^H Y(t), t \geq 0). \qquad \text{(Samorodnitsky, 2006)}$$

$\overset{d}{=}$ means that this equation is valid in all function's points except equality in distribution. If aggregated time series x_t^{Θ} have stationary increments, thus, partial sums $S_n = X_1 + X_2 + ... + X_n$, where *n=1,2,...,* when $n \geq 1$, o $X_i = Y(i) - Y(i-1), i = 1,2,...$ describe a stationary process $X = (X_1, X_2, ...)$, satisfying the equation $S_n = n^H S_1$. Here, the exponent *H* characterises significance of distribution of a stationary process *X* and is called *Hurst* coefficient. The value of *Hurst* coefficient characterises the type of time series memory. If *Hurst* coefficient *H=0.5*, then members of the sequence are random and every subsequent member does not depend on previous queue members; in an opposite case, we can state that previous events recorded in time series bear a constant impact on further processes, and this impact is the stronger the more event is closer to the past. Such series are invariant with regard to time.

If, then the process characterised by the time series is anti-dispersive, i.e. we can state that if increase is observed in one period, in other period decrease will definitely follow, and the probability is higher the closer *H* is to 0. In this case, correlation is negative and draws closer to 0.5. Such series usually bear a feature of high changeability and are formed of frequent increases and decreases.

If $0{,}5 < H < 1{,}0$, thus it is a persistent process with long-range memory, also called Markov dependence, i.e. if the process was bound to increase in the past, in the future it will retain this peculiarity with the bigger probability the closer *H* is to 1. Usually, such series are called trend resistant, when *H* gets closer to 0.5, more trends (noises) appear in the series. Therefore, while estimating self-similarity of a time series, the value of *Hurst* coefficient, i.e. an interval where it occurs, is very significant. For calculation of *Hurst* statistics, two methods for series estimation are usually applied: time analysis and estimators of frequency/wave features (Karagiannis, Faloutsos, Molle, 2003).

While investigating dependences of gradual selected sequence characteristics and special *m* size block, by applying specific statistics, the following methods are usually applied: absolute value (absolute moments), variance, R/S (rescaled adjusted range), variance of residuals (Taqqu, Teverovsky, 1998, Karagiannis, Faloutsos, Riedi, 2002).

Estimators of frequency/ wave features are grounded on frequency wavelet features. The following methods are usually applied for analysis: *periodograms* (Taqqu, Teverovsky, 1998, Karagiannis, Faloutsos, Riedi, 2002), *whittle* (Karagiannis, Faloutsos, Riedi, 2002), *Abby-Veich* (Karagiannis, Faloutsos, Molle, 2003).

Fractan (2003) programme calculates *Hurst* coefficient only by employing *R/S* statistics. The designed programme *SSE* (Self-Similarity Estimator) employs robust methods of time series calculation; they bear some errors occurring in empirical data, capacities of filtering and high calculation pace: *FamaRoll* (Fama, Roll, 1971), *McCulloc* (1986), *regression* (Belovas, Kabašinskas, Sakalauskas, 2005, Koutrouvell, 1981) and *moments'* (Koutrouvell, 1981, Press, 1972). For calculation of *Hurst* statistics, 11 different methods are applied. We will shortly discuss them.

The oldest and most popular is the *R/S* statistics method grounded on analysis of time sequences, employed in programmes *Fractan* (Fractan, 2003) and *Selfis* (Karagiannis, 2002). Here, formed and aggregated time queues x_t^{Θ} in the network node M, *Hurst* coefficient is calculated according to the formula $H = \log(R/S)/\log(n/2)$, where *H* – *Hurst* coefficient, *R/S* – r/s statistics acquired according to the formula:

$$R/S = \frac{R(n)}{S(n)} = \frac{Max(\sum_{i=1}^{\tau}(x_i^{\Theta} - \overline{x_t^{\Theta}})) - Min(\sum_{i=1}^{\tau}(x_i^{\Theta} - \overline{x_t^{\Theta}}))}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i^{\Theta} - \overline{x_t^{\Theta}})^2}}$$, (Samorodnitsky, 2006)

here $1 \le \tau \le n$, where *n* – number of sequence members, $\overline{x_t^{\Theta}}$ – value of an average row x_t^{Θ}, and $\sum_{i=1}^{\tau}(x_i^{\Theta} - \overline{x_t^{\Theta}})$ – the formed cumulative row describing the sum of changes during time τ. According to *Hurst* (1951), it can be stated that majority of phenomena taking place in nature can be attributed with the right expression: $M\left(\frac{R(n)}{S(n)}\right) \sim cn^H, n \to \infty$, where *c* – constant (Park, Willinger, 2000). It was estimated that when a number of queue members (amount of observations) increases, *Hurst* coefficient gets closer to the value 0.5, i.e. the memory effect decreases.

