International Journal of Computer Networks and Communications Security

VOL. 2, NO. 9, SEPTEMBER 2014, 279–284
Available online at: www.ijcncs.org
ISSN 2308-9830

A Dynamic Flooding Attack Detection System Based on Different Classification Techniques and Using SNMP MIB Data

SAHAR NAMVARASL[1], MARZIEH AHMADZADEH[2]

[1, 2] Shiraz University of Technology, Department of Computer Engineering & IT, Shiraz, Iran

E-mail: [1]sahar.namvarasl@gmail.com, [2]ahmadzadeh@sutech.ac.ir

ABSTRACT

Currently, the amount of exchanged data in network has increased dramatically and consequently, detection of malicious data is an important issue for network's users and administrators. DoS and DDoS attacks have always taken consideration of attackers and researchers, and distinguishing them from normal packet is difficult. Therefore, using data mining techniques along traditional mechanism such as firewall, improves the performance of intrusion detection systems. This paper introduces flooding attack detection system based on SNMP MIB data, which selects effective MIB variables and compares some different classification algorithms based on chosen variables. Finally, the proposed system, models detection mechanism, is using the algorithm with the highest accuracy. The advantage of this system is its ability to learn. System's detection model will be optimized after receiving the new data. While the behavior of attack changes, the system will be adapted easily.

Keywords: *Dos attack, SNMP, MIB, Intrusion Detection System, Data Mining.*

1 INTRODUCTION

Recent improvements in technologies such as wireless network caused significant growing number of users and huge amount of transmitted data on this media which brings many challenges especially in the scope of security. One of the most important aspects of security is rapid detection of attack in order to preventing more damage. Denial of Service (DoS) and Distributed Denial of Service (DDoS) are usually most attractive attack to attacker. Due to the features of DoS/DDoS attack, it's not easy to find differences between the attacked network packets and normal network packets. DoS attack flooded great number of packets to special victim IP and DDoS attack flooded packets by distributed attacker system. Attack is deployed by different data packet types, which generally TCP-SYN, UDP and ICMP flooding are the most commonly used ones. [1]

Simple Network Management Protocol (SNMP) [2] is one of the useful protocols in the scope of monitoring and controlling network devices. Currently, this protocol is developed in 4 different versions: SNMPv1, SNMPv2, SNMPv2c, and SNMPv3, which the first version was appeared in 1988. Due to simplicity and efficient use of resources, many administrators prefer SNMP compared with other network managing tools. Management Information Base (MIB) is a virtual database that stores the information gathered by SNMP. To gather SNMP information, manager sends requests to the agent, and the agent extracts required data from MIB and returns them to the manager. The structure of SNMP and its operations is shown in fig 1.

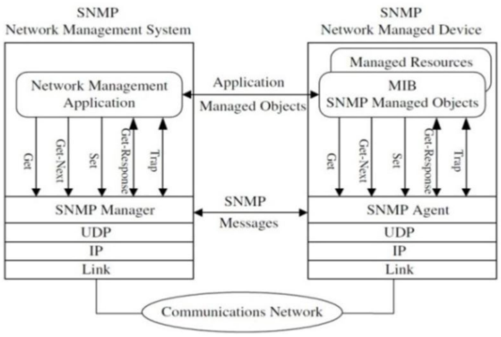

Fig. 1. The structure of SNMP.

The methods was previously used to detect DoS and DDoS attacks based on packet analysis, but because of huge number of transmitted packets over network , this is not suitable solution in practical. As a result, proposed mechanism in this paper is benefited with the features of SNMP MIB-II for analysis network behavior and find schema to separate normal and destructive packets.

Data mining [3] techniques provides substantial contributions to model network operations behavior. In this paper, appropriate model is constructed based on effective MIB variables, and the system which is able to optimize the model over time is introduced. Effective variables are obtained from comparing different classification algorithms, and detection model is made using classification algorithm.

The rest of paper is organized as follows: In chapter 2, different related works in this scope will be reviewed. Chapter 3 describes detection attack system completely. In chapter 4, the detail of implementation and analysis of system will be expressed and at the end, in chapter 5, the paper's conclusion will be discussed.

