`**Original Research Article**

Traffic Modeling Using Raw Packet Generator on Corporate Computer Network

Abstract

Fractal dimension is mathematically defined as a ratio of statistical complexity of network traffic. Its significant manifestation can affect the network performance. In this work, two models of corporate computer networks have been developed using optimized network engineering tool (OPNET) technology. The corporate network is modeled using raw packet generator (RPG) traffic with (H = 0.7, D = 1.3), under the influence of Pareto distribution. Autocorrelation function and power law were used to confirm the presence of fractal traffic on the networks. Average Hurst index (H) of 50 and 100 work stations were estimated using aggregate of variance, absolute moment, periodogram and R/S methods as 0.627, 0.608 and its corresponding fractal dimensions (D) were obtained as 1.371 and 1.391 respectively. These results obtained mean, there is a manifestation of fractal traffic and delay is minimized on the network.

***Keywords*:** Fractal. Traffic, Modeling, Hurst index, Dimension and Network

1.0 Introduction

In modern telecommunication networks being, global system of mobile communication (GSM), corporate computer or general packet radio service (GPRS) network. The traffic structure is becoming complex due to the increasing demand of the internet services. Corporate network is a group of computers that are connected on a local area network such as institutes, companies and ministries. This complexity of the traffic has become an area challenge to the communication engineers. It is noticed that, modern network traffic structure becomes self – similar and fractal in nature [1 – 10], which is likely associated with information transmission delay, over flow probability, information retransmission and may eventually lead to loss of vital information. When any of these conditions occur, it degrades the performance of the network. In this regards, there is a need to develop mathematical models that can capture the traffic behaviors, doing this will provide information about the networks and how to manage them properly.

Different mathematical models have been developed to investigate various properties and characteristics of information flow across telecommunication networks using Poison, Lognormal or Weibull distribution [11 – 16]. However, to date there is no systematic method that has addressed traffic irregularity on telecommunication networks. Therefore, this calls for constant traffic monitoring and modeling in order to maintain the network quality of service (QoS). It is generally known that, Lognormal and Weibull distribution are accepted as the methods of capturing the behaviors of the traffic on telecommunication networks. Furthermore, Lognormal distribution is used as one of the earliest models for investigating fractal traffic and it is also, used to simulate intervals between the request and the web – resources, size of the file transmitted, while Weibull distribution is used for modeling flow protocol blocks like file transfer protocol (FTP). In other hand, modern computer network has aggregated traffic from multiple sources or in other words, many technologies today are harmonized on a single network such as web browsing, audio and video streaming. It is observed by different authors that, lognormal or Weibull can no more fully capture the traffic behavior on the network. In recent studies, A work demonstrates that, local and wide area networks are statistically self – similar and also observed, Hurst parameter is the only parameter for evaluating self – similarity on a computer network [17]. Other authors, tested self – similarity on a web traffic in terms of Hurst index (H) and found that, H = 6.8 using least square fitting line and also noticed that, the network variance slowly decayed rather than exponential decay [18], Another work, addressed long range dependence and self – similarity in web traffic using popular technique called Hurst parameter technique [19]. A research work was carried out on how to calculate fractal dimension on a complex network, the work revealed that covering a network with a reasonable number of boxes can indicate fractal dimension and self – similarity on a network [20]. In another study [21], obtained fractal dimension using regression analysis and another work presents fractal dimensions of percolating networks. [22], worked on how routing protocols induce self – similarity on a wireless network, [23] investigated fractal and multifractal properties on bipartite network and found that, fractality exists on bipartite network and [24] used heuristic algorithm to reduce fractal traffic on a complex network In this regards, it is observed that, all these works mentioned above didn’t use RPG traffic model to minimize delay or confirms the presence of the self – similarity on the network using autocorrelation and power law before evaluating the Hurst index. Therefore, this work proposes, to model and detect fractal traffic on corporate computer network using RPG traffic under Pareto distribution influence to minimize transmission delay on the network and the following objectives shall be realized, firstly, develop corporate network of two different sizes with 50 and 100 work stations to differentiate the traffic density in terms of moderate and high traffic respectively, impose ON and OFF RPG traffic, with Hurst index (H = 0.7) or fractal dimension (D = 1.3), confirm the presence of fractal traffic on the network using autocorrelation and power law, evaluate Hurst index and responding fractal dimension (D), and finally, compare the values of H and D obtained with H = 0.7 and D = 1.3. These value is chosen arbitrary on a scale of $(0.5 \leq H \leq 1.0)$ or $(1 \leq D \leq 2)$ and signifies the level of traffic congestion as well as the network delay. Usually, as H or D approaches 1, network congestion increases [25 – 26]., The contribution of this work lies in optimizing the QoS of a corporate networks

