An Implementation of LoSS Detection with Second Order Statistical Model

Mohd Fo'ad Rohani [*], Mohd Aizaini Maarof [*], Ali Selamat [*] and Houssain Kettani[**]

[*] Faculty of Computer Science and Information Systems
University Teknologi Malaysia
81300 Skudai, Johor
{ foad,aizaini,aselamat}@.utm.my

[**] Department of Computer Science
Jackson State University
Jackson, Mississippi 39217
houssain.kettani@jsums.edu

Abstract

Recent studies have shown that malicious Internet traffic such as Denial of Service (DoS) packets introduces distribution error and perturbs the self-similarity property of network traffic. As a result, Loss of Self-Similarity (LoSS) is detected due to the abnormal traffic packets hence degrading the Quality of Service (QoS) performance. In order to fulfill the demand for high speed and accuracy for online Internet traffic monitoring, we propose LoSS detection with higher order statistical model and estimate the self-similarity parameter using the Optimization Method (OM). We test our approach using synthetic and real traffic data. For the former, we use Fractional Gaussian Noise (FGN) generator, while for the latter we use FSKSMNet simulation dataset. We investigate the behavior of self-similarity property for normal and abnormal traffic packets with different aggregation sampling level (*m*). The results show that normal Internet activities preserve exact self-similarity property while abnormal traffic perturbs the structure of self-similarity property. The results also demonstrate that fixed *m* is not sufficient to detect distribution error accurately. Accordingly, we suggest a multi-level aggregation sampling approach for future work to improve the accuracy of LoSS detection efficiently.

Keywords

Loss of Self-Similarity (LoSS) detection, Second Order Self-Similarity (SOSS) model, Self-similarity property

1. Introduction

In 1993 the work presented in [7], followed by [8] found evidence of long-range correlation in LAN traffic and brought the concept of self-similarity (and the related concept of Long-Range Dependence (LRD) into the field of network traffic and performance analysis. Self-similarity describes the phenomenon in which the behavior of a process is preserved irrespective of scaling in space or time. The knowledge of LRD states that network traffic always exhibit long-term memory, i.e., its behavior across widely separated times is correlated. This finding was in contrast to widely accepted Poisson model of the network traffic, which is memoryless and inter-arrival times are exponentially distributed. The founding challenged the validity of the Poisson assumption and shifted the community's focus from assuming memoryless and smooth behavior network traffic to assuming LRD and bursty behavior.

Previous works had pointed out several causes of the self-similarity phenomenon. One is the mixed behavior of TCP services model such as log-normal, log-extreme and Pareto distributions [13]. Another one is the mixture of actions from individual users, hardware and software in interconnecting networks [2]. The third reason is the heavy-tailed distribution of file sizes where huge transferred files occurred with non-negligible probability [2]. The work done in [3] and [12] showed that congestion due to uncontrolled self-similarity structure degrades Quality of Service (QoS) performance by drastically increasing queuing delay and packet loss. Protocol intensity distribution plays an important role to the interactions that produce self-similarity behavior [13]. Denial of Service (DoS) attacks with very high bit rate injection packets can dominate the traffic protocol and produce distribution error, hence disturb the property of self-similar behavior [14]. As a result, Loss of Self-Similarity (LoSS) behavior is detected [14] and as shown in [1] and [10], this can be used as a flag to alert security analysts of the possible presence of a malicious action, provided that the normal traffic background is self-similar which is a common network traffic attribute.

The work in [1] has presented a new technique for detecting the possible presence of new DoS attacks without a template of the background traffic. The method used LoSS definition with the self-similarity or Hurst parameter beyond normal self-similarity behavior ($0.5<H<1$) using the Periodogram and the Whittle methods. However, new methods of estimating Hurst parameter which is more accurate and faster had been developed such as the

Optimization Method (OM) [4], [5] which used the Second Order Self-Similarity (SOSS) statistical model. In this paper, we present a new LoSS detection method using SOSS statistical model and OM.

The sequel of this paper is as follows: Section 2 presents mathematical definitions and properties of SOSS and how to estimate its parameter. Section 3 on the other hand, discusses the concept of LoSS detection and related work. Section 4 discusses the datasets that were used in the simulations while Section 5 presents our experiment procedure and the results. Finally our conclusions and future work directions are summarized in Section 6.