2 RELATED WORK

In the field of DoS and DDoS attack, many researches have been done up to now. Many of them consisted of traditional approaches such as firewall and etc, which was not met all needs of a robust detection system. Thereupon, researchers attended to the artificial intelligence and data mining techniques. Iftikhar et al. at [4] introduced feature selection mechanism based on Principal Component Analysis (PCA). However, since this method might ignore some sensitive features, a method was proposed based on Genetic Algorithm and multilayer perceptron (MLP) - The neural network algorithm for mapping input to appropriate output. KDD-cup was used for dataset. As a result, they selected 12 features among 44 features and claimed that accuracy has improved to 0.99. As mentioned at [5, 6], PCA is not suitable for large dataset and this method is executable just for small dataset.

In [7] Singh and Silakari stated that PCA is not proper solution for non-linear dataset, therefore they presented an algorithm based on Generalized Discriminant Analysis (GDA), to generate small size of features and improve classification operation. They asserted that this method is premier than other classification method such as Self-Organizing Map (SOM) and C4.5. KDDCup99 was used for dataset in that research, also 4 different attacks were reviewed: DoS attack, User to Root Attacks, Remote to Local (User) Attacks, and probing. Finally their method accuracy was about 0.98.

Most of the researches in the scope of intrusion detection attack, offer the model by analyzing raw packet data, and processing vast amount of data especially while occurring DoS/DDoS attack is the main challenge for researchers. For this reason, the idea of attack detection based on statistical data gained from network management protocol was raised. MAID [8] was an intrusion detection system that monitored 27 different SNMP MIB variables and compared the behavior of normal and attack packet. Normal behavior of packet was modeled using probability density function (PDF), and was kept as reference PDF. They compared five similarity metrics by examining algorithm on actual network data and attack. They stated that KST is able to detect more attacks in all situations even at low traffic intensities.

D.Dutta and K.Choudhury at [9] claimed that their research was the first intrusion detection system which was integrated Digital Signature of Network Segment (DSNS) with Particle Swarm Optimization (PSO). They also benefited SVM to optimize clustering operation and better centroids selection. PSO [10] is a Swarm Intelligence algorithm, which despite the high Efficiency has low computational complexity.

At [11], J.Yu et al. Also presented a model based on SNMP and SVM. Unlike previous model that had just introduced a detection model, they proposed a two layer architecture. The first layer detected DoS/DDoS attack and the second layer detected these types of attack: TCP-SYNC, UDP and ICMP. Attack type identifying has the advantage of filtering the corresponding packet. Extended architecture of this model was proposed at [12] . Classification and association rule mining that performed by C4.5 algorithm was operated offline, while getting SNMP MIB variable and detection DoS/DDoS attack was done online. After getting Dataset and generated new packet data, Offline modules extracted model and valuable rule and passed the result to detection module. Function of Getting MIB module was to schedule operation of SNMP pooling. Authors asserted that accuracy value obtained for detection attack was about 99.13%.

3 PROPOSED SYSTEM

Proposed system in this paper is made of three main modules. The function of first module is selecting appropriate MIB variables based on algorithm which will be explained in the rest, and afterward extracting the most accurate model based on these variables (module2). The role of third module is detecting DoS or DDoS attack on real time, which operates using MIB data gathered from network.

Total number of MIB-II variables considered in this system, is 66 that classified in 4 categories: IP, TCP, ICMP, and UDP. Among these, the variables more effective to detect attack, should be selected. Therefore, classification is performed using different algorithms. In the rest, more details of these three modules will be explained.

This suggested system improves its performance, using 3 different algorithms for selecting effective variables and detecting attack model, instead of one. Therefore, during each iteration of system, algorithm with highest performance will be selected.

Forasmuch as the behavior of attack changes continuously (for example changing attack type), the accuracy of model will be reduced. So, the system repeats operations at special interval time. Also selected variable and model would be updated, if needed.

