I J C T A, 9(28) 2016, pp. 149-156

DATA MINING TECHNIQUE FOR DDOS ATTACK IN CLOUD COMPUTING

Sivamohan S* R.Veeramani* Liza K* and Krishnaveni S, Jothi B**

Abstract: Data Mining plays a vital role for implementation of Cloud security against various types of attacks. Cloud services can be vulnerable to Distributed Denial of Service (DDoS), which is one of the most common and damaging forms of attack on the cloud. The Detection of various attacks on the Cloud computing environment becomes an important task. This paper focus on the analysis of various detection techniques with corresponding merits and demerits on the previous attack against defense mechanisms based on cloud computing, this enable to understand effective and efficient technique to detect the DDoS attacks on cloud computing environment can be devised. In order to detect the DDoS attacks in cloud areas many approaches had been implemented by the researchers, the earlier approaches focused on the area of data mining techniques for Detection of DDoS attack on the Cloud Computing environment. This paper presents a comprehensive review of DDoS attacks, detection techniques and tools used in Cloud environment. This review paper aims to describe the DDoS and its classifications, and identify the effective parameters in detection of DDoS attacks in cloud.

Keywords: Data Mining techniques, Distributed Denial of Service (DDoS) attack, cloud security

1. INTRODUCTION

The Distributed Denial of service attack (DDoS)is one is one of the most common and damaging forms of attack on the cloud. The Denial of service attacks is attempts to unavailable the functioning of resources to the end users, hackers can send the unwanted messages continuously and make the traffic from multiple resources, hackers will send packets to the receiver which make harmful to the system and temporarily stop the services between client and server communication[1]. A Distributed Denial of Service attacks is type of DoS attack and various distributed attack sources. Usually, the attackers use a huge number of controlled bots dispersed in different locations to start on a great number of denial of service attacks to a lone target or several targets. Distributed Denial of Service attacks has been rigorously increased with the target system, these attacks become serious issues in the current industries and internet infrastructures, and including networking devices such as routers, firewalls, switches and network bandwidth. Now a days attackers are targeting the cloud computing environment, Cloud computing environment is a distributed computing model, anybody can access from anywhere also called as pay as you go model, cloud is always tempting the target for attackers, they can easily get in to the cloud as user, actually cloud providers are providing their binaries to their user,even user may be a hacker,they can easily modifies the code and make harmful to the end system, The growth of cloud computing provides the on-demand services through online is leading to a new requirement for its sustenance that is, its security. The cloud security has become a key area of research; as a result a new dimension is added to the field of information security. Cyber security plays a major role in cloud security because most of the cloud services are accessed through the cyber interface. This paper focus on the analysis of various detection techniques with corresponding merits and demerits on the previous attack against defense mechanisms based on cloud computing, this enable to understand effective and efficient technique to detect the DDoS attacks on cloud computing environment can be devised. In order to detect the DDoS attacks in cloud areas many approaches had been implemented

* Assistant Professor, Department of IT, SRM University, Ramapuram,
Emails: sivamohan7@gmail.com[1], gvmani.r@gmail.com[2], lizamk2006@yahoo.co.in[3]
** Assistant Professor, Department of SWE, SRM University, Kattankulathur, Email: vanimithila@gmail.com

by the researchers, the earlier approaches focused on the area of data mining techniques for Detection of DDoS attack on the Cloud Computing environment. This paper is structures in seven sections; Section 1 describes the basic introduction about DDos attacks whereas section 2 presents the background of Data mining and Dos Attack. Section 3 related work. section 4 dos attacks in cloud environment. section 5 describes the various effective parameters in DDoS detection. Section 6 covers the a taxonomy of defense mechanisms against DDoS flooding attacks**.** and the conclusion and future enhancements is drawn in the section 7.

