Review on Triangle Area Based MCA Technique and Anomaly Based Detection Technique for Detecting DOS Attacks

Varsharani Dudhanade
ME Computer Student from
Modern Education Society’s College of Engg., Savitribai Phule Pune University, India

Asst. Prof. Sharmila Wagh
Professor, of Department of Computer Engineering
Modern Education Society’s College of Engg., Savitribai Phule Pune University, India

Abstract: Systems, such as Web servers, database servers, cloud computing servers etc. are below threats from network attackers. The availability of such interconnected systems are severely degraded by increasing number of DOS attacks. Denial-of-Service (DoS) attacks cause serious impact on these computing systems such as router, host or entire network. In this paper, we introduce a DoS attack detection system using Multivariate Correlation Analysis (MCA) technique. Multivariate correlation analysis employs for accurate network traffic characterization by extracting the geometrical correlations between network traffic features. Therefore, this paper studies the techniques for detecting Denial-of-Service (DoS) attacks to network services and proposes an effective system for DoS attack detection. The proposed system applies the idea of Multivariate Correlation Analysis (MCA) to network traffic characterization and employs the principal of anomaly-based detection in attack recognition. Anomaly –based detection makes system capable of detecting known as well as unknown DoS attacks effectively by learning the genuine pattern of network traffic. Furthermore, a triangle area based technique is proposed to enhance and speed up the process of MCA. The KDD Cup 99 dataset employs to evaluate the effectiveness of proposed DOS attack detection system. The impact of both non-normalized data and normalized data on the performance of the proposed detection system is tested.

Keywords—Denial-of-Service attack, network traffic characterization, multivariate correlations, triangle area.

I. INTRODUCTION

A common method used in DENIAL-OF-SERVICE (DoS) attacks is to deluge the system or site with a flood of useless packets or messages that drastically slow down its response time or overwhelm its data handling capacity, resulting in online system crashes. DoS attacks severely decrease the availability of a sufferer, which can be a host, a router, or an entire network. The sufferer can be forced to stop providing services to other parties of network from a few minutes to several days. This causes serious damages to the services running on the sufferer. Thus effective mechanism for Denial-Of-Service (DOS) attack detection are highly demanded to protect services running on the sufferer. The work done until mainly concentrate on network based attack detection .the system based on this mechanism monitor traffic passing on the protected network. These mechanism prevent the online servers from monitoring attack and ensure that the servers facilitate with quality services in minimum delay while respond.

Moreover network-based detection systems are categorized in two main approaches such as misuse-based detection systems [1] and anomaly-based detection systems [2] Misuse-based detection system employs to identify known intrusions with the help of signature of previously define rules. Misuse-based detection system having low false positive rate and provides accuracy in detecting known attacks. This is not applicable in case of unknown attacks. As signature rules are generated manually, it becomes complicated to keep updating and to protect network security.

To overcome the drawbacks of misuse-based detection system as well as to archive novelty tolerant detection system and to develop a new advance concept such as anomaly based detection system. Anomaly-based system detects known as well as unknown attacks by monitoring network activities presenting significant deviation from legitimate traffic as suspicious object. The legitimate profile generation is based on statistical analysis. Moreover paper concentrate on feature correlation analysis. In this approach entire incoming group of traffic observed as legitimate or attack but not individual from group of traffic. To overcome with this problem paper employs approach based on triangle area to generate a better normalized feature. But this approach is depending on knowledge of malicious behavior. In [15] ten et al. proposed Denial –of –service (DOS) attack detection system uses multivariate correlation analysis (MCA) for accurate network traffic characterization by extracting geometrical correlation between network traffic features. In this paper a new MCA based detection system to protect online services against Dos attack which is build on previous work in [16]. In this paper Denial-of –service (Dos) attack detection system employs the triangle area based detection technique and anomaly based detection technique. Such system provides accurate characterization of traffic and detection of known and unknown attacks.

II. LITERATURE REVIEW

Z.Tan, A. Jamdagni,X. He, P. Nanda And R.P.Liu [14] illustrate that the effective detection of dos attack is essential to the protection of the online services. The DOS attack detection, mainly focuses on the development of the network
based detection mechanism [3]. The detection system employs two approaches namely misuse detection [1] and anomaly detection [2]. Misuse detection is used to identify the known attacks, using the signatures of predefined rules. [2] Anomaly detection is used to establish the usage profile of the system. During the training phase, the profiles for the legitimate traffic records are generated and the generated records are stored in the database. The trusted profile generation is build and handed over to the “attack detection” module, which compares the individual tested profile with the normal profile.