In the programme *Selfis*, besides discussed *R/S* statistics, six more methods of calculation of *Hurst* coefficient are employed; we will discuss them.

The method of absolute moments (Taqqu, Teverovsky, 1998, Karagiannis, Faloutsos, Riedi, 2002, Taqqu, Teverovsky, 1998, Ilnickij, 2004) is based on *N* length time sequence division into blocks of *m* length, while forming partial sequences $X^{(m)}(k)$, where *k=1,2,...,[N/m].* Then, the n moment of the sequence is calculated:

$$AM_n^{(m)} = \frac{1}{N/m}\sum_{k=1}^{N/m}| X^{(m)}(k) - \overline{X} |^n, kur\ \mathrm{X}^{(m)}(k) = \frac{1}{m}\sum_{i=(k-1)m+1}^{km} X_i$$. (Ilnickij, 2004)

The sequence $X^{(m)}$ behaves asymptotically like $Cm^{n(H-1)}$ for high *m*, thus, the obtained moment $AM_n^{(m)}$ is proportional to $m^{n(H-1)}$.

The method of aggregate variance formed for the sequence $X^{(m)}(k)$ calculates sample variance (Beran, 1994):

$$Var(X^m) = \frac{1}{N/m}\sum_{k=1}^{[N/m]}(X^{(m)}(k) - \overline{X})^2 .$$

The sequence $X^{(m)}(k)$ behaves asymptotically like m^{H-1}, if it has a finite variance, thus, for high *N/m* the sequence of variances asymptotically is proportional to m^{2H-2}.

According to the method of variance of residuals proposed by Peng (Karagiannis, Faloutsos, Riedi, 2002, Taqqu, Teverovsky, 1998), variances of residuals of linear dependence are calculated by the least square method for m-length sequence subsets:

$$\frac{1}{m}\sum_{j=1}^{m}(Y(j) - a - bj)^2, \mathrm{kur}\ Y(j) = \sum_{i=1}^{j} X_i$$. (Peng, Buldyrev, Stanley, Goldberger, 1994)

For all obtained variances proportional to m^{2H}, a common median is calculated and log-log-type dependence is estimated between *m* and incline angle *2H*, if it is linear, then *H* is estimated by regression.

By the method of periodograms an iterative function is described:

$$I(v) = \frac{1}{2\pi N}\left|\sum_{j=1}^{N} X(j)e^{ijv}\right|^2 \text{, (Taqqu, Teverovsky, 1998)}$$

where *v* – frequency, *N* – length of a sequence, *X(j)* – time queue. When *I(v)* has a finite variance, then an iterative function describes density of a sequence *X* which, in case of long-memory is proportional to $|v|^{1-2H}$, sequences close to the beginning of coordinates.

The method of whittle is based on minimisation of maximum probability of a periodogram when a function of spectral density is known (Kokoszka, Taqqu, 1996):

$$Q^*(\eta) = \sum_{j=1}^{[(N-1)/2]} \frac{I(v_j)}{f^*(v_j,\eta)}.$$

Here, η is a function value of a vector minimising the aim function Q^* by calculating which *Hurst* coefficient's value and the relied interval are obtained, when a function of spectral density is known.

By Abry-Veitch (Karagiannis, Faloutsos, Molle, 2003, Karagiannis, Faloutsos, Riedi, 2002, Abry, Veitch, 1998) method, *Hurst* coefficient is estimated while employing transformation of wavelet sequence.

$$\hat{H}(j_1,j_2) = \frac{1}{2}\left[\frac{\sum_{j=j_1}^{j_2} S_j j\eta_j - \sum_{j=j_1}^{j_2} S_j j \sum_{j=j_1}^{j_2} S_j \eta_j}{\sum_{j=j_1}^{j_2} S_j \sum_{j=j_1}^{j_2} S_j j^2 - \left(\sum_{j=j_1}^{j_2} S_j j\right)^2}\right]$$

Here $\eta_j = \log_2\left(\frac{1}{n_j}\sum_k |d_x(j,k)|^2\right)$, weights – $S_j = (n\ln^2(2))/2^{j+1}$, $|d_x(j,k)|^2$ – a measure of process energy during time $2^j k$ sequence $2^j v_0$, when v_0 is selected from the so called mother wavelet, and *n* – length of a partial sequence. The method is widely described in the article by P. Abry and D. Veitch (1998).

