Detection of Network Traffic Anomaly Based on Instantaneous Parameters Analysis

Xingmiao Yao Peng Zhang Jie Gao Guangmin Hu
Key Lab of Brodband Optical Fiber Transmission and Communication Networks
University of Electronic Science and Technology of China
Chengdu, Sichuan Province, China
yxm@uestc.edu.cn

Abstract—Identifying network traffic anomalies accurately and rapidly is critical to the efficient operation of any network. In this paper, a new algorithm is proposed based on instantaneous parameters (Instantaneous frequency and Instantaneous amplitude) analysis. The characteristic of traffic anomaly would be revealed more evidently through analyzing the instantaneous parameters of the original network flow data. The simulation shows that the proposed algorithm can identify network traffic anomaly effectively.

Key words: instantaneous parameter, network traffic anomaly detection, sliding window, variance ratio analysis

I. INTRODUCTION

Typical network traffic anomaly refers to situation when network traffic departs from its normal performance. It can be aroused by various reasons, e.g. overload of the network, invasion of the network, wrong operation of network equipments. Detection of network traffic anomaly can be widely applied to network management, prevention of network invasion and improvement of network performance, etc. Detection of network traffic anomaly with accuracy and real time has become the focus of the present anomaly detection technologies. Many methods have been proposed to solve this problem. Most of the them are to detect by anomalous traffic's statistical characteristics in time domain[3][4][5][7], e.g. using change rate of IP address or change rate of port. But their limitations restrict the wide application of these methods. Moreover, with the involvement of more complex factors, especially DDos network attack, simple time domain detection measures don't work efficiently. In recent years, some researchers successfully applied wavelet transform method to detection of network traffic anomaly. But, the complexion of wavelet algorithm restricts its application to those traffic anomaly detections with real time. Some researchers then switch to frequency analysis, hoping to find out more clear internal characteristics of the anomaly by this way, e.g. Chen-Mou Cheng proposes a method to identify DDos attack by analyzing TCP traffic's period characteristics with energy spectral density of traffic signal (but it can only identify part of DDos attack)[6].

A new detection of anomaly is proposed in this paper, which detects traffic anomaly by computing and analyzing network traffic signal instantaneous parameters (instantaneous frequency[8] and instantaneous amplitude) got by Generalized Hilbert Transform of original traffic data. This detection algorithm is simple, fast and with good real time. Our research is based on two hypotheses: anomalous traffic possesses some characteristics which are different from normal traffic; the value of anomalous traffic is far less than that of the normal traffic.

II. INSTANTANEOUS PARAMETERS

There are three important instantaneous parameters for non-stationary signals: instantaneous amplitude, instantaneous frequency and instantaneous phase. They are very crucial for many practical applications. G. Gabor proposes a analytic signal method[1] to compute these instantaneous parameters, which essentially computes the parameters by Hilbert Transform. As Hilbert transformation is sensitive to noise, this method can't obtain accurate instantaneous parameters for those signals with noise. On the contrary, instantaneous parameters got by Generalized Hilbert Transform[2] are hardly influenced by noise, which clearly reflect some essential characteristics of the signal.

A. Generalized Hilbert Transform

For any continuous time signal $x(t)$, define its Generalized Hilbert Transform $h(t)$ as

$$h_r(t) = \{2 * \sum_w \{\mathrm{Re}[X(t,w)]\}^n + \mathrm{Re}[X(t,0)]\}^{\frac{1}{n}} \quad (1)$$

$$h_i(t) = \{2 * \sum_w \{\mathrm{Im}[X(t,w)]\}^n\}^{\frac{1}{n}} \quad (2)$$

$$h(t) = h_r(t) + i * h_i(t) \quad (3)$$

Here $\sum_w$ is to sum all the positive frequencies. $h_r(t)$ is real part of L^n order Generalized Hilbert Transform, $h_i(t)$ is imaginary part L^n order Generalized Hilbert Transform, $h(t)$ is L^n order Generalized Hilbert Transform. $X(t,w)$ is Fourier transform of time signal $x(t)$ in time window whose center time is *t*, if gauss function is adopted as window function, else $X(t,w)$ is Gabor Transform.

