

Anomaly Detection of Network Traffic Based on Wavelet Packet

Jun Gao Guangmin Hu Xingmiao Yao
Key Lab of Broadband Optical Fiber Transmission
and Communication Networks,
UESTC, Chengdu 610054, China
{jgao,hgm,yxm}@uestc.edu.cn

Rocky K. C. Chang
Department of Computing
The Hong Kong Polytechnic University
Hung Hom, Kowloon, Hong Kong
{csrchang}@comp.polyu.edu.hk

Abstract—The rapid and accurate detection of network traffic anomaly is one of the preconditions to guarantee the effective work of the network. Aiming at the deficiency of present methods of network traffic anomaly detection, we propose a scale-adaptive method based on wavelet packet. By means of wavelet packet decomposition, our method can adjust the decomposition process adaptively, has the same detective ability to the anomaly of various frequency, especially the middle and high frequency ones which can not be checked out by the multi-resolution analysis. By means of adaptive reconstruction of the wavelet packet coefficient of different wavelet domains which anomaly, our method is able to confirm the characteristics of anomaly and enhance the reliability of detection. The simulation results prove that the method can detect the network traffic anomaly efficiently.

I. INTRODUCTION

The network traffic anomaly refers to the status that traffic behaviors depart from the normal behaviors. Many reasons, such as the misuse of network equipments, network operations anomaly, flash crowd, network intrusion and so on will cause network traffic anomaly. The characteristic of anomaly traffic is that it breaks out without any omen and can destroy networks and computers in a short time (For instance, the outburst of traffic behavior caused by specific attack programs or worm burst). Therefore, detecting traffic anomaly rapidly and accurately is one of the preconditions of ensuring the efficient network operation. The traffic in routers especially the routers in trunk network is very large and change continually, while the anomalous traffic is small, compared with the normal traffic and changes of the normal traffic. The ultimate aim of the anomaly detection algorithm is to detect the relatively small anomaly traffic from the relatively large background traffic. It is very difficult to implement this aims, so detection of anomalous traffic has become the attractive and valuable subject in the present academic and industrial circles.

Various schemes are proposed for the network traffic anomaly detection, such as the case based reasoning approach[1][2], the limit state machine approach[3][4], mode matching approach[5][6], statistical analysis approach[7][8], Hurst parameter analysis approach[9] and subspace analysis approach[10][11][12], etc. The achievements of former researches have greatly promoted the development of the anomaly detection and improved the detection results constantly. However, due to the complexity of network traffic

anomaly detection, there were problems in the real-time performance and the accuracy of detection still. Researchers have found that almost all the traffic time-varying signal were multi-scales [13], and the time-varying signals of the normal network traffics and that of the abnormal network traffics were different in frequency band range. That is to say, the difference between anomalous traffics and background traffics is various in different frequency bands. In certain frequency band, the energy of anomalous traffics is rather high in proportion to the total energy, so anomaly detection could be done easily. The wavelet transform can get arbitrary signal characteristic of time-frequency domain, which can help to explore the transient abnormal phenomenon from normal signals and demonstrate its components. Therefore, researchers put forward the wavelet analysis approach[14][15][16][17][18].

As a new technology, anomaly detection of network traffic based on wavelet analysis is taken more seriously recently years. V.Alarcon-Aquino presented an algorithm based on undecimated discrete wavelet transform and bayesian analysis[14]. This algorithm is able to detect and locate subtle changes in variance and frequency in the given time series, but its decomposition scale is limited and the algorithm is complicated. Anu Ramanathan presented a WADeS (Wavelet based Attack Detection Signatures) mechanism based on wavelet analysis to detect the DDoS attack[15], which makes wavelet transform for the traffic signals, then computes the variance of the wavelet coefficients to estimate the attack points. However, this method has very high computation complexity. Barford presented a method[16] presented a method which decomposes network traffic with decimated discrete wavelet transform , then synthesizes to Low, Mid, High frequency-parts, and finally detect anomaly with Deviation Scoring respectively. This algorithm is able to detect the flash crowds and short-term anomalies in postmortem. But it doesn't solve the problems of adaptive choice of detective scale and detection time-windows, has high computation complexity. Seong Soo Kim proposed a technique for traffic anomaly detection based on analyzing correlation of destination IP addresses in outgoing traffic at an egress router[17]. This technique can be employed for postmortem and real-time analysis of outgoing network traffic, but it has not the same detective ability towards various frequency anomaly since it is based on multi-resolution

