Designing an Approach for Network Traffic Anomaly Detection

Seyed Mahmoud Anisheh
Department of Computer and Electrical Engineering, Noushirvani University of Technology, P. O. Box 47144, Babol, Iran

Hamid Hassanpour
School of Information Technology and Computer Engineering, Shahrood University of Technology, P.O. Box 316**,** Shahrood, Iran

ABSTRACT
The aim of this research is to analyze aggregate network traffic for anomaly detection. The accurate and rapid detection of network traffic anomaly is crucial to enhance the effective operation of a network. It is often difficult to detect the time when the faults occur in a network. In this paper, a new algorithm is presented to monitor the aggregate network traffic to rapidly detect the time anomaly occurs in a network. This is accomplished by monitoring the statistical characteristics of the time series representing the network behavior. The technique analyzes the network behavior using fractal dimension and discrete stationary wavelet transform. In the proposed method, after applying discrete stationary wavelet transform on the signal representing the network traffic, the fractal dimension of the decomposed signal is calculated in a sliding window. Then, variations of signal fractal dimension are considered for anomaly detection. Performance of the proposed method is compared with that of three other existing methods using both synthetic signal and real data. The results indicate superiority of the proposed technique in terms of accuracy compared to existing methods.

Keywords anomaly detection, effective operation of the network, fractal dimension, wavelet transform

1. INTRODUCTION
The network traffic anomaly refers to the situation when network traffic departs from its normal behaviors [1-3]. Anomaly detection techniques attempt to provide normal activity profiles by considering various measures, and an anomaly is detected once the actual system behavior deviates from the normal profiles.

Anomaly in a network is either host-based or network-based [1-5]. The former systems run on a local monitored host and use its log files as information sources. This anomaly detection system has limitations in detecting distributed and coordinated attacks. In contrast, network-based anomaly detection try to protect the entire networks against intrusions by watching the network traffic either on designated hosts or particular sensors; therefore, it can protect simultaneously a group of computers employing different operating systems against remote attacks such as port scans.

Many reasons such as defective components, their transient failure, software failures, invasion of the network, and misuse of network equipments may cause network traffic anomaly behaviors [1-5]. Fast and accurate detection of network traffic anomaly is very important in network management and improvement of network performance [1-3]. It is often difficult to implement this aims, therefore detection of network traffic anomaly has become the interesting and important subject in the present academic and industrial researches. Signal processing techniques can be used to analyze and detect network anomalies due to their potential to find novel or unknown anomalies. Anomaly detection algorithms are based on the premise that the statistical characteristics of the Management information Base (MIB) variables change in response to fault occurrence [1-3].

In this paper, we present an anomaly detection algorithm based on discrete stationary wavelet transform (DSWT) and fractal dimension (FD). Wavelet technique has been used to exhibit the important underlying unadulterated form of the time series. This pre-processing step is used to increase the accuracy of the proposed method in anomaly detection. The main advantage of DSWT is the preservation of time information of the original signal sequence at each level compared to classical wavelet transform such as proposed in [6]. In the next step, the fractal dimension of the decomposed signal is calculated. For application of anomaly detection fractal dimension variations can be used as the change in statistical characteristics of a signal affect on its corresponding fractal dimension. Network traffic experiments have demonstrated that the proposed method is efficient in anomaly detection. In addition, a comparative study between four typical wavelet basis functions on the proposed method accuracy has been performed when applying wavelet techniques for detecting network anomaly.

This paper is organized as follows: the existing anomaly detection methods are described in Section 2. Section 3 presents the theoretical foundation of wavelet analysis and fractal techniques. The proposed anomaly detection method is explained in Sections 4. The performance evaluation of the proposed method and comparison results are provided in Section 5. Finally, conclusions of the paper are drawn in Section 6.