2. Method of generating ON/OFF traffic

Fractal traffic is generated by multiplexing sources of ON and OFF traffic under the influence of Pareto distribution [27], on a network with packet switching; ON traffic is represented as active period (transmitted traffic) and OFF traffic represented as inactive period (no traffic is transmitted) and the average value of Pareto distribution is given by Eq. (1)

$$E(X) = \frac{\alpha b}{\alpha - 1} \qquad (1)$$

the formula for generating Pareto distribution is also given as in Eq. (2)

$$X_{pareto} = \frac{b}{Z^{1/\alpha}} \qquad (2)$$

Where, Z stands for the distributed values between (0, 1]. i.e., the probability that traffic is transmitted can be determined by Eq. (3)

$$l_i = \frac{\overline{ON}}{\overline{ON+OFF}} \qquad (3)$$

The entire traffic that may be transmitted from different sources and can be obtained using Eq. (4)

$$L = \sum_{i=1}^{N} l_i \qquad (4)$$

The average value of the Pareto distribution may be computed using Eq. (5) as

$$E(X) = \int_b^q Xf(x)dX = \frac{\alpha b}{\alpha - 1}\left[1 - \left(\frac{b}{q}\right)^{\alpha-1}\right] \qquad (5)$$

$$\text{And } q = \frac{b}{S^{1/\propto}} \qquad (6)$$

By simplifying Eq. (6), the expression for determining the OFF period is obtained as

$$M_{OFF} = K\frac{T_{OFF}}{T_{ON}} \times \frac{1-S^{T_{ON}}}{1-S^{T_{OFF}}} \times \left(\frac{1}{L_I} - 1\right) \qquad (7)$$

If the α_{ON} and α_{OFF} are chosen to be the same, Eq. (7) will take the form of Eq. (8)

$$M_{OFF} = K \times \left(\frac{1}{L_I} - 1\right) \qquad (8)$$

In practice or real life, the probability of getting OFF period is higher than the probability of getting ON period and to obtain the ON period, OFF period has to be subtracted from 1 as given in Eq. (9).

$$M_{ON} = 1 - M_{OFF} \qquad (9)$$

Eq. (9) is the mathematical model that shall be used to generate the ON/OFF RGP traffic in the OPNET environment.

2.1 Method of simulation

The corporate network is developed as depicted in Fig 1. Firstly, an office topology of $100/100m^2$ is created; required numbers of components are dragged into the work space in OPNET environment, such as switch, work stations, application and profile configurations while the simulation matric parameters of the work stations are set as shown in Table 1. For example, ON state time is set to (10%), OFF state time (90%), interval time (1s), Pareto parameters (10, 0.8) that is, the packet size given by α_{ON} and α_{OFF} respectively and no segmentation is applied, profile configuration is set to define the all profile such as (H=0.7) and application configuration is set to support the profile. Then simulation is applied.

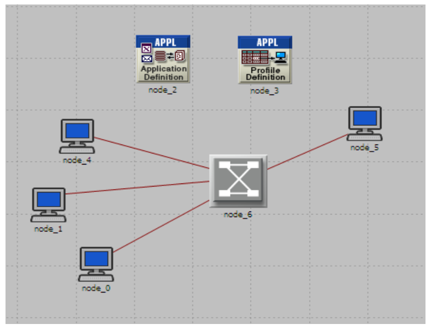

Fig. 1 Sample of corporate network

Table 1 Simulation Matric Parameters

Traffic Property	Description
Work station	Ethernet stations
Start time	Constant (5s)
ON time	Exponential (10%)
OFF time	Exponential (90%)
Inter arrival time	Exponential (1s)
Packet (byte)	Pareto (0.1 0.8)
Segmentation	No
Stop time	Never

2.2 **Confirmation of fractal traffic on the simulated network**

As earlier stated**,** an autocorrelation and power law shall be used to confirm the fractal nature of the traffic on the network, One observation about autocorrelation, is that as it coefficient decays to zero [28, 29] fractal traffic increases on the network, while in the case of power law, the tail of the distribution will become heavily distributed towards y – axis [30] and they can be computed using Eqs. (10) and (11) respectively.

$$r(k) = \frac{1}{N-\tau} \frac{\sum_i (Xx_i - \overline{X})(X_{i+k} - \overline{X})}{\sigma^2(X)} \quad (10)$$

