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 10, October 2015

A Survey on Denial-of-Service Attack Detection Using Multivariate Correlation Analysis

Deepashree Mulay[1], Ankita Dungarwal[2], Chetna Palve[3], Ravindra Tambe[4]

[1,2,3]B.E. Students, Dept. of CSE, SCSMCOE,Ahmednagar, Maharashtra, India

[4]Assistant Professor, Dept. of CSE, SCSMCOE, Ahmednagar, Maharashtra, India

ABSTRACT: In todays world the rate of using Interconnected System is very high. But now these interconnected systems are under threats from the hackers and attackers. A Denial-of-Service (DoS) Attack is an intrusive attempt, which has a motive to force a designated resource to be unavailable to its intended users. DoS Attacks have emerged as a type of most severe network intrusive behaviours and have posed serious threats to the infrastructures of computer networks and various network-based services. In this paper, present the DoS attack detection system that uses Multivariate Correlation Analysis (MCA) for exact characterization of network traffic by extracting the geometrical correlations between features of network traffic.

Proposed system of detecting DoS attacks based on MCA employs a principle of anomaly based detection in attack recognition. This makes the solution capable to find unknown and known DoS attacks effectively by learning patterns of legitimate network traffic only.

KEYWORDS: Denial-of-Service attack, multivariate correlations, network traffic characterization, triangle area, trace back Scheme.

I. INTRODUCTION

Denial-of-Service(DoS) attacks have become a major threat in todays networking field to current computer networks. Early DoS attacks were technical games played among underground attackers. For example, an attacker might want to get complete access or control of an IRC channel by performing DoS attacks against the channel owner. Attackers could get recognition in the underground community by taking down popular web sites. Because easy-to-use DoS tools, like Trinoo (Dittrich 1999), can be easily available and downloaded from the Internet, normal computer users can become DoS attackers as well. Generally, network-based detection systems are classified into two main categories, namely misuse based detection systems [11] and anomaly-based detection systems [6].Misuse-based detection systems by monitoring network activities and looking for matches with the existing attack signatures detects attacks. 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's a complicated and labour intensive task to keep signature database updated because signature generation is a manual process and heavily involves network security expertise. Research community, thus, started to explore a way to achieve novelty-tolerant detection systems and developed a more advanced concept, namely anomaly based detection.

In Proposed System, the Triangle area based techniques is used to detect the attack by the intruder over an intrusive network. This TAM technique is used to identify the attackers efficiently and supports a large scalability. Furthermore, use of a triangle-area-based technique enhances and speed up the process of MCA. This method is applied in a wide area of network to block the attackers which was much efficient and protect the data from the attackers.

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 10, October 2015

II. LITERATURE SURVEY

There are different Denial-of-Service Attack detection techniques proposed by the researchers over the time which have some advantages over and vice-versa. There are many techniques used such as K-map, combination of stateful and stateless signature with trace back technique, game-theoretic, Multivariate Correlation Analysis (MCA).

Suseela T. Sarasamma, Qiuming A. Zhu, and Julie Huff [16] put a new K-Map (Kohonen Net) multilevel hierarchical structure for an intrusion finding system is presented. Each step of the hierarchical map is organized as the simple winner takes all K-Map. Calculation capability is one of the most important advantage of this K-Map multilevel hierarchical. Apart from other statistical inconsistency detection techniques such as K-means clustering or probabilistic analysis, nearest neighbour approach that engage distance measurement in a feature interval to recognize the outlines our request does not carry any costly point to point calculations in organizing the data into clusters. Its one more advantage is network size reduced. It uses the grouping efficiency of the K-Map to detect anomalies on selected dimensions of data set. It randomly selected data subsets that contain both the attacks and normal records from a KDD Cup data are used to train the hierarchical net.

The paper [16] illustrate the multilevel hierarchical Kohonen Net or Kohonen self-ordering map (K-Map) for implementing an inconsistency based intrusion detection system (IDS sensor). We did our testing and training using the pre-processed KDD Cup 99 data set. Main objective was detecting different types of attacks as possible. The experiment was done in two levels.

Detecting different kinds of DoS attacks effectively, various techniques have evolved. These techniques have shown some constraint likewise they are applicable for certain network traffic. Zhiyuan Tan, ArunaJamdagni, Xiangjian He, Priyadarsi Nanda1, and Ren Ping Liu proposed technique which runs analysis on original feature space (first-order statistics) and extracts the multivariate correlations between the first-order statistics [1]. The extracted multivariate correlations, namely the second-order statistics, preserve significant discriminative information for accurate characterizations of network traffic records. And these multivariate correlations can be the high-quality potential features for theDoS attack detection. The effectiveness of the proposed technique is evaluated using KDD CUP 99 dataset and experimental analysis shows encouraging results.