analysis. Lan Li proposes an energy distribution based on wavelet analysis to detect the DDoS attack[18]. Research finds the energy distribution variance changes markedly causing a "spike" when traffic behaviors affected by DDoS Attack.

Generally speaking, two problems exist in the present anomaly detection methods based on wavelet transform: (1) Almost all the algorithms use multi-resolution analysis, so the poor resolution in middle-high frequency leads to the insufficiency of comprehensiveness in anomaly detection. That is to say, the detection method only works well on low frequency anomaly. (2) An anomaly probably distributes in many discontinuous frequency bands, so the detection is unauthentic if only in one scale. To solve the problems described above, this paper proposes a new network anomaly detection method based on wavelet packet transform. our method can adjust the decomposition process adaptively, and has the same detective ability to middle and high frequency as well as low frequency anomaly. By means of adaptive reconstruction of the wavelet packet coefficient of different wavelet domains which anomaly, our method is able to confirm the characteristics of anomaly and enhance the reliability of detection.

The rest of paper is organized as follows: In Section II we describe our method for anomaly detection. In Section III we illustrate the simulation results of our detection method, and Section IV concludes the paper.

II. ANOMALY DETECTION OF NETWORK TRAFFIC METHOD BASED ON WAVELET PACKET

A. A Statistical Detection Algorithm

The statistical detection algorithm is improved by the deviation scoring algorithm[16] proposed by P. Barford. Two detection windows are applied to the deviation scoring algorithm, one is *HisWin*, the other *DetWin*, as it is shown in Figure 1. Both slide with time, and are for real-time update. At time t , we compute the variance V_1 in the detection window $(t - DetWin, t)$, and the variance V_2 in the historical window $(t - HisWin, t)$. We set $ratio = V_1/V_2$, then, to some extent, the parameter $ratio$ represents the deviation of the sample variance in the detection window compared to the historical normal data variance. If the traffic is anomalous in the detection window, there must be an increase in the magnitude of $ratio$. The deviation scoring algorithm is originally used in reconstruction signal, while we made the first improvement by applying such algorithm directly in wavelet packet coefficient. At the initial anomaly detection of time t , through the wavelet packet decomposition for the network traffic signals during $(t - HisWin, t)$, we get the wavelet packet coefficient on each scale, which we detect by means of deviation scoring algorithm, then we get to the conclusion that whether there is an anomaly as long as the value of $ratio$ is beyond the alert threshold. At the confirmation of anomaly of time t , we reconstruct the wavelet packet coefficient on those scales and detect again by means of deviation scoring algorithm.

If we only use deviation scoring to detect anomalies, the low frequency anomalies which last for a long-time may be invalid. The main reason is that, under certain condition the low

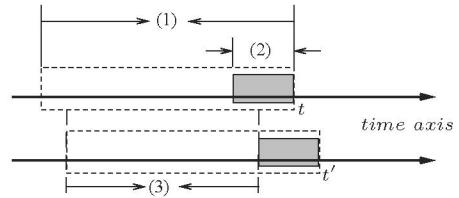


Fig. 1. The statistic detection algorithm based on sliding window.(1) HisWin, (2) DetWin, (3) The length of the redundancy data