2. RELATED WORK
There are a number of anomaly detection methods in the literature. In generalized likelihood ratio (GLR) method [3], two successive windows, namely *R* and *S*, are slided along the signal.

$$R = \{r_1, r_2, ..., r_p, r_{p+1}, ..., r_{N_R}\} \quad (1)$$

$$S = \{s_1, s_2, ..., s_p, s_{p+1}, ..., s_{N_R}\} \quad (2)$$

Where, N_R represents the length of the sliding windows. For each window, the observations are modeled as autoregressive

(AR) order *p* process. The following function, namely G function, is used to detect anomaly points:

$$G = N'_R\left(\ln\frac{\hat{\sigma}_p}{\hat{\sigma}_R}\right) + N'_S\left(\ln\frac{\hat{\sigma}_p}{\hat{\sigma}_S}\right) \quad (3)$$

where $N'_R = N_R - p$ and $\hat{\sigma}$ denotes variance estimate. Local maxima of *G* function above the threshold value are considered as anomaly points. In [7], the authors proposed an anomaly detection method, namely WGLR, which combines wavelet transform and GLR method. The authors have used the wavelet transform for its strong ability in detecting abrupt failure points. It could also extract the transient characteristic of the signal in short time. It has been shown that the accuracy of this method is superior to GLR method. In another research, Error performance Detection (EPD) method is introduced for anomaly detection [7]. This technique is based on the prediction error of the traffic model and regards the error as a statistical variable. The points where error values exceed the predefined threshold are regarded as anomaly points.

In [1], an anomaly detection method has been proposed which detects traffic anomalies by computing and analyzing network traffic signal instantaneous parameters (frequency and amplitude) obtained by Generalized Hilbert Transform of original traffic data. It has been revealed that anomaly points would be more evident through analyzing the instantaneous parameters of the original network flow data.

3. OVERVIEW OF DISCRETE STATIONARY WAVELET TRANSFORM AND FRACTAL DIMENSION

Since this research is based on discrete stationary wavelet transform and fractal dimension, they are briefly reviewed in this section.

3.1 Discrete Stationary Wavelet Transform

The wavelet transform can provide arbitrary signal characteristic of time-scale domain, which can help to extract the transient abnormal phenomenon from normal signals, while Fourier transform does not have this characteristic [9-13]. Discrete wavelet transform (DWT) is the decomposition of a time domain signal *x* into the so-called approximates and details signals. The decomposition of the signal into different frequency bands is obtained by low-pass and high-pass filtering of the time domain signal. In this approach, the obtained results at each level are half the size of the original sequence. The stationary wavelet transform (SWT), on the other hand, pads the corresponding high-pass and low-pass filters with zeros to retain length of the signal [14]. Consequently, the major advantage of SWT is the preservation of time information of the original signal at each level. This property is useful for applications such as breakdown point detection [14].

3.2 Fractal Dimension

Fractal dimension value is an index for measuring the complexity of an object [15, 16]. Its applications have been considered in different fields such as criminology, epidemiology, economy, social and behavioral sciences [15, 16]. It has been shown that FD is a promising method in transient detection [17-19]. In addition, in this approach there is no need to have a prior knowledge about the characteristics of the transient. Several algorithms have been suggested to compute the fractal dimension of waveforms such as Higuchi [20], Petrosian [21] and Katz [22] methods. Selection of FD algorithm depends on the application [23]. Higuchi's method presents the most accurate estimation of the FD, but it is slower than Katz's method. Petrosian's method has less accuracy in calculating FD of a signal compared to Katz's method [23]. In this paper, we have used Katz's algorithm to calculate the FD of a time series. In this method, FD of a time series is defined as below [22]:

$$FD = \frac{\log(L)}{\log(d)} \quad (4)$$

where *L* is the length of the time series, *d* is diameter estimation of the distance between the first data point and the data with the highest distance. By normalizing the distance, by *a* as the average distance between the two successive data points, the following equation is obtained.

$$FD = \frac{\log(L/a)}{\log(d/a)} \quad (5)$$

The above equation is known as Katz's method to calculate FD of the time series.

4. PROPOSED ANOMALY DETECTION METHOD

The proposed anomaly detection method consists of three steps as described below:

I. The analyzing signal is initially decomposed into different frequency bands using discrete stationary wavelet transform. Wavelet technique has been used to reveal the important underlying unadulterated form of the data since details have been removed during filtering. In this paper, DSWT has been used to preserve the time information of the original signal sequence at each level. This is the main advantage of using DSWT in pre-processing step instead of using classical wavelet transform such as proposed in [6].

II. Two successive windows are slided along the signal. For each window, FD is computed using the Katz algorithm. As mentioned before, the changes in the statistical characteristics of the signal are reflected on the signal fractal dimension. Therefore, for application of anomaly detection, variations of signal fractal dimension can be considered as follows:

$$G_k = \left|FD_{k+1} - FD_k\right| \quad k = 1,\ldots,N \quad (6)$$

where, *N* is the number of samples in the *G* function.