2. BACKGROUND

2.1 Data Mining

Data mining is an one of the powerful field in computer Science, also called as Knowledge Discovery in Databases" process, or KDD.It is the process of discovering hidden value in large data sets, it involves different methods such as artificial intelligence, machine learning, statistics and data base systems.[3]. The data mining is a process of extracting the information from a various data set and transforms it into correct format for future use. Data mining process includes database and data management aspects, data preprocessing, inference considerations, complexity considerations, post-processing of discovered structures, and online updating. Roots of Data Mining [2] are statistics, Artificial Intelligence & Machine Learning, Databases, Pattern discovery, visualization, business Intelligence etc. There are various Data mining techniques are Clustering – It is the task of discovering groups and structures in the data that are in some way or another "similar", without using known structures in the data. Classification – It is the task of generalizing known structure which can be applied to new data. For example, an email program might attempt to classify an email as genuine or spam. Regular algorithms are decision tree learning, Naive Bayesian classification, neural networks (soft computing) and support vector machines. Regression - Attempts to find a function which models the data with the least error. Association Rule Learning - Searches for relationships between variables.

2.2 Denial of Service Attack

DoS attacks have become a major threat to current computer networks.To have a better understanding on DoS attacks. This statistics involves an outline of existing DoS, DDOS attacks and major defense techniques over the internet. In this observation we describe host based and networks based DoS attack methods.DOS attacks are categorized according to the major attack types. Current available techniques are also reviewed, including variety of defense products in deployment and representative defense approaches in research survey. The different DoS attacks and defenses in 802.11 protocol based wireless networks are explored physical,network layers and MAC. Some of the major attack names keywords are Denial of Service (DoS), Distributed Denial of Service (DDoS), Internet Security, Wireless Security, Scanner, Spoofing. DOS attack history and incidents DoS attacks started at around early '90s. At the first stage they were quite "primitive", involving only one attacker exploiting maximum bandwidth from the victim, denying others the ability to be served[1]. This was done by using different flooding attack methods.

3. RELATED WORK

Security of Information has always been crucial for the sustenance of future development since early days. Earlier information used to be gathered manually and proper means of preserving this information were not available. With the advancement of technology, there has been vast growth in the ways of preserving the information but simultaneously security is becoming a major concern due to the various security threats. These information security issues may arise in a desktop computer, office environment, on a network or

in a cloud. The literature review shows that data mining is key ingredient in the solution to information security problems. The author in [1] discusses the development of data mining and its application areas. Soft computing framework data mining is presented in paper [2] where soft computing approaches like fuzzy logic, neural network are discussed. Data mining provides a number of algorithms that can help detect and avoid security attacks [3].The author in [4] presents a survey on various data mining techniques for intrusion detection wherein the types of intrusion attacks like network and host based are also summarized. One of the intrusion detection technique known as anomaly detection has been discussed in detail [5]. Paper [6] specifies the measurement criterions for intrusion detection. Fraud detection is another area of focus as the number of online transactions is rising exponentially. Various types of frauds like computer fraud are given in [7] with the respective techniques to overcome the situation. A number of methods are proposed for privacy preserving through data mining in [8], for example Anonymity. In paper [9], author talks about the sensitivity of data which may risk an individual's privacy. This data can be general data, user specific or authentication data. PETRE in [10] specifies aspects of cloud computing and the top cloud computing companies with their respective key features. The cloud security issues have been addressed via a trusted third party in [11]. Data mining techniques can also be used for the analysis of various firewall policy rules [12]. Security framework for mobile cloud computing is proposed in [13]. In [14], the authors have identified the following types of attacks which are a major threat to cloud implementation denial of service attack, Cross virtual machine side-channel attack, malicious insiders' attack, Attacks targeting shared memory, and Phishing attack. Table 1 briefs the review of variety of work done in the area cloud computing security with the help of data mining techniques. Paper [15] details the need of mobile cloud computing. As the mobiles are getting cheaper with the availability of internet facility, a mobile can also be considered as an entity in a cloud.