According to Samorodnitsky, a self-similar symmetric process that is described by formula and attributed with infinite variance is an α-stable process (Samarodnitsky, Taqqu, 1994), if for every random process *Y(t)* heavy tails can be described according to the formula:

$$P(|Y(t)| > x) \sim cx^{-\alpha} \text{ (Samarodnitsky, Taqqu, 2006),}$$

here $x \to \infty, \text{o}\, c > 0$, thus, when $1 < \alpha < 2$ the mean is finite, and when $0 < \alpha \le 1$ – infinite.

While estimating any stable random value $S_\alpha(\beta,\sigma,\mu)$, it is recommended to estimate four stability parameters:

- *α* – stability index $\alpha \in (0,2]$, also called a tail index, defining burstiness of a process,
- *β* – asymmetry index $\beta \in [-1,1]$, defining shift of a process with regard to zero,
- *σ* – measure index, σ>0 and defines amount of process elements,
- *μ* – position index $\mu \in R$.

Robust time series estimation laws are peculiar with resistance to errors and high calculation pace. In this work, we employ empirical quantum methods in order to estimate parameters of aggregated queues.

One of the oldest employed methods (Fama, Roll, 1971) method based on estimators of stable law parameters *α*, *σ*, *β*, *μ*, when *β=0*, *μ=0*, o $1 < \alpha \le 2$. A stability index is estimated by evaluating a time queue stability index under a condition:

$$S_{\hat{\alpha}}\left(\frac{\hat{x}_p - \hat{x}_{1-p}}{2\hat{\sigma}}\right) = p$$

It was estimated that *p=0.95*, *0.96*, *0.97* were selected best.

In his works, McCulloch (McCulloch, 1986) has improved methods of estimation of stable values designed by FamaRoll, worked out interpolate tables. Two functions calculated by employing time queue quartiles are defined:

$$v_\alpha = \frac{x_{0.95} - x_{0.05}}{x_{0.75} - x_{0.25}} \; ir \; v_\beta = \frac{x_{0.95} + x_{0.05} - 2x_{0.05}}{x_{0.95} - x_{0.05}}$$

Stable parameters are calculated by interpolating values according to given rows.

The regression method for estimation of stable value parameters was proposed by Kotrouvelis (1981), I. Belovas, A. Kabašinskas and L. Sakalauskas described his applications for financial series more widely (2005). For calculation of α and σ the following sums are employed: $s_1 = \sum_{k=1}^{K} y_k w_k$, $s_2 = \sum_{k=1}^{K} y_k$, $s_3 = \sum_{k=1}^{K} w_k$, $s_4 = \sum_{k=1}^{K} w_k^2$, where the parameters employed in the sums are calculated as follows: $w_k = \log|t_k|$, $y_k = \log(-\log(|\phi_n(t_k)|^2, t_k = \frac{\pi k}{25}$. α and σ are calculated by applying formulas:

$$\alpha = \frac{K_{S_1} - s_2 s_3}{K_{S_4} - s_3^2}, \; \sigma = \tilde{\sigma}\left(0.5\exp\left\{\frac{s_4 s_2 - s_1 s_3}{K s_4 - s_3^2}\right\}\right)^{1/\alpha},$$

here $\tilde{\sigma}$ an absolute deviation, *K* recommended value 10.

For calculation of *β* and *μ* the following sums are employed: $s_5 = \sum_{l=1}^{L} u_l^2$, $s_4 = \sum_{k=1}^{K} w_k^2$, $s_6 = \sum_{k=1}^{K} q_l g_n(u_l)$, $s_7 = \sum_{l=1}^{L} q_l u_l$, $s_8 = \sum_{l=1}^{L} q_l^2$, $s_9 = \sum_{l=1}^{L} u_l g_n(u_l)$, where the parameters employed in the sums are calculated as follows: $u_l = \frac{\pi l}{50}$, $q_l = |\sigma u_l|^\alpha \tan\left(\frac{\pi\alpha}{2}\right) sign(u_l)$, o *L=15*. *β* and *μ* are calculated by applying formulas:

$$\beta = \frac{s_5 s_6 - s_7 s_9}{s_5 s_8 - s_7^2}, \; \mu = \tilde{\mu} + \sigma\left(\frac{s_8 s_9 - s_6 s_7}{s_5 s_8 - s_7^2}\right) - h\pi.$$

A more thorough description of the calculation methods can be found in the article by Kotrouvell (1981).