III. The local maxima of *G* function that are above the threshold represent the time instants of fault occurrence. Threshold value can be chosen automatically according to computational results.

In the proposed approach, a criterion has also been employed to choose a proper length for the sliding window which increases the accuracy of the proposed method. For an analyzing window with length *l*, the energy of the corresponding *G* function, G^l, is calculated as below:

$$E_{G^l} = \frac{\sum_k \left|G_k^l\right|^2}{N} \quad (7)$$

Where N is the number of samples in G^l. The proper window length is the minimum point of the normalized energy function, E_{G^l}, versus window length [6].

5. NETWORK TRAFFIC EXPERIMENT

Algorithms of the proposed method and existing techniques were implemented using MATLAB from Math Works, Inc. The performances of these techniques have been evaluated using both synthetic signal and real data.

5.1 Synthetic Signal

For the purpose of evaluating the performance of the proposed method, it was applied on a synthetic signal with network traffic behavior. Firstly, it is required to generate network traffic data. It has been shown that local-area network traffic can be modeled using statistically self-similar processes. In [8], an algorithm for generating approximate sample paths for a type of self-similar process known as fractional Gaussian noise (FGN) has been presented. In this algorithm, Giving Hurst parameter (H) and number of samples in synthetic data (N), synthetic traffic data can be generated.

Figure 1.a shows a three-component signal that has been generated using the method proposed in [8]. In this experiment, H and N are chosen as H=0.8 and N=200, respectively. In order to make anomaly in the signal, values of one component is chosen zero after the 100th data point. Therefore, due to the changes in statistical characteristics of the synthetic signal, it is expected that this point be detected by the proposed method. The original signal is initially decomposed using a one level DSWT and the approximate sub-band for the signal in Figure 1.a is shown in Figure 1.b. In this experiment, DSWT is performed with Daubechies wavelet of order 1.

In this experiment, using equation (7) leads to achieving an optimum window length with the length of 38 samples. FD of the decomposed signal and G function are computed using the optimum window length and the results are demonstrated in Figure 1.c and 1.d, respectively. The threshold value for the proposed method is chosen as $\overline{G}+2\sigma_G$, where $\overline{G}$ and σ_G represents the mean and standard deviation of G function, respectively. It has been experimentally chosen to have a better accuracy in anomaly detection. As can be seen from the result, the anomaly position has been successfully detected using the proposed method.

The existing approaches, the method in [1], WGLR and EPD methods [7], are also applied on the signal in Figure 1.a and the results are shown in Figure 2. This Figure indicates that the existing methods have more false detection rate (marked with the arrows in the plots) compared to the proposed method.

For the purpose of evaluating the performance of the proposed method, the true positive (TP), miss or false negative (FN) and false alarm or false positive (FP) ratios are used as defined below [24]:

$$TP = \frac{N_t}{N}, \qquad FN = \frac{N_m}{N}, \qquad FP = \frac{N_f}{N}, \qquad (8)$$

where N_t, N_m, N_f and N represent the number of true, missed, falsely detected and actual number of detected boundaries, respectively. An efficient anomaly detection approach should have a high value for TP ratio and low values for both FN and FP ratios. The proposed method and the three other existing approaches are applied on a set of 100 synthetic signals similar to the one shown in Figure 1.a, but anomaly occurs at random position in the signal. In order to compare the performance of different methods, Table 1 has been presented to address the detecting result belongs to different algorithms.

This table reveals that the proposed method has a better accuracy compared to other existing methods. For example, the FP ratio for the proposed method is 24 times lower than the method in [1], 12 times lower than WGLR method [7] and 44 times lower than EPD method [7]. As can be seen from the table, EPD method offers a high true-positive rate at the expense of a high false positive rate and WGLR method has the lowest TP ratio.

Table 1. Results of anomaly detection using proposed method, and three other existing approaches with 100 realizations for synthetic signals.

