www.ijecs.in
International Journal Of Engineering And Computer Science ISSN:2319-7242
Volume 7 Issue 1 January 2018, Page No. 23390-23493
Index Copernicus Value (2015): 58.10, 76.25 (2016) DOI: 10.18535/ijecs/v7i1.04

Review on EM-CURE Algorithm for Detection DDOS Attack

Miss Priyanka P. Narode[1], Prof I.R.Shaikh[2]

[1]Computer Engineering Department, Savitri Bai Phule University,Pune,India,
[2]Computer Engineering Department, Savitri Bai Phule University,Pune,India,

Abstract:
Distributed Denial of Service attack (DoS attack) is a cyber attack where the perpetrator seeks to make a machine or network resource unavailable to its intended users by temporarily or indefinitely disrupting services of a host connected to the internet. Denial of service is typically accomplished by flooding the targeted machine or resource with superfluous requests in an attempt to overload systems and prevent some or all legitimate requests from being fulfilled. It is necessary to analyze the fundamental features of DDoS attacks because these attacks can easily vary the used port/protocol, or operation method because they are designed to restricted applications on limited environments.DDoS attack detection very difficult because the non-existence of predefined rules to correctly identify the genuine network flow. A combination of unsupervised data mining techniques as IDS are introduced. The Entropy Method concept in term of windowing the incoming packets is applied with data mining technique using Clustering Using Representative (CURE) as cluster analysis to detect the DDoS attack in network flow. The data is mainly collected from datasets. The CURE DDoS attack detection technique based on entropy gives a promising way to analyze this attack and construct an efficient detection model using a clustering data mining techniques. This approach has been evaluated and compared with several existing approaches in terms of accuracy, false alarm rate, detection rate, F. measure and Phi coefficient.

Keywords: Distributed Denial of Service (DDoS), Data Mining, Intrusion Detection, Clustering, Network security .

1. Introduction

With the rapid progress of the Internet technology and growth of network infrastructure, many service have been implemented online. In that , the DDoS attack has become one of the most representative threats because its impact and frequency have grown to significant levels. Distributed Denial of Service(DDoS), is a relatively simple, yet very powerful technique to attack. Internet resources as well as system resources. Distributed multiple agents consume some critical resources at the target within the short time and deny the service to legitimate clients.

Distributed Denial of Services (DDoS) attacks are among the major security threats launched using internet services. It has been very critical for any organization to protect their computing environment from unauthorized access or malicious attacks. Detection suffers from efficiently differentiating the normal stream and abnormal stream of traffic. The attack itself often uses legitimate requests to flood the target and this makes it hard to distinguish an attack traffic from legitimate traffic, fast real time detection is difficult because of huge amount of data involved in current computer networks. I can used different solutions for detecting and preventing DDoS attack. Some forms of Intrusion Detection Systems (IDS) for detection and preventing these threats, and identifying intruders from the legitimate users. The analysis of intrusion, the intrusion detection approaches falls into three categories that include: Misuse Intrusion Detection (MID), Anomaly Intrusion Detection (AID), and a hybrid that combines the strategies of both of MID and AID. Misuse detection is based on pattern matching with the known signatures that extracted from the known attacks. However, misuse detection needs to be updated periodically to include the new types of attacks. An efficient method using Entropy Method with CURE (EM-CURE) is introduced in intrusion detection for inspecting and monitoring of the traffic packet header features. The EM-CURE cluster analysis is implement using training and testing datasets.

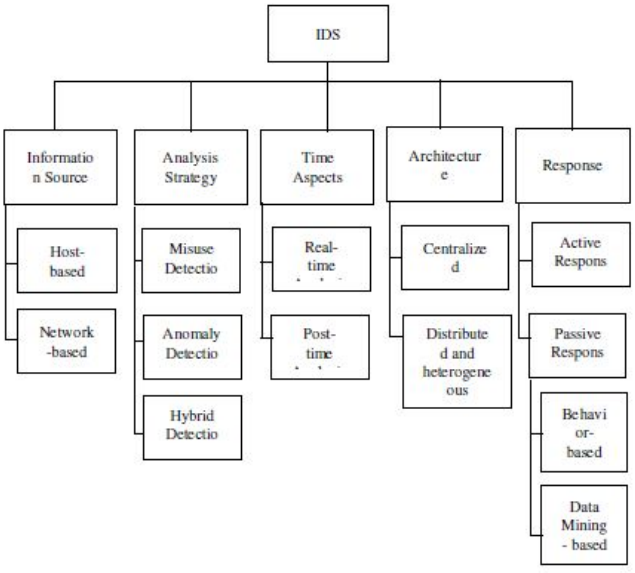

Figure 1: IDS-DDOS Attack Detection Strategies.
IDS-DDOS strategies based on data mining. The significance of designing and implementing EM-CURE cluster analysis is to analyze and detect DDOS attacks