Shuyuan Jin, Daniel S. Yeung proposed technique which discusses these effects of multivariate correlation analysis on the DDoS attack detection and proposes a covariance analysis model for detecting SYN flooding attacks [12]. The simulation results of this method shows that this method is highly accurate in detecting malicious network traffic in DDoS attacks of different intensities. This method can effectively differentiate between normal traffic and attack traffic. Indeed, this method can detect even a very subtle attacks only slightly different from normal behaviors. The linear complexity of this method makes it real time detection practical.

Mihui Kim, Hyunjung Na, KijoonChae, Hyochan Bang, and Jungchan Na proposed a technique which putforth a combined data mining approach for modeling the network traffic pattern of normal and diverse attacks [13]. The automatic feature selection mechanism is used in this approach for selecting the important attributes. And the classifier is build with the theoretically selected attribute from the neural network. Then, our experimental results show that our approach can provide the best performance on the real network, in comparison with it by heuristic feature selection and any other single data mining approaches.

AikateriniMitrokotsa, ChristosDouligeris proposed a technique which putforth an approach that detects Denial of Service attacks using Emergent Self-Organizing Maps [14]. This approach is based on classifying “normal” traffic against “malicious” traffic in the sense of Denial of Service attacks. The approach allows the automatic classification of events that are contained in logs and visualization of network traffic. Extensive simulations result shows the effectiveness of this approach compared to previously proposed approaches regarding false alarms and detection probabilities.

Haining Wang Danlu Zhang Kang G. Shin proposed a technique which putforth a simple and robust mechanism, called Change-Point Monitoring (CPM), to detect denial of service (DoS) attacks [15]. The core of CPM is based on the

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 10, October 2015

inherent network protocol behaviors, and is an instance of the Sequential Change Point Detection. Making the detection mechanism insensitive to sites and traffic patterns, a non-parametric Cumulative Sum (CUSUM) method is applied, thus making the detection mechanism robust, more generally applicable and its deployment much easier. CPM does not require per-flow state information and only introduces a few variables for recording the protocol behaviors. The statelessness and low computation overhead of CPM make it immune for any flooding attacks. As a case study, the efficiency of CPM is evaluated by detecting a SYN flooding attack. The most common DoS attack. The evaluation results shows that CPM has short detection latency and high detection accuracy.

III. RELEVANT THEORY

A. Denial-Of-Service

Denial-of-Service(DoS) attacks are one of the major type of aggressive and menacing intrusive behavior to online servers. DoS attacks severely degrade the service 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 that runs on the victim. Thus, effective detection of DoS attacks is necessary to the protection of online services. The main focus of the Work on DoS attack detection is on the development of network based detection mechanisms. Detection systems based on these mechanisms monitorsnetwork traffic transmitting over the protected networks. These mechanisms release the protected online servers from monitoring attacks and ensures that the servers can dedicate themselves to provide quality services with much possible minimum delay in response. Moreover, network-based detection systems are loosely coupled with operating systems runs on the host machines which they are protecting. As a result, by comparing the configurations of network based detection systems are less complicated than that of host-based detection systems. Denial of Service(DoS) attacks are a class of attacks on targets, which aims at exhausting target resources, thereby denying service to valid users. The target resources could be in terms of space and/or time. For example, servers providing SSL service could be time-attacked by making them perform a lot of expensive cryptographic operations (public key decryption in this case) thereby preventing them from serving their genuine clients. Alternately, servers could also be space-attacked by exhausting their bandwidth or connection buffers with lot of bogus packets/requests.

B. KDD Cup 1999 Data

KDD Cup-99 is the data set used for The Third International Knowledge Discovery and Data Mining Tools Competition. This was held in conjunction with KDD-99 The Fifth International Conference on Knowledge Discovery and Data Mining. The task included in that competition was to build a network intrusion detector, a predictive model which should be capable of distinguishing between ``bad" connections, called intrusions or attacks, and ``good" normal connections. Standard set of data was included in database to be audited, which includes a wide variety of intrusions simulated in a military network environment.