2. SOSS Statistical Model

Let $X = \{X(t), t \in ¢\}$ be a second-order stationary process with constant mean μ, finite variance σ^2, and autocorrelation function $\rho(k)$ that depends only on the integer *k*. Their definitions are given as follows:

$$\mu = E[X(t)], \;\; \sigma^2 = E[(X(t)-\mu)]^2$$

$$\rho(k) = E[(X(t)-\mu)(X(t+k)-\mu)]/\sigma^2$$

Let $X^{(m)} = \{X^{(m)}(t), t \in ¢^+\}$ denote the aggregate process of X at aggregation $m, (m \in ¢^+)$. That is, for each $m, X^{(m)}$ is given by $X^{(m)}(t) = \dfrac{1}{m}\sum_{l=m(t-1)+1}^{mt} X(l)$, $t \in ¢^+$.

Let $\gamma^{(m)}(k)$ and $\rho^{(m)}(k)$ denote the variance and autocorrelation function of $X^{(m)}$ respectively. X is called *exactly second-order self-similar (ESOSS)* with self-similarity parameter $H = 1 - \dfrac{\beta}{2}$, if

$$\rho(k) @ \frac{1}{2}[(k+1)^{2-\beta} - 2k^{2-\beta} + (k-1)^{2-\beta}] \, , \, 0 < \beta < 1, \; k \in ¢^+.$$

X is called *long-range dependence (LRD)* with $H = 1 - \dfrac{\beta}{2}$, 0<*β*<1, if its autocorrelation function satisfies $\rho(k) : ck^{-\beta}, k \to \infty$, where *c* is a positive constant. X is called *asymptotical second-order self-similar (ASOSS)* with $H = 1 - \dfrac{\beta}{2}$, 0<*β*<1, if $\lim_{m\to\infty} \rho^{(m)}(k) = \rho(k), \; k \in ¢^+$.

Exact self-similar process occurs when $\rho(k) = \rho^{(m)}(k)$ for all m≥1. Thus, *second order self-similarity* captures the property of correlation structure preserving under time aggregation and represented by $\rho(k) = \dfrac{1}{2}[(k+1)^{2H} - 2k^{2H} + (k-1)^{2H}]$ for ESOSS or $\lim_{m\to\infty} \rho^{m}(k) = \dfrac{1}{2}[(k+1)^{2H} - 2k^{2H} + (k-1)^{2H}]$ for ASOSS. In second-order stationary for 0<*H*<1, *H*≠0.5, autocorrelation function $\rho(k)$ holds $\rho(k) : \; H(2H-1)k^{2H-2}$, $k \to \infty$. In particular, if 0.5<*H*<1, $\rho(k)$ asymptotically behaves as $ck^{-\beta}$, $\rho(k) : \; c_r k^{-\beta}$ for 0<β<1 where c_r > 0 is a constant, and *β=2-2H.* More details about the SOSS statistical model can be found at [7], [8] and [11].

There are several methods to estimate H. In this paper we will be using the OM which was developed in [4][4], [5] and was shown to be comparatively fast and accurate with respect to other methods. The method is based on how near sample autocorrelation measure fits to *ESOSS* model. The estimation method defines error fitting function $E_K(\beta)$ as $E_K(\beta) = \dfrac{1}{4K}\sum_{k=1}^{K}(\rho(k) - \rho_n(k))^2$ where *ρ(k)* denotes the autocorrelation function of the model with parameter *β* that OM would like to fit the data to, $\rho_n(k)$ is the sample autocorrelation function of the data and *K* is the largest value of *k* for which $\rho_n(k)$ is to be computed to reduce edge effects. The estimation of parameter *β* is based on optimizing $E_K(\beta)$ with threshold value ≤ 10^{-3} is chosen from experiment [4].