$\overline{X}$ – Selective medium range X, $\sigma^2(X)$ – Number of sample variance of X, $k \in Z+ = \{0,1,2\ldots\}$

Power law expresses functional relationship between two quantities. Where a change in one quantity results in the change of the other quantity.

$$y = mX^p \quad (11)$$

2.3 **Methods of estimating Hurst index**

There are many methods of estimating Hurst index these include method of Absolute Moments, Variance of Residuals and Abry - Veicth Estimator, Whisttle and Peiodogram methods. Furthermore, the following methods shall be considered to evaluate the Hurst index of the traffic obtained from OPNET environment..

I) The Variance method is given by

$$\sigma^2(X_i^m) \sim am^{-\beta},\ \ \beta \to \infty \qquad (12)$$

It is possible to evaluate β for all values of m by taking the log of both side of Eq. (12) as shown in Eq. (13)

$$log[\sigma^2(X_i^m)] \sim\sim \beta log(m) + log(a), m \to \infty \qquad (13)$$

Variance method uses the logarithmic sample variance to equalize the level of aggregation, which is expected to give a straight line with a slope ≥ -1.Where $H = 1 + \beta/2$ and X^m is the variance of the combine processes, m is the size interval, β is the slope of the straight line and a is the finite positive constant

II) R / S plot method, is the ratio of rescale adjusted range given by

$$M\left[\frac{R(n)}{S(n)}\right] \sim cn^H, \text{ as } n \to \infty \qquad (14)$$

Equation (14) can be further evaluated by estimating, H, and also by taking the log of both side of Eq. (14).

$$log\,\langle M\left[\frac{R(n)}{S(n)}\right]\rangle \sim Hlog(n) + \log(c),\ \ n \to \infty \qquad (15)$$

It is expected that, logarithmic samples of the R / S statistics in the Eq. (15) with the number of the aggregated series may give a straight line with a slope$H$.

III) Periodogram Method. The method plots the logarithm of the spectral density of the time series verves the logarithm of frequency. The slope will provide an estimate as given by N in Eq. (16)

$$I(v) = \frac{1}{2\pi N}\left|\sum_{j=1}^{N} X(je^{ijv})\right|^2 \qquad (16)$$

Where v is the frequency and N is the Length of time series,

IV) Whittle Estimation, is done based on minimizing the likelihood function, which is applicable to the period of time series, evaluated H dependence on the confidence interval.

2.4 **Hurst index and fractal dimension calculation**

Usually, H is examined in three classes such as anti-persistent, random and persistent in the range of 0 – 0.4, 0.5 and 0.6 – 1.0 respectively [31] and fractal dimension can be obtained from Eq. (17) given by

$$D = 2 - H \qquad (17)$$

where D ranges from $1 \leq D \leq 2$, which translates that, as D approaches 1 there is a manifestation of fractal traffic on the network as earlier mentioned. .

3.0 **Results and discussions**

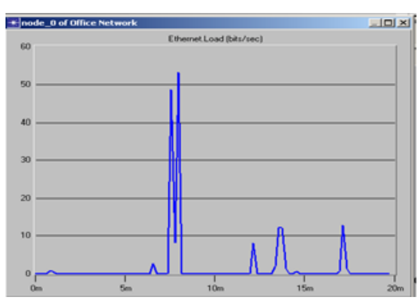

Fig. 2a Pareto (10, 0.8)

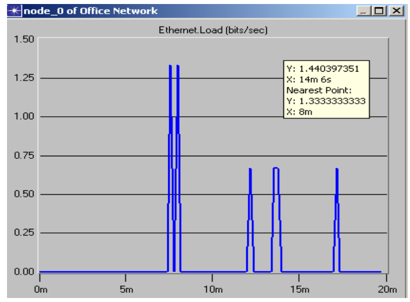

Fig. 2b Pareto (10, 1.2)

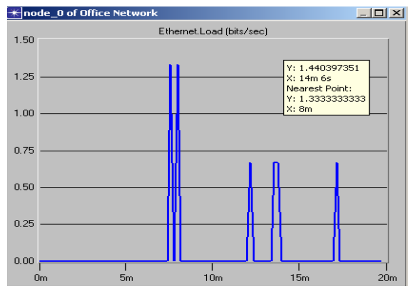

Fig. 2c Pareto (10, 1.6)

