The Scale Self-Adjusting Multi-Resolution Network Traffic Anomaly Detection

Qi Xing-jian, Hu Guang-min, Yang Dan, Li Zong-lin
School of Communication and Information Engineering
University of Electronic Science and Technology of China
Chengdu, Sichuan, China

***Abstract*—Identifying network traffic anomalies accurately and rapidly is very critical to efficient operation of computer network. In this paper, to improve the existing anomaly detection, we propose a novel multi-resolution network traffic anomaly detection approach based on S Transform with self-adjusting scale. By introducing S Transform, we can decompose network traffic signal into a group of different frequency sub-bands according to the traffic signal's characteristics. By means of self-adjusting reconstruction of the signal from different frequency sub-bands, our method is able to confirm the anomaly characteristics and enhances the reliability of detection. By means of self-adaptive window selection, we are able to determine the length of detection window according to the spectrum characteristics of the corresponding signal. The simulation results prove that the method can detect the network traffic anomaly efficiently and rapidly, and excels the existing multi-resolution anomaly detection methods.**

***Keywords*—traffic anomaly detection, multi-resolution analysis, S Transform, deviation scoring**

I. Introduction

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 maloperation, flash crowd, network intrusion, etc, could cause network traffic anomaly. One characteristic of anomaly traffic is that it breaks out without any omen and brings a breakdown to networks and computers in a short time. Therefore, detecting traffic anomaly rapidly and accurately is one of the preconditions of ensuring the efficient network operation.

Many schemes are proposed for network traffic anomaly detection, such as the case-based reasoning approach, the limit state machine approach, the mode matching approach, the statistical analysis approach, the Hurst parameter analysis approach and subspace analysis, etc. These approaches can somewhat satisfactorily detect anomalies, but because of the great complexity of network traffic, they are often inaccurate and can not meet the real-time need.

Researchers have found that almost all the traffic time-varying signals are of multi-scales [1], and the time-varying signals of the normal network traffics and that of the abnormal network traffics were different in frequency ranges. Namely, the power difference of anomalous traffics and background traffics varies with frequency bands. In certain frequency bands, the energy of anomalous traffics is rather high in proportion to the total energy, which will make the anomaly detection easier. Being able to distill the signal's characteristics from arbitrary time and frequency range, the multi-resolution analysis method is good at detecting signal anomaly, and has lately become a worldwide hotspot. In 2000, V.Alarcon-Aquino presented an algorithm based on undecimated discrete wavelet transform and bayesian analysis [2]. 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 [3] based on wavelet analysis to detect the DDoS attack, which conducts wavelet transform on the traffic signals, then computes the variance of the wavelet coefficients to estimate the attack points. However, MALLAT based wavelet can only decompose the low frequency components and works well on low frequency signal, but works poor on medium and high frequency components.

Researchers then proposed wavelet packet analysis based anomaly detection algorithms. Their key advantage over MALLAT decompose is that, they are able to decompose high frequency band as well as low frequency band, so the algorithms are able to overcome MALLAT wavelet's shortcoming. Nevertheless, some problems do exist. First, it is hard to determine a good decomposition strategy and decomposition level count in particular. Second, the binary decomposition of MALLAT wavelet and the binary decomposition tree of wavelet packet analysis are both fixed scale decomposition. It takes great computational cost to make their decomposition deep enough to focus on certain frequency range, and their decomposition levels are very much limited by the detection window length. Such factors make the wavelet packet based detection algorithm unstable and may cause high false alarm rate.

In this paper, a self-adjusting multi-resolution network anomaly detection method based on S Transform is proposed aiming at the previous mentioned problems. By means of this method, the decomposition is able adjust itself to the spectrum characteristics of both normal and anomalous traffic signals. By reconstructing time series from different frequency band and using deviation scoring algorithm, we are able to confirm traffic anomalies. Using a double-threshold mechanism, we are able to enhance the reliability of anomaly detection. By means of self-adaptive detection window selection, we are able choose different window length according to different frequency scale. Finally, the simulation results show that the method proposed in this paper acquires good performance on anomalies of various frequency bands, and excels the existing multi-resolution network traffic anomaly detection methods.