3. LoSS Detection

The ESOSS model preserves the second order distribution property at all levels of time scale aggregation. It is equivalent to distribution ratio of higher scale to lower scale such as $a^{-H} = \dfrac{x(t)}{x(at)}$ where *x(t)* is distribution at higher scale, *x(at)* is distribution at lower scale and parameter *a,H* >0. It has been proven that in the existence of DoS attacks, the self-similarity property is disturbed hence LoSS is detected as shown in [1], [14] and [15]. The LoSS detection in [14] used the abrupt change property of a^{-H} as indicator to the existence of distribution error. However, the work was not suggesting at what level of '*a*' to be used for revealing the abrupt change of a^{-H} significantly. Alternatively, instead of using distribution ratio, the work in [1] defined LoSS as Hurst value beyond normal range of LRD which is 0.5≤ H≤0.99 using Periodogram or Whittle method. The results show that the method can detect new DoS attack pattern without specific normal template. The results also demonstrate that the method has high detection rate with an average of 60% to 84% which depends on the intensity of the attack packets. Recently, a new method of estimating Hurst parameter which is more accurate and faster was developed in [4] and [5]. The method is known as OM and it is based on the SOSS statistical model. Therefore, we propose a new approach of LoSS detection based on SOSS statistical model in order to improve detection accuracy.

The foundation of second order statistical model is the stationary concept of higher order distribution. Internet traffic is considered as normal behavior when the traffic is near to self-similarity model while otherwise it is considered as abnormal behavior [15]. In the existence of malicious traffic such as DoS packets, they introduce distribution error and shift the stationary property toward non-stationary as shown in [1], [14] and [15], hence LoSS is detected. Data insufficient probability and detection loss probability are two important attributes that can influence the correctness of anomaly detection [15]. The data insufficient probability is to identify minimum requirement window size to obtain reliable self-similarity measurement, while detection loss probability is probability where non-stationary data is detected. LoSS is detected if it fulfils two conditions where it must be longer than minimum window size and it must be non-stationary. Experiments have shown in [7] and [8] windows sizes from 15-30 minutes are practical and sufficient for modern LANs Ethernet Internet traffic to comply with data insufficient probability. Self-similarity tests become more sensitive as the window size get smaller and consequently generate false alarms if it gets too small.

A current study on self-similarity measurement used the OM in [4] and [5] that made this window size estimation more realistic because of the increased calculation speed. In addition this method also provides a technique to identify whether the data tend toward the self-similarity model according to the curve-fitting error value calculated. To this end, we define normal behavior of self-similarity traffic as Hurst value with *MIN(H,OM)≥0.5* and *MAX(H,OM)*≤0.99 with condition of data insufficient probability and fitting error $E_K \le 10^{-3}$. Otherwise if $E_K > 10^{-3}$, LoSS is detected and we refer this as abnormal behavior for Internet traffic. The ESOSS process refers to $\rho(k) = \rho^{(m)}(k)$ for all $m \ge 1$. Thus, SOSS captures the property of correlation structure which is preserved under time aggregation. Therefore it is required to study the effect of different aggregation sampling value such as m=10, 100 and 500 to detect any changes of self-similarity property in order to reveal any hidden distribution error accurately.

4. Data Preparation

We prepare two sets of data to investigate the pattern of normal and abnormal self-similarity behavior. The former used Fractional Gaussian Noise (FGN) that will generate synthetic traffic which is ESOSS and the latter used FSKSMNet Internet traffic simulation on September 29, 2006 at Faculty of Computer Science and Information (FSKSM) LANs.

4.1 Synthetic Self-Similar Generator with FGN

We generate at random synthetic trace that will exhibit exact self-similarity behavior by using Fractional Gaussian Noise (FGN) model developed in [6] for 0.5<H<1. The length of the trace is equivalent to 15-30 minute at normal traffic Ethernet LAN as used in [7] and [8]. The synthetic trace is sampled with aggregation level *m* for *m*=10, 100 and 500. Then, this dataset is used to investigate the self-similarity property at different aggregation level *m*.

4.2 Simulation of Internet Traffic FSKSMNet

We set an Internet Monitoring Laboratory (InMonLab) with baseline 100BaseFX Fast Ethernet as LAN backbone of FSKSM and connected to main university Gigabit backbone. Network design at FSKSM is constructed with ten proxies of Virtual LANs. There are seven VLAN segments for undergraduate students and one VLAN segment for postgraduate students. Administrators and academic staffs are allocated with one VLAN segment each. The number of students that are currently enrolled at FSKSM is more than one thousand and the number of staff is about one hundred and fifty. We capture internet protocol packets with *tcpdump* software and each of capturing session is about 30 minutes.