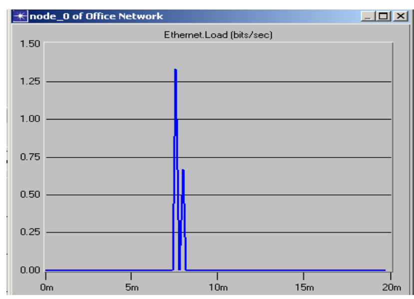

Fig. 2d Pareto (10, 1.8)

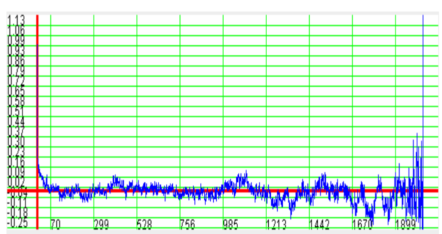

Fig. 3 Autocorrelation coefficients

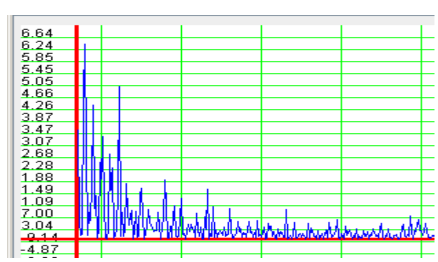

Fig.4 Power law

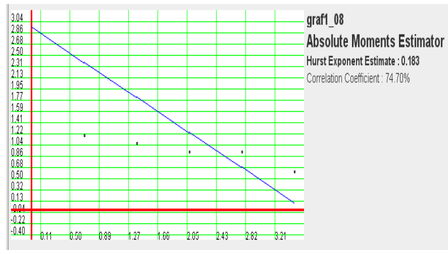

Fig. 5a Aggregate variance estimator

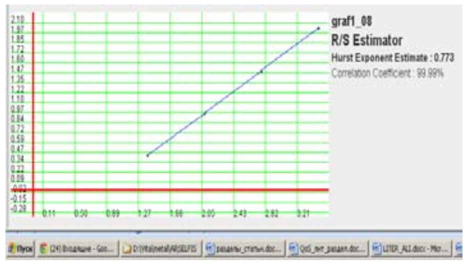

Fig. 5b R/S method estimation

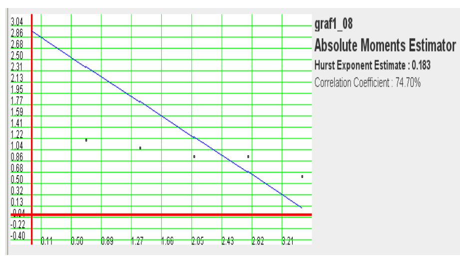

Fig. 5c Absolute moment estimator

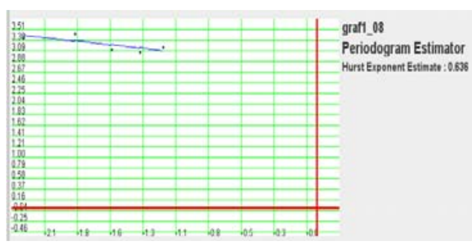

Fig. 5d Periodogram method

Figs 2 (a –d) show the traffic generated after the simulation using different Pareto parameters, for example, (10, 0.8), (10, 1.2), (10, 16) and (10, 1.8). Fig. 3 depicts the variation and weak decay of the autocorrelation coefficients, which signifies presence of fractal traffic on the network. Fig. 4, shows the power law characteristics, as it can be seen, the heavy tail is skewed towards y – axis which also indicates presence of fractal traffic on the network. Figs. 5 (a –d) show how Hurst index is estimated using Eq. (13, 15, 16). Average the several methods are considered because each method has its accuracy and error. The Hurst index estimated are summarized in Tables 2 and 4 for 50 and 100 work stations while the corresponding fractal dimensions is computed using Eq. (17) and are also, summarized in Tables 3 and 5 . The study developed a model of a corporate computer network in terms of fractal traffic. Moreover; fractal properties vary depending on the characteristics of ON-OFF period.