978-1-4244-1674-5/08 /$25.00 ©2008 IEEE

II. S TRANSFORM

Generally speaking, the energy of power spectral density (PSD) of normal traffic in each frequency band is relatively well-proportioned, but the energy of the anomalous traffic is concentrated in certain frequency bands. Researchers used multi-resolution analysis to detect anomaly just based on the differences between normal and anomalous traffic signals in the frequency domain. Since the scale of the present detection algorithms based on multi-resolution analysis is binary and fixed, the frequency resolution is relatively poor in high frequency. However, traffic anomalies can be caused by many reasons, and as a result, the anomalies traffic signal may occur in low frequency band as well as in high frequency band. Therefore, these approaches cannot detect anomaly traffic of various frequency bands effectively.

Aiming at this problem, this paper proposes a detection algorithm based on S Transform. S Transform is able to adjust the decomposition scale to meet various needs, and enables us to adjust analysis scales according to anomaly characteristics.

S Transform is a time-frequency analysis approach. It is an extension of both wavelet transform and Short-Time Fourier Transform (STFT). The STFT of signal *h(t)* is defined as follows:

$$STFT(\tau,f)=\int_{-\infty}^{\infty} h(t)g^*(t-\tau)e^{-2\pi f t i}dt \qquad (1)$$

In (1), *g(t)* is a window function. When time τ changes, the window specified by *g(t)* shifts on t-axis, and the signal *h(t)* is analyzed gradually. With *g(t)* being a normalized Gaussian window, we get:

$$g(t)=\frac{1}{\sqrt{2\pi}\,\sigma}\exp\left(-\frac{t^2}{2\sigma^2}\right) \qquad (2)$$

Coefficient σ defines the window length. In order to implement high frequency resolution for low frequency band and high time resolution for high frequency band, we let $\sigma = 1/|f|$, and rewrite (1) as follows:

$$STFT(\tau,f)=\int_{-\infty}^{\infty} h(t)\left\{\frac{|f|}{\sqrt{2\pi}}\exp\left[\frac{-f^2(\tau-t)^2}{2}\right]\exp(-2\pi f t i)\right\}dt \qquad (3)$$

Equation (3) is the very representation of S Transform of signal *h(t)*.

In order to confirm the detection results, the decomposed components of S spectrum indicating anomalies have to be reconstructed into time series. The inverse of S Transform is able to implement a lossless reconstruction of time series using the following equation:

$$h(t)=\int_{-\infty}^{\infty}\left\{\int_{-\infty}^{\infty} S(\tau,f)d\tau\right\}\exp(2\pi f t i)df \qquad (4)$$

S Transform possesses the strongpoint that STFT and wavelet transform possess, and is free of their shortcomings, thus has now become a hotspot on signal processing. S Transform has the following advantages over STFT and wavelet transform: (1) the resolution of S Transform is relative to the frequency scale; (2) the result of S Transform is directly related to Fourier Transform; (3) the original signal can be reconstructed from several frequency band of S spectrum, and the reconstructed signal is free of spectrum leak; etc. Besides, S Transform has a key advantage over wavelet and wavelet packet decomposition as well: the S spectrum could be decomposed to arbitrary scale, that is, the frequency analysis scale and decomposition strategy is not fixed and can be arbitrary. Therefore, we are able to distill and focus on certain frequency band among the enormous background traffic flow.

III. SCALE SELF-ADJUSTING MULTI-RESOLUTION NETWORK TRAFFIC ANOMALY DETECTION METHOD

A. *Self-Adjusting Decomposition and Reconstruction Based on S Transform*

As illustrated in Fig. 1, the detection consists of flour elements: Network Traffic Sampling, S Transformation on sampled signal, Decomposition and Reconstruction of sub-band time series, and Anomaly Detection.

First, we acquire the S spectrum matrix by applying S Transform to the sampled signal. Second, we apply the self-adjusting decomposition to the S spectrum matrix. Finally, we can process the reconstructed signal using the deviation scoring algorithm.

