ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

A System for Denial-of-Service Attack Detection Based on Multivariate Correlation Analysis

Arjun H[1], S G Maknur[2].

M.Tech Student, Dept. of Computer Science and Engineering, STJIT, Ranebennur, India.

Head of Department, Dept. of Information of Science and Engineering, STJIT, Ranebennur, India

ABSTACT: Interconnected systems, such as Web servers, database servers, cloud computing servers etc., are now under threads from network attackers. As one of most common and aggressive means, Denial-of Service (DoS) attacks cause serious impact on thesecomputing systems. In this paper, we present a DoS attack detection system that uses Multivariate Correlation Analysis (MCA) for accurate network traffic characterization by extracting the geometrical correlations between network traffic features. Our MCA-based DoS attack detection system employs the principle of anomaly-based detection in attack recognition. This makes our solution capableof detecting known and unknown DoS attacks effectively by learning the patterns of legitimate network traffic only. Furthermore, a triangle-area-based technique is proposed to enhance and to speed up the process of MCA. The effectiveness of our proposed detection system is evaluated using KDD Cup 99 dataset, and the influences of both non normalized data and normalized data on the performance of the proposed detection system are examined. The results show that our system outperforms two other previously developed state-of-the-art approaches in terms of detection accuracy.

KEYWORDS*:* Denial of Services, Detecting Attack System, Multivariate Correlation Analysis, triangle Area Map Generations.

I. INTRUDUCTION

Denial of service (DoS) attacks has become a major threat to current computer networks. Early DoS attacks were technical games played among underground attackers. For example, an attacker might want to get control of an IRC channel via performing DoS attacks against the channel owner. Attackers could get recognition in the underground community via taking down popular web sites. Because easy-to-use DoS tools, such as Trinoo, can be easily downloaded from the Internet, normal computer users can become DoS attackers as well. They sometime coordinately expressed their views via launching DoS attacks against organizations whose policies they disagreed with. DoS attacks also appeared in illegal actions. Companies might use DoS attacks to knock off their competitors in the market. Extortion via DoS attacks were on rise in the past year. Attackers threatened online businesses with DoS attacks and requested payments for protection. Known DoS attacks in the Internet generally conquer the target by exhausting its resources that can be anything related to network computing and service performance, such as link bandwidth, TCP connection buffers, application/service buffer, CPU cycles, etc. Individual attackers can also exploit vulnerability, break into target servers, and then bring down services. Because it is difficult for attackers to overload the target's resource from a single computer, many recent DoS attacks were launched via a large number of distributed attacking hosts in the Internet. These attacks are called **Distributed Denial of Service (DDoS)** attacks. In a DDoS attack, because the aggregation of the attacking traffic can be tremendous compared to the victim's resource, the attack can force the victim to significantly downgrade its service performance or even stop delivering any service, Compared with conventional DoS attacks that could be addressed by better securing service systems or prohibiting

unauthorized remote or local access, DDoS attacks are more complex and harder to prevent. Since many unwitting hosts are involved in DDoS attacks, it is challenging to distinguish the attacking hosts and take reaction against them.

In recent years, DDoS attacks have increased in frequency, sophistication and severity due to the fact that computer vulnerabilities are increasing fast, which enable attackers to break into and install various attacking tools in many computers. Wireless networks also suffer from DoS attacks because mobile nodes (such as laptops, cell phones, etc.) share the same physical media for transmitting and receiving signals; and mobile computing resources (such as bandwidth, CPU and power) are usually more constrained than those available to wired nodes. In a wireless network, a single attacker can easily forge, modify or inject packets to disrupt connections between legitimate mobile nodes and cause DoS effects.

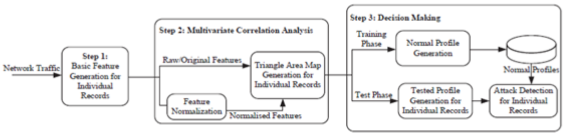
Network Traffic
Step 1:
Basic Feature Generation for Individual Records
Step 2: Multivariate Correlation Analysis
Raw/Original Features
Triangle Area Map Generation for Individual Records
Feature Normalization
Normalised Features
Step 3: Decision Making
Training Phase
Normal Profile Generation
Normal Profiles
Test Phase
Tested Profile Generation for Individual Records
Attack Detection for Individual Records

Figure 1: Framework of the proposed denial-of service attack detection system.

II. RELATED WORK

1) Bro: A System for Detecting Network Intruders in Real time.

