IJSEAT International Journal of Science Engineering and Advance Technology

Improved MCA Based DoS Attack Detection

Lakshmi Prasanna Kumar Relangi[1], M. Krishna Satya Varma[2]
[1] M.Tech (IT), S.R.K.R.Engineering College, Bhimavaram, A.P., India.
[2]Asst Professor, Dept. of Information Technology, S.R.K.R.Engineering College, Bhimavaram, A.P., India.

***Abstract* — A denial of service (DoS) attack is a malicious attempt to make a server or a network resource unavailable to users, usually by temporarily interrupting or suspending the services of a host connected to the Internet 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 these computing 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 capable of 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-Service attack, network traffic characterization, multivariate correlations, triangle area.

I. INTRODUCTION

In computing,a denial-of-service (DoS) attack is an attempt to make a machine or network resource unavailable to its intended users, such as to temporarily or indefinitely interrupt or suspend services of a host connected to the Internet. A distributed denial-of-service(DDoS) is where the attack source is more than one–and often thousands–of unique IP addresses. DENIAL-OF-SERVICE (DoS) attacks are one type of aggressive and menacing intrusive behavior to online servers. DoS attacks severely degrade the availability of a victim, which can be a host, a router, or an entire network. They impose intensive computation tasks to the victim by exploiting its system vulnerability or flooding it with huge amount of useless packets. The victim can be forced out of service from a few minutes to even several days. This causes serious damages to the services running on the victim. Therefore, effective detection of DoS attacks is essential to the protection of online services. Work on DoS attack detection mainly focuses on the development of network-based detection mechanisms. Detection systems based on these mechanisms monitor traffic transmitting over the protected networks. These mechanisms release the protected online servers from monitoring attacks and ensure that the servers can dedicate themselves to provide quality services with minimum delay in response. Moreover, network-based detection systems are loosely coupled with operating systems running on the host machines which they are protecting. As a result, the configurations of network based detection systems are less complicated than that of host-based detection systems.

Generally, network-based detection systems can be classified into two main categories, namely misuse based detection systems [1] and anomaly-based detection systems [2]. 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. Research community, therefore, started to explore a way to achieve novelty-tolerant detection systems and developed a more advanced concept, namely anomaly based detection. Owing to the principle of detection, which monitors and flags any network activities presenting significant deviation from legitimate traffic profiles as suspicious objects, anomaly-based detection techniques show more promising in detecting zero-day intrusions that exploit previous unknown system vulnerabilities [3]. Moreover, it is not constrained by the expertise in network security, due to the fact that the profiles of legitimate behaviors are developed based on techniques, such as data mining [4], [5], machine learning [6], [7] and statistical analysis [8], [9]. However, these proposed systems commonly suffer from high false positive rates because the correlations between features/attributes are intrinsically neglected [10] or the techniques do not manage to fully exploit these correlations. Recent studies have focused on feature correlation analysis. Yu et al. [11] proposed an algorithm to dis-criminate DDoS attacks from flash crowds by analyzing the flow correlation coefficient among suspicious flows. A covariance matrix based approach was designed in [12] to mine the multivariate correlation for sequential samples. Although the approach improves detection accuracy,

it is vulnerable to attacks that linearly change all monitored features. In addition, this approach cans only label an entire group of observed samples as legitimate or attack traffic but not the individuals in the group. To deal with the above problems, an approach based on triangle area was presented in [13] to generate better discriminative features. However, this approach has dependency on prior knowledge of malicious behaviors. More recently, Jamdagni et al. [14] developed a refined geometrical structure based analysis technique, where Mahalanobis distance was used to extract the correlations between the selected packet payload features. This approach also successfully avoids the above problems, but it works with network packet payloads. In [15], Tan et al. proposed a more sophisticated non-payloadbased DoS detection approach using Multivariate Correlation Analysis (MCA). Following this emerging idea, we present a new MCA-based detection system to protect online services against DoS attacks in this paper, which is built upon our previous work in [16]. In addition to the work shown in [16], we present the following contributions in this paper. First, we develop a complete framework for our proposed DoS attack detection system. Second, we propose an algorithm for normal profile generation and an algorithm for attack detection respectively. Third, we proceed a detailed and complete mathematical analysis of the proposed system and investigate further on time cost . As resources of interconnected systems (such as Web servers, database servers, cloud computing servers etc.) are located in service providers' Local Area Networks that are commonly constructed using the same or alike network underlying infrastructure and are compliant with the underlying network model, our proposed detection system can provide effective protection to all of these systems by considering their commonality. 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. Our proposed DoS detection system is evaluated using KDD Cup 99 dataset [17] and outperforms the state-of -the- art systems shown in [13] and [15].