The traffic signal can be virtually characterized by its power spectrum. On this basis, we can have the decomposition adjust to the traffic's power spectrum, and can enhance the detection efficiency and accuracy. Thus in this paper, we decompose the S spectrum into components that equal by power.

In order to control the decomposition level, this paper introduces a Double-Threshold mechanism. This mechanism defines an alarm threshold α and a decomposition threshold β ($\alpha > \beta$). When the detection result of the current decomposed component is above α, alarm is to be generated and further decomposition on this component is not needed. When detection result is between α and β, possible anomaly exists and further decomposition should be applied. When detection result is below β, there is no anomaly and no need to perform further decomposition either. The decomposition is to be carried out recursively on the basis of the Double-Threshold mechanism, until either alarm is produced or no anomaly is confirmed. Combining this mechanism and the randomicity of S spectrum decomposition, we can have the decomposition

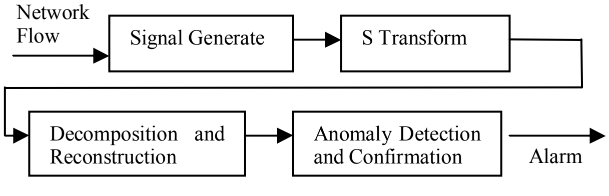

Network Flow
Signal Generate
S Transform
Decomposition and Reconstruction
Anomaly Detection and Confirmation
Alarm

Figure 1. S Transform based network traffic anomaly detection model

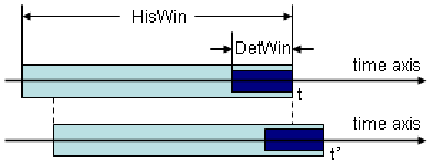


Figure 2. Two detection windows are applied to the deviation scoring algorithm

scale and depth be self-adjusting, and can acquire remarkable results.

B. Statistical Detection Algorithm

The statistical detection algorithm is a modified version of the deviation scoring algorithm [6] proposed by P. Barford. Two windows, *HisWin* and *DetWin*, are applied to the deviation scoring algorithm, as shown in Fig. 2, both of which slide and update with time. At time t, we compute the variance *V* in the historical window (*t - HisWin, t*), and the variance *U* in the detection window (*t - DetWin, t*). If the traffic is anomalous in the detection window, there must be an increase in the magnitude of *ratio = U / V*. We can apply deviation scoring algorithm to the reconstructed time series, and compute the ratio defined above, and the ratio is what is used to confirm anomaly.

There are three key coefficients that should be determined in advance:

a) HisWin. In order to determine the length of the historical window, the system resource consumption and the detection sensitivity should both be taken into account. Anomalies are sudden and temporary, whereas normal traffics are time-lasting and relatively stable. As a result, the longer the historical window is, the closer to normal traffic the behavior can it characterize, and the less effected by temporary phenomenon it is. Nevertheless, longer window will consume more system resources, and is less sensitive to normal traffic's changes.

b) DetWin. The algorithm works best when the detection window length is close to the duration of anomalies, which is never fixed and, most importantly, unknown. In order to determine the detection window length, we propose a self-adaptive selection method on the basis of the signal's frequency range, which will be discussed soon after.

c) Confirmation threshold. We refer to historical traffic and select that $ratio_{threshold} = \bar{x} + 3\sigma$. $\bar{x}$ refers to the mean value of traffic flow, and σ refers to the deviation standard of traffic flow.

C. The Self-Adaptive Selection of Detection Window Length

Generally, it is hard to determine the detection window length. Selecting the window length at random won't bring in satisfying results. Therefore, we propose a window length selecting method based on the center frequency of the concerned frequency band.

The signal to be detected is reconstructed from the decomposed component of S spectrum, thus it is easy to compute the center frequency f_c of the signal. Then we get the approximate period of the reconstructed signal: $T = 1/f_c$. Take it that the sampling rate of the signal is Δ, and then the number of samples within one period is:

$$L = T/\Delta = 1/(f_c \cdot \Delta) \tag{5}$$

Let the window length be integer times of the signal period, and we get:

$$DetWin = k \cdot L = k/(f_c \cdot \Delta) \tag{6}$$

As we can see, the window length computed based on signal's center frequency varies with frequency scale, that is, longer window for lower frequency band and shorter window for higher frequency band. On one hand, it is a feasible approach to compute the detection window length using (6). On the other hand, the window length determined using (6) varies with signal's center frequency, and can adjust to different frequency range.