AUTHORS**:** V. Paxson.
We describe Bro, a stand-alone system for detecting network intruders in real-time by passively monitoring a network link over which the intruder's traffic transits. We give an overview of the system's design, which emphasizes high-speed (FDDI-rate) monitoring, real-time notification, clear separation between mechanism and policy, and extensibility. To achieve these ends, Bro is divided into an “event engine” that reduces a kernel filtered network traffic stream into a series of higher level events, and a “policy script interpreter” that interprets event handlers written in a specialized language used to express a site's security policy. Event handlers can update state information, synthesize new events, record information to disk, and generate real-time notifications via *syslog*. We also discuss a number of attacks that attempt to subvert passive monitoring systems and defenses against these, and give particulars of how Bro analyzes the four applications integrated into it so far: Finger, FTP, Port mapper and Telnet.

2) Anomaly-based Network Intrusion Detection: Techniques, Systems and Challenges.

AUTHORS: P. Garca-Teodoro, J. Daz-Verdejo, G. Maci-Fernndez, and E. Vzquez.

The Internet and computer networks are exposed to an increasing number of security threats. With new types of attacks appearing continually, developing flexible and adaptive security oriented approaches is a severe challenge. In this context, anomaly-based network intrusion detection techniques are a valuable technology to protect target systems and networks against malicious activities. However, despite the variety of such methods described in the literature in recent years, security tools incorporating anomaly detection functionalities are just starting to appear, and several important problems remain to be solved. This paper begins with a review of the most well-known anomaly-based intrusion detection techniques. Then, available platforms, systems under development and research projects in the area are presented. Finally, we outline the main challenges to be dealt with for the wide scale deployment of anomaly-based intrusion detectors, with special emphasis on assessment issues.

3) DDoS attack detection method using cluster analysis.

AUTHORS: K. Lee, J. Kim, K. H. Kwon, Y. Han, and S. Kim.

Distributed Denial of Service (DDoS) attacks generate enormous packets by a large number of agents and can easily exhaust the computing and communication resources of a victim within a short period of time. In this paper, we propose a method for proactive detection of DDoS attack by exploiting its architecture which consists of the selection of handlers and agents, the communication and compromise, and attack. We look into the procedures of DDoS attack and then select variables based on these features. After that, we perform cluster analysis for proactive detection of the attack. We experiment with 2000 DARPA Intrusion Detection Scenario Specific Data Set in order to evaluate our method. The results show that each phase of the attack scenario is partitioned well and we can detect precursors of DDoS attack as well as the attack itself.

III. EXISTING SYSTEM

Generally, network-based detection systems can be classified into two main categories, namely misuse-based detection systems and anomaly-based detection systems. Misuse-based detection systems detect attacks by monitoring network activities and looking for matches with the existing attack signatures. In spite of having high detection rates to known attacks and low false positive rates, misuse-based detection systems are easily evaded by any new attacks and even variants of the existing attacks. Furthermore, it is a complicated and labor intensive task to keep signature database updated because signature generation is a manual process and heavily involves network security expertise.

Disadvantage:

- Most existing IDS are optimized to detect attacks with high accuracy. However, they still have various disadvantages that have been outlined in a number of publications and a lot of work has been done to analyze IDS in order to direct future research.
- Besides others, one drawback is the large amount of alerts produced.

IV. PROPOSED SYSTEM

In this paper, we present a DoS attack detection system that uses Multivariate Correlation Analysis (MCA) for accurate network traffic characterization by extracting the geometrical correlations between network traffic features. Our MCA-based DoS attack detection system employs the principle of anomaly-based detection in attack recognition.

The DoS attack detection system presented in this paper employs the principles of MCA and anomaly-based detection. They equip our detection system with capabilities of accurate characterization for traffic behaviors and detection of known and unknown attacks respectively. A triangle area technique is developed to enhance and to speed up the process of MCA. A statistical normalization technique is used to eliminate the bias from the raw data.

Advantages:

- The MCA-based DoS attack detection system which is the former technique extracts the geometrical correlations hidden in individual pairs of two distinct features within each network traffic record.
- And it offers more accurate characterization for network traffic behaviors.

V. DESIGN OF THE SYSTEM

INPUT DESIGN

The input design is the link between the information system and the user. It comprises the developing specification and procedures for data preparation and those steps are necessary to put transaction data in to a usable form for processing can be achieved by inspecting the computer to read data from a written or printed document or it can occur by having people keying the data directly into the system. The design of input focuses on controlling the amount of input required, controlling the errors, avoiding delay, avoiding extra steps and keeping the process simple. The input is designed in such a way so that it provides security and ease of use with retaining the privacy. Input Design considered the following things:

- ➢ What data should be given as input?
- ➢ How the data should be arranged or coded?
- ➢ The dialog to guide the operating personnel in providing input.
- ➢ Methods for preparing input validations and steps to follow when error occur.