3.1 Module 1

Due to large number of MIB variables, using all of them for classification is not practical and wastes lots of resources. So, using a mechanism which can select effective variable without reducing system performance. Therefore in this paper, it is suggested to use 3 different classification algorithms. Each algorithm forms a set of variables, which a subset or whole of them could be chosen. Variable of algorithm with highest accuracy and lowest cost will be considered as effective variable.

So far, several algorithms have been presented for classification, some of them such as decision tree algorithms and rule based algorithms eliminate variables that have no effect on the result of prediction. C4.5 and AttributeSelectionClassifier (attribute selection and classification algorithm), are the decision tree and RIPPER is rule based algorithm that was considered for this module. The studies done on different dataset showed that these 3 algorithms have better performances in different situations.

To evaluate cost of each algorithm cost matrix was used. The value of this matrix depends on network situation and can be filled by network administrator.

3.2 Module 2

The task of this module is constructing intrusion detection model with effective variables. There is some suggested models based on BP, Bayesian and C4.5 [12, 13]. Neural network, Bayesian network and C4.5 are 3 algorithms which have been selected for this module. By comparing the results of different classification algorithms, it is proved that these are the algorithms with high performance. Appropriate model is constructed using the most accurate model with lowest cost.

3.3 Module 3

So far a model for analyzing SNMP MIB-II data and detecting attacks has been achieved. As regards, since the structure and the behavior of DoS and DDoS attack is continuously changing, in this paper learning mechanism with novel dataset is used. One third of novel dataset is chosen from initial dataset and the remaining is acquired from new DoS/DDoS attacks that system detects and new data packets. Once enough dataset records are gathered, module 1 and module 2 operations have been carried out and effective variables and model is updated, if required. As a result, the proposed system behavior changes during the variation of attack behavior. Explained structure is shown in fig 2.

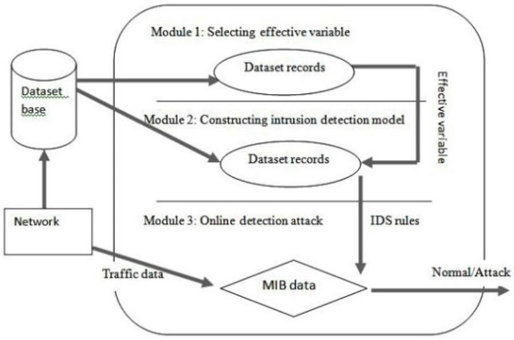
Module 1: Selecting effective variable
Dataset records
Dataset base
Effective variable
Module 2: Constructing intrusion detection model
Dataset records
Network
Module 3: Online detection attack
IDS rules
Traffic data
MIB data
Normal/Attack

Fig. 2. The proposed system architecture

With this mechanism, the system continuously improves its behavior and even though the dataset or initial model is inappropriate, system performance will be optimal over time.

The size of Dataset in this part is obtained using trial and error, which could vary depends on resources conditions. Small dataset make module 1 and 2 repeat continuously, results in wasting the resources and reducing the accuracy of model, and large dataset causes model to be updated late. Here, the size is considered as the initial dataset size.

Another point of designing this system is interval rate for pooling SNMP data. Long time interval causes belated detection attack, whereas too short interval occupies network resources. In this paper, according to the experimental result and analysis in [11], optimal interval rate is considered 15 second.

4 EXPERIMENTAL AND ANALYSIS

For analyzing the system a LAN network was considered with 4 PCs (CPU: Core i32.93 GHZ; memory 4G; Hard disk 500G) and one switch. One pc had been considered for pooling SNMP MIB data and implementing modules. The OS of every 4 systems was Linux ubuntu 14.04. This test is shown in Fig 3. The Dataset was formed using real packet streaming over LAN, during 10 days. In order to compensate small size of LAN, hping3 was used [14]. Hping3 generates different packet type (TCP, UDP, and ICMP) with random packet size.