Trace	Class	Capture	Total Packet	DoS Packet
FGN-1	Synth-etic	≈ 30min Ethernet LAN	180,000 (at m=10ms)	
F-Net2	N	12.15pm-12.45pm	IP=3846328: TCP(97.94%), UDP(1.91%), ICMP(0.11%), IGMP(0.01%), Others(0.03%)	
F-Net3	N	1.45pm-2.15pm	IP=4197509: TCP(97.87%), UDP(1.69%), ICMP(0.12%), IGMP(0.01%), Others(0.31%)	
F-Net4	AB	3.45pm-4.15pm	IP=8932254: TCP(93.73%), UDP(1.20%), ICMP(0.04%), IGMP(0.003%), Others(5.021%)	TCP SYN(58.5%)

Table 1 Synthetic FGN and Simulation of FSKSMnet on September29, 2006

We divide our Internet traffic simulation activities into normal and abnormal traffic. For normal Internet activities we define as legal Internet activities as set by faculty network policy. We do not disturb normal Internet activities and do passive sniffing at main router inside InMonLab. For abnormal traffic, we inject at certain rate Denial of Service (DoS) flooding packets into FSKSM network infrastructure such that they will disturb the normal behavior of self-similarity pattern. We launch TCP SYN packets from Packet Injection node to Honeynet server. However we limit the time stressor for each session to less than two minute to minimize the bandwidth effect. Table 1 shows the detail of simulation traces of synthetic FGN and real traffic of FSKSMNet at FSKSM LANs for 30 minute each slot. From Table 1, *F-Net 2* and *F-Net3* traces represent normal traffic activities which do not contain DoS packets. We label normal traffic as class N. The total of normal IP packet is about four million with TCP >95%, UDP < 2%, ICMP, IGMP and others <1%. We simulate abnormal traffic *F-Net4* with TCP SYN attack and label as AB. The abnormal traffic contains almost doubles IP packets as compared to normal which is more than eight million. We sample the traces with *m*=10, 100 and 500 in order to investigate how normal and abnormal traffic preserve self-similarity property. Table 2 shows window sizes of different sampling *m* values for synthetic FGN and real Internet traffic FSKSMnet datasets. We assume that all the simulation traffic traces have fulfilled the minimum windows requirement. Then the estimated Hurst parameter can be used to classify whether or not the traffic follows the ESOSS model.

Level	m	Window Size	
		FGN	FSKSMnet
1	10	180000	173998
2	100	18000	17399
3	500	3600	3479

Table 2 Aggregation (*m*) and Window Size

5. Empirical Analyses

5.1 LoSS Behavior Detection

The purposes of our experiments are two folds. Firstly to identify the normal traces which follow ESOSS model and secondly to investigate how self-similarity property is preserved at different level of sampling *m* for normal and abnormal Internet traffic. Throughout this experiments, we set threshold fitting error equal to 10^{-3}, estimate Hurst using OM with *K*=200.

Table 3 shows the result of *Hurst* estimation for Synthetic FGN and FSKSMnet Sep29, 2006 traffic while Table 4 shows variance *Hurst* and error for different level *m*. It is shown clearly in Table 3 that synthetic *FGN-1* and *F-Net3* follow normal self-similar behavior for all *m*. However for *F-Net2* and *F-Net4* they have two different classes; at lower *m*=10 and 100, the traces are categorized as normal self-similar behavior while at higher *m*=500 the traces deviate from normal behavior.

