Continuous LoSS Detection Using Iterative Window Based On SOSS Model and MLS Approach

Mohd Fo'ad Rohani[1], Mohd Aizaini Maarof[2], Ali Selamat[3] and Houssain Kettani[4]
[1,2,3]*Faculty of Computer Science and Information Systems*
Universiti Teknologi Malaysia
81300 Skudai, Johor, Malaysia
[4]*Department of Electrical and Computer Engineering and Computer Science*
Polytechnic University of Puerto Rico
P. O. Box 192017
San Juan, Puerto Rico 00919, USA
Email : {foad, aizaini, aselamat}@utm.my, hkettani@pupr.edu

Abstract

This paper proposes a continuous Loss of Self-Similarity (LoSS) detection using iterative window and Multi-Level Sampling (MLS) approach. The method defines LoSS based on Second Order Self-Similarity (SOSS) statistical model. The Optimization Method (OM) is used to estimate self-similarity parameter since it is fast and more accurate in comparison with other estimation methods known in the literature. The probability of LoSS detection is introduced to measure continuous LoSS detection performance. The proposed method has been tested with real Internet traffic simulation dataset. The results demonstrate that normal traces have probability of LoSS detection below the threshold at all sampling levels. Meanwhile, abnormal traces have probability of LoSS that imitates normal behavior at sampling levels below 100ms but exceeds the threshold at sampling levels larger than 100ms. Our results show the possibility of detecting anomaly traffic behavior based on obtaining continuous LoSS detection monitoring.

I. INTRODUCTION

The need for continuous monitoring of anomaly traffic detection in Ethernet network is very crucial to providing uninterrupted Quality of Service (QoS) performance. This can be achieved by continuously detecting Loss of Self-Similarity (LoSS) occurrences in traffic when the packets are treated as a time series [1], [4], [12], [13]. Malicious traffic such as Denial of Service (DoS) packets have tendency to contribute to the deviation from the self-similarity model [12]. Consequently, LoSS is detected [1], [12] and a high percentage of LoSS detection will alert a signal of poor QoS performance [3], [9] due to uncontrolled self-similarity structure. Implementation of LoSS detection with Second Order Self-Similarity (SOSS) statistical model has been introduced due to high speed and accuracy needs [4]. Previous works have used fixed sampling and fixed window to detect LoSS [1], [4]. Fixed sampling, however, is insufficient to reveal the self-similarity distribution error efficiently [10], [11]. On the other hand, dynamic window is needed for continuous anomaly traffic detection. In this paper, we propose a continuous LoSS detection using iterative window and Multi-Level Sampling (MLS) approach while estimating the self-similarity parameter or Hurst (*H)* with the Optimization Method (OM). We have evaluated the LoSS detection performance using probability of LoSS detection measurement. This paper is organized as follows: Section 2 discuses in brief the self-similarity model and the estimation of the parameter *H*. The proposed LoSS detection method is discussed in Section 3 while the experimental and empirical analyses are presented in Section 4. Finally, our conclusion and future work are summarized in Section 5.

II. SOSS MODEL AND ESTIMATION METHOD

Let define a second-order stationary process $X = \{X(t), t > 0\}$ with constant mean μ, finite variance σ^2 and autocorrelation $\rho(k)$ as follow:

$$\mu = E[X(t)], \quad \sigma^2 = E[(X(t) - \mu)]^2 \qquad (1)$$

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

Let $X^{(m)} = \{X^{(m)}(t), t > 0\}$ denote the aggregate process of X at aggregation level $m > 0$.
Thus, we have:

978-1-4244-1692-9/08/$25.00 ©2008 IEEE

$$X^{(m)}(t) = \frac{1}{m} \sum_{w=m(t-1)+1}^{mt} X(w)\ ,\ t > 0\ . \qquad (3)$$

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) if $\rho(k) = \rho^{(m)}(k)$ for all $m \geq 1$. In ESOSS, the autocorrelation structure is preserved for all m such that:

$$\rho(k) = \frac{1}{2}[(k+1)^{2-\beta} - 2k^{2-\beta} + (k-1)^{2-\beta}] \qquad (4)$$

where $k>0$ and $0<\beta<1$. X is called Asymptotical Second-Order Self-Similar (ASOSS) if

$$\lim_{m\to\infty} \rho^{m}(k) = \frac{1}{2}[(k+1)^{2-\beta} - 2k^{2-\beta} + (k-1)^{2-\beta}] \qquad (5)$$

where $k>0$, $m>0$ and $0<\beta<1$. X is called Long-Range Dependent (LRD) if its autocorrelation function satisfies: $\rho(k) = ck^{-\beta}$ where $k \to \infty$, $c>0$ and $0<\beta<1$.

There are several methods to estimate H. In this paper we use OM that was developed in [5], [6] which was proven relatively fast and accurate compared to other methods such as the wavelet method. The OM defines Curve-Fitting Error (*CFE*) function as $E_K(\beta)$ such as:

$$E_K(\beta) = \frac{1}{4K} \sum_{k=1}^{K} (\rho(k) - \rho_n(k))^2 \qquad (6)$$

where $\rho(k)$ denotes the autocorrelation function of the model with parameter β that we 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 such that it minimize the edge effect for the calculation of $\rho_n(k)$. If the minimum of $E_K(\beta)$ is less than 10^{-3}, then the data fits the model and the minimizer $\hat{\beta}$ is picked to be the estimate of the parameter β [5].

III. LoSS DETECTION WITH ITERATIVE WINDOW AND MULTI-LEVEL SAMPLING

We have considered a multi-level sampling approach with sampling level (m) in the range of $10ms \leq m \leq 1000ms$ in order to investigate the proposed LoSS detection method [2], [10]. We define iterative window as the window size is continuously incremented in a defined fixed size window denoted by ΔW. The proposed LoSS detection consists of initializing the window and LoSS detection processes. Initialization window is done to fulfill the minimum window size requirement before the estimation of H [4]. This involves window size that meets the *CFE* criterion below the threshold value. The process of LoSS and SOSS detection are only continued if initialization window has been established. LoSS is detected if *CFE* is above the threshold value, otherwise SOSS is detected. If initialization window fails (IF) even though enough capturing time is given such as 30 minutes as used in [7] and [8], then we declare the detection of malicious traffic behavior. The algorithm for initialization window process is shown in Figure 1.

```
set stepSize, ΔW
set window, W = W + ΔW
while (W < Wmax)
    estimate H and CFE
    if (CFE < Threshold) && (0.5 < H < 1)
        Inialization Window success
        SOSS is detected
        proceed with LoSS detection
        else
        increment ΔW++
        set W = W + ΔW
    end
end
if (W ≥ Wmax) && (errCheck = 0)
    Initialization Window failed
    suspect suspicious behavior
end
```

Figure 1. Initialization window process

On the other hand, continuous LoSS and SOSS detection with iterative window is based on whether $CFE > 10^{-3}$ for LoSS and $CFE \leq 10^{-3}$ for SOSS. The detection algorithm is shown in Figure 2.

```
while (W < Wmax)
    increment ΔW++
    update W=W+ ΔW
    estimate H and CFE
    if (CFE < Threshold) && (0.5 < H <1)
            SOSS is detected
            else
            LoSS is detected
    end
end
```

Figure 2. LoSS and SOSS detection process

We introduce measurement probability of LoSS detection to assess the effectiveness of the proposed detection algorithm. Thus, we define iterative window update in a continuous hunting mode as $W_i = \{W_1, W_2, W_3, ..., W_N\}$ for $i = 1, 2, 3, ..., N$. For each of the updated window W_i, if LoSS is detected then we update LoSS window equal to W_i or else update SOSS window with W_i. Suppose we have the

updated LoSS and SOSS window as follows: $LoSS(i) = \{L_1, L_2, L_3, ..., L_K\}$ for $i = 1, 2, 3, ..., K$ and $SOSS(j) = \{S_1, S_2, S_3, ..., S_M\}$ for $j = 1, 2, 3, ..., M$. Then, we define the probability of LoSS (P_L) and the probability of SOSS (P_S) detection using equations (8) and (9):

$$P_L = P(LoSS) = \sum_{i=1}^{K} \frac{L_i}{\sum_{j=1}^{K} L_j + \sum_{k=1}^{M} S_k}, \quad (8)$$

$$P_S = P(SOSS) = \sum_{i=1}^{M} \frac{S_i}{\sum_{j=1}^{K} L_j + \sum_{k=1}^{M} S_k} \quad (9)$$