C. Multivariate Correlation Analysis

The coefficient of multiple correlations gives a measure of how well a given variable can be predicted using a linear function of a set of other variables. It is measured by the squareroot of determination, but under the particular assumptions the best possible linear predictors are used and the intercept is included, whereas the coefficient of determination is defined for more general cases, including nonlinear prediction which the predicted values have not been derived from a model-fitting procedure. The multiple correlation takes values between zero and one; a higher value indicates a better predictability of the dependent variable from the independent variables, with a value indicating that the predictions are exactly correct and a value of zero indicating that no linear combination of the independent variables is a better predictor than is the fixed mean of the dependent variable. To describe these statistical properties, a novel Multivariate Correlation Analysis(MCA) approach is presented. By using the MCA approach triangle area for extracting the correlative information between the features within an observed data object (traffic record). The details are present as follows. We had applied the concept of triangle area to extract the geometrical correlation between the ith and jth features in the vi. When the two Triangle area maps compared, we can imagine the map into two images symmetric along their main diagonals. Any differences were identified on the upper triangles of the images, and those can be found on their lower triangles as well. Thus, to perform a quick comparison of the two TAM, we can choose to

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 10, October 2015

investigate either the upper triangles or the lower triangles of the TAM only. This produces the same result as comparing using the entire TAM (thus, the correlations residing in a traffic record (vector vi) can be represented effectively and correctly by the upper triangle or the lower triangle of the respective.

IV. PROPOSED ARCHITECTURE

The overview of proposed Denial-of-service(DoS) attack detection system architecture is given in this section, where the system framework and detection mechanism are discussed. The working process of the system is explained in detail. The complete detection process is made up of three levels as shown in Fig.

Level1. Multivariatecorrelationanalysis level.
Level 2. Normal profile generation level.
Level 3. Attack Detection level.

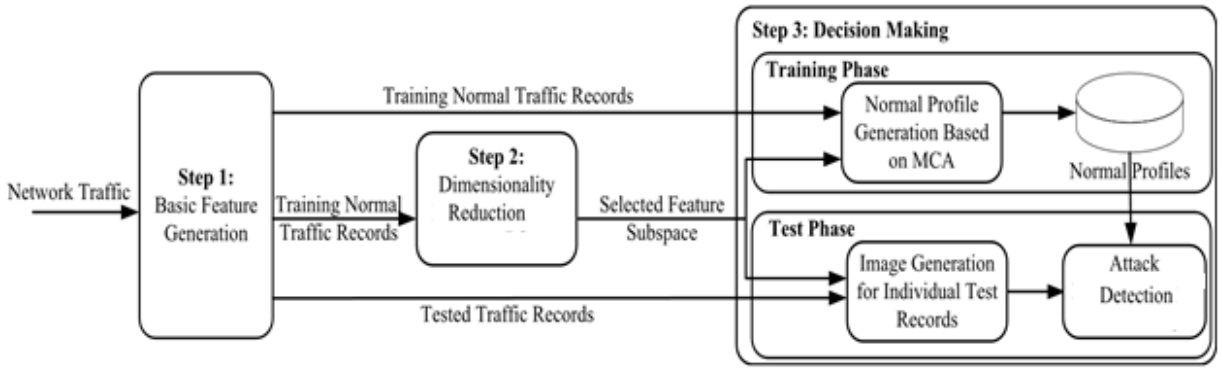

Fig.1.A System Framework of Denial-of-Service Attack Detection

The framework consists of three Levels:

Level 1: In this level using the network traffic ingress the basic features are generated to internal network where proposed servers resides in. And those servers are used to form the network traffic records for well-defined time period. Monitoring and analysing network for reducing the malicious activities only on relevant inbound traffic.
To provide a best protection for a targeted internal network. This also enables our detector to provide protection which is the best fit for the targeted internal network because the detectors uses the legitimate traffic profiles which are developed for a smaller number of network services.

Level 2: In this step the Multivariate Correlation Analysis is applied in which the Triangle Area Map Generation module is applied to extract the correlation between two separate features within individual traffic record. The distinct features are come from level 1or "feature normalization module" in this step. All the extracted correlation are stored in a place called Triangle area Map(TAM), are then used to replace the original records or normalized feature record to represent the traffic record. It's differentiating between legitimate and illegitimate traffic records.

Level 3: The anomaly based finding mechanism is adopted in decision making. Decision making involves two phases as

- Training phase.
- Test phase

Normal profile generation module is work in "Training phase" to generate a profiles for various types of 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 at last the tested profiles are handed over to

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 10, October 2015

"Attack Detection" module it compares tested profile with stored normal profiles. This distinguishes the Dos attack from legitimate traffic.

