A Bidimensional Wavelet Transform based Algorithm for DoS Attack Detection

Renato Preigschadt de Azevedo*, Bruno Mozzaquatro*, Cristian Cappo†,
Raul Ceretta Nunes*, Christian E. Schaerer† and Alice Kozakevicius*
* *Technology Center, Federal University of Santa Maria, UFSM, Santa Maria, Brazil*
Email: {rpa.renato, brunomozza, ceretta, alicek}@inf.ufsm.br
† *Polithecnic Faculty, National University of Asuncion, UNA, San Lorenzo, Paraguay*
Email: ccappo@pol.una.py

***Abstract*—The network traffic analysis is a fundamental area in the network dependable management, since network anomalies may affect availability and quality of service (QoS). This paper proposes a fast and accurate algorithm for detecting network traffic anomalies generated by Denial of Service (DoS) attacks. Our algorithm is based on bidimensional wavelet transform (Wavelet 2D), which explores correlation between network traffic descriptors. From two different databases, the experiments show that our algorithm achieves high detection rate for DoS attack (from 95% in an own university database up to 100% in the DARPA database).**

***Keywords* - Anomaly Detector; Availability; Bidimensional Wavelet Transform; Denial-of-Service**

I. Introduction

The use of web systems is growing every day, causing the necessity of providing access to networks with high availability and quality. Network Intrusion Detection System (NIDS) is a set of tools that enables network analysis and protection against intrusion. Designing efficient NIDS is a challenging issue, because the increasing network traffic amount and the absence of a well established probability distribution.

The presented work proposes an algorithm based on the bidimensional discrete Wavelet transform [1] for traffic anomaly analysis. The algorithm is oriented to detect DoS attacks in computer networks considering the different network descriptors. With a lightweight algorithm our approach permits the anomaly detection without need of previous training, differently from algorithms based on statistical or machine learning methods, witch requires a training phase and complex analysis for performing the detection [2].

II. Algorithm for anomaly detection

In the network traffic, anomalies caused by either attacks or problems in network structure can occur. DoS attacks may cause the disruption of the service and often generate anomalies in some network descriptors.

The algorithm proposed on this work fills the matrix of the 2D direct Wavelet transform with previous chosen network descriptors. Each network descriptor generates a sliding window composed of n samples of wavelet signal variable. Thus, each network descriptor is stored as a line of the matrix. The 2D direct Wavelet transform is applied on the matrix and a hard threshold is calculated for all sub-bands of Wavelet coefficients. An alarm is generated if after the threshold operation remains non zero Wavelet coefficients.

The Figure 1 shows the algorithm proposed in this work. When new network data is available (line 1) the 2D Wavelet transform is applied on the matrix (line 2). In next step each sub-band of Wavelet coefficients is normalized (line 3). The Wavelet threshold operation is shown on lines 4 and 5. The algorithm traverses all the sub-bands of Wavelet Coefficients to verify the existence of alarms (line 6). Last instruction (line 7) returns the information about alarms.

input : New observation $x[p][n]$
output: Alarm
1 **for** *each new sample* **do**
2 Execute the 2D Wavelet Transform on $x[p][n]$
3 Normalize the Wavelet coefficients
   `/* Calculate the Standard Deviation of Wavelet coefficients */`
4 $\sigma = \sqrt{\frac{1}{K-1}\sum_{l=0}^{K-1}(d_{j,k}-\mu)^2}$
5 Calculate the Hard Threshold value: $\tau = C * \sigma$
6 Verifies the coefficients looking for alarm
7 Return $alarm$
8 **end**

Figure 1. Algorithm for anomaly detection via 2D Direct Wavelet Transform.

III. Experiments and Results

In order to validate the anomaly detection mechanism two different databases were used, a synthetic database (DARPA 1999 [3]) and one generated from the border router of the Federal University of Santa Maria (UFSM).

The DARPA database contains a set of five weeks of network traffic. In this work we analyze the second week of data because it has a number of attacks identified by type and time allowing an adequate analysis of the proposed algorithm performance. There are 8 DoS attacks on the second week of DARPA database.

The UFSM database contains all descriptors available in the border router from one second sample rate. From the UFSM database we have selected one week without attacks, were we have added 20 DoS attacks modeled from DARPA database. The 8 DoS attacks present in the DARPA database were added at different locations in the UFSM database, resulting in 20 attacks.