Method	TP ratio	FN ratio	FP ratio
Proposed Method	0.9667	0.0333	0.2667
Instantaneous Parameters	0.7667	0.2333	6.3000
WGLR	0.7000	0.3000	3.1333
EPD	0.9333	0.0667	11.8667

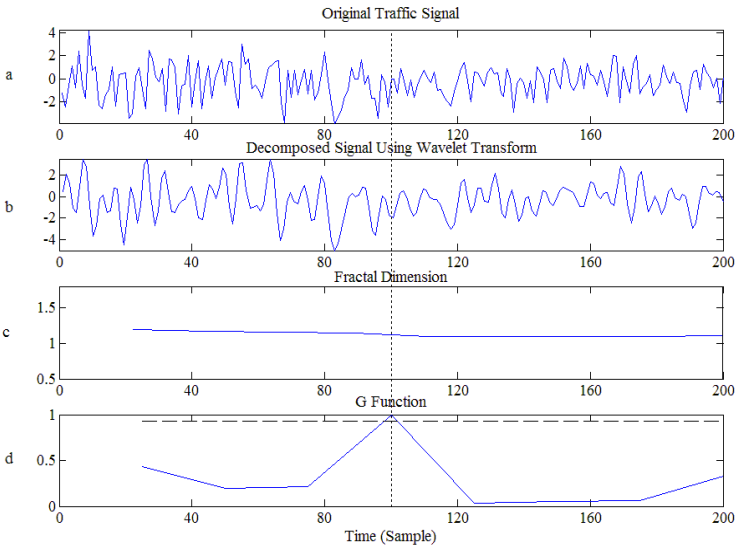

Fig 1: Anomaly detection in synthetic signal using proposed approach. (a) Original traffic signal. (b) Approximate signal after applying one level DSWT, (c) Fractal dimension computed using optimal window length, (d) Computed G function.

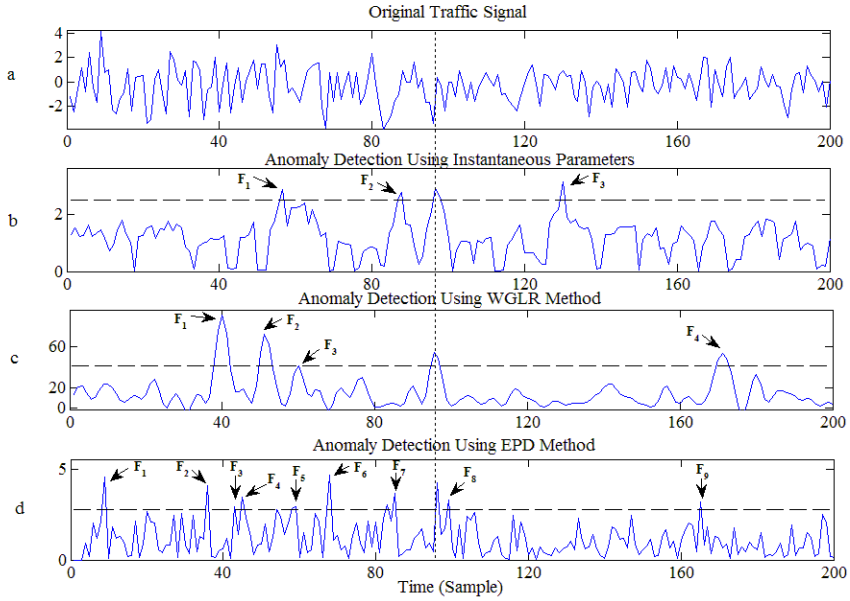

Fig 2: Anomaly detection in synthetic signal using the exiting approaches, (a) Original traffic signal, (b) Output of method based on instantaneous parameters, (c) Output of WGLR method, (d) Output of EPD method.

In following, impact of wavelet basis functions on the anomaly detection accuracy is considered. Table 2 reveals the results of applying proposed method on a set of synthetic data when four different wavelet basis functions, namely Daubechies1, Coiflets1, Symlets2 and Discrete Meyer are used. On the basis of experiment we can see that Daubechies1 basis function achieves the best results compared to other three wavelet basis functions.

Table 2. Results of anomaly detection using proposed method when four different wavelet basis functions are used.

Wavelet Basis	TP ratio	FN ratio	FP ratio
Daubechies1	0.9667	0.0333	0.2667
Symlets2	0.9000	0.1000	0.3333
Discrete Meyer	0.9333	0.0667	0.4667
Coiflets1	0.7333	0.2667	1.1000

5.2 Real Traffic Signal

To further evaluate the performance of the proposed method, it was applied on the real network traffic data. The experiment has been conducted on local area network (LAN) of Noushirvani University of Technology. A network topology is the pattern in which nodes (i.e., computers, printers, routers or other devices) are connected to a LAN or other network via links [3, 7]. The network topology of Noushirvani University of Technology is shown in Figure 3. In order to monitor the network traffic, we utilize PRTG software that uses SNMP protocol. The SNMP provides an organized structure to MIB variables as well as a mechanism for communication [3, 7]. We have collected the data values for two MIB variables, incoming and outgoing link utilization in Mbps with sampling interval of 10 seconds. Figures 4.a-d shows samples of recorded traffic signals from links 1-4 (links are shown in Figure 3), respectively where the lengths of the signals are 15 minutes.