Table 2 Hurst parameter with 50 work stations

Time setting (ON/OFF)	A	R/S	P	AM	Ave (H)
Pareto (10,0.8)	0.647	0.773	0.636	0.628	0.671
Pareto (10,1.2)	0.640	0.699	0.576	0.608	0.631
Pareto (10,1.6)	0.637	0.669	0.518	0.546	0.593
Pareto (10,1.8)	0.598	0.671	0.625	0.559	0.613

Aggregate variance estimator (A), adjusted rescale estimator (R/S), Periodogram estimator (P) and Absolute moment estimator (AM)

Table 3 Fractal dimension with 50 work stations

Time setting (ON/OFF)	A	R/S	P	AM	Ave (D)
Pareto (10,0.8)	1.353	1.227	1.364	1.372	1.329
Pareto (10,1.2)	1.360	1.301	1.424	1.392	1.369
Pareto (10,1.6)	1.363	1.301	1.482	1.454	1.386
Pareto (10,1.8)	1.402	1.329	1.375	1.441	1.398

Table 4 Hurst parameter with 100 work stations

Time setting (ON/OFF)	A	R/S	P	AM	Ave (H)
Pareto (10, 0.8)	0.607	0.673	0.536	0.646	0.616
Pareto (10, 1.2)	0.590	0.601	0.555	0.618	0.591
Pareto (10, 1.6)	0.677	0.569	0.618	0.646	0.628
Pareto (10, 1.8)	0.508	0.682	0.535	0.659	0.596

Table 5 Fractal dimension with 100 work stations

Time setting (ON/OFF)	A	R/S	P	AM	Ave (D)
Pareto (10, 0.8)	1.393	1.327	1.464	1.354	1.376
Pareto (10, 1.2)	1.410	1.399	1.445	1.382	1.409
Pareto (10, 1.6)	1.333	1.431	1.382	1.354	1.375
Pareto (10, 1.8)	1.492	1.318	1.465	1.341	1.404

As shown, in Tables 2 – 5, there is a moderate manifestation of fractal properties on the corporate computer network. However, the network developed was modeled with RPG traffic of H = 0.7 or equivalent, D = 1.3 and the results in Tables 1 and 3 show that, average value of H = 0.627 and 0.608 for 50 and 100 work stations respectively and Table 2 and 4 show that, D = 1.371 and 1.391 for 50 and 100 work stations respectively. In comparison, H obtained in this work, is less than 0.7 or D is greater than 1.3. This means, the modeling results shows that, fractal dimension is minimized by 5. 5% and 7.0% for 50 and 100 work stations respectively. And it is worthy to say network congestion is minimized as well as delay.

4.0 **Conclusion**

Two models of computer networks with 50 and 100 work stations have been developed using OPNET technology to capture traffic behavior on the networks using Pareto distribution in terms of RPG traffic (ON and

OFF). After the simulation, auto – correlation, power law were used to confirm the presence of fractal traffic on the network before evaluating the Hurst index using aggregate of variance, absolute moment, periodogram and R/S methods and its corresponding fractal dimensions. The results of finding observed following

- There is a moderate manifestation of fractal traffic on the network
- Increasing number of work station does not induce fractal traffic on the network
- Transmission delay is minimized
- Pareto distribution model is capable of capturing traffic behavior on a modern computer network. Therefore it is recommended that, Pareto distribution models are better than Poison, Lognormal and Weibull models in analyzing traffic on modern computer network

References

[1]. Bo. R. and Lowen S. (2000). Fractal Traffic Models for Internet Simulation *Proc. of 5th IEEE Symposium* No. 150: 200- 206

[2]. Keltani H and Aubner J. A. (2002). A Noval Approach to the Estimation in self-similar Traffic .*Procedures of the 27th Annual IEEE Conference on Local Computer Network*, Tampa USA Nov. 6-8: 160-165

[3]. Willinger W., Paxson V., Reedi R. H. and Taggu M. S. (2003). Long Range Dependence and Data Network Traffic Theory and Application of Long Range Dependence, 373-407

[4]. Afshin S. and Fatemeh J .(2010). Network Traffic detection based on fractal dimension. Vol. 1: 27 - 32

[5]. Can I., Cleveland W. S., Giao Y., Jeffrey K., Smith F. D. and Weigle M. C. (2004). Stochastic models for generating synthetic HTTP source traffic. *Proceeding of IEEE INFOCOM*, 1547 - 1558

[6]. Millan G., Kaschel H. and Lefraine G. (2010). Discussion of Self – Similarity Teletraffic with long range dependence at the network layer levels. *International Journal of Computer Communications and Control*, 5(5): 799 – 812.

[7]. Ruogu Yan and Yingfeng Wang. (2012). Hurst Parameter for Security Evaluation of LAN Traffic. *Information Technology Journal,* Vol. 11: 269-275

[8] Byung C. K. and Sunghwan J. (2013). Effective Immunization of online networks a self – similar section approach. *Information Technology Manag*.