The method of moments for estimation of stable value parameters was proposed by S. J. Press (1972), I. Belovas, A. Kabašinskas ir L. Sakalauskas described his applications for financial series more widely (Belovas, Kabašinskas, Sakalauskas, 2005). This method is based on calculation of a time series empirical characteristic function: $\hat{\phi}(t) = \frac{1}{n}\sum_{i=1}^{n} e^{itx_i}$, where *n* – number of series elements, x_i – the *i*-th series element, *t* – a selected series value. The stable parameters are suggested to be calculated by employing the following formulas (Press, 1972):

$$\hat{\alpha} = \frac{\log \frac{\log|\hat{\phi}(t_1)|}{\log|\hat{\phi}(t_2)|}}{\log\left|\frac{t_1}{t_2}\right|}, \; \log\hat{\sigma} = \frac{\log|t_1|\log(-\log|\hat{\phi}(t_2)|) - \log|t_2|\log(-\log|\hat{\phi}(t_1)|)}{\log\left|\frac{t_1}{t_2}\right|},$$

$$\hat{\beta} = \frac{\frac{\bar{u}(t_4)}{t_4} - \frac{\bar{u}(t_3)}{t_3}}{(|t_4|^{\alpha-1} - |t_3|^{\alpha-1})\hat{\sigma}^\alpha \tan\left(\frac{\pi\hat{\alpha}}{2}\right)} \; \hat{\mu} = \frac{|t_4|^{\alpha-1}\frac{\bar{u}(t_4)}{t_4} - |t_3|^{\alpha-1}\frac{\bar{u}(t_3)}{t_3}}{|t_4|^{\alpha-1} - |t_3|^{\alpha-1}}.$$

Here t_1, t_2, t_3, t_4 – selected values satisfying equations $t_1 \neq t_2$ and $t_3 \neq t_4$, and $\bar{u}(t) = \arctan\left(\frac{\sum_{i=1}^{n} \sin(tx_i)}{\sum_{i=1}^{n} \cos(tx_i)}\right)$. According to the suggestion by Kotrouvelis, the following values are to be selected best: $t_1 = 0.2$, $t_2 = 0.8$, $t_3 = 0.1$, $t_4 = 0.4$ (Koutrouvell, 1981).

2. Testing Analysis methods by Simulation

Suitability of employed algorithms for analysis was tested by the way of computer-based simulation. For simulation of random flows, formulas were applied for generation of standard stable values $S_\alpha(\beta,1,0)$, when $\alpha \neq 1$ (Weron, 2004):

$$X = \mu + \sigma \cdot S_{\alpha\beta} \frac{\sin\{\alpha(V + B_{\alpha\beta})\}}{\{\cos(V)\}^{1/\alpha}} \left[\frac{\cos\{V - \alpha(V + B_{\alpha\beta})\}}{W}\right]^{(1-\alpha)/\alpha},$$

here $B_{\alpha\beta} = \frac{\arctan\left(\beta \tan\frac{\pi\alpha}{2}\right)}{\alpha}$, $S_{\alpha\beta} = \left\{1 + \beta^2 \tan^2\left(\frac{\pi\alpha}{2}\right)\right\}^{1/(2\alpha)}$, o *V* uniformly distributed on $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ and an independent exponential random variable *W* with mean 1. The programme has a foreseen possibility to indicate amount of generated files and amount of elements in every file.