4. DDOS ATTACKS IN CLOUD ENVIRONMENTS

The denial of service attacks happens in cloud computing infrastructure. This might occur obviously or for a while it may occur due to botnets or professional attacker. This attack may be prepared for several reasons. This may happen when the service request to cloud environments [6]. Distributed denial of attack has several types such as UDP flood, ICMP flood, SYN flood, Ping of Death, slowloris, HTTP flood attacks. UDP Flood Attack: This type of denial of service attack happens in User Data Gram protocol.It establish a session less connection by user datagram protocol. It enters into any one of the port in host computer with one or more numerous UDP packets. This session-less service roots the port that will require to confirm the port whether the packets will be reached or nor [5] [6]. ICMP Flood Attacks: Internet Control Message Protocol Attacks. Normally ping request packets are send to destination host to check whether the host is connected are not. This is identified by ping replies. Attackers send more number of packets without waiting for replies. SYN Flood Attacks: In general TCP which follows a three way

Handshaking. The client send synchronization request to destination host server. The servers respond to the clients by sending synchronization acknowledgements. Then the client will launch the Synchronization Acknowledgement. This attacks will occur by the client will send one or more SYN request to single or extra numbers of host destination server and does not respond to the Acknowledgment request. The host servers system that continuously was waiting for a reply. This will cause a Synchronization Flood DDoS attacks. Ping of Death Attacks: The attacker approved away this attack by throw a malware attacking ping request packets to one or more computer. HTTP flood Attacks: The attacker carried out these attacks by GET and POST methods. It tries to complete a resource request with greatest number of resources. These will fallout in HTTP Flood DDoS attacks.

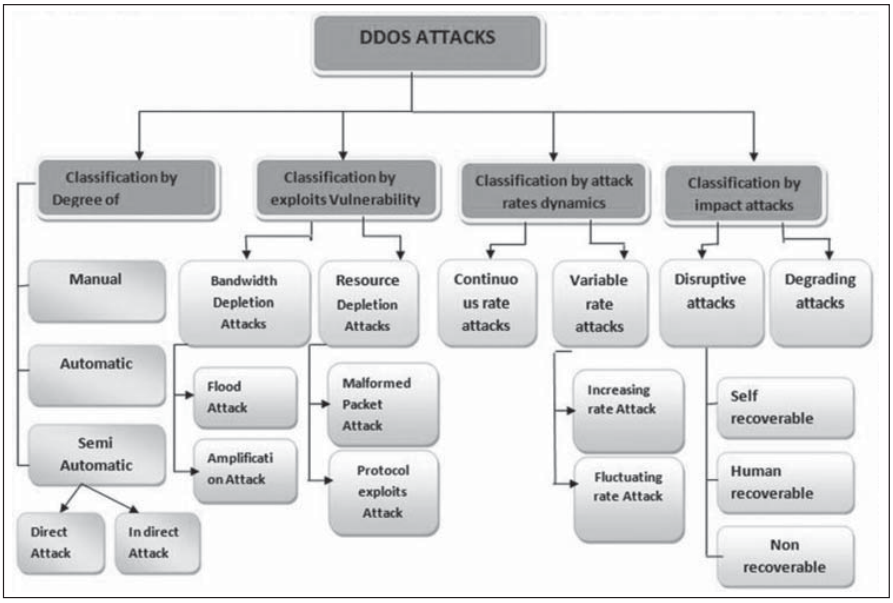

Figure 1. Classification of DDoS Attacks

Table: 1 Data Mining Techniques for Information Security in Cloud

Area	*Types*	*Detection*
Intrusion Detection	Network Based Host Based	Anomaly ID Misuse ID Data mining Based
Fraud Detection	Management Fraud Customer Fraud Network Fraud Computer Based Fraud	**Avoidance** Data fusion based Immunological Approach based Outlier detection Self Organizing Maps
Privacy Preserving	Data Privacy User Privacy	K-Anonymity (Identity disclosure) Perturbation Approach Cryptography Randomized Response Condensation Approach
Detecting Information Leakage	Buffer Overflow attack Data Mining	Brute Force method Exploratory data analysis **Avoidance** Legitimacy tags External Leakage
Firewall	Basic Distributed Network	Anomaly Detection Generalization Association rule mining Frequency based technique
Data Security Enhancement	Multi-level Security model Encryption-Blind signatures Biometric encryption Anonymous	K-Anonymity (Identity disclosure) Perturbation Approach Cryptography