D. Anomaly Detection Flow

The flow chart of anomaly detection is shown in Fig. 3.

1. Apply S Transform to traffic signal, and compute its S spectrum.
2. Decompose the S spectrum to the nth level. Divide the current frequency range of S spectrum into three components, that is, low frequency band, medium frequency band, and high frequency band, and their energy should be equal and each possesses one third of the current frequency range's total energy.
3. Reconstruct the time series of the sub-bands of the nth level, and apply the deviation scoring algorithm to them.
4. When a sub-band's detection result is below decomposition threshold or above alarm threshold, stop further decomposition on this component.
5. When a sub-band's detection result is above decomposition threshold and below alarm threshold, further decomposition should be applied to it.
6. When several sub-bands alarm at a same time range, they may indicate a same anomaly. Put these sub-bands together and confirm the anomaly on the reconstructed signal once more.

IV. SIMULATION RESULTS

A. Simulation Background

In our simulation experiments, we adopt the data [7] 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 sources as attack sources, which send a huge volume of traffic to a victim at the same time. The simulation topology is shown in Fig. 4. Ax(x = 1 ~ 8) are the data sources of DDoS attack, V is the victim host and R is the router before the Victim host. We use

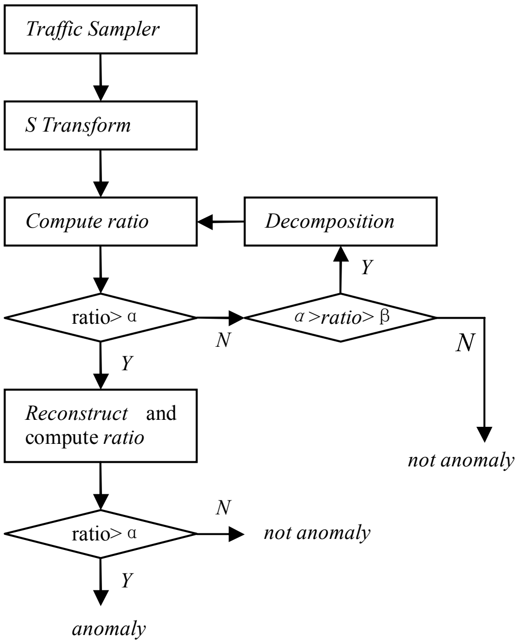


Figure 3. The flow chart of anomaly detection

the DDoS attacks as anomalous traffics, which were generated using NS2. The 8 attack hosts start low-frequency, middle-frequency and high-frequency attacks to the victim randomly within 100ms. Attack 1 (starting time: 30-34s) added the background traffic is low-frequency attack, attack 2 (starting time: 80-82s) medium-frequency, and attack 3 (starting time: 160-161s) high-frequency. The time interval of sampling is 10ms and the traffic data are shown in Fig. 5(a). If we detect traffic anomaly using deviational numeric (picking 30 from the detection window) of traffic signal in time domain traffics, we can only detect low-frequency attack whose last time is the longest and whose amplitude is the greatest. The result is shown in Fig. 5(b). Background traffic and attack power spectrum are shown in Fig. 5(c)(d)(e)(f). It can be seen from the illustration that background traffic has wide frequency domination and well-distributed energy, while the frequency of attack flow is relatively narrow and the energy centralizes in certain ranges. All the simulation experiments in this paper are carried out on a computer with 2.4GHz Pentium 4 processor and 512M memory.