IV. EXPERIMENTS AND EMPIRICAL ANALYSES

We have simulated the FSKSMNet Internet traffic traces collected on September 29, 2006 at Faculty of Computer Science and Information Systems (FCSIS) [11] in order to evaluate the proposed LoSS detection method. The network infrastructure consists of ten VLANs with 100BaseFX Fast Ethernet backbone which is connected to university Gigabit backbone. The simulation is divided into normal and abnormal traffic. Normal traffic is defined as Internet activities that strictly follow FCSIS network policy. On the other hand, abnormal traffic contains simulated injection of DoS flooding packets at controlled rate. Each of simulation traces is about 30 minutes and the details are shown in Table 1. FNet -1 and FNet-2 are normal traces while FNet-3 contains UDP, TCP SYN and TCP RST, while FNet-4 has TCP SYN and UDP flooding packets.

TABLE 1. FSKSMNET DATASET [11]

FSKSMNet-Normal		*FSKSMNet-Abnormal*	
Trace	*Total Packet*	*Trace*	*Total Packet*
FNet-1	IP=4197509: TCP(97.87%), UDP(1.69%), ICMP(0.12%), IGMP(0.01%), Others(0.31%)	FNet-3	IP=7468026: TCP(85.60%), UDP(14.35%), ICMP(0.04%), IGMP(0.01%), Others(0.005%)
FNet-2	IP=7371721: TCP(92.17%), UDP(0.93%), ICMP(0.07%), IGMP(0.004%), Others(6.83%)	FNet-4	IP=9707011: TCP(69.17%), UDP(30.77%), ICMP(0.04%), IGMP((0.005%), Others(0.02%)

The results of continuous LoSS detection for FSKSMNet traces are shown in Figures 3 and 4 respectively. Zero probability of LoSS is detected at all sampling levels of *m* for normal trace FNet-1. This illustrates that the FNet-1 trace strictly follows ESOSS model as clearly shown in Figure 3(a). However, there is a tendency for normal Internet activities to have a small portion of LoSS detection occurs at higher value of *m* such as above 500ms as illustrated by normal trace FNet-2 in Figure 3(b). Meanwhile, the results also demonstrate that LoSS is hardly detected for malicious traces of FNet-3 at small value of *m* such as lower than 100ms as shown in Figure 4(a). However, the LoSS occurrence is revealed clearly at higher value of *m* such as larger than 100ms.

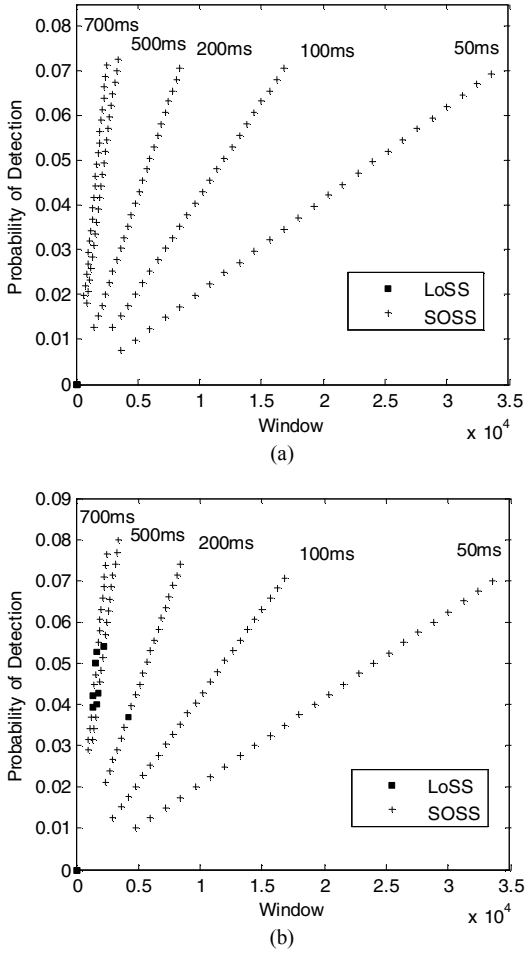

Figure 3. LoSS detection for normal traces: FNet-1 (a) and FNet-2 (b)