Trace	m=10		m=100		m=500	
	Hurst	Error	Hurst	Error	Hurst	Error
FGN-1	0.87	0.00004	0.86	0.00003	0.84	0.00014
F-Net2	0.82	0.00042	0.87	0.00029	0.82	0.00278
F-Net3	0.89	0.00050	0.93	0.00011	0.93	0.00027
F-Net4	0.97	0.00042	0.96	0.00033	0.90	0.01265

Table 3 *Hurst* estimation for FGN and FSKSMnet

In our simulation *F-Net4* contains DoS packets and known as abnormal traffic. These illegal packets have extremely high bit rate transfer and their structure become dominant hence introduce error to self-similar model. Similarly, legal Internet traffic *F-Net2* has also been identified as contained structure that contributes error to the model. However the error can only be revealed at certain transition of *m*. Table 3 shows the error is hidden at m=10 and 100, and exposed at m=500.

5.2 Autocorrelation Structure of Normal and Abnormal Internet Traffic Behavior

We use the ESOSS autocorrelation structure $\rho(k)$ to examine in details how self-similarity structure is preserved at different level of sampling *m*. We divide our observation into two categories that are normal-normal and normal-abnormal patterns.

Case I: Normal-Normal

We define normal-normal behavior as the $\rho(k)$ structure preserved the LRD property for all *m*. Figure 1(a) and (c) show the $\rho(k)$ structure of *FGN-1* and *F-Net3* traces preserved LRD structure for m=10,100 and 500. The traces have produced normal behavior pattern in two ways. First, at all *m* the *FGN-1* and *F-Net3* traces follows second order statistical model (fitting error<10^{-3}) and second, the deviation of multi-level sampling *m Hurst* (*Var(m-H)*<$5.5\text{x}10^{-4}$) and fitting error (*Var(m-Error)*<10^{-6}) are small as shown in Table 4.

Trace	Var(m-H)	Var(m-Error)
FGN-1	0.000233333	0.00000000370
F-Net2	0.000833333	0.00000196443
F-Net3	0.000533333	0.00000003843
F-Net4	0.001433333	0.00005022723

Table 4 Variance of *Hurst* and Fitting Error for different *m*

Case II: Normal-Abnormal

We define normal-abnormal behavior as the $\rho(k)$ structure is preserved at lower *m* however its lost LRD structure at higher *m*. As shown in Figure 1(b) and (d), at lower *m*=10 and 100, the $\rho(k)$ structure of *F-Net2* and *F-Net4* traces do follow LRD property however at *m*=500 they don't. Table 2 shows the abnormal traces of *F-Net2* and *F-Net4* have fitting error < 10^{-3} for *m*=10 and 100. In contrast, the error exceeds threshold for *m=500*. The disturbance of $\rho(k)$ structure is shown clearly in Figure 1(b) and (d). In addition, multi-level sampling (*m*) *Hurst* (*Var(m-H*)

$>5.5\text{x}10^{-4}$) and *fitting error* (*Var(m-Error)* $>10^{-6}$) are large as shown in Table 4.

It is the possible for legal Internet activities to contribute error and disturb self-similarity property as shown by *F-Net2*. However, the error can only be revealed effectively if we consider LoSS analysis at different level of *m*. Therefore fixed sampling *m* is not enough to detect traffic anomaly behavior accurately.