Z. Tan, A. Jamdagni, X. He, P. Nanda illustrate [3] Dos attack generally consists of efforts to temporarily or indefinitely suspend or delay services of a connected to the internet. To maintain the consistency and the availability of network services, research community has put a lot of efforts to the development of intrusion detection techniques. Benefiting from the principal of detection, which monitors and flags any network activity presenting any significant deviation from their normal profiles as a suspicious and show more advanced in detecting zero day intrusion. Therefore recent works in Dos attack mainly focus on anomaly-based techniques and various detection techniques have been proposed. However, some of these proposed techniques often suffer high false positive rate since the dependencies and correlation of the features are intrinsically neglected. While many IP trace back techniques have been proposed, they all have shortcomings that limit their usability in practice. The IP trace back method permits the routers to encode particular data on the attack packets. The routers encode the data based on some programming probability. The victims can build a set of paths which are navigated or traversed by the attack packets when they receive marked packets in sufficient number. As a result of this, the location of attacker can be recognized by the help of victim.

It has three key features. First, our model for anomaly detection operates on aggregate traffic, without flow separation or deep-packet inspection. Both of these characteristics are essential for a practical and deployable anomaly detection system. Flow separation, per-flow anomaly detection, and deep-packet inspection are difficult or impossible for most backbone routers, which have tens to hundreds of thousands of active flows per minute [8]. Since our approach only considers packet headers and timing information, it is robust to traffic concealment via encryption or tunneling. While it is true that the source and destination IP address of each packet are always available at the routers, port numbers are not available without flow-separation. Some prior work [21] uses features related to the source and destination port numbers and so will not be able to detect anomalies in aggregate or VPN tunneled traffic. Note that operating on aggregate traffic is sufficient to detect anomalies; we assume that responses such as filtering can involve heavier weight, per-flow analysis.

Shuyuan Jin, denial so young, xizhao wang [1] illustrate that covariance matrices employs multiple network attack detection. Network based detection system categorized in two different approaches namely misuse based detection and anomaly based detection. Misuse based detection have low false positive rate as well as accurate detection of known attacks. This is not applicable in case of unknown attack. It uses signature of predefined rule. Anomaly system detect known as well as unknown attacks by monitoring network activities presenting significant deviation from legitimate traffic as suspicious object. The paper proposes the covariance matrices to find out the impact of coherent relations and feature depending on multiple attacks. The effectiveness of intrusion detection system by evaluating substantial percentage of known and unknown attack. As a second order statistics a covariance matrix keep two types of information: first is information in group of samples and second is correlation information among the observed features. Furthermore, the covariance based detection employs performance improvement by using group of samples in the detection and efficiency differentiate different classes where mean based detection approaches fails.

P.garcia-teodoro, j. Diaz-Verdejo and E. Vazquez [2] puts the adaptive security oriented approaches anomaly based intrusion system protect online or protected system against malicious behavior. Intrusion detection system architectures is based four functional modules:

1. Analysis boxes: processing modules for analyzing event and detecting potentials behavior.
2. Database boxes: This element intended to store information from E blocks.
3. Response boxes: the main function of this type of block is the execution.

The intrusion detection system can be either host based or network based. Host mainly analyzes event related to OS information such as system calls. On the other hand, network related events such as traffic volume, IP address protocol, and payload length etc. intrusion detection system are classified as either signature or anomaly based detection system. Signature based defined pattern within analyzed pattern, within analyzed data. Anomaly system detects known as well as unknown attacks by monitoring network activities presenting significant deviation from legitimate traffic as suspicious object.

Chih-fong tsai , chai-ying lin [13] described a hybrid learning model based on triangle area based nearest neighbors (TANN) in order to detect attack with accuracy. This method illustrate that the technique of triangle area based nearest neighbors (TANN) by combining unsupervised and supervised learning technique to detect attacks. The classification technique is used as component and then clustering technique. This is because clustering is unsupervised learning method. Like supervised it is not able to distinguish data with accuracy. Hence, initially classifier is trained and then it provides output. Then this output is given as a input to cluster for the purpose of improving clustering performance. The proposed TANN is composed of 3 steps: 1) cluster’s center extraction, 2) new data formation by triangle area and KNN training, 3) training and testing based on new data. The centroids from given dataset having capabilities of distinguish between similar and dissimilar data or classes. The k-means used to extract a number of centers, where each centroid indicates as a one particular type of attack. Then the triangle area can be calculated by using two cluster centers and one data point from dataset. Therefore triangle area represents the new features for evaluating similar attack. Then KNN classifier used features of triangle area to detect attack.
V. Paxson [1] puts growing Internet connectivity comes growing opportunities for attackers to illicitly access computers over the network. The problem of detecting such attacks is termed network intrusion detection, a relatively new area of security research. We can divide these systems into two types, those that rely on audit information gathered by the hosts in the network they are trying to protect, and those that operate “stand-alone” by observing network traffic directly, and passively, using a packet filter. In this paper we focus on the problem of building stand-alone systems, which we will term “monitors.” Though monitors necessarily face the difficulties of more limited information than systems with access to audit trails, monitors also gain the major benefit that they can be added to a network without requiring any changes to the hosts. For our purposes—monitoring a collection of several thousand heterogeneous, diversely administered hosts—this advantage is immense of packet filter drops. If an application using a packet filter cannot consume packets as quickly as they arrive on the monitored link, then the filter buffers the packets for later consumption. However, eventually the filter will run out of buffer, at which point it drops any further packets that arrive. From a security monitoring perspective, drops can completely defeat the monitoring, since the missing packets might contain exactly the interesting traffic that identifies a network intruder. Given our first design requirement—high-speed monitoring—then avoiding packet filter drops becomes another strong requirement.