This needs the expertise in the targeted detection algorithm and it is manual task. Particularly, two levels (i.e., the Training Phase and the Test Phase) are included in Decision Making. The Normal Profile Generation module is operated in a Training Phase [1] to generate profiles for various types of legal records of traffic, and the normal profiles generated are stored in the database. The tested profile generation module is used in a TestPhase to build profiles for the each observed traffic documentation. Next, the profiles of tested are passed over to an attack detection part, which calculates the tested profiles for individual with the self-stored profiles of normal. A threshold based classifier is employed in the attack detection portion module to differentiate DoS attacks from appropriate traffic.

V. PROPOSED ALGORITHM

A. *Basic Feature Generation for Individual Records:*

With Administrator of the system having access to questions posted by normal-user, he can generate the Triangle Area Map (TAM) for the system which consists of irrelevant questions.
i.e TAM(Tr1,Tr2…….Tn);
where Tr1 = (f1,f2,f3……fn)
and f1 = {x1,x2….xn} represents the dataset of questions.

B. *Multivariate Correlation Analysis:*

Here we check the question posted by normal-user is in normalized form and out of Triangle Area Map(TAM).
1) Feature Normalization
Step1: NormalizationStepOne:

Here we normalize the data (question) according to

i) Question Length i.e. checkQuestionLength(question)
Which return the length of question.

ii) WordCount(i.e. To complete a question we require minimum three words) i.e. checkWordCount(question)
Which return the word count.

iii) Consecutive words (words appearing one after) must not be same i.e. checkSameWords(question)
Which returns a Boolean value true or false.

Step 2:NormalizationStepTwo:

Here we normalize the data (question) according to

i) Punctuation Marks (question must not contain only punctuation marks)
i.e. checkPunctuation(question)
which returns a Boolean value true or false.

ii) Complete Questiion(A question that starts with wh or do must end with ?)
i.e. checkQuestionMark(question)
which returns a Boolean value true or false.

Step 3:NormalizationStepThree:

Here we normalize the data (question) according to

i) Only numbers (A question must not contain only numbers)
i.e. checkOnlyNumbers(question)
which return a Boolean value true or false.

ii) Only spacesb or tabs (A question must not be blank)
i.e. checkOnlySpacesorTabs(question)
which returns a Boolean value true or false.

2) Triangle Area Mapping
Here we map the question posted by normal-user to the Triangle area map generated in Procedure 1

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 10, October 2015

i.e. mapQuestion(question)
which returns a Boolean value true or false.

C. ***Decision Making:***
Step 1: Training Phase
Here we check the data with the normalized data that we created in Procedure 2
If data doesn't match then normal profiles are generated and question is passed to expert-user for answering.
i.e. insertProfile(question)
which will insert the profile.
Step 2: Test Phase
Here we detect the attack by comparing the data with dataset and if match is found then we deny the service for the user for individual records.
i.e. block(user)
which blocks the user and denies service to him.

VI. RESULT ANALYSIS

Many threshold frequency were set in comparison. The result reveals that at a certain threshold the server gets enumerous request from the user, at the instance the server goes to sleep mode for long time period and crashes. Now this particular threshold is set as a limit to detect the intrusive networks attack by user. When such attack is detected , then the service to the user is denied i.e. the user is blocked .

VII. CONCLUSION AND FUTURE WORK

The system implemented here is useful and efficient technique for the detection of dos attack. We have extracted important features from MCA (analysis technique) with use of triangle area map method. The application also has implemented the detection of sql injection attack to preserve the data needed for the analysis purpose and stored in the databases like the normal profiles generated and any other application related data.

The future work will be consist of packing methods implemented in the application as the services which will run on the server and prevent any website from the attacks listed above. Here, we have used the analysis technique just to prevent dos attack but in future it can be the package or more techniques which can prevent different kind of attacks on the system and which can be implemented directly through the package of attack detection and prevention services.

REFERENCES

[1] A Jamdagni, Z. Tan, X. He, P. Nanda, and R. P. Liu, " A System For Denial-of-Service Attack Detection Based On Multivariate Correlation Analysis " IEEE Transactions on parallel and distributed systems VOL-25 NO:2 YEAR 2014

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

[3] S. Yu, W. Zhou, W. Jia, S. Guo, Y. Xiang, and F. Tang, "Discriminating DDoS Attacks from Flash Crowds Using Flow Correlation Coefficient," IEEE Transactions on Parallel and Distributed Systems, vol. 23, pp. 1073-1080, 2012.