frequency anomalies will keep stable when rising to a certain magnitude, and when the detection window is far shorter than the duration of the low frequency anomalies, the deviation scoring will abruptly change in the both beginning and end stage of anomalies, as well it will remain in the middle stage. At this time, a long-time anomaly is taken as two short-time ones. To solve this problem, we define a mean scoring to detect this kind of anomaly. At time t , we compute the mean E_1 of the detection window $(t - DetWin, t)$, and the mean E_2 of the historical window $(t - HisWin, t)$. Set $ratio_E = E_1/E_2$ as the mean scoring, which represents the variation of the sample in the detection window compared to the historical normal data. Generally speaking, the mean scoring will be stably larger than 1 at the beginning of the low frequency anomalies. Hence, combining with this mean scoring, we can accurately decide the long-time low frequency anomalies. We will illustrate it in step 3 of simulation.

B. Detection Model and The Double Thresholds Mechanism

The model of network traffic anomaly detection is shown in Figure 2. It can be divided into five parts in general, such as, wavelet packet analysis, initial anomaly detection, reconstruction of wavelet packet and confirmation of anomaly. The initial anomaly detection applies the double thresholds mechanism: they are two thresholds—alert threshold (T_a and T_{Ea}) and decomposition threshold (T_d and T_{Ed}), where $T_a > T_d$, $T_{Ea} > T_{Ed}$. If alert threshold is reached, we consider it as an anomaly, and if the decomposition threshold is reached, we consider it as could-be anomaly.

At first, we make multi-scale 1-level decomposition for the traffic signals and make wavelet packet decomposition for the coefficient of node [1,0] to the level 3. Now, we can detect the coefficient anomaly under different scales by means of statistic detection algorithm of the sliding window. When it is found that anomaly reaches the alert threshold in some certain scale of the former n levels, reconstruction detection starts immediately; if it is still anomalous, it alarms. When it is found that anomaly reaches the decomposition threshold in certain level of the n -th scale, decomposition continues to detect on the level $(n + 1)$, decomposition will end until the anomaly reaches the alert threshold or the anomaly is below the decomposition threshold. The wavelet packet decomposition tree of our method is shown in Figure 3.

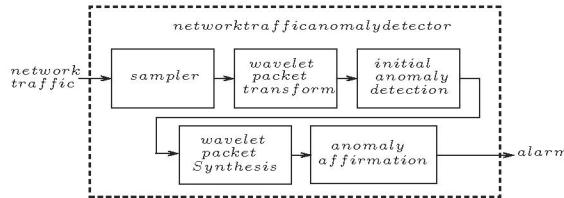


Fig. 2. The model of network traffic anomaly detection

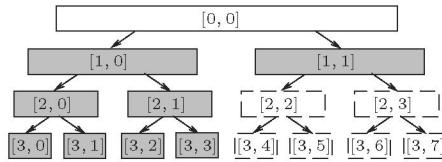


Fig. 3. The wavelet packet decomposition tree(level=3), and the fuscous panes are the coefficient of our method

C. Detection Steps

Functions of each module and the steps of detection are as follows, and the flow chart of detection algorithm is shown in Figure 4.

1. Generation of detection signal: We take packets gathered by router per unit time as the traffic signal, the sample interval is T_0 sec. If $f(n)$ is the value of the n -th sample, then:

$$f(n) = \begin{cases} 0, & n = 0. \\ (T_0 \cdot (n - 1), T_0 \cdot n] \text{ packet numbers,} & \text{otherwise.} \end{cases} \quad (1)$$

2. Wavelet packet analysis: Even the weak high-frequency anomaly could be detected from the node $[1,1]$, so it is enough to make multi-scale decomposition to level 1, and then make wavelet packet decomposition from the node $[1,0]$. With the increase of the levels of the wavelet packet decomposition, the number of the wavelet packet coefficient halves. If the length of the detection series is N , the length of the wavelet packet coefficient series from level j would be $N/2^j$. Besides, since the wavelet packet decomposition is based on 2-abstract DWT, with the increase of the levels, the nodes of each level increase by way of 2^j . Therefore, the levels decomposed in the initial period are limited, and afterwards they will continue self-adaptive decomposition with the specific detection situations.