Aruna jamdagni, zhiyuan tan, priyadarshani nanda, r, ping hill [14] proposed the principle component analysis technique that employs data preprocessing, mahalanobis map for extracting features from incoming packets. It also propose iterative feature selection engine for feature selection purpose. This system detects payload based attacks in real time system. The real time payload intrusion detection system. Have 3 tier IFSENG and MDM components, which facilitate efficient and effective attack detection of network traffic. The proposed 3 tier IFSENG is used for reducing dimensionality of data, because data used in payload – base intrusion detection receive high dimensionality which influence detection The partial efficiency. Information captured of payload by mahalanobis distance map to improve performance of proposed model.

Weiming Hu, Wei Hu, and Steve Maybank [7] describes that it is necessary to develop fast machine learning-based intrusion detection algorithms with high detection rates and low false-alarm rates. The proposed network intrusion detection system is based on adaboost algorithm. In that decision rules are provided for categorical and continuous feature classification. The adaboost algorithm is the machine learning algorithm is very simple. It has been used for pattern recognition such as face and thumb recognition. The adaboost algorithm correct misclassification made by weak classifier.

M. Tavallaee, E. Bagheri, L. Wei [21] describes the rapid development of business and other transaction systems over the Internet makes computer security a critical issue. In recent times, data mining and machine learning have been subjected to extensive research in intrusion detection with emphasis on improving the accuracy of detection classifier. But selecting important features from input data lead to a simplification of the problem, faster and more accurate detection rates. In this paper, we presented the relevance of each feature in KDD '99 intrusion detection dataset to the detection of each class. Rough set degree of dependency and dependency ratio of each class were employed to determine the most discriminating features for each class. Empirical results show that seven features were not relevant in the detection of any class. In this paper, selection of relevance features is carried out on KDD '99 intrusion detection evaluation dataset. Empirical results revealed that some features have no relevance in intrusion detection.

### III. FRAMEWORK OF PROPOSED SYSTEM

![Fig. 1 framework of dos attack detection system](image-url)
In Step 1, basic features are generated from entrance network traffic to the internal network. Instead of monitoring malicious traffic, concentrate on related inbound traffic. Step 2 is 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. The feature normalization module is used to normalize traffic record. The intrusion in traffic changes the correlation. Hence, changes can be identified as intrusive activity. All the extracted features in TAM are then replace the normalized features to indicate traffic records. In Step 3, the anomaly based detection mechanism [3] is employs in Decision Making. It detects DOS attacks without having relevant knowledge. In training phase the normal profile generation module employs to generate various legitimate traffic records. In testing phase the tested profile generation module employs to build observed traffic records. This tested profile is handed over to detection module, which compares individual tested record with normal profile record.

JIN et al. [12] illustrate that detecting group of ingress network traffic samples require higher probability in classifying ingress traffic as compared to sample-by-sample detection mechanism. This restrict technique of detecting group based mechanism to limited scenario, because attack occur unpredictably and it’s tough to obtain group of sequential samples from same distribution. To remove the restriction this paper investigates ingress traffic samples individually, which provide benefits that are not provided by group based detection mechanism. Detection mechanism such as: attacks can be detected in prompt manner as compared with group-based detection mechanism, illegitimate traffic samples can be individually labeled, the probability of correctly classifying network traffic samples is higher than the group-based detection mechanism.

A. Detection Mechanism

The [12] presents a threshold-based anomaly detector. In threshold-based anomaly detector normal profiles are generated using legitimate network traffic records, such legitimate network traffic records used for comparison with incoming traffic records. The dissimilarity between incoming traffic record and respective normal profile is calculated by proposed detector. If the dissimilarity is greater than predetermined threshold, the traffic record is flagged as an attack. Otherwise, it is flagged as a legitimate traffic record.

B. Objectives of proposed work

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

1. To accurate characterization for traffic behaviors
2. To detect of known and unknown attacks respectively.
3. To enhance and to speed up the process of MCA.
4. To eliminate the bias from the raw data by using statistical normalization technique.

IV. CONCLUSION

In this way we have studied the approaches such that MCA based denial of service attack detection employs triangle area based technique and anomaly based detection technique in distributed system. The former method expresses geometrical correlations hidden in single pairs of two distinct features within the every record of network traffic and offers more correct characterization for behaviors of network traffic. The latter technique facilitates that system to be able to differentiate both unknown and known denial of service attacks from proper network traffic. DOS attack detection methods have been extensively studied.

REFERENCES