Generalized Hilbert Transform develops traditional Hilbert Transform from two aspects. One is to introduce window function, the other is to introduce the concept of L^n. Emphasizing the "instantaneous" characteristic of the original signal, the former better shows the anomaly of the network; smoothing filtering the original signal, the latter guarantees the reliability and accuracy of process.

B. Instantaneous Amplitude

In Generalized Hilbert Transform, instantaneous amplitude, according to equation (3), is defined as

$$a_g(t) = \sqrt{h_r^2(t) + h_i^2(t)} \quad (4)$$

C. Instantaneous Frequency

In Generalized Hilbert Transform, instantaneous frequency is defined as $\theta(t) = \arctan\frac{hi(t)}{hr(t)}$

$$\omega(t) = \frac{\mathrm{d}\theta(t)}{\mathrm{dt}} \quad (5)$$

III. Detection of Network Traffic Anomaly Based on Instantaneous Parameter Analysis

Detection of network traffic anomaly regard network traffic signal as one dimensional time-varying signal (that we called as traffic signal). Our idea of detection is: to detect the anomaly by the difference of anomalous signal statistical characteristics with normal situations, that is, analyzing the abrupt change. The step is to process the traffic signal by Generalized Hilbert Transform, compute the instantaneous amplitude and instantaneous frequency respectively, and show the characteristics of traffic anomaly in time and frequency domain by variance analysis of instantaneous amplitude and frequency, then set appropriate threshold to detect anomalous traffic in time- frequency domain, and unite the detection results from the two areas.

A. Detection Algorithm

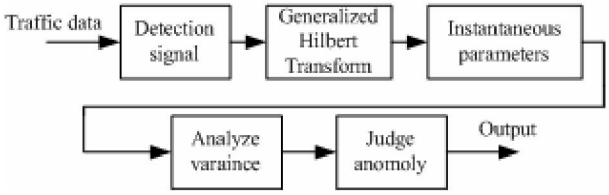

Figure 1 Traffic anomaly detection model

Step1 Detection signal generate module: packages which pass router in unit time is used as detection signal, unit time interval is T_0 sec. Firstly compute original network traffic package information collected by router to produce detection signals. The rule is as follows:

$$f[n] = \begin{cases} 0, n = 0 \\ packages \quad in \quad time \quad [T_0 \cdot (n-1), T_0 \cdot n], otherwise \end{cases}$$

$f[n]$ is the value of the *n*th sample.

Step2 Instantaneous parameters computing module: firstly process the detection signal by Generalized Hilbert Transform, then compute instantaneous parameters of network traffic.

Step3 Variance analysis module: deal with instantaneous parameters data of detection window and historical window by variance analysis method.

Step4 Judge anomaly model: According to the results of variance analysis and experience of historical traffic, set an alarm threshold in time domain and frequency domain respectively. Anomaly is then decided, if alarm threshold is achieved. Unite the anomaly result of time domain and frequency domain.

B. Variance Statistic Detection Algorithm

Statistical characteristic of the normal traffic, which is variance of the traffic signal instantaneous parameters in this paper, is obtained by monitoring the normal network and

analysis of historical network traffic data before anomaly detection,. In real time detection, we detect the present network traffic signal to decide whether there is an anomaly or not. We adopt sliding window's sample variance detection algorithm, shown in figure 2. $HisWin$ which is based on historical variance and detection window $DetWin$ are involved in. sliding with time, both of the two windows are real time.