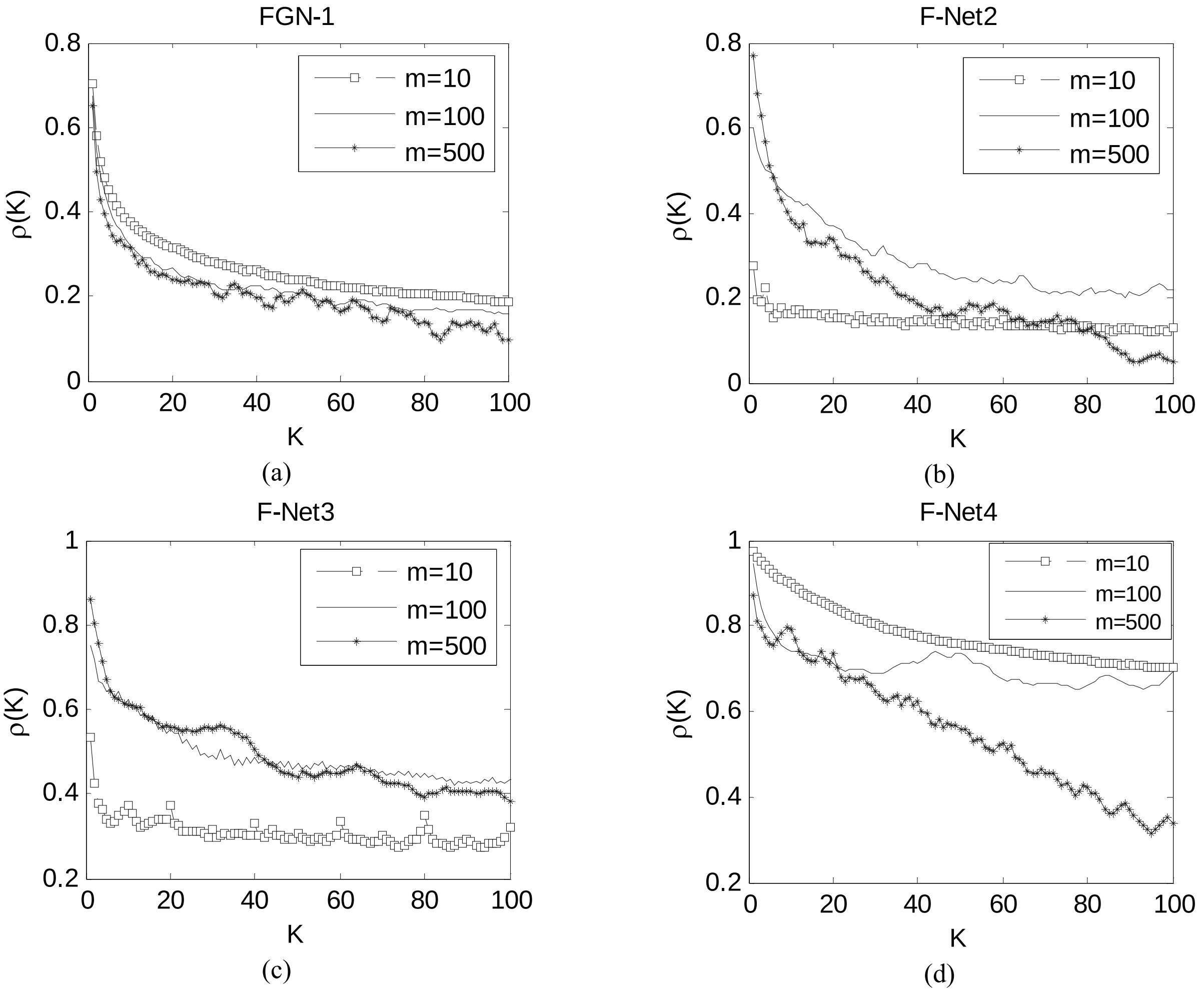

Figure 1 Self-Similarity Property of FGN (a) and FSKSMnet on September 2006 (b), (c), (d)

From the observation we can define normal and abnormal Internet traffic behavior based on LoSS detection. We consider three parameters, i.e. estimated *Hurst, fitting error* and aggregation sampling level *m*, to detect LoSS accurately.

We define normal behavior as follows:

$\{(H\in 0.5<H<1)\cap$ *fitting error* <Threshold (at normal *m*)$\}$ $\cap$ $\{(H\in 0.5<H<1)\cap$ *fitting error* < Threshold (at higher *m*)$\}$

On the other hand, abnormal behavior is detected when Internet traffic deviates from self-similarity property such that:

$\{(H\in 0.5<\text{H}<1)\cap$ *fitting error* >Threshold (at normal *m*)$\}$ $\cup$ $\{(H\in 0.5<H<1)\cap$ *fitting error* >Threshold (at higher *m*)$\}$

Normal aggregation *m* is referred to *m*=10ms and 100ms which is used for Hurst estimation in [4], [5], [7], [8] and [15] while higher aggregation level *m* we set randomly equal to 500ms.

6. Conclusion and Future Work

This paper presents the implementation of LoSS detection with second order statistical model. From our results, legal and illegal Internet traffic activities can contribute distribution error which deviate autocorrelation structure from self-similarity model. This can be shown clearly when fitting error exceeds from the threshold value.

