International Journal of Computing and Business Research (IJCBR)

ISSN (Online) : 2229-6166

Volume 2 Issue 1 2011

DETECTION OF DDOS ATTACKS USING DATA MINING

Kanwal Garg [1] , Rshma Chawla [2]

[1] Assoc.Prof., M.M. Institute of Computer Technology & Business Management,
M. M. University, Mullana- Ambala.
Email id: gargkanwal@yahoo.com

[2] Lecturer, M.M. Institute of Computer Technology & Business Management,
M. M. University, Mullana- Ambala.
Email id: rshmabedi@gmail.com

ABSTRACT

Distributed Denial of Service (DDoS) attacks are large-scale cooperative attacks launched from a large number of compromised hosts called Zombies is a major threat to Internet services. Popular web sites such as Yahoo, CNN, and Amazon, are among the well-known victims of DDoS attacks. Large number of companies transacting online are mainly facing the considerable loss as they are being targeted to DDoS attacks. Therefore, keeping this problem in view author presents various significant areas where data mining techniques seem to be a strong candidate for detecting and preventing DDoS attack.

KEY WORD: Distributed Denial of Service attack, Data mining, Zombies.

1. Introduction

Today, the number of attacks against large computer systems or networks is growing at a rapid pace. One of the major threats to cyber security is Distributed Denial-of-Service (DDoS) attack .In which the victim network element(s) are bombarded with high volume of fictitious attacking packets originated from a large number of Zombies. The aim of the attack is to overload the victim and render it incapable of performing normal transactions [Kim.et.al. 2004]. To protect network servers, network routers and client hosts from becoming the handlers, Zombies and victims of distributed denial-of-service (DDoS) attacks data mining approach can be adopted as a sure shot weapon to these attacks.

The recent rapid development in data mining has made available wide variety of algorithms, drawn from the fields of statistics, pattern recognition, machine learning, and database. These algorithms made it possible to achieve the ultimate aim of writing this paper. The central theme of this paper is to explore areas where data mining techniques extensively gathers the audited data to compute patterns which predict the actual behavior that can be used for detecting or tracing various DDoS attack.

This paper has been divided into five sections. Section 1 defines the overview of problem .Section 2 highlights DDoS Attacks. Section 3 portrays a basic idea of data mining. Section 4 highlights some application areas where data mining protect over resources against DDoS attacks. Section 5 finally concludes by discussing the outcome of study.

2. DDoS Attack

Distributed Denial-of-Service (DDoS) attack is the one in which the victim's network elements are bombarded with high volume of fictitious attacking packets that originate from a large number of machines [Kim et.al., 2004]. A successful attack allows the attacker to gain access to the victim's machine, allowing stealing of sensitive internal data and possibly cause disruption and denial of service (DoS) in some cases.

The number of DDoS attacks grew 20 % last year - a major decrease in the rate of attacks from 2007 to 2008, when these devastating attacks increased 67 percent, according to a report.[1] According to a report Internet Service Providers (ISPs) are most worried about botnet-driven distributed denial-of-service (DDoS) attacks[2].DDoS attacks launched by the group Anonymous took down the Web sites of U.K. record label Ministry of Sound and its legal firm Gallant Macmillian on 3rd Oct, 2010[3] contributes some latest DDoS attacks

Out of the various categories of DoS attacks such as flooding, software exploit, protocol based etc Distributed Denial of service attack is the most prominent. In fact, DDoS attack uses series of Zombies to initiate a flood attack against an unsafe single site. DDoS attack is initiated in 2-phases [Mirkovic and Reiher 2004] [Dietrich et al. 2000] i.e. Recruiting phase and Action phase.

In Recruiting phase attacker initiates the attack from the master computer and tries to find some slave (Zombies) computers to be involved in the attack. A small piece of software is installed on the Zombies to run the attacker commands. The Action phase continued through a command issued from the attacker resides on the master computer toward the Zombies computers to run their pieces of software. The mission of the piece of software is to send dummy traffic designated toward the victim. The result is a massive flood of packets that crashes the host or swamp down the entire network operations shown in Figure 2.1. Very few networks or hosts can effectively cope with such a scale of attacks today. Most of the handler and Zombie are completely unaware of the fact that they were being used for launching of a DDoS attack.