In order to make anomaly signal, link 1 has been disconnected temporally. This failure region has been shown in Figure 5.a using an oval shape. In order to verify the performance of the proposed method to detect the fault, we have applied it on the signal in Figure 5.a. This signal has been decomposed using a one level DSWT, and the approximate sub-band is depicted in Figure 5.b. FD and G function of the decomposed signal are computed using the optimum window length following equation (7) and the results are demonstrated in Figure 5.c and 5.d, respectively. After thresholding, the local maximum in Figure 5.d represents anomaly point. This figure reveals that if a fault occurs in the network, the local maximum of the G function indicates failure point nicety.

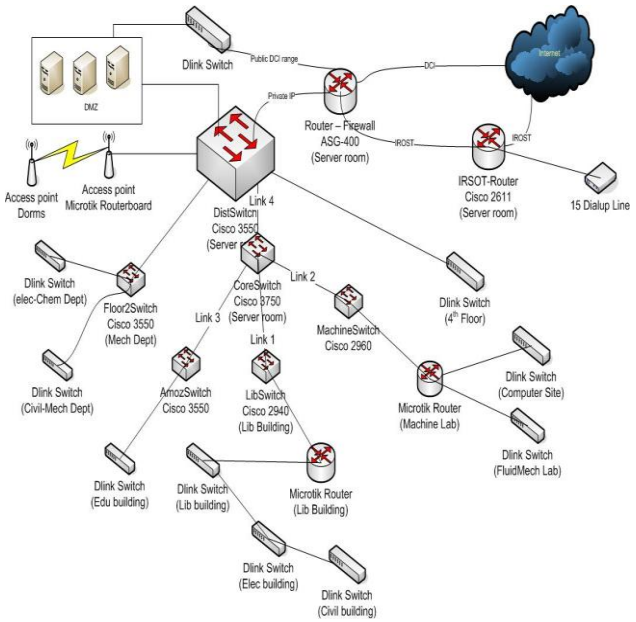

Fig 3: Network Topology of Noushirvani University of Technology.

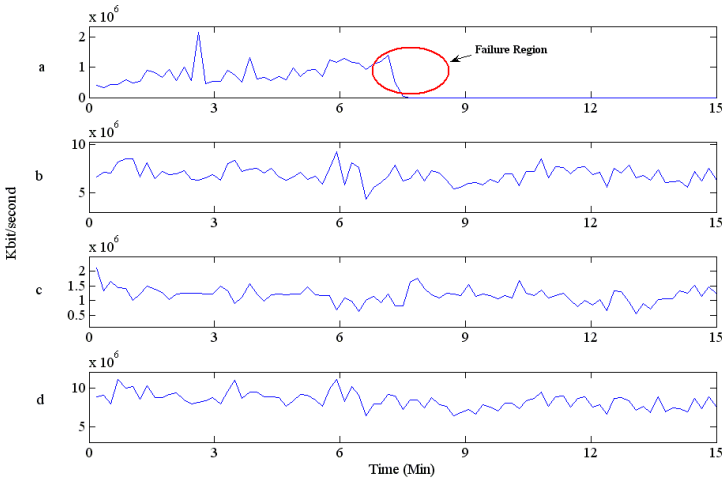

Fig 4: Samples of recorded traffic signals. The lengths of the signals are 15 minutes and sampling interval is 10 seconds.