3. Initial anomaly detection: Make the initial detection on the wavelet packet coefficient of each scale, and check whether there is any anomaly at some moments on this scale, by the means of statistical detection algorithm. If $\text{ratio} > T_a$ or $\text{ratio}_E > T_{Ea}$ (that is anomaly), go to step 4; if $\text{ratio} < T_a$, $\text{ratio} > T_d$ or $\text{ratio}_E < T_{Ea}$, $\text{ratio}_E > T_{Ed}$ (that is could-be anomaly), we further decompose the wavelet packet coefficient and then detect them by step 3 again; if $\text{ratio} < T_d$, $\text{ratio}_E < T_{Ed}$, the could-be anomaly can be removed. Therefore, the decomposition levels are totally self-adaptive.

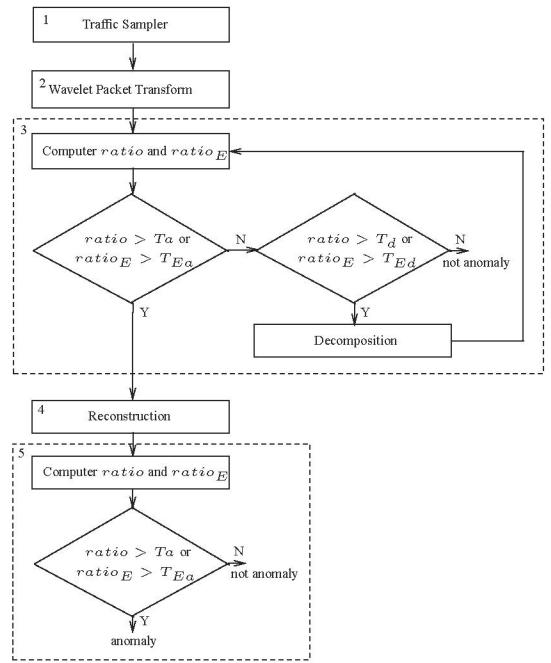


Fig. 4. The flow chart of the detection algorithm

4. Reconstruction of wavelet packet: The energy of anomalous signals mainly locates on those scales, where we detect the anomaly in step3. Therefore, we selectively reconstruct the wavelet packet coefficient on those scales. So in this way, the reconstructed signals can be distinct from original signals to greater degree.

5. Confirmation of anomaly: The anomaly we get from initial detection module may be a mis-detection, so we need to redetect the reconstructed signal to reduce the ratio of mis-detection. Furthermore, 2-abstract DWT has poor performance in time domain. We can get better performance in time location if we detect reconstructed signal. In our method, we set an alert threshold. If the detection result in reconstructed signal reaches the alert threshold, we regard it as an anomaly; otherwise, we recognize it as a mis-detection.

III. RESULTS

A. Simulation Backgrounds

In our simulation experiments, we adopt the data[19] as the background traffic, which were gathered by Lawrence Berkeley lab in University of California, Berkeley Institute; According to the principle of DDoS attack, we simulate 8 data source as attack source, which send a huge volume of traffic to a victim at the same time. We use the DDoS attack as anomalous traffics, which was produced in NS2 simulation. And then we simulate that eight attack source host computers launch attacks to the victim host computer during 100ms at random, if each attack source host computer sends data packets by exponential distribution with the average value of 10ms, it is a kind of low-frequency attack; if each attack source host computer sends 7,5,2,0,3 data packets respectively every 10ms,