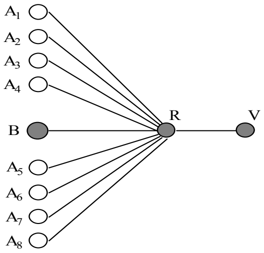


Figure 4. The topology of simulated network

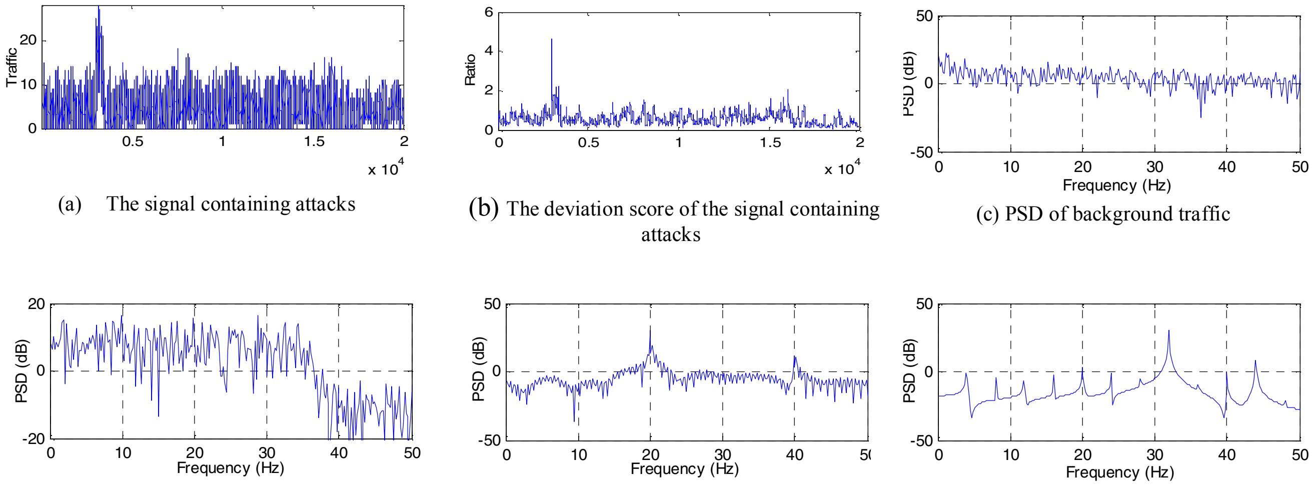
B. Detection Result

The threshold value in each scale is determined through the research of historical traffic before detection.

1. Decompose the traffic signal into three sub-bands, V_H for high frequency band, V_M for medium frequency band, and V_L for low frequency band. Apply the initial anomaly detection one by one to the three sub-bands. The deviation score of the signal reconstructed from sub-band V_L reaches the alarm threshold at point 3000 ~ 34000, as shown in Fig. 6(a), which indicates that the Attack 1 that added into the background traffic has been detected out and alarm should be generated. The deviation score of V_M reaches decomposition threshold at point 8000 ~ 8200, as shown in Fig. 6(b), so V_M needs further decomposition. V_H reaches the decomposition threshold at point 16000 ~ 16100 too, as shown in Fig. 6(c), and needs further decomposition as well.

2. Decompose V_M to 2nd level, and we acquire three more sub-bands, V_{MH}, V_{MM} and V_{ML}. V_{MM} reaches the alarm threshold at point 8000 ~ 8200, as shown in Fig. 7(b), and it is the Attach 2 that added into the background traffic. The deviation scores of V_{MH} and V_{ML} are below decomposition threshold, as shown in Fig. 7(a)(c), which means there is no anomaly component in these two sub-bands.

3. Decompose V_H to 2nd level, and acquire three more sub-bands, V_{HH}, V_{HM} and V_{HL}. V_{HL} reaches the alarm threshold at point 16000 ~ 16100, as shown in Fig. 7(d), and it is the Attach 3 that added into the background traffic. The deviation scores of V_{HM} and V_{HH} are below decomposition threshold, as shown