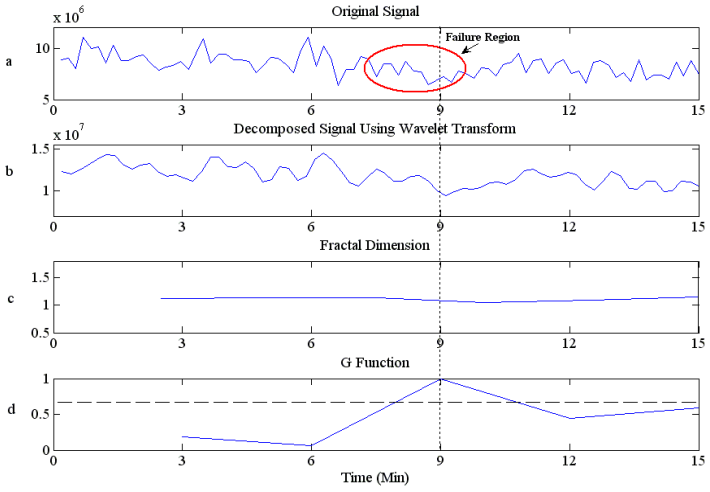

Fig 5: Anomaly detection in real signal using proposed method. (a) Original signal. (b) Approximate signal after applying one level DSWT, (c) Fractal dimension computed using optimal window length, (d) G function.

The existing methods are also applied on the signal in Figure 5.a and the results are shown in Figure 6. This figures indicate that existing methods have more false boundary detection rate (marked with the arrows in the plots) compared to the proposed method. The results of anomaly detection using the proposed method and the three other existing approaches on a set of 50 real data are shown in Table 3. This table confirms that the proposed method has a better accuracy compared to other existing methods.

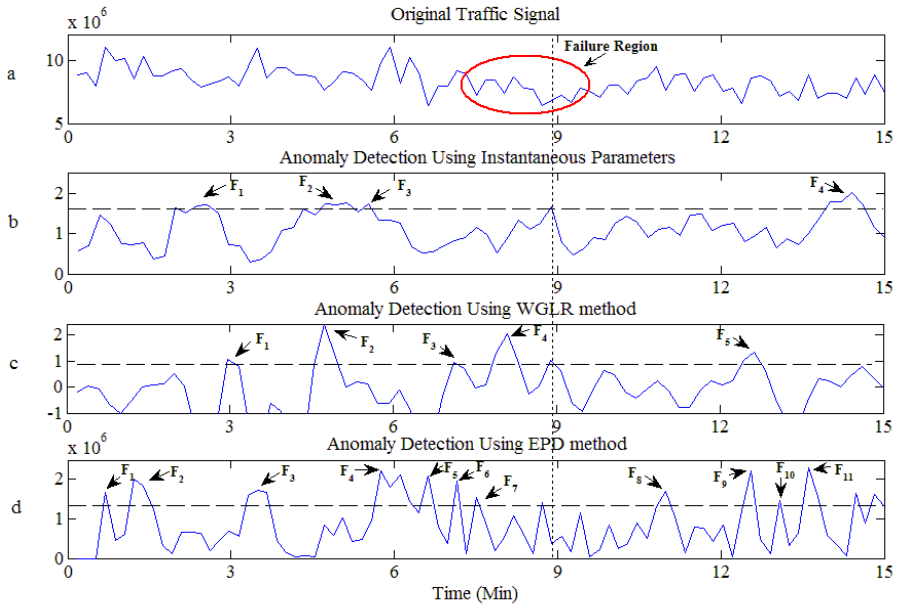

Fig 6: Anomaly detection in real signal using exiting approaches, (a) Original traffic signal, (b) output of method based on instantaneous parameters, (c) output of WGLR method, (d) output of EPD method.

Table 3. Results of anomaly detection using proposed method and three other existing approaches with 50 Realizations for real signals.

Method	TP ratio	FN ratio	FP ratio
Proposed Method	0.9333	0.0667	0.3000
Instantaneous Parameters	0.7000	0.3000	8.4000
WGLR	0.6333	0.3667	3.9000
EPD	0.8667	0.1333	12.1667

In this paper a new anomaly detection method using discrete stationary wavelet transform and fractal dimension has been introduced. Wavelet transform reveals the important underlying unadulterated form of the data and fractal dimension is a powerful method in transient detection. A comparative study between four typical wavelet basis functions on anomaly detection accuracy has been performed. Performance of the proposed method was compared with that of three other existing anomaly detection methods. Simulation results indicated superiority of the proposed method in anomaly detection. The proposed method had the highest value for TP ratio and the lowest value for both FN and FP ratios compared to the existing methods.

7. REFERENCES

[1] Yao, X., Zhang, P., Gao, J. and Hu, G. 2006 .Detection of Network Traffic Anomaly Based on Instantaneous Parameters Analysis, *International Conference on Communication Technology, ICCT '06.* , 1-4.