Another indication for malicious or suspicious behavior presence in the traces is that the failure to obtain the initialized window as illustrated by abnormal trace FNet-4 in Figure 4(b). Figure 4(b) shows that the algorithm was not able to obtain the initialized window or initialization window failed (IF) when the value of *m* was larger than 100ms. This is an additional alert signal that reveals the presence of abnormal traffic packets in the network despite the

capability to imitate SOSS model behavior at lower sampling such as value of *m* is less than 100ms.

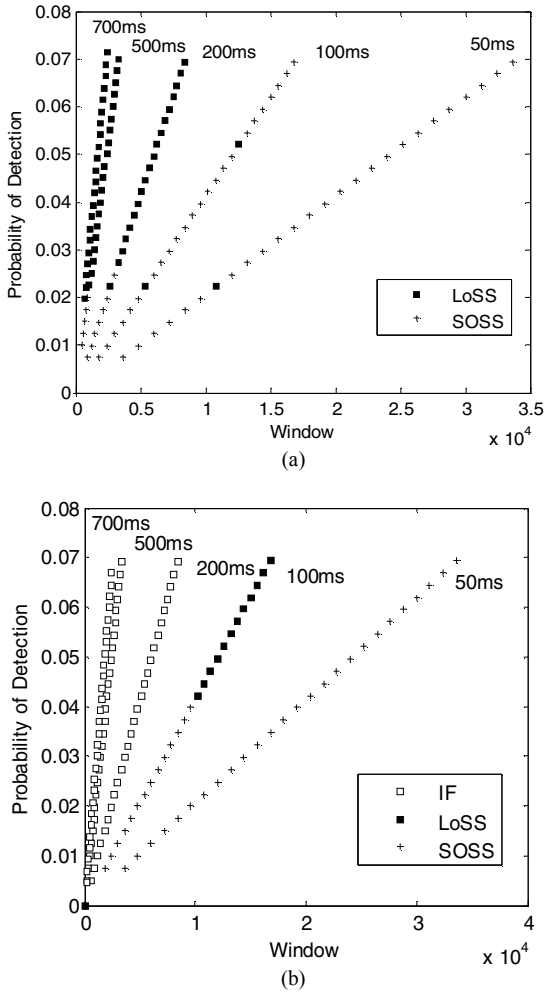

Figure 4. LoSS detection for abnormal traces: FNet-3 (a) and FNet-4 (b)

The probability of continuous LoSS detection for normal and abnormal traffic is shown in Figure 5. The LoSS probability for the normal FNet-1 trace is less than 0.1 for all values of *m,* which is much smaller than the threshold 0.5. Similarly, the normal FNet-2 trace has probability of LoSS value less than 0.1 at sampling level *m* below 500ms but increase slowly to 0.4 at higher value of *m* larger than 500ms. However, the probability of LoSS for malicious traces FNet-3 and FNet-4 have similar patterns with FNet-2. They hide the self-similarity distribution error at value of *m* below 100ms but the error is exposed clearly at value of *m* larger than 100ms. The result in Figure 5 also illustrates that both malicious traces FNet-3 and FNet-4 have exceeded the critical probability of LoSS threshold at 0.5. This will give a clear alarm signal to network security analyst that a severe ESOSS leakage is occurred during the continuous LoSS detection process. We define ESOSS leakage as in any level of *m*, there is at least one sampling level where LoSS can be detected. Another word, the leakage demonstrates the breach indication of ESOSS model property. Our assumption is that if probability of LoSS value estimated below the threshold, then the traffic behavior will be dominated by SOSS model. Otherwise, the majority of updated windows are detected with LoSS which alerts a large amount of ESOSS leakage occurrence. This can be shown in Figure 5 where FNet-3 and FNet-4 have critical ESOSS leakage warning compared to FNet-1 and FNet-2 at sampling level *m* larger than 100ms. Therefore, probability of LoSS measurement can be used as a technique to reveal the abnormal traffic behavior in a continuous LoSS detection at multi-level sampling approach.