Numbers of mechanisms are given to either defend or prevent against DDOS attacks such as starting from increasing the resourced at defender side, implementing authentication policies at routers, filters, firewalls with hardware security appliances, Learning based mechanisms, agents based detection at host level or at immediate level etc but none of them has proved to be the best, addressing all the challenges and still there exist a gap amongst the security requirements &

[1] http://www.darkreading.com/security-services/167801101/security/perimeter-security/222301511/index.html
[2] http://www.csoonline.com/article/518514/report-isps-fear-many-more-ddos-attacks-in-2010
[3] http://news.cnet.com/8301-1009_3-20018427-83.html

existing mechanisms. Therefore, a mechanism that is strong and reliable is desired. Hence the key idea is to use data mining techniques to discover consistent and useful patterns of system features that describe program and user behavior of attack.

3. Data Mining

Data mining is becoming a persistent technology in activities as diverse as using historical data to predict the success of a marketing campaign, looking for patterns in network traffic to discover illegal activities or analyzing sequences [Sundari and Thangadura, 2010]. From this outlook, the approach is gaining importance in the field of DDoS attacks.

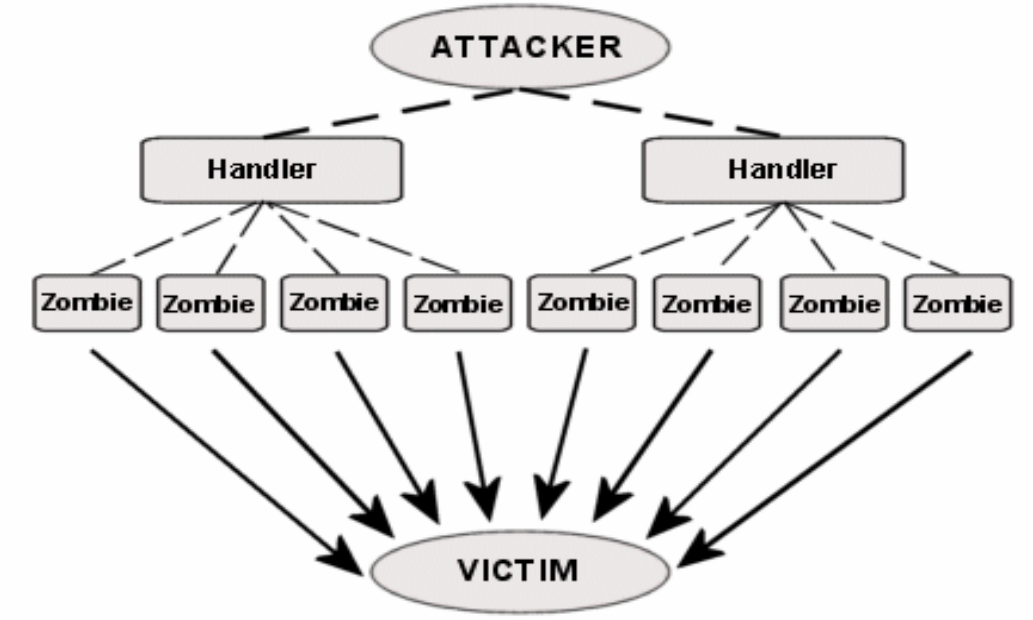

Figure 2.1: Architecture of DDoS attack[4]

Data mining is, at its core, pattern finding. Data miners are proficient at using specialized software to find regularity (and irregularities) in large & complex data sets. Data mining

[4] Source: http://www.fnokd.com/wp-content/uploads/2007/08/ddos_attack.gif

applications are computer software programs or packages that enable the extraction and identification of patterns from stored data.[5] A data mining application is typically a software interface which interacts with a large database containing Network traffic parameters or other important data. Data mining is widely used by companies and public bodies for marketing, detection of fraudulent activities such as DDoS attacks.