(a) The signal containing attacks

(b) The deviation score of the signal containing attacks

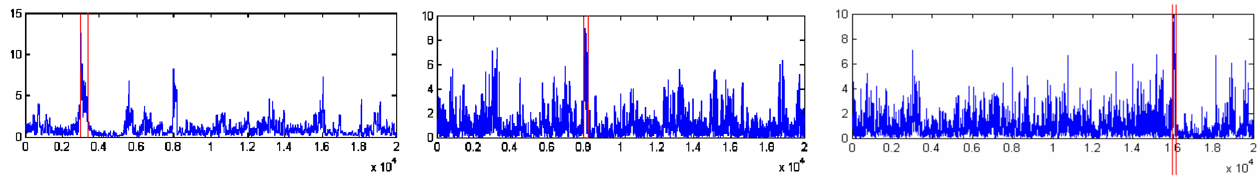
(c) PSD of background traffic

(d) PSD of low frequency attack

(e) PSD of medium frequency attack

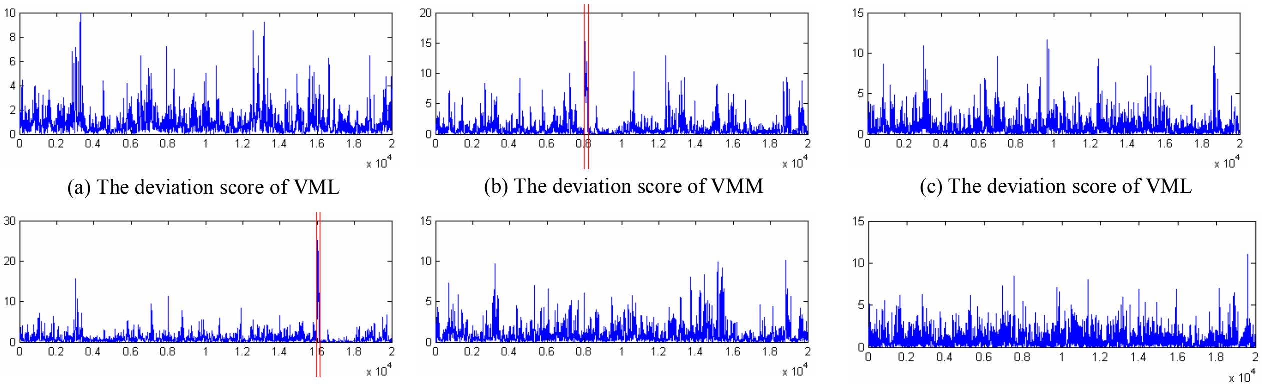
(f) PSD of high frequency attack

Figure 5. The simulated traffic



(a) The deviation score of VL (b) The deviation score of VM (c) The deviation score of VL

Figure 6. The 1st level decomposition and their detection results



(a) The deviation score of VML (b) The deviation score of VMM (c) The deviation score of VML

(d) The deviation score of VHL (e) The deviation score of VHM (f) The deviation score of VHH

Figure 7. The 2nd level decomposition and their detection results

in Fig. 7(e)(f), which means there is no anomaly in these two sub-bands.

The decomposition goes to the 2nd level when we successfully detected out the three attacks that are added to the background traffic.

C. Comparison with CWT

Literature [9] proposed that Adaptive threshold or Cumulative sum can be adopted to detect traffic signals roughly to find out suspicious anomaly time as much as possible, and a precise detection on continuous wavelet transform coefficient of anomaly time point found during the rough detection is then performed. Fig. 8 shows the results detected in the way proposed in literature [8]. To reduce the influence of rough detection on the final results, all of the results of rough detection are alarmed, i.e. analyzing continuous wavelet transform coefficient at each time point during the precise detection. The figure above is the analyzing results of CWT coefficient of original traffic, and the below is the detecting results added by the same anomaly traffic mention in this paper. Shown as the red area in the figure, maximum coefficient of attack 1 with a wide band is increased when attack is added while those of attack 2 and 3 with narrow bands change little, which are even less than the maximum coefficient got at some time in original traffic, which is shown in green area in the figure. Therefore, the following threshold division can't distinguish normal traffic and anomaly traffic.