[2] Tran, D., Ma W. and Sharma, D. 2006. Network Anomaly Detection using Fuzzy Gaussian Mixture Models, *International Journal of Future Generation Communication and Networking*, 37-42.

[3] Marina, T., and Ji, C. 1998. Adaptive Thresholding for Proactive Network Problem Detection, *IEEE International Workshop on Systems Management*, 5, 108-116.

[4] Sotiris, V.A., Tse, P.W. and Pecht, M.G. 2010. Anomaly Detection Through a Bayesian Support Vector Machine, *IEEE Transactions on Reliability*, 59, 277 – 286.

[5] Lu, W. and Ghorbani, A. 2008. Network Anomaly Detection Based on Wavelet Analysis, *EURASIP Journal on Advances in Signal Processing*, 2009, 1-16.

[6] Anisheh, S. M. and Hassanpour, H. 2009. Adaptive Segmentation with Optimal Window Length Scheme Using Fractal Dimension and Wavelet Transform, *International Journal of Engineering*, 22, 257-268.

[7] Luv, J., Li, X. and Li, T. 2007. Research on Network Traffic Anomaly Detection Algorithm, *12th IEEE Symposium on Computers and Communications*, 95-100.

[8] Paxson, V. 1997. Fast approximate synthesis of fractional gaussian noise for generating self- similar network traffic [J]. *Computer communication review*. 10, 5-18.

[9] Salagean, M. and Firoiu, I. 2010. Anomaly detection of network traffic based on Analytical Discrete Wavelet Transform, *8th IEEE International Conference on Communications (COMM)*, 49 – 52.

[10] Kim, S. S. and Reddy, A. 2004. Detecting traffic anomalies at the source through aggregate analysis of packet header data, *Proceedings of Networking*.

[11] Lan, L. and Gyungho, L. 2005. Ddos attack detection and wavelets, *Telecommunication Systems*, 435-451.

[12] Thangavel, M., Thangaraj, P. and Saravanan, K. 2010. Defend against Anomaly Intrusion Detection using SWT Mechanism, *International Journal of Innovation, Management and Technology*, Vol. 1, No. 2, 209-213.

[13] Bachmann, G., Narici, L. and Beckenstein, E. 2002. *Fourier and Wavelet Analysis*, Springer.

[14] Nason, G.P. and Silverman, B.W. 1995. The stationary wavelet transform and some statistical application, *Lecture Notes in Statistics*, vol.103, pp. 281-299.

[15] Falconer, J. 2003. *Fractal Geometry-Mathematical Foundations and Applications*, John Wiley and Sons.

[16] Tykierko, M. 2008. Using invariants to change detection in dynamical system with chaos, *Physica D: Nonlinear Phenomena*, 237, 6-13.

[17] Paramanathan, P. and Uthayakumar, R. 2007. Application of fractal theory in analysis of human electroencephalographic signals, *Computers in Biology and Medicine*, 38. 372 – 378.

[18] Li, Y., Le Fan, Y. and Ye Tong, Q. 2007. Endpoint detection in noisy environment using complexity measure, *Proc. IEEE International Conference*, 3, 1004-1007.

[19] Liu, M., He, Y., Meng, Q. and Wang, Z. 2010. Research on Anomaly Detection of Network Traffic Based on Fractal Technology and Vector Quantization, *IEEE International Workshop on Education Technology and Computer Science (ETCS)*, 2, 428 – 431.

[20] Higuchi, T. 1988. Aproach to an irregular time series on the basis of the fractal theory. *Physica D*, 31, 277-283.

[21] Petrosian, A. 1995. Kolmogorov Complexity of Finite Sequences and Recognition of Different Preictal EEG Patterns, *Proc. IEEE Symposium on Computer-Based Medical Systems*, 5, 212-217.

[22] Katz, M. J. 1988. Fractals and the Analysis of Waveforms, *Comput. Biol. Med.*, 18, 145-156.

[23] Esteller, R., Vachtsevanos, G., Echauz, J. and litt, B. 2001. A comparison of fractal dimension algorithms using synthetic and experimental data, *IEEE Trans. Circuits Syst.*, 48, 177-183.

[24] Malarvili, M., Hassanpour, H., Mesbah, M. and Boashash, B. 2005. A histogram-based electroencephalogram spike detection, *Proc. IEEE Int Symposium on Signal Processing and its App*,1, 207-210.