However, our result shows that distribution error was hidden at certain level of *m* such as m=10 or 100, but revealed at higher level of *m* such as m=500. We believe this can be possible reason why anomaly detection of Internet traffic behavior which based on LoSS model gives high false alarm detection rate. Therefore, for our future work we will consider a multi-level aggregation sampling approach in order to increase the accuracy of LoSS detection base on second order self-similar statistical model. We will also consider a wider range of Internet traffic traces and different type of illegal Internet activities to test the robustness and reliability of our methods toward development of efficient Internet traffic anomaly detection systems.

Acknowledgements

This work was funded by Universiti Teknologi Malaysia (UTM). We would like to thanks to Assoc. Prof. Dr. Sulaiman Mohd Noor, Mr. Mohd Hamri of CICT, UTM and Dr. Md. Asri Ngadi, Mr. Firoz of Unit IT, FSKSM for their helps in conducting the simulation of real traffic FSKSMNet dataset.

References

[1] W. H. Allen and G.A. Marin, "The LoSS technique for detecting new Denial of Service attacks," SoutheastCon, 2004. Proceedings. IEEE, pp. 302-309, 26-29 March 2004.

[2] M.E. Crovella and A. Bestavros, “Self-similarity in World Wide Web traffic: Evidence and possible causes networking,” IEEE/ACM Transactions on Volume 5, Issue 6, pp. 835 – 846, December 1997.

[3] A. Erramilli, O. Narayan, and W. Willinger, “Experimental queueing analysis with long-range dependent packet traffic,” IEEE/ACM Trans. Networking, 4:209–223, 1996.

[4] H. Kettani, “A Novel Approach to the Estimation of the Long-Range Dependence Parameter,” University of Wisconsin – Madison : PhD. Thesis (2002).

[5] H. Kettani and J. A. Gubner, "A Novel Approach to the Estimation of the Long-Range Dependence Parameter," IEEE Transactions on Circuits and Systems II, Volume 53, Issue 6, pp. 463-467, June 2006.

[6] S. Ledesma and D. Liu, "Fractional Gaussian noise power spectrum synthesis using linear approximation for generating self-similar network traffic," ACM Computer Communication Review, vol.30, no.2, pp. 4-17, April 2000.

[7] W. Leland, M. Taqqu, W.Willinger and D.Wilson, “On the self-similar nature of Ethernet traffic,” Proc. of ACM SIGCOMM 23(4) (1993), pp. 183–193.

[8] W. Leland, M. Taqqu, W. Willinger and D. Wilson, “On the self-similar nature of Ethernet traffic (extended version),” IEEE/ACMTransactions on Networking 2(1) (1994), pp. 1–15.

[9] M. Li, “Change trend of averaged Hurst parameter of traffic under DDOS flood attacks”, Computers & Security, Volume 25, Issue 3, pp. 213-220, May 2006.

[10] M. Li, W. Jia, and W. Zhao, "Decision analysis of network-based intrusion detection systems for denial-of-service attacks," Proceedings of IEEE International Conferences on Info-tech and Info-net (ICII 2001), Vol. 5, Beijing, PRC, 29 Oct. - 1 Nov. 2001, pp. 1-6.

[11] C. Park, F.H. Campos, J.S. Marron, D. Rolls and F.D. Smith, “Long-Range-Dependence in a changing Internet traffic mix,” Statistical and Applied Mathematical Sciences Institute (SAMSI) Technical Report 2004-9, 26 March 2004.

[12] K. Park, G. Kim and M. Crovella, “On the effect of traffic self-similarity on network performance”, SPIE International Conference on Performance and Control of Network Systems, November 1997.

[13] V. Paxson and S. Floyd, "Wide-area traffic: The failure of Poisson modeling," IEEE-ACM Transactions on Networking, 3(3), June 1995.

[14] W. Schleifer and M. Mannle, "Online error detection through observation of traffic self-similarity," IEE Proceedings on Communications, 148(1), Feb. 2001.

[15] M. Yazid, A. Hanan and M. Aizaini, “Iterative window size estimation on self-similarity measurement for network traffic anomaly detection”, International Journal of Computing and Information Science, (IJCIS), vol. 2(2), pp. 83-91, 2004.