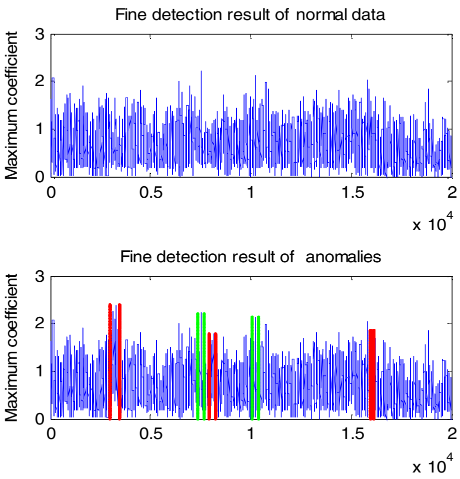


Figure 8. The detection result using CWT

Literature [9] analyzed wavelet transformation coefficient of traffic signal at alert time, which can reduce error alarm of AT and CUSUM. But for precise detection is conducted only on traffic at alert time during rough detection, accuracy of the whole detection can then be guaranteed only if the rough detection has detection failure rate as low as possible. During the continuous wavelet transformation, decomposition scale of this method is (log_2N -1) * *12*, i.e. the traffic signal needs to decompose 168 times at one time when its length is 20000 sampling points. While in our simulation, the same traffic signal is decomposed to the second level, and the number of decomposition is far less than CWT. So the computational complexity of our algorithm is much less than CWT. Simulation result also proves that the selection of detecting window in this paper can successfully detect attacks of all frequency, both wide and narrow.

V. Conclusion

In this paper, we propose a scale self-adjusting network traffic anomaly detection method based on S Transform. It can select different time-frequency resolution to decompose and adjust to the characteristics of traffic signal. According to the simulation results, this mechanism is proved to be feasible, and it possesses the following merits: (1) It can effectively detect the long-time durative anomalous traffic and the short-time sudden-changing one, as well as the middle/high frequency anomaly traffic which probably can't be detected by the network traffic anomaly detection methods based on multi-resolution analysis. (2) By means of self-adjusting selection of decomposition scale, and the Double-Threshold mechanism, our method can avoid the blindness of wavelet packet decomposition, and can also reduce computational complexity. (3) By means of self-adjusting reconstruction of different S spectrum sub-bands which contain anomalies, our method is able to confirm the characteristics of anomaly and enhance the reliability of detection. (4). Using the center frequency of different frequency band to compute the size of the corresponding detection windows, we have found out a solution to the problem of detection windows selection.

Acknowledgment

The work described in this paper was supported by NSFC (Project No.60572092), National Key Basic Research Program of China "973 Project" (2007CB307104 of 2007CB307100) and Program for New Century Excellent Talents in University.

References

[1] B.R.Bakshi, "Multi-scale analysis and modeling using wavelets," Journal of Chemometrics, 13, (3), 1999

[2] V.Alarcon-Aquino, and J.A.Barria. "Anomaly Detection in Communication Networks Using Wavelets," IEEE Proc-Commun, vol.148, No.6, December 2001.

[3] Anu Ramanathan, "WADeS: A Tool for Distributed Denial of Service Attack Detection", TAMU-ECE-2002-02, Master of Science Thesis, August 2002.

[4] K Ramchandran, and M Vetterli. "Best wavelet packet bases in a rate-distortion sense," IEEE Trans on Image Processing, 1993,2(2):160-175.

[5] R.G. Stockwell, L. Mansinha, and R.P. Lowe, "Localization of the complex spectrum: The S transform," IEEE Trans. Signal Process. 44 (1996) 998-1001.

[6] P. Barford, J. Kline, D. Plonka, and A. Ron. "A signal analysis of network traffic anomalies," In InternetMeasurement Workshop, 2002.

[7] Lawrence Berkeley National Laboratory, The Internet Traffic Archive, http://ita.ee.lbl.gov/index.html.

[8] Dainotti Alberto, Pescape Antonio, and Ventre Giorgio, "Wavelet-based Detection of DoS Attacks," GLOBECOM '06.

[9] A. Magnaghi, T. Hamada, and T. Katsuyama, "A Wavelet-based framework for proactive detection of network misconfigurations," ACM SIGCOMM'04 Workshops, 2004.