5. EFFECTIVE PARAMETERS IN DDOS DETECTION

1. **Real-time (R) or non-real time (N):** A defense mechanism with real-time detection enjoys a good performance in high speed traffics. Offline methods in high speed traffics face problems due to the generated overhead by delay in processing. It causes failure or lower speed in **detection.**
2. **Scalability**: A scalable defense mechanism can effectively handle its attack detection and response duties even if both the number of attackers and the amount of attack traffic increases.
3. **Unknown attacks detection**: New attacks detection is challenging for defense systems. The observed methods are not capable of detecting the unknown attacks.
4. **Defense strength**: The strength of a defense mechanism can be measured by various metrics depending on how well it can prevent, detect, and stop the attacks.

   These metrics could be defined based on the decision or prediction that each defense mechanism makes [6, 7].

   (a) **Accuracy:** Ratio of the correct outcomes of the defense mechanism (true positives and true negatives) over the total outcomes of the defense mechanism. ((TP+TN))/ ((P+N))

   (b) **Sensitivity** or true positive rate: Ratio of true positives over total desired positive outcomes. TP/ ((FN+TP))

   (c) **Specificity** or true negative rate: Ratio of true negatives over total desired negative outcomes. TN/ ((TN+FP))

   (d) **Precision** or positive predictive value (PPV): Ratio of true positives over the total positive outcomes of the defense mechanism. TP/ ((FP+TP))

   (e) **Reliability** or False positive rate: Ratio of false positive outcomes of the defense mechanism over total positive outcomes of the defense mechanism. FP/ ((FP+TN))

   (f) **False negative rate:** Ratio of false negative outcomes of the defense mechanism over total negative outcomes of the defense mechanism. FN/ ((TP+FN))
5. **Request response time**: It refers to the average response time of each successful HTTP. The response time will increase with the increase of attack rate since those bad HTTP requests also consume the processing capacity of (DDoS defense system) nodes. When the bad requests are filtered at network-layer, the average response time will decrease dramatically [7, 8].
6. **Availability:** The signs of DDoS attack can be abnormal consumption of server resources such as memory and bandwidth that can be caused lack of access.
7. **Request dropping probability**: Low level of request dropping probability is more appropriate
8. **Throughput:** It stands for the client's request per second or the average end-to-end throughput of a legitimate client who sends one request per second to download a file of 100 Kbytes. The client of directly accessed base server will suffer from high request dropping probability and large response time.
9. **Delay in detection/response**

   (a) One-way delay [9]

   (b) Request-response delay [9,10]

   (c) Delay variation (Jitter) [9,10]
10. **System performance degradation**: Defense mechanism causes system performance degradation such as memory storage and lack of CPU cycles.

5.1 TYPE OF DATASETS

All methods must be tested and analyzed, so the best way is to use dataset. Table 2 depicts type of datasets that to be using for testing.

1. **Benchmark Datasets:** Only a few benchmark intrusion datasets are publicly available but they are not for DDoS attacks. KDDcup99 intrusion data set, DARPA Intrusion Detection Data Sets
2. **Simulated Datasets:** Simulate the environment using available tools. ns2, Qualnet, OMNeT++, CloudSim
3. **Private Datasets:** The best approach for testing any intrusion detection system or DDoS attack detection method is to create a real network test bed with a large number of host and network components.