4. Various Application areas of Data mining in DDoS attacks

Recently, data mining has become an important component for DDoS attack prevention. Different data mining approaches like classification, association rule, clustering, and outlier detection are the few techniques frequently used to analyze network traffic or data to gain knowledge that helps in controlling intrusion. Various applications where data mining approach can be used in prevention and detection of DDoS attacks are discuss below:

4.1 Intrusion Detection

Intrusion detection is the process of observing the events occurring in a computer system or network and analyzing them for instances which violates related security policies or practices. Intrusion detection techniques can be classified as misuse detection and anomaly detection. Misuse detection systems, e.g., IDIOT [Kumar and Spafford, 1995] and STAT [Ilgun et.al., 1995], use patterns of well-known attacks or weak spots of the system to match and identify known intrusions. Anomaly detection systems, e.g., IDES [Lunt et. al., 1992] flag observed activities that deviate significantly from the established normal usage profiles as anomalies, i.e., possible intrusions. Today the main reason of using Data Mining for intrusion detection systems is the enormous volume of existing and newly appearing network data that requires processing. Literature also provides evidence where data mining techniques are used for intrusion detection.

[5] http://www.wisegeek.com/what-are-data-mining-applications.htm

Various applications where data mining approach can be used in Intrusion detection of DDoS attacks are discussed here:

An overview of real time data mining-based intrusion detection systems (IDSs) is presented by researcher that focused on problems related to deploying a data mining-based IDS in a real time environment also discussed a distributed architecture for estimating cost-sensitive models in real time. Adaptive learning algorithms are used to improve usability that facilitates model creation and incremental update. Unsupervised anomaly detection algorithms are used to reduce the reliance on labeled data. Author [Lee et. al., 2002] gives an architecture consisting of sensors, detectors, a data warehouse, and model generation components. Presented architecture facilitates the sharing and storage of audit data and the distribution of new or updated models which improves the efficiency and scalability of the IDS.

Another similar example of Intrusion detector has been countered by [Brahmi et.al. 2010] which explains that it is a novel distributed multi-agent IDS architecture, called MAD-IDS. MAD-IDS integrate the mobile agent methodology and the data mining techniques to accommodate the special requirements in distributing IDS. Author expressed that the data mining techniques and in particular the unsupervised clustering algorithm and the generic association rule mining are capable of discovering anomalous connections, as well as, generating an informative summarize. The last result of this architecture is to the point and intuitive detection rules that can be periodically supplied to the Misuse Detection Agent to update its signature database allowing the detection of known attacks.

Author described an experimental system, based on the Common Intrusion Detection Framework (CIDF), where multiple IDSs can exchange attack related information to detect distributed interruptions. Above system also comprises an ID model builder, where a data mining engine can receive audit data of a novel attack from an IDS, compute a new detection model, and then distribute it to other IDSs[Lee et.al. 2000].

[John et.al, 2007] builds a Fuzzy Intrusion Recognition Engine (FIRE), which is an anomaly-based intrusion detection system that uses fuzzy logic to evaluate whether malevolent activity is taking place on a network. The FIRE system applies basic data mining techniques to TCP packet data for extracting metrics that are not obvious in the raw data. These metrics are then evaluated as fuzzy sets.

DDoS attack detection model presented by [Zhong et.al. 2010] was based on data mining algorithm. FCM cluster algorithm and Apriori association algorithm used to extracts network traffic model and network packet protocol status model. Here threshold is set for detection model. There are many other IDSs & detection systems where data mining techniques were applied. Now we are moving towards other application area i.e. IP Traceback describes in next section.

4.2 IP Traceback

DDoS is rapidly growing problem. IP Traceback is the ability to trace IP packets from source to destination. This is a significant step towards identifying and thus stopping attackers. The IP Traceback is an important mechanism in defending against DDoS attacks. Lot of techniques and methodologies are used to trace the DDoS attacks. Some are given below:

An approach suggested by [Sager 1998] and [Stone 2000] is called ‘Logging’ that is to log packets at key routers and then use data mining techniques to determine the path that the packets traversed. This scheme has the functional property that it can trace an attack long after the attack has completed. However, it also has obvious drawbacks, including potentially enormous resource requirements (possibly addressed by sampling) and a large scale inter provider database integration problem.

The data mining techniques are providing very efficient way for discovering useful knowledge from the available information. [Nalavade and Meshram 2010] proposed a system which uses packet marking mechanisms along with Intrusion Prevention Systems for efficient IP traceback. Above author considers that data mining techniques can be applied to the data collected from the packet marking scheme for detecting attack and therefore, resultant database of knowledge can be further used by network Intrusion prevention systems for decision making.