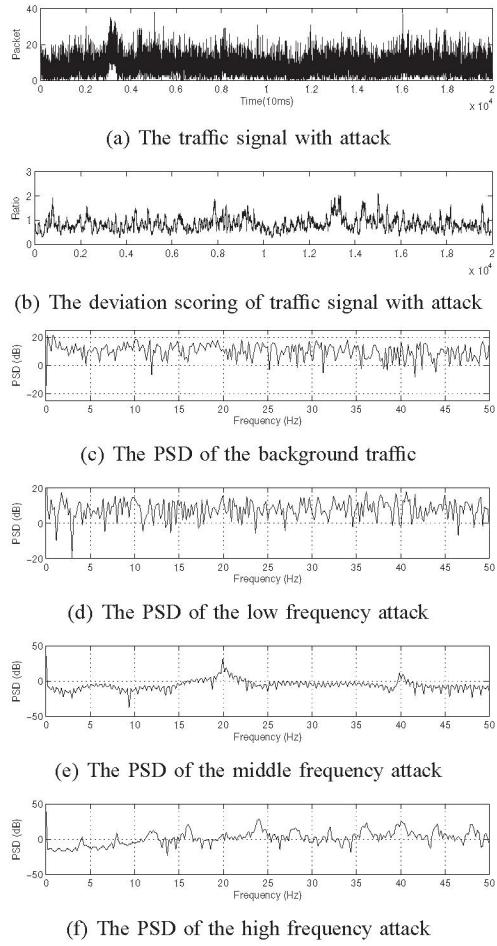


Fig. 5. The traffic signal, detection result and PSD

it is a sort of middle-frequency attack; if each attack source host computer sends 2,10,1,8,5 data packets respectively every 10ms, it is a kind of high-frequency attack. Attack 1 (starting time: 40-44s) added the background traffics is low-frequency attack; Attack 2 (starting time: 80-82s) is middle-frequency attack; Attack 3 (starting time: 160-161s) is high-frequency attack. The time interval of sampling is 10ms and the traffic data is shown in Figure 5(a). If we get deviational numeric (picking 30 from the detection window) in time domain traffics to detect, it cannot detect any attack at all, the result is shown in Figure 5(b), The choice of Wavelet packet filter is important in detection performance in the decomposition and reconstruction of traffic data. After weighing the effect of linearity, symmetry, vanishing moment, and localization, we select *Daubechies(db6)* filter. All the simulation experiments in this paper are operated on the computer with 2.4GHz Pentium 4 processor and 512M memory.

B. The Network Traffic Anomaly Detection Algorithm

Through the research of historic traffics, make sure the threshold value in each scale before detection.

1. Make 1-level multi-scale decomposition to the traffic signals and initial anomaly detection on the wavelet packet

coefficient in each corresponding level. As it is shown in Figure 6(a), Attack 3 surpasses the alert threshold at [1,1], so it is detected out. Then decompose node [1,0] to level 3; as it is shown in Figure 6(c) and 6(d), attack 1 and 2 surpass the alert threshold at [3,0] and [3,2] respectively, thus they are required to be further decomposed and detected. If we only use multi-scale decomposition, attack 2 hidden at [2,1],as is shown in Figure 6(b), can not be further detected, which will leads a mis-detection.

2. Decompose the coefficient [3,0],[3,2] of could-be attack into level 4, as in Figure 7. When finding attack 1 and 2 surpass the alert threshold at [4,0],[4,5], we consider there is an attack. Certain energy of attack 3 is concentrated at [4,4], so we should take it into consideration when making reconstruction.

3. Make reconstruction detection to ensure attack. Reconstruct [4,0] for attack 1, [4,5] for attack 2, [1,1] and [4,4] for attack 3, as the detection result shown in Figure 8(a),8(c) and 8(d), all of the three surpass the alert threshold. We find that only by using deviation scoring, we cannot detect out attack 1 completely. As is shown in Figure 11(a), a long-time anomaly is taken as two short-time ones. But combined with Mean Scoring, we can detect out that it is a durative attack, as in Figure 8(b).

