A Frame Work for Geometrical Structure Anomaly Detection Model

Aruna Jamdagni[1,2], Zhiyuan Tan[2], Ren P. Liu[1], Priyadarsi Nanda[2], Xiangjian He[2]

[1]CSIRO ICT Centre, [2]Faculty of Engineering & Information Technology, UTS

1. Introduction

The growth of Internet offers quality and convenience to human life, but at the same time provides a platform for hackers and criminals. The Internet security hence becomes an important issue. Intrusion Detection System (IDS) is designed to detect intrusion and also to prevent a system from being compromised. In this paper, we present a novel Geometrical Structure Anomaly Detection (GSAD) model. GSAD employs pattern recognition techniques previously used in human detection [2].

2. GSAD Architecture used in anomaly detection

In this section, we give a comprehensive introduction to the GSAD which exploits the geometrical structure in payload-based anomaly detection.

The intrusion detection is based on a statistical analysis of Mahalanobis Distances Map among characters in network traffic and distinguishes abnormal traffic from normal ones with patterns. The GSAD Architecture contains the following 5 components as shown in figure 1.

Payload feature classifier is the front-end of GSAD, which takes input from the tcpdump or network analyzer.

Payload feature analyst is the first key constituent of GSPM. It is responsible for payload feature analysis and prepares raw data for the following analysis phase.

Payload geometrical structure model is another key component of GSPM. It detects anomaly using Mahalanobis Distance Map.

Attack recognizer handles the recognition of attacks from the input network traffic.

Acknowledge/Communication is for generating alarms.

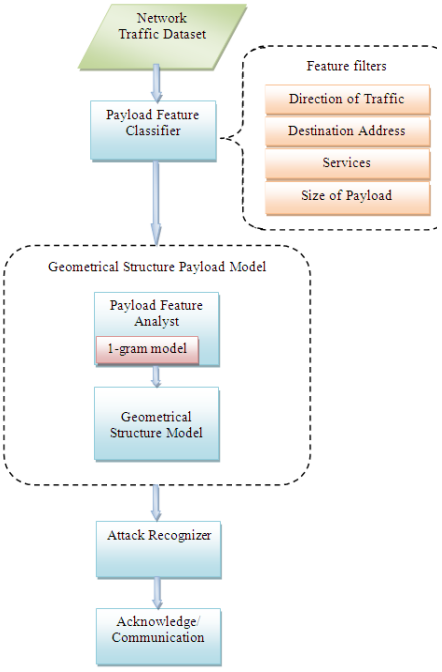

Fig 1: GSAD architecture

3. GSAD Model Characteristics

GSAD models the normal behavior of the network traffic. The most significant contribution of GSAD is the integration of geometrical structures in payload-based anomaly detection systems. Two models are fused into the GSAD: 1-Gram Payload Model [1], which is based on Language Independent Categorization of Text; and the Geometrical Structure Model (GSM), which is based on image processing technique [2]. The 1-gram payload model calculates the average frequency of each ASCII character (0-255). Mean and standard deviation of each byte's frequency are used to characterize network traffic behavior. The GSM detects similarity between the behavior of new input traffic and that of the normal traffic. The model takes into account the correlations among different features (256 ASCII characters). For each network packet, a feature vector X is defined as $X = [x_0\ x_1\ \ldots x_{255}]$. The average value of features in the 1-gram model is $\mu = \frac{1}{256}\sum_{i=0}^{255} x_i$ and the covariance value of each feature is $\sum_i = (x_i - \mu)(x_i - \mu)'\ (0 \le i \le 255)$. The Mahalanobis distance between two indicated characters $d_{(i,j)}$, and Distance map D are:

$$d_{(i,j)} = \frac{(x_i - x_j)(x_i - x_j)'}{\Sigma_i + \Sigma_j} \quad (0 \le i, j \le 255)$$

$$D = \begin{bmatrix} d_{(0,0)} & d_{(0,1)} & \cdots & d_{(0,255)} \\ d_{(1,0)} & d_{(1,1)} & \cdots & \vdots \\ \vdots & \vdots & \ddots & \\ d_{(255,0)} & d_{(255,1)} & \cdots & d_{(255,255)} \end{bmatrix}$$

Then, the weight w is calculated using expression: $w = \sum_{i,j=0}^{255,255} \frac{(d_{obj(i,j)} - \bar{d}_{nor(i,j)})^2}{\sigma^2_{nor(i,j)}}$. If the weight is larger than a threshold, the input packet is considered as an intrusion.

4. Experimental Results and Analysis

We tested GSAD model on the 1999 DARPA IDS data set which is considered as standard data set for the evaluation of intrusion detection systems. Experiments were conducted to identify crashiis attack using 150 bytes of packet payload length. In this experiment, we used the inside network traffic. We trained the GSAD model on the week1 and week 3 (attack free) data set, and evaluate the model on week 2 data set, which contains 43 instances of 15 different attacks.

a. MDM for attack free packets b. MDM for attack packets

Fig 2: Crashiis attack: geometrical structure models

The x and y axis in the Fig 2 represents the features used in the calculation of MDM. The result clearly indicates that the geometrical structure model of attack packets in Fig 2.b is very different from that of normal packets as shown in Fig 2.a. These provide strong evidences for distinguishing attacks from normal packets.

5. Conclusions

In this paper we present the frame work for GSAD model, based on geometrical structure of packet payload. This is used for network intrusion detection. The key features are to compute byte distribution and geometrical structure for normal traffic, conditioned to service type and payload length. The experiments for crahiis attack show some promising results. In our future work we aim to evaluate the performance of our model and validate our results. We also plan to test this model on 1999 DARPA IDS dataset for variable length payload, protocols and services.

6. References

[1] Ke Wang and S. Stolfo, “Anomalous payload-based network intrusion detection”, *In Recent Advances in Intrusion Detection, RAID*, pages 203–222, September 2004.

[2] Akira Utsumi and Nobuji Tetsutani, "Human Detection using Geometrical Pixel Value Structures", *In Proceeding of 5th International Conference on Automatic Face and Gesture Recognition (FGR 2), pp*. 34-39, 2002.

Aruna Jamdagni Zhiyuan Tan Ren P. Liu Priyadarsi Nanda Xiangjian He