5. Conclusion

DDoS attacks are quite complex methods of attacking a computer network, ISP, Individual system make it ineffectual to legitimate network users. These attacks are an aggravation at a minimum, and if they are against a particular system, they can be brutally destroying. Loss of network resources costs money, delays work, and interrupts communication between various legal network users. The drastic consequences of a DDoS attack make it important that strict and productive solutions and security measures must be made to prevent these types of attacks. Detecting, preventing, and mitigating DDoS attacks is important for national and individual security. This paper discussed various detection algorithms which are using data mining concepts & algorithms for DDoS detection & prevention. But with the improvement in technology new areas are emerging where data mining techniques can be utilized for handling DDoS attacks that are to be discuss in future.

References

[1] T. Lunt, A. Tamaru, F. Gilham, R. Jagannathan, P. Neumann, H. Javitz, A. Valdes, and T. Garvey. “A real-time intrusion detection expert system (IDES)” - final technical report. Technical report, Computer Science Laboratory, SRI International, Menlo Park, California, February 1992.

[2] Yoohwan Kim, Wing Cheong Lau, Mooi Choo Chuah And Jonathan H. Chao "Packetscore: Statistical-Based Overload Control Against Distributed Denial-Of-Service Attacks" IEEE INFOCOM 2004 ,The 23rd Annual Joint Conference of the IEEE Computer and Communications Societies, Hong Kong, China, March 7-11, 2004. IEEE, 2004.

[3] K. Ilgun, R. A. Kemmerer, and P. A. Porras. "State transition analysis: A rulebased intrusion detection approach". IEEE Transactions on Software Engineering, 21(3):181–199, March 1995

[4] Wenke Lee, Rahul A. Nimbalkar, Kam K. Yee, Sunil B. Patil, Pragneshkumar H. Desai, Thuan T. Tran, and Salvatore J. Stolfo **"**A Data Mining and CIDF Based Approach for Detecting Novel and Distributed Intrusions" RAID 2000, LNCS 1907, pp. 49–65, 2000. c_Springer-Verlag Berlin Heidelberg 2000

[5] John E. Dickerson, and Julie A. Dickerson "Fuzzy Network Profiling for Intrusion Detection" Electrical and Computer Engineering Department Iowa State University Ames, Iowa, 50011

[6] S. Kumar and E. H. Spafford. "A software architecture to support misuse intrusion detection". In Proceedings of the 18th National Information Security Conference,pages 194–204, 1995.

[7] P. Chatterjee, D. Joffman, and T. Novak. "Modeling the clickstream: Implications for Web-based advertising efforts," Marketing Science. 22, pp. 520-541 (2003).

[8] Rui Zhong, and Guangxue Yue "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, 2-4, April.2010, pp. 062-065

[9] G. Sager, "Security Fun with OCxmon and cflowd," presented at the Internet 2 Working Group, Nov. 1998.

[10] R. Stone, "CenterTrack: An IP overlay network for tracking DoS floods," in Proc. 2000 USENIX Security Symp., pp.199–212, July 2000.

[11] J. Mirkovic and P. Reiher, "A Taxonomy of DDoS Attack and DDoS Defense Mechanisms," ACM SIGCOMM Computer Communications Review, Volume 34, Number 2, April 2004, pp. 39-53.

[12] Dietrich, S., Long, N., and Dittrich, D. 2000. Analyzing distributed denial of service attack tools: The shaft case. In Proceedings of 14th Systems Administration Conference. New Orleans, Louisiana, USA, 329-339.

[13] Wenke Lee, Salvatore J. Stolfo , Philip K. Chan Eleazar Eskin , Wei Fan , Matthew Miller , Shlomo Hershkop and Junxin Zhang "Real Time Data Mining-based Intrusion Detection"2002,ieexlore.ieee.org.

[14] Imen Brahmi, Sadok Ben Yahia, and Pascal Poncelet " MAD-IDS: Novel Intrusion Detection System using Mobile Agents and Data Mining Approaches"In intelligent & security informatics,2010-springer.

[15] P.Sundari, Dr.K.Thangadurai " An Empirical Study on Data Mining Applications" Global Journal of Computer Science and Technology, Vol. 10 Issue 5 Ver. 1.0 pp23-27 July 2010