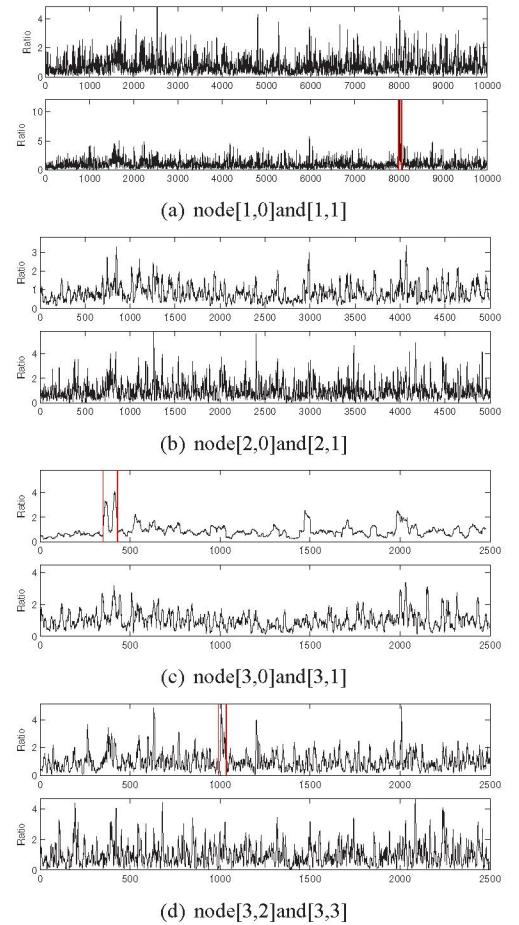


Fig. 6. The deviation scoring of 1,2,3-level wavelet packet coefficient

IV. CONCLUSION

In this paper, we propose a new mechanism of network traffic anomaly detection based on wavelet packet analysis. It can select different time-frequency resolution to decompose adaptively according to the characteristics of traffic signal. Using our method, we can locate time-frequency domains and get the faint signal effectively. According to simulation results, this mechanism is proved to be feasible, and it possesses some merits as follows: (1). It can effectively detect the long-time durative anomalous traffic and the short-time suddenly changing one, and also it can effectively detect middle-high frequency attack traffic which can not be checked out by the network traffic anomaly detection based on multi-resolution analysis. By means of decomposing threshold, it has avoided the blindness of wavelet packet decomposition, and solved the problem of decomposing scale's self-adaptive selection. (2). by means of adaptive reconstructing the wavelet packet coefficient in different wavelet domains which include anomaly, our method is able to confirm the characteristics of anomaly and enhance the reliability of detection.

ACKNOWLEDGMENTS

This work is supported by Sichuan Youth Science and Technology Foundation (No.04ZQ026-028) and National Science Foundation of China (No.60572092).