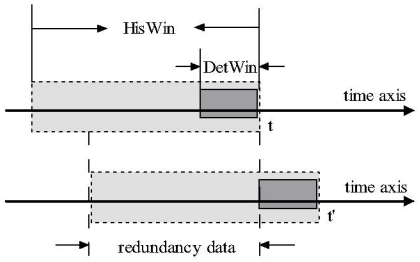

Figure 2 Sliding window variance detection

With the change of the time, we compute variance V_1 of detection window of $(t-DetWin,t)$ and variance V_2 of historical window $(t-HisWin,t)$ at the present time t. Set $ratio=(V_1/V_2)^2$, parameter $ratio$ reflects departure of the sample of detection window from historical normal data. If there is an anomaly at present time, it will influence the results of detection window and $ratio$ will consequentially increase in a parameter value. The time when the value of $ratio$ is more than that of alarm threshold is regarded as anomaly.

Three parameters are involved in the detection algorithm:

(1) Historical window: the bigger the historical window size, the nearer sample variance approaches variance of total signal and the more accurate the results. However, too large historical window will result in increase of the cost of system's storage and computation. Therefore, we should balance both of the window size.

(2) Detection window: it is ideal if the size of detection window is equal to the lasting time of possible anomaly. But usually, lasting time of anomalous traffic varies in a certain range. In order to detect all the anomalies in traffic, we select the time of the longest anomalous traffic as detection window.

(3) Judge threshold: according to the analysis of historical traffic, select anomalous threshold $ratio_{threshold}=\bar{x}+3\sigma$. $\bar{x}$ is average value of traffic, σ is variance of traffic.

IV. Simulation Results

In detection of traffic anomaly simulation, background traffic adopts the data collected by Lawrence Berkeley laboratory of University of California, Berkeley. According to theory of DDos attack, we only need to simulate lots of Constant Bit rate (CBR) as attack resource in simulation topology and make these CBR send information to the attacked side simultaneously in a short time. During our simulation, 8 attacking source CPU simulated by OPNET send attacking data to the attacked CPU simultaneously and each attacking source CPU attack in exponential distribution whose mean value is 10ms. Figure 3 is traffic data with Attack 1 (time range: 3000-3400), Attack 2 (time range: 5000-5400) and Attack 3(time range: 14000-14400).

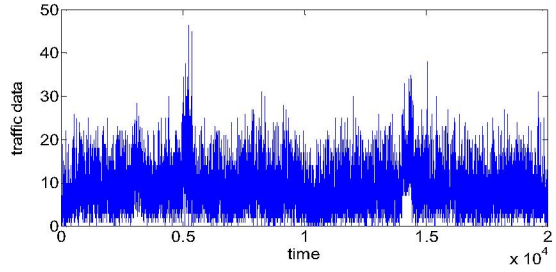

Figure 3 Traffic data with attack

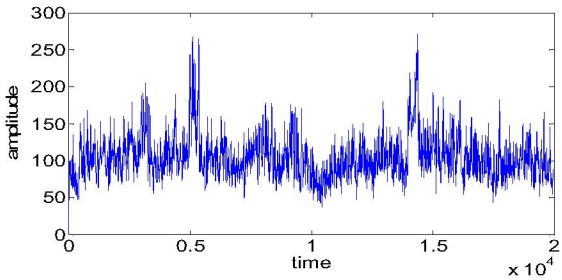

Figure 4 Instantaneous amplitude of traffic data

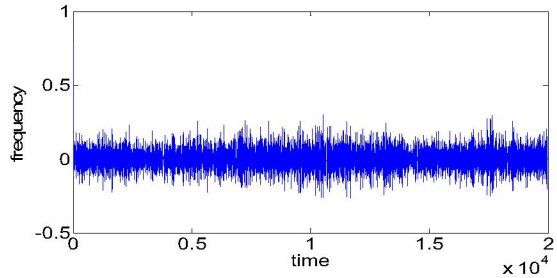

Figure 5 Instantaneous frequency of traffic data

In figure 4 and figure 5, instantaneous amplitude and instantaneous frequency got by Generalized Hilbert Transform of traffic signal are computed by the above fast algorithm. In the process of Generalized Hilbert Transform of traffic signal, order is selected as n=5 and window function is gauss window function.