II **Problem Statement**

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 these computing systems. This makes our solution capable of detecting known and unknown DoS attacks effectively by learning the patterns of legitimate network traffic only.

III Related Work

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 capable of 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. Advantages: The results show that our system outperforms two other previously developed state-of-the-art approaches in terms of detection accuracy. To find various attacks from the user to avoid Network Intrusion.

The whole detection process consists of three major steps as shown in Fig. 1. The sample-by-sample detection mechanism is involved in the whole detection phase (i.e., Steps 1, 2 and 3) In Step 1, 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. The detailed process can be found in [17]. 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 coming from the first step or the traffic record normalized by the "Feature Normalization" module in this step (Step 2). 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. This provides higher discriminative information to differentiate between legitimate and illegitimate traffic records. Our MCA method and the feature normalization technique are explained respectively. In Step 3, the anomaly-based detection mechanism [3] 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.

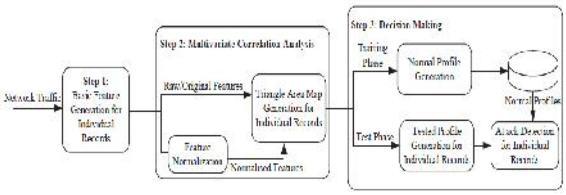

Fig. 1. Framework of the proposed denial-of-service attack detection system

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 [20] 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. The above explanation shows that our MCA approach supplies with the following benefits to data analysis. First, it does not require the knowledge of historic traffic in performing analysis. Second, unlike the Covariance matrix approaches proposed in [12] which is vulnerable to linear change of all features, our proposed triangle-area-based MCA withstands the problem. Third, it provides characterization for individual network traffic records rather than model network traffic behavior of a group of network traffic records. This results in lower latency in decision making and enables sample-by-sample detection. Fourth, the correlations between distinct pairs of features are revealed through the geometrical structure analysis. Changes of these [19] structures may occur when anomaly behaviors appear in the network. This provides an important signal to trigger an alert.

Unlike Euclidean distance and Manhattan distance, it evaluates distance between two multivariate data objects by taking the correlations between variables into account and removing the dependency on the scale of measurement during the calculation.

Require: $X^{normal}_{TAM_{lower}}$ with g elements
$\overline{TAM^{normal}_{lower}} \leftarrow \frac{1}{g}\sum_{i=1}^{g} TAM^{normal,i}_{lower}$
Generate covariance matrix Cov for $X^{normal}_{TAM_{lower}}$ using (12)
for $i = 1$ to g **do**
  $MD^{normal,i} \leftarrow MD(TAM^{normal,i}_{lower}, \overline{TAM^{normal}_{lower}})$ {Mahalanobis distance between $TAM^{normal,i}_{lower}$ and $\overline{TAM^{normal}_{lower}}$ computed using (14)}
end for
$\mu \leftarrow \frac{1}{g}\sum_{i=1}^{g} MD^{normal,i}$
$\sigma \leftarrow \sqrt{\frac{1}{g-1}\sum_{i=1}^{g}(MD^{normal,i} - \mu)^2}$
$Pro \leftarrow (N(\mu,\sigma^2), \overline{TAM^{normal}_{lower}}, Cov)$
return Pro

Fig. 2. Algorithm for normal profile generation based on triangle-area-based MCA.

Attack Detection

To detect DoS attacks, the lower triangle ($TAM^{observed}_{lower}$) of the TAM of an observed record needs to be generated using the proposed triangle-area-based MCA approach. Then, the MD between the $TAM^{observed}_{lower}$ and the $\overline{TAM^{normal}_{lower}}$ stored in the respective pre-generated normal profile Pro is computed using (15). The detailed detection algorithm is shown in Fig. 3.

Require: Observed traffic record $x^{observed}$, normal profile $Pro : (N(\mu,\sigma^2), \overline{TAM^{normal}_{lower}}, Cov)$ and parameter α
Generate $TAM^{observed}_{lower}$ for the observed traffic record $x^{observed}$
$MD^{observed} \leftarrow MD(TAM^{observed}_{lower}, \overline{TAM^{normal}_{lower}})$
if $(\mu - \sigma * \alpha) \leq MD^{observed} \leq (\mu + \sigma * \alpha)$ **then**
  return Normal
else
  return Attack
end if

Fig. 3. Algorithm for attack detection based on Mahalanobis distance.