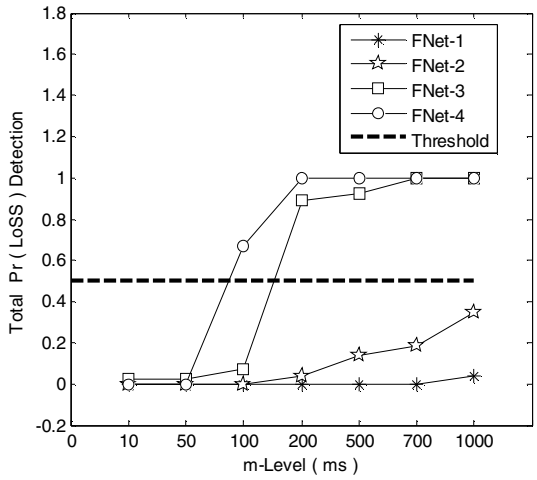

Figure 5. Probability of continuous LoSS detection

V. CONCLUSION AND FUTURE WORK

In this paper, we proposed a multi-level sampling approach for continuous LoSS detection using iterative window. The LoSS detection method is based on SOSS model and the Optimization Method is used to estimate *H* parameter. The proposed method used *CFE* criterion that above the threshold as LoSS while below the threshold as SOSS detection. Meanwhile, the probability of LoSS introduced in this paper can continuously measure the LoSS detection performance effectively. Our results show that at all sampling levels, the normal traces have a small probability of LoSS detection value below the threshold. On the other hand, the malicious traces have probability of LoSS detection value that imitate normal behavior at lower sampling level such as below 100ms. However, they exposed the probability of LoSS detection value clearly when sampling level is larger than 100ms. This is a

promising approach to obtain a continuous monitoring of LoSS detection method significantly in order to detect anomaly traffic behavior based on SOSS model. In future, we plan to test the reliability and robustness of the proposed method with more datasets.

ACKNOWLEDGMENT

The authors are grateful to Dr. Sulaiman Mohd Noor at CICT, UTM and Mr. Firoz at Unit IT, FSKSM for their helps in preparing the simulation of FSKSMNet dataset.

REFERENCES

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

[2] C. Cairano-Gilfedder and R.G. Clegg, "A decade of Internet research -- advances in models and practices," *BT Technology Journal 23*, Vol. 4, pp. 115-128, Oct. 2005.

[3] A. Erramilli, O. Narayan and W. Willinger, "Experimental queuing analysis with long-range dependent packet traffic," *IEEE/ACM Trans. Networking*, Vol.4, pp. 209–223, 1996.

[4] M.Y. Idris, A.H. Abdullah and M.A. Maarof, "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.

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

[6] 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, Vol. 53(6), pp. 463-467, June 2006.

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

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

[9] 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.

[10] M.F. Rohani, M.A. Maarof, A. Selamat and H. Kettani, "Uncovering Anomaly Traffic Based on Loss of Self-Similarity Behavior Using Second Order Statistical Model," *International Journal of Computer Science and Network Security (IJCSNS),* Vol.7 No.9, pp 116-122, September 2007.

[11] M.F. Rohani, M.A. Maarof, A. Selamat and H. Kettani, "Loss of Self-Similarity Detection with Second Order Statistical Model and Multi-Level Aggregation Approach," *Proceedings of the International Conference on Robotics, Vision, Information and Signal Processing ROVISP2007,* pp. 152-156, 28-30 November 2007.

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

[13] W. Yan, E. Hou and N. Ansari, "Anomaly Detection and Traffic Shaping under Self-similar Aggregated Traffic in Optical Switched Networks", Proceedings of ICCTZ003, pp. 378-381, 2003.