For variance detection algorithm, we select the length of historical window as 20000, detection window as 450. Considering the reliability and real time of detection results, we select sliding step of detection window as 10. A ratio is

computed every 10 samples the detection window slides. Instantaneous amplitude and Instantaneous frequency got by Hilbert Transform of signal are shown in figure 6 and figure 7 respectively.

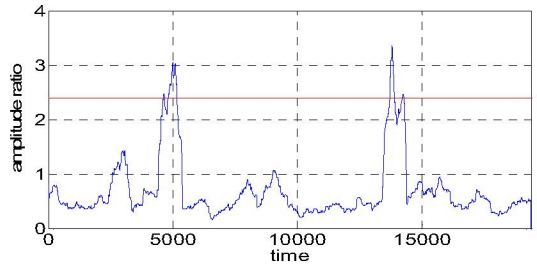

Figure 6 Ratio value of instantaneous amplitude

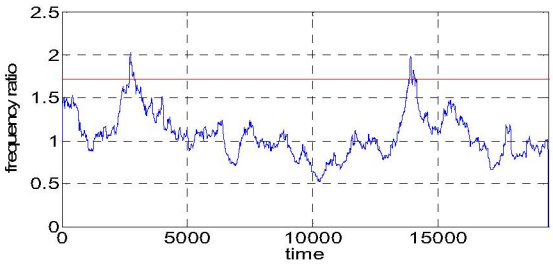

Figure 7 Ratio value of instantaneous frequency

From figure 6, it can be clearly seen that ratio computed by instantaneous amplitude got by Generalized Hilbert Transform of signal can show anomaly of Attack 2 and Attack 3. Figure 7 shows that ratio computed by instantaneous frequency detect the anomaly of Attack 1 and Attack 3. Uniting these results, we get all the anomaly results by Attack 1, Attack 2 and Attack 3.

V. CONCLUSION

This paper proposes a new mechanism which computes instantaneous parameters to detect network traffic anomaly by Generalized Hilbert Transform of traffic signal. Simulation proves its practicability. It has the following advantages:

It adopts two parameters, which detects more anomalies and avoids omitting the real existing anomaly.

Emphasizing "instantaneous" characteristic of traffic signal, smoothing filtering the traffic signal and getting rid of the disturbance of noise, process traffic signal by Generalized Hilbert Transform guarantees the reliability and accuracy of signal process.

It adopts sample variance detection algorithm with simple idea and operation, which can deal with network traffic in a short time and achieve real time detection.

Compared with other methods of traffic anomaly detection, it is simpler, possesses better identifying ability, and can be better adapt to network with higher real time.

REFERENCES

[1] D. Gabor, Theory of communication. J Inst Elect Eng(London), pp429-457,1993.

[2] Yi Luo et, Saleh Al-Dossary, Marhoon Maher, etc. "Generalized Hilbert Transform and its Applications in geophysics", The Leading Edge,Vol. 22 , No.3, pp198-202, March,2003.

[3] Marina Thottan, Chuanyi Li, "Anomaly Detection in IP Networks", IEEE Transactions on Signal Processing, Vol.51, No.8, pp2191-2204, August 2003.

[4] Marina K. Thottan, Chuanyi Ji, "Properties of Network Faults", Proceedings of the IEEE/IFIP Networks Operations and Management Symposium, April 2000.

[5] Matthew V.Mahoney, "Network Traffic Anomaly Detection Based on Packet Bytes", SAC2003, Melbourne, Florida, USA, 2003.

[6] Chen-Mou Cheng, H.T.Kung, Koan-Sin Tan, "Use of Spectral Analysis in Defense Against Dos Attacks". Proceedings of IEEE GLOBECOM, 2002.

[7] Michele Basseville,Igor V.Nikiforov, Detection of Abrupt Changes: Theory and Application. Prentice-Hall.

[8] Boualem Boashash, "Estimating and Interpreting the Instantaneous Frequency of a Signal-Part 1: Fundamentals", Proceedings of the IEEE, Vol. 80, No. 4, pp 520-538, April 1992.