For simulation, series attributed with self-similarity were selected with parameters: *α=1.8 (H=0.56)*, *β=0*, *σ=1*, *μ=0*. Obtained time series were estimated by *Fractan* and *Selfis. Fractan* measures auto-correlation coefficient, *R/S* statistics, fractality, visualises data (draws dependence graphs and attractors). *Selfis* measures *Hurst* coefficient by employing time analysis methods for investigated gradual dependence of selected sequence characterisations and a special m size block by applying specific statistics. The following methods are usually employed for analysis: *absolute value* (*absolute moments*), *variance*, *R/S* (*rescaled adjusted range*), *variance of residuals* (Taqqu, Teverovsky, 1998). Estimations of frequency/wave features are grounded on frequency features of wavelets. The following methods are usually employed for analysis: *periodograms*, *whittle*, *Abby-Veich*.

As we can see in Table 1, *Hurst* coefficient varies from *0.61* to *0.79*; thus, the process of passed data described by aggregated series is a persistent long-range memory process. It can be stated approximately two thirds of received series are attributed with long-term memory. After analysing results of analysis of *Hurts* coefficient obtained by *Fractan* programme, we can state the following:

1. obtained results factually do not depend (matched 98% of the results) on the way of series aggregation (the methods of sum and medium were employed);
2. obtained results do not depend on a size of selected time interval $\Delta t \in$ *[100ms, 500ms, 1s]* and *89.4%* of obtained results matched;
3. obtained results do not depend on a data flow (minimum, medium, maximum ones were selected) and 88% of obtained results matched.

These conclusions show that analysed aggregated series describe a self-similar process attributed with short-range or long-range memory.

In Table 2 one can see that approximately 43% of measured series are attributed with long-range memory, and the medium of coefficient *0.56* shows a weak dependence, approximately 47% of series are attributed with short-range memory, and the medium of coefficient *0.28* shows medium dependence. With regard to obtained results, we can state that the programme *Selfis* estimated that the vast majority of series are attributed with short-range or long-range memory; this proves self-similarity of series. It can be observed that obtained results depend on neither the method of aggregation of series nor time interval; this also proves their self-similarity. Analysis by applying Selfis programme shows the following:

4. obtained results factually do not depend (matched 98% of the results) on the way of series aggregation (the methods of sum and medium were employed);
5. a very weak dependence of obtained results from a size of a selected time interval $\Delta t \in$ *[100ms, 500ms, 1s]* (matched 65.63% of the results) was estimated;

6. a very weak dependence of obtained results from a data flow when minimum, medium, maximum flow is selected (matched 63.5% of the results) was estimated.

Table 1

Distribution of Hurst coefficient values

	0.5<H<1.0									$0 \le H < 0.5$									$1.0 \le H$								
x_t^{Σ}	81.92%									11.30%									6.78%								
	100ms			**500ms**			**1000ms**			**100ms**			**500ms**			**1000ms**			**100ms**			**500ms**			**1000ms**		
	81.36%			76.27%			88.14%			10.17%			15.25%			8.47%			8.47%			8.47%			3.39%		
	Minimal	**Medium**	**Maximal**	**Minimal**	**Medium**	**Maximal**	**Minimal**	**Medium**	**Maximal**	**Minimal**	**Medium**	**Maximal**	**Minimal**	**Medium**	**Maximal**	**Minimal**	**Medium**	**Maximal**	**Minimal**	**Medium**	**Maximal**	**Minimal**	**Medium**	**Maximal**	**Minimal**	**Medium**	**Maximal**
	88.24%	70.37%	93.33%	70.59%	74.07%	86.67%	76.47%	85.19%	93.33%	11.76%	14.81%	0.00%	11.76%	18.52%	13.33%	23.53%	11.11%	0.00%	0.00%	14.81%	6.67%	17.65%	7.41%	0.00%	0.00%	3.70%	6.67%
x_t^{Δ}	80.23%									12.43%									7.34%								
	100ms			**500ms**			**1000ms**			**100ms**			**500ms**			**1000ms**			**100ms**			**500ms**			**1000ms**		
	83.05%			76.27%			81.36%			13.56%			15.25%			8.47%			3.39%			8.47%			10.17%		
	Minimal	**Medium**	**Maximal**	**Minimal**	**Medium**	**Maximal**	**Minimal**	**Medium**	**Maximal**	**Minimal**	**Medium**	**Maximal**	**Minimal**	**Medium**	**Maximal**	**Minimal**	**Medium**	**Maximal**	**Minimal**	**Medium**	**Maximal**	**Minimal**	**Medium**	**Maximal**	**Minimal**	**Medium**	**Maximal**
	88.24%	81.48%	80.00%	70.59%	74.07%	86.67%	88.24%	77.78%	93.33%	11.76%	14.81%	13.33%	11.76%	18.52%	13.33%	5.88%	11.11%	0.00%	0.00%	3.70%	6.67%	17.65%	7.41%	0.00%	5.88%	14.81%	6.67%

It can be stated that the investigated series describe the self-similar process attributed with short-range or long-range memory.