6. A TAXONOMY OF DEFENSE MECHANISMS AGAINST DDOS FLOODING ATTACKS.

This section elaborate the taxonomy of defense mechanisms against DDoS flooding attacks, it consist of two types of DDoS flooding attacks against defense mechanisms, it can be classify in to two broad categories one is based on the deployment location of Defense mechanism where the defense mechanism is implemented in location based. Another one is the point in time that defense take place, We classify the Deployment location of defense mechanism in to further two subtypes one is defense mechanisms against network/transport-level and defense mechanisms against application level, In defense mechanisms against network/transport-level is implemented in centralized and hybrid defense mechanism, the centralized Defense mechanism against DDoS flooding attacks classified into four categories: source-based, destination-based, network-based, and hybrid and the defense mechanisms against application-level DDoS flooding attacks into two categories: destination-based, and hybrid based on their deployment location [11]. The second classification defense mechanisms against DDoS flooding attacks is the point of time when the DDoS defense mechanisms should act in response to a possible DDoS flooding attack, we classify DDoS flooding attacks into three categories three points of defense against the flooding attack that is before the attack (attack prevention), during the attack (attack detection), and after the attack (attack source identification and response) [12].

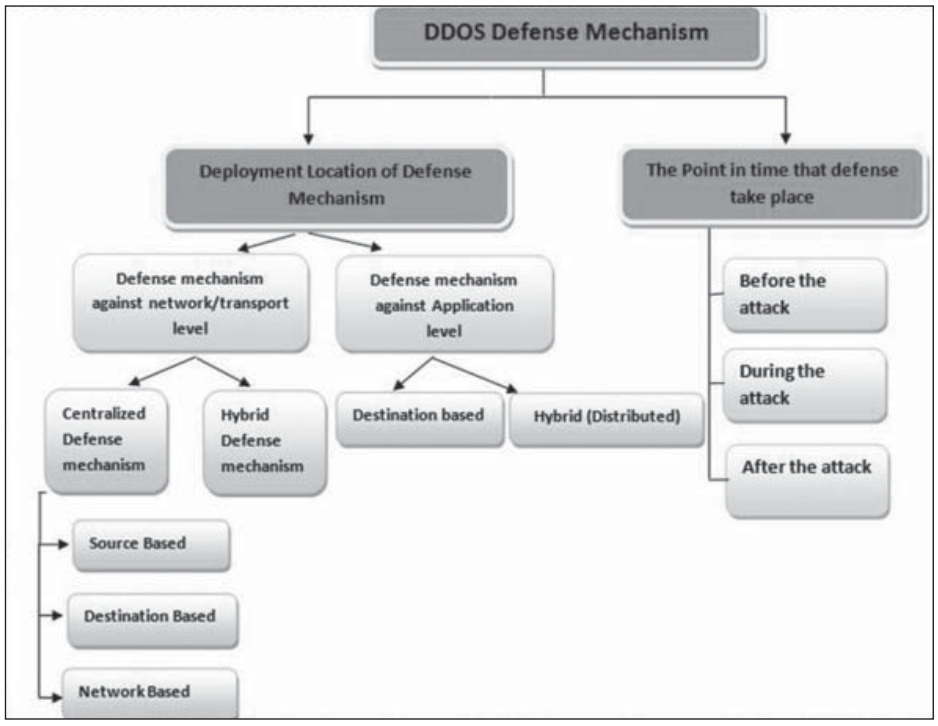

Figure 2. Different Types of DDoS Defense mechanisms

Table 2.
Evaluation of some defense mechanisms against DDoS flooding attacks in the cloud

Techniques	*Advantage*	*Disadvantage*	*Effective Detection Parameters*			
			Accuracy	*Scalability*	*System Performance*	*Implementation Complexity*
Cooperative IDS	Increasing confidence in Proportion to an ordinary IDS	Consuming more computing time in proportion to an ordinary IDS	High	Medium	Low	low
Cloud Trace back Model	Overcoming direct DDoS Attacks, Identifying the attacker in a successful attack	Collecting the dataset is difficult for the neural net	Low	Medium	Low	High
Confidence based Filtering	Low storage capacity, High speed filtering attack Packets, Reducing the overhead of the server	The accuracy of the model is less than the other models	Medium	Medium	High	Low
CLASSIE	Reducing false positive rates of attacks, Reducing the overhead of the server	Detecting the attacks at application level	Medium	Low	High	Medium
Filtering Tree	Filtering the attacks at different levels	Detecting the attacks at application level	Low	Medium	Low	High
Information theory based metrics	Easy deployment and decreases of negative rate	Probability of information loss due to entropy compression	Medium	Medium	High	Low