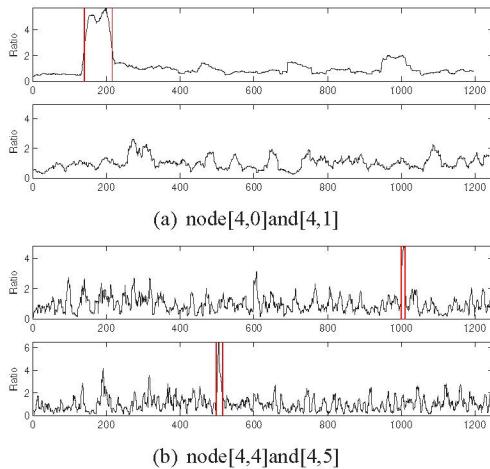


Fig. 7. The deviation scoring of level 4 wavelet packet coefficient

REFERENCES

- [1] L. Lewis and G. Dreo, "Extending trouble ticket systems to fault diagnosis," *IEEE Network*, vol. 7, pp. 44–51, 1993.
- [2] L. Lewis, "A case based reasoning approach to the management of faults in communication networks," *Proc.IEEE INFOCOM*, vol. 3, pp. 1422–1429, 1993.
- [3] I. Katzenbach and M. Schwarz, "Schemes for fault identification in communication networks," *IEEE/ACM Trans. Networking*, vol. 3, pp. 753–764, 1995.
- [4] I. Rouvellou and G. Hart, "Automatic alarm correlation for fault identification," *Proc. IEEE INFOCOM*, pp. 553–561, 1995.
- [5] F. Feather and R. Maxion, "Fault detection in an ethernet network using anomaly signature matching," vol. 23, pp. 279–288, 1993.
- [6] S. Papavassiliou, M. Pace, and L. Ho, "Implementing enhanced network maintenance for transaction access services: Tools and applications," vol. 1, 2000.
- [7] M. Thottan and J. Chuanyi, "Anomaly detection in ip networks," *IEEE TRANSACTIONS ON SIGNAL PROCESSING*, vol. 51, no. 8, 2003.
- [8] C.-M. Cheng, H. Kung, and K.-S. Tan, "Use of spectral analysis in defense against dos attacks," *Proceedings of IEEE GLOBECOM 2002*, 2002.
- [9] W. Allen and G. Marin, "On the self-similarity of synthetic traffic for the evaluation of intrusion detection systems," *Proceedings of the 2003 Symposium on Applications and the Internet (SAINT03)*, 2003.
- [10] A. Lakhina, M. Crovella, and C. Diot, "Diagnosing network-wide traffic anomalies," *ACM SIGCOMM*, 2004.
- [11] A. Lakhina, K. Papagiannaki, M. Crovella, C. Diot, E. Kolaczyk, and N. Taft, "Structural analysis of network traffic flows," *ACM SIGMETRICS*, 2004.
- [12] A. Lakhina, M. Crovella, and C. Diot, "Characterization of network-wide anomalies in traffic flows," *BUCS-2004-020*, Boston University, Tech. Rep., 2004.
- [13] B. BR, "Multi-scale analysis and modeling using wavelets," *Journal of Chemometrics*, vol. 13, 1999.
- [14] V. Alarcon-Aquino and A. Barria, "Anomaly detection in communication networks using wavelets," *IEEE Proc-Commun*, vol. 148, no. 6, 2001.
- [15] A. Ramanathan, "Wades: A tool for distributed denial of service attack detection," *TAMU-ECE-2002-02, Master of Science Thesis*, 2002.
- [16] P. Barford, J. Kline, D. Plonka, and A. Ron, "A signal analysis of network traffic anomalies," *Internet Measurement Workshop*, 2002.
- [17] S. S. Kim and A. Reddy, "Detecting traffic anomalies at the source through aggregate analysis of packet header data," *Proceedings of Networking*, 2004.
- [18] L. Lan and L. Gyungho, "Ddos attack detection and wavelets," *Telecommunication Systems*, pp. 435–451, 2005.
- [19] *Internet Traffic Archive*, <http://ita.ee.lbl.gov/index.html>.

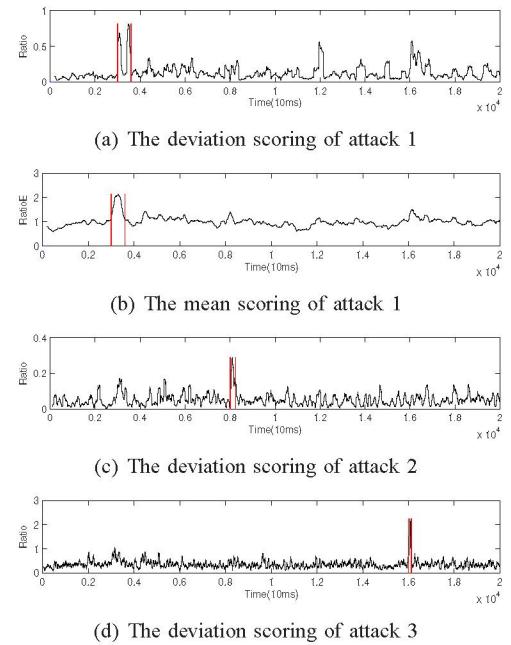


Fig. 8. The deviation scoring and Mean Scoring of the reconstructed signal