[4] Z. Tan, A. Jamdagni, X.He, P.Nanda, and R. P. Liu, "Denial-of-Service Attack Detection Based on Multivariate Correlation Analysis," IEEE Transactions on Neural Information Processing, 2011, pp. 756-765.

[5] G. Thatte, U. Mitra, and J. Heidemann, "Parametric Methods for Anomaly Detection in Aggregate Traffic," IEEE/ACM Transactions on Networking,, vol. 19, no. 2, pp. 512-525, 2011.

[6] P. Garca-Teodoro, J. Daz-Verdejo, G. Maci-Fernndez, and E. Vzquez, "Anomaly-based Network Intrusion Detection: Tech-niques, Systems and Challenges," IEEE Transactions on Computers and Security, vol. 28,pp. 18-28, 2009

[7] C. Yu, H. Kai, and K. Wei-Shinn, "Collaborative Detection of DDoS Attacks over Multiple Network Domains," IEEE Transactions on Parallel and Distributed Systems, vol. 18, pp. 1649-1662, 2007.

[8] Zhiyuan Tan D.E.Denning, "An Intrusion-detection Model," IEEE Transactions on Software Engineering, pp. 222-232, 1987.

[9] M. Tavallaee, E. Bagheri, L. Wei, and A. A. Ghorbani, "A Detailed Analysis of the KDD Cup 99 Data Set," The The Second IEEE International Conference on Computational Intelligence for Security and Defense Applications, 2009, pp. 1-6.

[10] 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.

[11] 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.

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 10, October 2015

[12] ShuyuanJin,Daniel S. Yeung, „A Covariance Analysis Model for DDoS Attack Detection. IEEE Communications Society 0-7803-8533-0/04/$20.00 (c) 2004 IEEE Hong Kong RGC project research grant number B-Q571

[13] Mihui Kim, Hyunjung Na, KijoonChae, Hyochan Bang, and Jungchan Na, „A Combined Data Mining Approach for DDoS Attack Detection. ICOIN 2004, LNCS 3090, pp. 943–950, 2004 Springer-Verlag Berlin Heidelberg 2004

[14] AikateriniMitrokotsa, ChristosDouligeris, „Detecting Denial of Service Attacks Using Emergent Self-Organizing Maps. 2005 IEEE International Symposium on Signal Processing and Information Technology

[15] Zhiyuan Tan1; Aruna Jamdagni1; Xiangjian He1, Priyadarsi Nanda1, and Ren Ping Liu, Multivariate Correlation Analysis Technique Based on Euclidean Distance Map for Network Trac Characterization".

[16] 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.

BIOGRAPHY

DeepashreeDnyaneshwarMulayis a B.E. Student in the Computer Engg. Department, Shri. ChhatrapatiShivajiMaharaj College of Engg, SavitribaiPhule Pune University. She perceiving Bachelor ofEngg. in Computer Engg. degree in 2015 from SPPU, Ahmednagar, MS, India. Her research interests are Computer Networks, Artificial Intelligence, HCI, Cloud Computing, Algorithms, etc.

AnkitaTarachandDungarwalis a B.E. Student in the Computer Engg. Department, Shri. ChhatrapatiShivajiMaharaj College of Engg, SavitribaiPhule Pune University. She perceiving Bachelor ofEngg. in Computer Engg. degree in 2015 from SPPU, Ahmednagar, MS, India. Her research interests are Computer Networks, WSN, Cloud Computing, Pervasive Computingetc.

ChetnaSampatPalveis a B.E. Student in the Computer Engg. Department, Shri. ChhatrapatiShivajiMaharaj College of Engg, SavitribaiPhule Pune University. She perceiving Bachelor ofEngg. in Computer Engg. degree in 2015 from SPPU, Ahmednagar, MS, India. Her research interests are Computer Networks, Robotics, Cloud Computing, Pervasive Computing etc.

Ravindra Tambeis Assistant Professor in the Computer Engg. Department, Shri.Chhatrapati Shivaji Maharaj College of Engg, Savitribai Phule Pune University.He completed Bachelor of Engg. in Computer Engg.degree from SPPU, Ahmednagar, MS, India. His research interests are Computer Networks, Robotics, Cybernetis, Cyber Securityetc.