7. CONCLUSION AND FUTURE WORK

This paper presents a comprehensive review of DDoS attacks, detection techniques and tools used in Cloud environment. This review paper aims to describe the DDoS and its classifications, and identify the effective parameters in detection of DDoS attacks in cloud. DDoS attack is an attempt to make a machine or network resources unavailable to legitimate user. In result of DDoS attack, network consumption leads to cost, delay and interruption in communication between various legal network users. Data Mining techniques provide very efficient way for discovering useful knowledge from various source information. This paper is based on survey of Data Mining techniques in DDoS attack detection and focuses on various researches in the form of method, algorithm and protocol. Furthermore, this paper figure out the overall possibilities for determines the DDoS attack using Data Mining technique. This Review paper provides opportunities for developing an advanced detection algorithm. It can improve detection rate resulting from existing work. It can analyze algorithm by using different types of DDoS attacks and data set.

References

1. Mihui Kim, Hyunjung Na, Kijoon Chae, Hyochan Bang, and Jungchan Na (2004). A Combined Data Mining Approachfor DDoS Attack Detection. ICOIN 2004, LNCS 3090, cSpringer-Verlag Berlin Heidelberg. pp. 943–950.
2. Lin, S. C., & Tseng, S. S. (2004). Constructing detectionknowledge for DDoS intrusion tolerance. Expert Systemswith Applications, 27. pp 379–390.

3. Yoohwan Kim, Wing Cheong Lau, Mooi Choo Chuah AndJonathan H. Chao (2004). Packetscore: Statistical-Based Overload Control Against Distributed Denial -Of-ServiceAttacks. IEEE INFOCOM, The 23rd Annual JointConference of the IEEE Computer and Communications Societies, Hong Kong, China.

4. Keunsoo Lee, Juhyun Kim, Ki Hoon Kwon, Younggoo Han,Sehun Kim (2008). DDoS attack detection method using Cluster analysis. Expert Systems with Applications 34. pp1659–1665.

5. P.Sundari, Dr.K.Thangadurai (2010). An Empirical Study on Data Mining Applications. Global Journal of Computer Science and Technology,Vol.10 Issue 5 Ver. 1.0. pp 23-27.

6. G. Sager (1998).Security Fun with Ocxmon and Cflowd.presented at the Internet 2 Working Group.

7. R. Stone (2000). CenterTrack: An IP overlay network fortracking DoS floods. in Proc. USENIX Security Symp. pp.199–212.

8. K.C.Nalavade, and B.B.Meshram (2010). Identifying the Attack Source by IP Traceback. Springer, ICT 2010, CCIS 101. pp. 292-296.

9. Kanwal Garg, Rshma Chawla (2011). Detection of DDoS attacks using Data Mining, International Journal of Computing and Business Research (IJCBR). Pp. 2229-6166.

10. Jignesh Vania, Arvind Meniya and Harikrishna Jethva (2013). Association Rule Based Data Mining Approach to HTTP Botnet Detection. IJAIEM, Volume 2, Issue 4, ISSN.pp 2319 –4847.

11. Rui Zhong, and Guangxue Yue (2010). DDoS Detection System Based on Data Mining. ISBN 978-952-5726-09-1(Print) Proceedings of the Second International Symposium on Networking and Network Security (ISNNS ’10)Jinggangshan, P. R. China.pp 062 -65

12. Shaveta Gupta, Dinesh Grover and Abhinav Bhandari(2014). Detection Techniques against DDoS Attacks:A Comprehensive Review, International Journal of Computer Applications, Volume 96–No.5. pp 0975-8887.