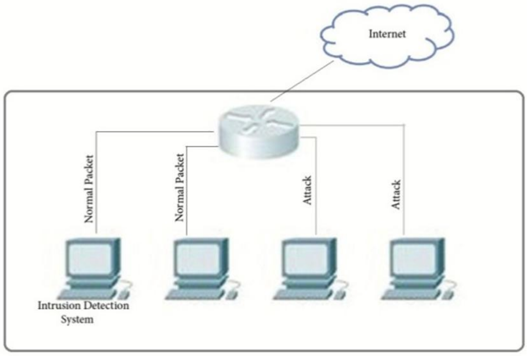

Fig. 3. Testbed network

Hping3 is a command line packet generator, scriptable security tools, which is compatible with Linux. The most important feature of hping3 is its ability to send packets with different options just by one line command. Therefore, users able to change any feature of packet and mange how to generate and transmit packets, in addition to its being easy to learn.

There are some tools for simulating DoS/DDoS attack, such as hping3, Stacheldraht, TFN2K and etc. here, hping3 was used to simulate the attack due to its capabilities. Two systems were responsible for generating attacks in specific periods of time with 3 different attack types: TCP-Sync, UDP and ICMP.

As a result, a dataset with about 4600 records consisting of normal and attack records was achieved and analyzed. Weka (Machine Learning Lab introduced by The University of Waikato) [15] is the tool, used to accomplish classification and clustering. Weka is an open source software based on java, compatible with Linux, which implements collection of machine learning algorithms, and supports large size data. In proposed system, cost matrix is considered as shown in table 1.

Table 1. The considered Cost Matrix of the system

	PREDICTED CLASS		
ACTUAL CLASS		Class=Yes (attack)	Class=No
	Class=Yes (attack)	-2	5
	Class=No	2	1

The cost of each algorithm is defined as:

$$\text{Cost}= C(YES|YES)(TP)+C(NO|YES)(FN) + C(YES|NO)(FP)+C(NO|NO)(TN) \qquad (1)$$

In the above equation C(i|j) is the cost of classifying class j which is classified as class i. TP is the total number of attack traffic MIB records which is classified as attack and TN is the total number of normal records which is classified as normal. Also FN indicates the MIB records of attack traffic which misclassified as normal and FP is the total number of MIB records which misclassified as normal.

For the performance evaluation is used accuracy rate according to the formula 2.

$$\text{Accuracy rate}=\frac{\sum_{i=1}^{n} Ti}{N} * 100 \qquad (2)$$

Where T_i is an individual MIB record which is classified correctly and N indicates the total number of MIB records.

The result of first module operation is shown in table 2. Classification corresponding to C4.5 algorithm had highest accuracy rate, and the variables were considered as effective variables.

Table 2. The result of Module 1

Classificatio n Algorithm	SNMP MIB-II variables	Accuray rate (%)	Cost
C4.5	ipInReceives, ipInDelivers, ipOutRequests, icmpInMsgs, icmpOutMsgs, tcpOutRsts	98.72	353
RIPPER	ipInReceives, ipInDelivers, ipOutRequests, icmpInMsgs, icmpOutDestUnreachs , tcpInSegs	95.92	781
AttributeS-election	ipInReceives, ipForwDatagrams,ipIn Delivers, ipOutRequests, icmpInMsgs, tcpInSegs	97.97	627

The performance of three classification algorithms (Neural network, Bayesian network and C4.5) performance which used to implement module 2 is shown in table 3. The most accurate algorithm (Neural network) was selected for third module.

Table 3. The result of module 2

Classification Algorithm	Accuracy rate (%)	Cost
Neural network	99.03	310
Bayesian network	98.83	317
C4.5	98.72	353

To evaluate the system performance better in dealing with new attack behavior, novel DoS attack which involved just UDP flooding was performed. The accuracy rate before and after updating model is summarized in table 4.

Table 4. Comparing the performance of model after and before updating

	Classification Algorithm	Accuracy rate (%)
Before updating	Neural network	99.027
After updating	Neural network	99.035

Since running time and resource usage of the system depends on the size of dataset that used to update variable set, an acceptable system performance can be achieved by choosing an optimal size.