EVALUATION OF THE MCA-BASED DOS ATTACK DETECTION SYSTEM

The evaluation of our proposed DoS attack detection system is conducted using KDD Cup 99 dataset [17]. Despite the dataset is criticized for redundant records that prevent algorithms from learning infrequent harmful records [20], it is the only publicly available labeled benchmark dataset, and it has been widely used in the domain of intrusion detection research. Testing our approach on KDD Cup 99 dataset contributes a convincing evaluation and makes the comparisons with other state-of-the-art techniques equitable. Additionally, our detection system innately withstands the negative impact introduced by the dataset because its profiles are built purely based on legitimate network traffic. Thus, our system is not affected by the redundant records. 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.

IV Conclusion

This paper has presented MCA-based DoS attack detection system which is powered by the triangle-areabased 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 and shown in Section 6. The proposed system achieves equal or better performance in comparison with the two state-of-the-art approaches. To be part of the future work, we will further test our DoS attack detection system using real world data and employ more sophisticated classification techniques to further alleviate the false positive rate.

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] D. E. Denning, “An Intrusion-detection Model,” IEEE Transactions on Software Engineering, pp. 222-232, 1987.
[4] K. Lee, J. Kim, K. H. Kwon, Y. Han, and S. Kim, “DDoS attack detection method using cluster analysis,” Expert Systems with Applications, vol. 34, no. 3, pp. 1659-1665, 2008.
[5] A. Tajbakhsh, M. Rahmati, and A. Mirzaei, “Intrusion detection using fuzzy association rules,” Applied Soft Computing, vol. 9, no. 2, pp. 462-469, 2009.
[6] 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.
[7] W. Hu, W. Hu, and S. Maybank, “AdaBoost-Based Algorithm for Network Intrusion Detection,” Trans. Sys. Man Cyber. Part B, vol. 38, no. 2, pp. 577-583, 2008.
[8] C. Yu, H. Kai, and K. Wei-Shinn, “Collaborative Detection of DDoS Attacks over Multiple Network Domains,” Parallel and Distributed Systems, IEEE Transactions on, vol. 18, pp. 1649-1662, 2007.
[9] G. Thatte, U. Mitra, and J. Heidemann, “Parametric Methods for Anomaly Detection in Aggregate Traffic,” Networking, IEEE/ACM Transactions on, vol. 19, no. 2, pp. 512-525, 2011.
[10] S. T. Sarasamma, Q. A. Zhu, and J. Huff, “Hierarchical Kohonenen Net for Anomaly Detection in Network Security,” Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on, vol. 35, pp. 302-312, 2005.
[11] S. Yu, W. Zhou, W. Jia, S. Guo, Y. Xiang, and F. Tang, “Discriminating
DDoS Attacks from Flash Crowds Using Flow Correlation Coefficient,” Parallel and Distributed Systems, IEEE Transactionson, vol. 23, pp. 1073-1080, 2012.
[12] S. Jin, D. S. Yeung, and X. Wang, “Network Intrusion Detection in Covariance Feature Space,” Pattern Recognition, vol. 40, pp. 2185- 2197, 2007.
[13] C. F. Tsai and C. Y. Lin, “A Triangle Area Based Nearest Neighbors Approach to Intrusion Detection,” Pattern Recognition, vol. 43, pp. 222-229, 2010.
[14] 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.
[15] Z. Tan, A. Jamdagni, X. He, P. Nanda, and R. P. Liu, "Denialof Service Attack Detection Based on Multivariate Correlation Analysis," Neural Information Processing, 2011, pp. 756-765.
[16] Z. Tan, A. Jamdagni, X. He, P. Nanda, and R. P. Liu, "Triangle- Area-Based Multivariate Correlation Analysis for Effective Denialof- Service Attack Detection," The 2012 IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications, Liverpool, United Kingdom, 2012, pp. 33-40.
[17] S. J. Stolfo, W. Fan, W. Lee, A. Prodromidis, and P. K. Chan, "Costbased modeling for fraud and intrusion detection: results from the JAM project," The DARPA Information Survivability Conference and Exposition 2000 (DISCEX '00), Vol.2, pp. 130-144, 2000.
[18] G. V. Moustakides, "Quickest detection of abrupt changes for a class of random processes," Information Theory, IEEE Transactions on, vol. 44, pp. 1965-1968, 1998.
[19] A. A. Cardenas, J. S. Baras, and V. Ramezani, "Distributed change detection for worms, DDoS and other network attacks," The American Control Conference, Vol.2, pp. 1008-1013, 2004.
[20] W. Wang, X. Zhang, S. Gombault, and S. J. Knapskog, "Attribute Normalization in Network Intrusion Detection," The 10th International Symposium on Pervasive Systems, Algorithms, and Networks (ISPAN), 2009, pp. 448-453.