[9]. Domauska J., Domauska A. and Czachorsk T (2014). A new investigation of long range dependence in network traffic. *29th International Symposium on Computer and Information Science. Information Science and System part III.* Springer, Heidelberg, 137 – 144

[10]. Dymora P., Mazurek M. and Strzalka D. (2013). Computer network traffic analysis with the use of statistical self – similarity factor. *ANN UMCS Information AI*, 13(2): 69 – 81

[11]. Paxson V, and Floyd S. (1997)). Wide Area Traffic: the Failure of Poisson Models,. *IEEE/ACM Transaction Network*, Vol. 3: 226-224

[12]. Taqqu M. S., Willinger W., Sherman R. (1997). Proof of a fundamental result in self-similar traffic modeling. *SIGCOMM Computer. Communications. Rev*., 27(2.):: 5 - 23.

[13]. Clegg R. G. (2004). Statistics of dynamic networks. PhD thesis. Dept of Mathematic University of York .

[14]. Olejruk R. and Chackler R. (2011). The nature of HTTP traffic in wireless LANs metody. *Informaty Stosowanaj 4,* 175 - 180

[15]. Danladi A, Gnatushenko V. V. and Abdullahi M. (2014). Neural network random error prediction on wireless local area network and application of one-dimensional multi resolution wavelet De-noising technique. *Proceedings of World Congress on Engineering and Computer Science.* San Francisco, USA 22 - 24 October, Vol 1:

[16]. Damauska J., Damauska A. and Czachorski T. (2015). Estimating the intensity of LAN in real and synthetic traffic trace. *CCIS*, Vol 522: 11 - 22

[17]. Xiarong C., Kun X. and Dong W. (2009). Estimation of network traffic Hurst parameter using HHT and wavelet transform. WICOM' *5th International Conference on Wireless Communication networking and Mobile Computing IEEE* Beijing China, 24 – 26

[18]. Gagandeep K., Vikas S. and Gupta J. P. (2010). Characteristics analysis of web traffic with Hurst index. *Proceeding of World Congress on Engineering and Computer Science.* San Francisco, USA. October, 20 – 23. Vol 1

[19]. Grazyna S. and Adam D. (2016). Investigating long range dependence in e – commerce web traffic. *23rd International Conference CN*, 42 – 55

[20]. Chaoming S., Lazaros K. G., Shlomo H. and Heman A. M. (2007). How to calculate the fractal dimension of a complex network: the box covering algorithms. *Journal of Statistical Mechanism theory and experiments*, Vol. 2007

[21]. Vlastimil H. and Petr S. (2014). Surface evaluation by estimation of fractal dimension and statistical tools. *The Scientific World Journal,* Vol. 2014: 1- 10

[22]. Ming L., Wei Z., Weijia J., Dongyang L. and Chi – Hung C. (2003). Modeling autocorrelation function of self – similar teletraffic in communication networks based on optimal approximation in Hilber space, *Applied Mathematical Modeling,* 27(3): 155 – 168

[23] Danladi A., Silikwa N. W. and Michael Y. Routing protocols source of self – similarity on the network. In press. *Alexandra Engineering Journal*.

[24] Jin-Long L., Jian W. and Xian-Hua. Fractal and multifractal analyses of bipartite networks. Scientific Reports, 2017, 7: 1-29

[25]. Ming L. (2008). Modeling autocorrelation functions of long range dependent teletraffic series based on optimal approximation in Hilber space: A further study. *Applied Mathematical Modeling*, 31(3): 625 – 631

[26]. Ming L. and Lim S. C. (2008). Modeling network traffic using generalized Cauchy Process. *Physics A: Statistical Mechanics and its Application*, 387(11): 2584 – 2594

[27]. Arrost L. S. and David S. (2003). Power law area relationships and self – similar species distribution within finite areas. *Ecology Letter*, 7(1): 60 – 68

[28]. Daniel A. M. (2011). A verification of selected properties of telecommunication traffic generated by OPNET Simulator. Erasmus Exchange Project Work. University of Ljubljana, Faculty of Engineering, 1 – 55

[29]. Chaomings S., Shlomo H. and Hernan A. M. (2006). Origin of fractality in the growth of complex networks. *Nature Physics*, 2(4): 275 - 281

[30]. Guo L. and Cai X. (2009). The fractal dimensions of complex networks. *CHIN. Physics Letters*, 26(8): 1 – 3

[31]. Roland M. (2013). Networks and fractals. BSc Thesis, Budapest University of Technology and Economics, Institute of Mathematics, Department of Stochastic, 1 – 55.