OBJECTIVES

1. Input Design is the process of converting a user-oriented description of the input into a computer-based system. This design is important to avoid errors in the data input process and show the correct direction to the management for getting correct information from the computerized system.
2. It is achieved by creating user-friendly screens for the data entry to handle large volume of data. The goal of designing input is to make data entry easier and to be free from errors. The data entry screen is designed in such a way that all the data manipulates can be performed. It also provides record viewing facilities.
3. When the data is entered it will check for its validity. Data can be entered with the help of screens. Appropriate messages are provided as when needed so that the user will not be in maize of instant. Thus the objective of input design is to create an input layout that is easy to follow.

OUTPUT DESIGN

A quality output is one, which meets the requirements of the end user and presents the information clearly. In any system results of processing are communicated to the users and to other system through outputs. In output design it is determined how the information is to be displaced for immediate need and also the hard copy output. It is the most important and direct source information to the user. Efficient and intelligent output design improves the system's relationship to help user decision-making.

1. Designing computer output should proceed in an organized, well thought out manner; the right output must be developed while ensuring that each output element is designed so that people will find the system can use easily and effectively. When analysis design computer output, they should Identify the specific output that is needed to meet the requirements.
2. Select methods for presenting information.
3. Create document, report, or other formats that contain information produced by the system.

The output form of an information system should accomplish one or more of the following objectives.

- Convey information about past activities, current status or projections of the Future.
- Signal important events, opportunities, problems, or warnings.
- Trigger an action.
- Confirm an action.

VI. MODULE DESCRIPTION

Number of Modules:

After careful analysis the system has been identified to have the following modules:

1. Feature Normalization.
2. Multivariate Correlation Analysis.
3. Decision Making Module.
4. Evaluation of Attack detection.

MODULES DESCRIPTION:

1. Feature Normalization Module:

In this module, basic features are generated from ingress network traffic to the internal network where protected servers reside in and are used to form traffic records for a well-defined time interval. Monitoring and analyzing at the destination network reduce the overhead of detecting malicious activities by concentrating only on relevant inbound traffic. This also enables our detector to provide protection which is the best fit for the targeted internal network because legitimate traffic profiles used by the detectors are developed for a smaller number of network services.

2. Multivariate Correlation Analysis:

In this Multivariate Correlation Analysis, in which the "Triangle Area Map Generation" module is applied to extract the correlations between two distinct features within each traffic record coming from the first step or the traffic record normalized by the "Feature Normalization" module in this step. The occurrence of network intrusions cause changes to these correlations so that the changes can be used as indicators to identify the intrusive activities. All the extracted correlations, namely triangle areas stored in Triangle Area Maps (TAMs), are then used to replace the original basic features or the normalized features to represent the traffic records.

3. **<u>Decision Making Module:</u>**

In this module, the anomaly-based detection mechanism is adopted in Decision Making. It facilitates the detection of any DoS attacks without requiring any attack relevant knowledge. Furthermore, the labor-intensive attack analysis and the frequent update of the attack signature database in the case of misuse-based detection are avoided. Meanwhile, the mechanism enhances the robustness of the proposed detectors and makes them harder to be evaded because attackers need to generate attacks that match the normal traffic profiles built by a specific detection algorithm. This, however, is a labor-intensive task and requires expertise in the targeted detection algorithm. Specifically, two phases (i.e., the “Training Phase” and the “Test Phase”) are involved in Decision Making. The “Normal Profile Generation” module is operated in the “Training Phase” to generate profiles for various types of legitimate traffic records, and the generated normal profiles are stored in a database. The “Tested Profile Generation” module is used in the “Test Phase” to build profiles for individual observed traffic records. Then, the tested profiles are handed over to the “Attack Detection” module, which compares the individual tested profiles with the respective stored normal profiles. A threshold-based classifier is employed in the “Attack Detection” module to distinguish DoS attacks from legitimate traffic.

4. **<u>Evaluation of Attack detection:</u>**

During the evaluation, the 10 percent labeled data of KDD Cup 99 dataset is used, where three types of legitimate traffic (TCP, UDP and ICMP traffic) and six different types of DoS attacks (Teardrop, Smurf, Pod, Neptune, Land and Back attacks) are available. All of these records are first filtered and then are further grouped into clusters according to their labels.

VII. SIMULATION RESULT

The relationship between DR and FPR is clearly revealed in the ROC (Relation of Correlation) curves. The DR increases when larger numbers of false positive are tolerated. In Fig. 2a, the ROC curve for analyzing the original data using our proposed detection system shows a rising trend. The curve climbs gradually from 86.98% DR to 89.44% DR, and finally reaches to 95.11% DR. Likewise, the ROC curve for analyzing the normalized data presents a resembling pattern but jumps dramatically from 99.97% DR to 99.99% DR after experiencing slow progress as shown in Fig. 2b. Then, the curve remains in a high level of DR around 100.00%. It is shown clearly in Fig. 2 that our detection system always enjoys higher detection rates while working with the normalized data than with the original data. The worst performance (99.96% DR and 1.25% FPR) of our system shown in Fig. 2b is even much better the best performance (95.11% DR and 1.26% FPR) in term of detection rate shown in Fig. 2a.