5 CONCLUSION

This paper proposed intrusion detection system, based on SNMP MIB data. The purpose of this research is to introduce the DoS/DDoS attack detection which able to improve the performance after receiving novel attack. It is also notable that while the behavior of attack changes, the model will be updated. The system performed in three steps: (1) Selecting effective variable. (2) Generating the most accurate model. (3) Detecting real time attack and updating dataset. Finally system was tested using actual network data and the accuracy rate of 99.03% was calculated. After receiving enough number of novel DoS/DDoS attack, the model repeated module 1 and 2 operation to optimize the detection system. To implement module 1, three classification algorithms was used: C4.5, RIPPER and attribute Selection Classification. Effective variables had been generated using the most accurate algorithm. The most accurate classification algorithm of second module formed detection model. Classification algorithm of module 2 consisted of: Neural network, Bayesian network and C4.5 which was implemented by Weka. As a result, this system is not limited to a particular algorithm and is able to select best model among exiting algorithm. This process will be done over time, with new data receiving and system continuously improves its performance. When the behavior of attack changes, model will be update and prevent more damage in future attacks. System overhead is acceptable and according to limitation of resource, with reducing second dataset size, will be less.

6 REFERENCES

[1] J. Mirkovic and P. Reiher, "A taxonomy of DDoS attack and DDoS defense mechanisms," ACM SIGCOMM Computer Communication Review, vol. 34, pp. 39-53, 2004.

[2] J. Ding, Advances in network management: CRC press, 2010.

[3] T. Pang-Ning, M. Steinbach, and V. Kumar, "Introduction to data mining," in Library of Congress, 2006.

[4] A. Iftikhar, A. Azween, A. Abdullah, A. Khalid, and H. Muhammad, "Intrusion detection using feature subset selection based on MLP," Scientific Research and Essays, vol. 6, pp. 6804-6810, 2011.

[5] H. M. Imran, A. Abdullah, M. Hussain, S. Palaniappan, and I. Ahmad, "Intrusion Detection based on Optimum Features Subset and Efficient Dataset Selection," International Journal of Engineering and Innovative Technology (IJEIT), vol. 2, pp. 265-270, 2012.

[6] K. Delac, M. Grgic, and S. Grgic, "Independent comparative study of PCA, ICA, and LDA on the FERET data set," International Journal of Imaging Systems and Technology, vol. 15, pp. 252-260, 2005.

[7] S. Singh, S. Silakari, and R. Patel, "An efficient feature reduction technique for intrusion detection system," in Machine Learning and Computing, 2009. International Conference on, 2011.

[8] J. Li and C. Manikopoulos, "Early statistical anomaly intrusion detection of DOS attacks using MIB traffic parameters," in Information

Assurance Workshop, 2003. IEEE Systems, Man and Cybernetics Society, 2003, pp. 53-59.

[9] D. Dutta and K. Choudhury, "Network Anomaly Detection using PSO-ANN," International Journal of Computer Applications, vol. 77, 2013.

[10] J. Kennedy, "Particle swarm optimization," in Encyclopedia of Machine Learning, ed: Springer, 2010, pp. 760-766.

[11] J. Yu, H. Lee, M.-S. Kim, and D. Park, "Traffic flooding attack detection with SNMP MIB using SVM," Computer Communications, vol. 31, pp. 4212-4219, 2008.

[12] J. Yu, H. Kang, D. Park, H.-C. Bang, and D. W. Kang, "An in-depth analysis on traffic flooding attacks detection and system using data mining techniques," Journal of Systems Architecture, vol. 59, pp. 1005-1012, 2013.

[13] J. B. Cabrera, L. Lewis, X. Qin, W. Lee, and R. K. Mehra, "Proactive intrusion detection and distributed denial of service attacks—a case study in security management," Journal of Network and Systems Management, vol. 10, pp. 225-254, 2002.

[14] http://www.hping.org/hping3.html, July 2014

[15] http://www.cs.waikato.ac.nz/ml, July 2014