2. Literature Survey

1. W. Cerroni, G. Monti, G. Moro, and M. Ramilli use an unsupervised technique to distinguish effectively the normal behavior from malicious network flow using k-means clustering. The proposed technique was tested using the web server as a real test bed for the experimental attacks.
2. M. Suresh and R. Anitha have been found Fuzzy c means techniques to efficiently detect the DDoS attacks with a better accuracy. DDoS attacks increases the serious damage, the rapid detection of the DDOS attack they use the various machine learning models, like Navies Bayes, C4.5, SVM, KNN, K-means and Fuzzy c-means clustering are developed for efficient detection of DDoS attacks. Then the experimental results show that Fuzzy c-means clustering gives better accuracy in identifying the attacks and it is fast compared to the other algorithms.
3. H. Om and A. Kundu proposed k-Means and two classifiers, K-nearest neighbors and Naïve Bayes, for anomaly detection. In their proposed hybrid detection solution model, certain attributes have been selected based on entropy. This hybrid method successfully reduced the false alarm rate.
4. J. Mazel, P. Casas, and P. Owezarski Detecting DDoS attack using cluster analysis in different shapes and size. Network attacks such as denials-of-service and worms spreading, network traffic anomalies can have serious detrimental effects on the performance and integrity of the network.Unsupervised detection is accomplished by means of robust data-clustering techniques, combining Sub-Space Clustering and multiple Evidence Accumulation algorithms to blindly identify anomalous traffic flows. They developed a completely unsupervised method to detect and characterize network anomalies.
5. V. Rajyaguru, V. R Tamma, B. S. Manoj, and M. Sarkar has been described the application of clustering technique in detecting multiple DDoS attacks has. Denial of Service (DoS) attack is one of the popular attack that prevents WLAN users from accessing the wireless network resources. The performance of the different clustering analysis techniques is extensively evaluated with real traffic from two different datasets. They use the clustering techniques on wireless traffic datasets for detecting Clear-to-Send (CTS) -based DoS attacks. The k-means clustering technique is used to achieve high detection rates and low false positive rates with relatively small values of k (i.e., number of clusters). Clustering able to identify attacks with high degree of accuracy. Under the cases of Naïve CTS duration attack as well as the sophisticated CTS duration attack.

2. Design of the System

The cluster analysis consists of CURE algorithm which is efficient for large data scales. CURE is more robust to identify outliers and clusters with non-spherical shapes. In CURE, a collection of well-scatter points is used to identify the cluster shape. The general overview of CURE algorithm is presented in Figure: 2

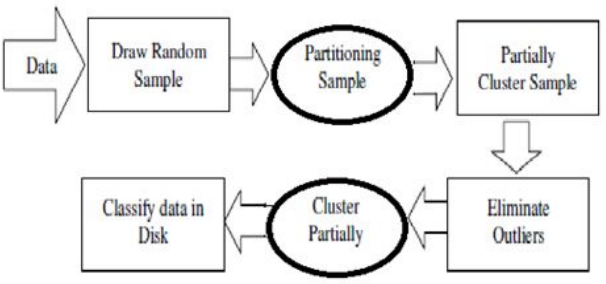

Figure:2 Overview of Cure Algorithm Steps.

Basic CURE algorithm includes steps as follows to large database:

1. Gets a sample after sampling in the database;
2. Data partition and partition clustering: we mark p partitions from the sample, the scale of every partition is n / p. Clustering begins at every partition first, so this strategy can accelerate algorithm; then we use layered algorithm to cluster aim at every partition;
3. Delete exceptional point: we delete clusters that increase slowly; we delete very small cluster at last phase;
4. Total clustering: we use c points to represent initial clustering result in order to ensure disposal in limited memory, then begin to cluster to total database.

I should notice that use data sample or use data structure to compress database aim at large database clustering algorithm. I want to find balance point by some strategy between resource environment and algorithm precision.