Between *Hurst* coefficient and the parameter alfa, the proportion *H=1/α*, when *1<α<2, β=0* was set by G. Samarodnitsky (2006). In order to more precisely estimate obtained calculation results, values of aggregated *Hurst* coefficient were divided into five intervals: *0<H* of $H \geq 1.0$ – non-defined values, *0<H<0.5* – a series describes the self-similar process with short-range memory, *H=0.5* – a series describes noise, $0.5 < H \leq 0.6$ – a series described the self-similar process with weakly expressed long-range memory, *0.6<H<1.0* – a series described the self-similar process with long-range memory. In order to highlight suitability of applied methods for estimation of time series in the worked out programme, obtained results by every method are displayed in a graph, and calculation results obtained by *Fractan* and *Selfis* programmes are generalised by applying percentage estimations.

Table 2

Distribution of Hurst coefficient values

	$0 \le H < 0.5$									**0.5<H<1.0**									**H=0.5**								
x_t^{Σ}	47.07%									42.82%									8.63%								
	100ms			**500ms**			**1000ms**			**100ms**			**500ms**			**1000ms**			**100ms**			**500ms**			**1000ms**		
	Minimal	**Medium**	**Maximal**	**Minimal**	**Medium**	**Maximal**	**Minimal**	**Medium**	**Maximal**	**Minimal**	**Medium**	**Maximal**	**Minimal**	**Medium**	**Maximal**	**Minimal**	**Medium**	**Maximal**	**Minimal**	**Medium**	**Maximal**	**Minimal**	**Medium**	**Maximal**	**Minimal**	**Medium**	**Maximal**

	x_t^{Δ}			
42.86%	Minimal	100ms	47.05%	42.86%
44.90%	Medium			52.38%
34.29%	Maximal			51.43%
64.29%	Minimal	500ms		64.29%
48.35%	Medium			48.35%
40.00%	Maximal			40.00%
47.62%	Minimal	1000ms		47.62%
53.57%	Medium			45.24%
47.62%	Maximal			31.43%
46.43%	Minimal	100ms	45.05%	42.86%
45.92%	Medium			33.33%
60.00%	Maximal			37.14%
25.00%	Minimal	500ms		25.00%
42.86%	Medium			43.96%
54.29%	Maximal			54.29%
52.38%	Minimal	1000ms		47.62%
38.10%	Medium			44.05%
40.48%	Maximal			57.14%
10.71%	Minimal	100ms	7.90%	14.29%
9.18%	Medium			14.29%
5.71%	Maximal			11.43%
10.71%	Minimal	500ms		10.71%
8.79%	Medium			7.69%
5.71%	Maximal			5.71%
0.00%	Minimal	1000ms		0.00%
8.33%	Medium			10.71%
11.90%	Maximal			2.86%

Conclusions

Applications of Hurst statistics for calculation of Hurst coefficient analyzed and tested by using freeware programme Fractan, the methods of frequency/wave features estimators - Periodograms, Whittle, Abby-Weich have been analyzed and tested by using freeware programme Selfis.

The carried out research show that simulated random flow (α=1.8, H=0.56) investigated by Fractan and Sefis have a Hurst coefficient of which varied from 0.53 to 0.70; this corresponds to variation of the stability index from 1.43 to 1.89.

References

1. Abry, P., Veitch, D. (1998). Wavelet analysis of long-range dependence traffic. IEEE Transactions on Information Theory 4(1) 2–15.

2. Belovas, I., Kabašinskas, A., Sakalauskas, L. (2005). Vertybinių popierių rinkos stabiliųjų modelių tyrimas. Konferencijos “Informacinės technologijos 2005” medžiaga, Technologija, Kaunas, pp. 439-462.

3. Beran, J. (1998). Statistics for Long-Memory Processes. Capman&Hall/CRC, USA, 1998.

4. Erramilli, A., Narayan, O, Willinger, W. (1996). Experimental Queuing Analysis with Long-Range Dependent Packet Traffic. [interactive]. [Accessed on 2020-03-20]. Accessed via the Internet: <http://www-net.cs.umass.edu/cs691s/narayan96.ps>

5. Fama, E. F., Roll, R. (1971). Parameter Estimates for Symmetric Stable Distributions. Journal of the American Statistical Association; Jun 71, Vol. 66 Issue 334, p331, 8p.