The experiments were performed in order to evaluate the ability of our algorithm in detecting DoS attacks. The Wavelet family selected for the execution of tests was the Haar Wavelet, since its filter size is the smallest of the entire family of orthonormal Wavelets of Daubechies. This feature turns the transformation even faster and simpler to implement, besides having the excellent ability to detect signal peaks [4]. The Haar Wavelet transform is then applied for 3 transformation levels. After the setup, the algorithm follows in round of Wavelet transformation after each update of the sliding windows.

In the experiments, the size of the sliding window was tested for 256, 128 and 64 samples per analysis. The data from DARPA database was served as benchmark to the effectiveness of our detection algorithm.

The metrics used for the evaluation of a detection algorithm are common in intrusion detection [5]: number of True Positives (TP), number of False Positives (FP), number of False Negatives (FN) and the Detection Rate (DR) where

$$DR = \frac{TP}{\text{Number of attacks}}.100$$

The Table I resumes the results obtained applying the algorithm for the DARPA database, and table II resumes the results from the UFSM database.

Table I
METRICS OBTAINED APPLYING THE ALGORITHM FOR THE DARPA DATABASE

Window	Samples	Number of DoS	TP	FP	DR
256	385379	8	8	14	100%
128	385379	8	7	8	87,5%
64	385379	8	4	5	50%

Table II
METRICS OBTAINED APPLYING THE ALGORITHM FOR THE UFSM DATABASE

Window	Samples	Number of DoS	TP	FP	DR
256	604800	20	19	13	95%
128	604800	20	18	7	90%
64	604800	20	14	5	70%

Table I shows that the detection rate of the algorithm is highly dependent of the window size, being 100% with a window size of 256, and only 50% with a 64 window size. The Table II shows that in a real world scenario the effectiveness is also dependent of window size. One attack was not detected in the UFSM database because the perturbation inserted by the attack was not too big in the total amount of traffic at that moment. Note that this results are obtained without training phase and prior knowledge of the network.

The average time required for processing the data with a sliding window of 256 samples was 2.13 ms, being feasible the use of proposed algorithm to analyze traffic in a production environment. This time was measured on a standard personal computer (intel dual-core with 4GB of RAM) with operating system Windows 7 and an implementation done in Java, version 1.6.

IV. CONCLUSIONS AND FUTURE WORKS

The network traffic analysis is a fundamental area in the network dependable management, since the Internet is very sensitive to attacks and especially DoS and distributed DoS attacks.

This paper proposed an algorithm for detect anomalies in network traffic occasioned by DoS attacks. The algorithm is based in bidimensional Wavelet transform with adaptive threshold. The experiments performed on two databases shows that the use of 2D Wavelet transform reduces the necessity of use of additional algorithm for provides correlation among different available network descriptors. The detection rate was 100% in the DARPA database, and 95% in the UFSM database. As result, the proposed algorithm has a fast method that not need a training phase and correlation the different network descriptors, obtaining a good rate of DoS attack detection without prior knowledge of the network. The performed test shows it is possible to apply the algorithm in online environments, because the time for processing each sliding window was a little more than 2 milliseconds.

As future work, other Wavelets families will be analyzed in order to further improve the effectiveness and the performance of the proposed algorithm.

REFERENCES

[1] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 11, no. 7, pp. 674–693, 1989.

[2] N. Samaan and A. Karmouch, "Network anomaly diagnosis via statistical analysis and evidential reasoning," *Network and Service Management, IEEE Transactions on*, vol. 5, no. 2, pp. 65–77, June 2008.

[3] Darpa, "Mit lincoln laboratory: Information systems technology," MIT, Tech. Rep., 1999. [Online]. Available: http://www.ll.mit.edu/mission/communications/ist/corpora/ideval/

[4] E. Stollnitz, A. DeRose, and D. Salesin, "Wavelets for computer graphics: a primer.1," *Computer Graphics and Applications, IEEE*, vol. 15, no. 3, pp. 76–84, May 1995.

[5] G. Gu, P. Fogla, D. Dagon, W. Lee, and B. Skorić, "Measuring intrusion detection capability: an information-theoretic approach," in *Proceedings of the 2006 ACM Symposium on Information, computer and communications security*. New York, NY, USA: ACM, 2006.