3. Basic CURE Algorithm

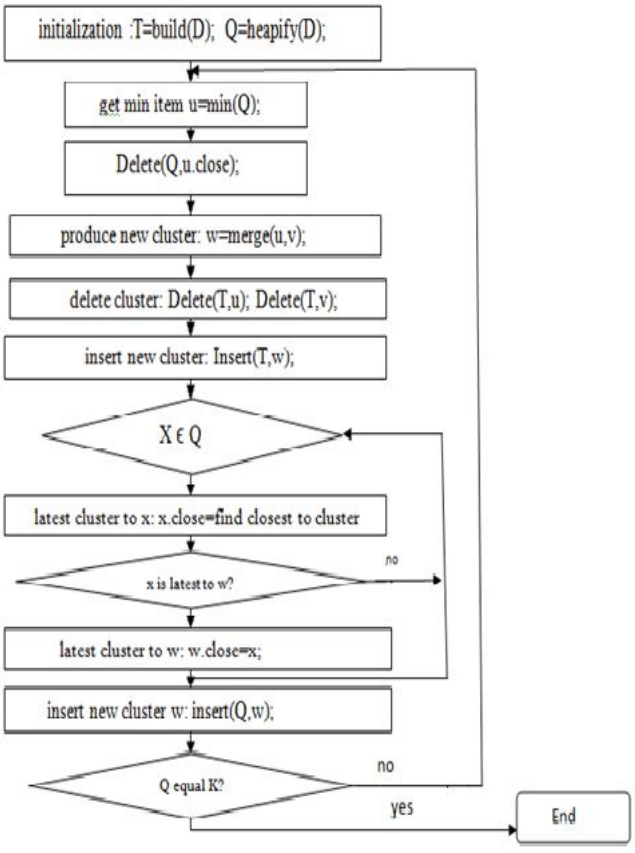

Figure:3 Flow of basic CURE Algorithm.

Input : A set of points S

Output : k clusters

1. For every cluster u (each input point), in u.mean and u.rep store the mean of the points in the cluster and a set of c representative points of the cluster (initially c = 1 since each cluster has one data point). Also u.closest stores the cluster closest to u.
2. All the input points are inserted into a k-d tree T
3. Treat each input point as separate cluster, compute u.closest for each u and then insert each cluster into the heap Q.(clusters are arranged in increasing order of distances between u and u.closest).
4. While size(Q) > k
5. Remove the top element of Q(say u) and merge it with its closest cluster u.closest(say v) and compute the new representative points for the merged cluster w.
6. Remove u and v from T and Q.
7. For all the clusters x in Q, update x.closest and relocate x
8. insert w into Q
9. repeat

4. Proposed System

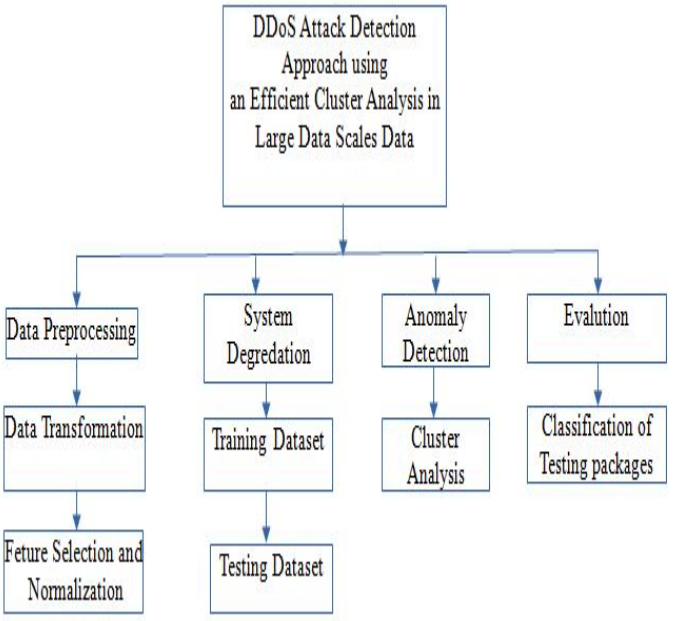

Figure 4: Proposed System Structure

The Praposed System Structure is shown in Figure:4 which mainly focuses on following modules:

1. Data Preprocessing
2. System Degredation
3. Anomaly Detection
4. Evaluation

1.Data Preprocessing

In the preprocessing phase, the Database features are specified and then the proposed entropy windows are applied. A major step in traffic pre-processing is data transformation. Information theory is a crucial step to convert the data from one format to another format.