6. Fractan 4.4. (2003). Time series analysis program. [interactive]. [Accessed on 2020-03-20]. Accessed via the Internet: <http://www.softoman.ru/windows/education/other/5105.html>

7. Gallos, L. K., Song, C., & Makse, H. A. (2007). A review of fractality and self-similarity in complex networks. *Physica A: Statistical Mechanics and its Applications*, *386*(2), 686-691.

8. He, G., Gao, Y., Hou, J. C., Park, K. (2004). A case for exploiting self-similarity of network traffic in TCP congestion control. Computer Networks: The International Journal of Computer and Telecommunications Networking, vol. 45.

9. Hurst, H. E. (1951). The long-term storage capacity of reservoirs. Transactions of the American Society of Civil Engineers, 116, 770-799.

10. Ilnickij, J. (2004). Research of the Network Server In Self-Similar Traffic Environment. Telecommunications and Electronics, vol.4.

11. Kaj, I. (2002). Stochastic Modeling in Broadband Communications Systems. SIAM, Philadelphia, USA.

12. Karagiannis, T. (2002). The SELFIS Tool [interactive]. [Accessed on 2007-12-15]. Accessed via the Internet: http://alumni.cs.ucr.edu/~tkarag/Selfis/Selfis.html

13. Karagiannis, T., Faloutsos, M. Riedi, R.H. (2002). Long-range dependence: now you see it, now you don't! Global Telecommunications Conference, GLOBECOM '02. IEEE, vol 3, 2165- 2169.

14. Kleinrock, L. (2002). Creating a Mathematical Theory of Computer Networks. Operations Research. Volume 50, Issue 1 Pages: 125 – 131 ISSN:0030-364X

15. Koutrouvell, I. A. (1981). An iterative procedure for the estimation of the parameters of the stable law. Communications in Statistics - Simulation and Computation 10, 17-28.
16. Leland, E., Taqqu, S., Willinger, W., Wilson, D. V. (1994). On the Self-Similar Nature of Ethernet Traffic. IEEE/ACM Transactions on networking, vol. 2, No 1.
17. McCulloch, J. H. (1986). Simple consistent estimators of stable distribution parameters. Communications in Statistics - Simulation and Computation, Volume 15, Issue 4, pages 1109 – 1136.
18. Park, K., Willinger, W. (2000). Self-Similar Network Traffic and Performance Evaluation. John Wiley & Sons, USA.
19. Peng, C. K., Buldyrev S. V., Stanley H. E., Goldberger A. L. (1994). Mosaic organization of DNA nucleotides. Phys. Rev. E, 49:1685 – 1689.
20. Press, S. J. (1972). Estimation in univariate and multivariate stable distribution. J. Amer. Statist. Assoc. 67 842-846.
21. Samarodnitsky, G., Taqqu, M. (1994). Stable Non-Gausian Random Processes. New York: Capman and Hall.
22. Samorodnitsky G. (2006). Long Range Dependence. Foundations and Trends in Stochastic Systems. Vol. 1, No. 3, 163-257.
23. Taqqu, M.S., Teverovsky, V. (1998). Estimating long-range dependence in finite and infinite variance series. In: R. Adler, R. Feldman and M.S. Taqqu (eds), A Practical Guide to Heavy Tails: Statistical Techniques for Analyzing Heavy-Tailed Distributions. Birkhauser, Boston, pp. 177.217.
24. Weron, R. (2004). Computationally intensive Value at Risk calculations, in J. E. Gentle, W. HÄardle, Y. Mori (eds.) Handbook of Computational Statistics, Springer, Berlin, 911-950.
25. Петров, В. В. (2003). То, что вы хотели знать о самоподобном телетрафике, но стеснялись спросить. *М.: МЭИ, ИРЭ.*
26. Петров, В. В. (2004). Структура телетрафика и алгоритм обеспечения качества обслуживания при влиянии эффекта самоподобия: Дис. канд. техн. наук: 05.12.13: Москва, 199 с. РГБ ОД, 61:05-5/1891

Received: 25 March 2020
Accepted: 26 November 2020