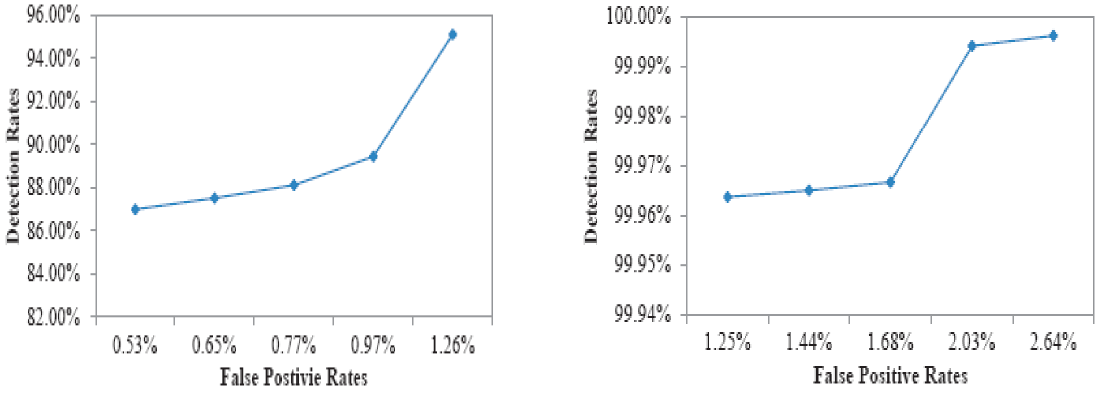

Fig a. ROC curves for analyzing original data . Fig b.ROC curve for analyzing normalized data .

Fig 2.ROC curves for the detection of DoS attacks

VIII. CONCLUSION AND FUTURE WORK

This paper has presented a MCA-based DoS attack detection system which is powered by the triangle-area based MCA technique and the anomaly-based detection technique. The former technique extracts the geometrical correlations hidden in individual pairs of two distinct features within each network traffic record, and offers more accurate characterization for network traffic behaviors. The latter technique facilitates our system to be able to distinguish both known and unknown DoS attacks from legitimate network traffic. Evaluation has been conducted using KDD Cup 99 dataset to verify the effectiveness and performance of the proposed DoS attack detection system. The influence of original (non-normalized) and normalized data has been studied in the paper. The results have revealed that when working with non-normalized data, our detection system achieves maximum 95.20% detection accuracy although it does not work well in identifying Land, Neptune and Teardrop attack records. The problem, however, can be solved by utilizing statistical normalization technique to eliminate the bias from the data. The results of evaluating with the normalized data have shown a more encouraging detection accuracy of 99.95% and nearly 100.00% DRs for the various DoS attacks. Besides, the comparison result has proven that our detection system outperforms two state-of-the-art approaches in terms of detection accuracy. Moreover, the computational complexity and the time cost of the proposed detection system have been analyzed.

REFERENCES

[1] V. Paxson, “Bro: A System for Detecting Network Intruders in Realtime,” *Computer Networks*, vol. 31, pp. 2435-2463, 1999

[2] P. Garca-Teodoro, J. Daz-Verdejo, G. Maci-Fernndez, and E. Vzquez, “Anomaly-based Network Intrusion Detection: Techniques, Systems and Challenges,” *Computers & Security*, vol. 28, pp. 18-28, 2009.

[3] A. Tajbakhsh, M. Rahmati, and A. Mirzaei, “Intrusion detection using fuzzy association rules,” Applied Soft Computing, vol. 9, no. 2, pp. 462-469, 2009.

[4] J. Yu, H. Lee, M.-S.Kim, and D. Park, “Traffic flooding attack detection with SNMP MIB using SVM,” Computer Communications, vol. 31, no. 17, pp. 4212-4219, 2008.

[5] A. Jamdagni, Z. Tan, X. He, P. Nanda, and R. P. Liu, “RePIDS: A multi tier Real-time Payload-based Intrusion Detection System,” Computer Networks, vol. 57, pp. 811-824, 2013.

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my Institution Sri Taralabalu Jagadguru Institute of Technology, Guide and Staff members for their continuous support for this survey and finally my friends for their coordination in this work.

BIOGRAPHY

ARJUN H received Bachelor degree B.E from ProudaDevaraya Institute of Technology, Hospete, Vishweshwaraya Technological University, Belgaum, Karnataka, India . He is now pursuing Masters Degree M.Tech in Computer science and engineering department at Sri TaralabaluJagadguru Institute of Technology,Ranebennur, Karnataka, India.

S G Maknur, HOD. Dept. of Information science and Engineering, ,Sri Taralabalu Jagadguru Institute of Technology, Ranebennur, Karnataka, India.