Entropy is used mainly with data clustering because it provides a good clustering of data since it can be used with network traffic to convert the network flow to numerical (quantitative) data type. Formally, Entropy considers the measure of uncertainty in random variables ($x_1, x_2, x_3 .... x_n$) from an information source that has (n) different values. The probability of x_i appears in information sample identified by x_i, then the entropy (H) is defined for the random variable (X)

$$H = -\sum_{t=0}^{n} P(x_i) \log_2 P(x_i) \quad \text{.....(1)}$$

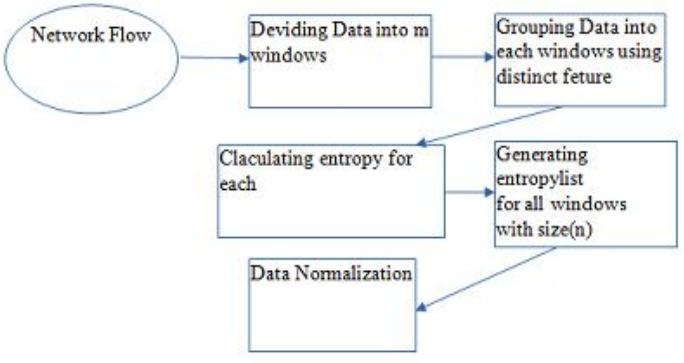

Figure 5: Block Diagram of Pre-processing using entropy Window

2.System Degredation The EM-CURE cluster analysis is implemented using training and testing features of datasets. In that, I can select the training and testing values are randomly . The number of points that is supposed to be included within a certain cluster determines and captures the shape and extent of the clusters.

3.Anomaly Detection

Anomaly Intrusion Ditection is an established profile of the system's normal behavior. AIDs is used to detect attack if there are any deviations from the established normal system profiles

4.Evaluation

The performance evaluation for a proactive DDoS attack detection system can be measured by a confusion matrix as shown in Table(1)

Actual	Predicted	
	Normal	Attack
Normal	TN	FP
Attack	FN	TP

Table 1: Confusion Matrix.

True Positive (TP): the number of the malicious packets correctly classified as malicious.

False Positive (FP): the number of normal traffic falsely classified as malicious.

False Negative (FN): it occurs when the malicious traffic is classified as normal traffic.

True Negative (TN): the number of benign packets correctly classified as benign.

By using all this values of TP,FP,FN,TN I will calculate the accuracy rate, detection rate, false alarm rate, F-measure and Phi coefficient to indicate the superiority of the proposed approach.

5. Conclusion

I have studied in this paper, Intrusion Detection Systems (IDS). In computing world it can play an integral role to protect essential information that may have been exposed to unauthorized access. The evaluation of Entropy Method with CURE for effectively detecting the DDoS attacks, I used data set is as the attack data and based packet information, The CURE DDoS attack detection technique using a clustering data mining techniques based on entropy gives a promising way to analyze this attack and construct an efficient detection model using a clustering data mining techniques, calculate the frequency of attack packet during the network flow.

6. References

1. Wesam Bhaya and Mehdi EbadyManaa, "DDoS Attack Detection Approach using an Efficient Cluster Analysis in Large Data Scale," in Annual Conference on New Trends in Information & Communications Technology Applications,(NTICT'2017) 7 - 9 March 2017.
2. W. Cerroni, G. Monti, G. Moro, and M. Ramilli, "Network Attack Detection Based On Peer-To-Peer Clustering Of SNMP Data," in international Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustnes, 2009, vol. 22, no. 213110, pp. 417-430.
3. M. Suresh and R. Anitha, "Evaluating Machine Learning Algorithms for Detecting DDoS Attacks," in International Conference on Network Security and Applications, 2011, pp. 441-452.
4. H. Om and A. Kundu, "A Hybrid System for Reducing the False Alarm Rate of Anomaly Intrusion Detection System," in International
5. Conference on Recent Advances in Information Technology, 2012, pp. 131-136.
6. J. Mazel, P. Casas, and P. Owezarski, "Sub-space Clustering and Evidence Accumulation for Unsupervised Network Anomaly Detection," in
7. International Conference in Traffic Monitoring and Analysis, 2011, vol. 6613, pp. 15-28.
8. V. Rajyaguru, V. R Tamma, B. S. Manoj, and M. Sarkar, "On Detecting CTS Duration Attacks Using K-means Clustering in WLANs," in
9. International Conference on Advanced Networks and Telecommunciations Systems, 2012, pp. 12-14.
10. R. Suganya, "Denial-of-Service Attack Detection Using Anomaly with Misuse Based Method," in International Journal of Computer Science and Network Security, vol. 16, no. 4, pp. 124-128, 2016.
11. T. Johnson Singh, Khundrakpam; Thongam, Khelchandra; De, "Entropy Based Application Layer DDoS Attack Detection Using Artificial Neural Networks," MDPI, vol. 18, pp. 1-17, 2016.
12. Shao Xiufeng, Cheng Wei, "Improved CURE Algorithm and Application of Clustering for Large-scale Data,".