ISSN: 2229-6956(ONLINE)
DOI: 10.21917/ijsc.2014.0112

COLLABORATIVE NETWORK SECURITY MANAGEMENT SYSTEM BASED ON ASSOCIATION MINING RULE

Nisha Mariam Varughese
Department of Computer Science and Engineering, Mount Zion College of Engineering, India
E-mail: nishamecse@gmail.com

Abstract

Security is one of the major challenges in open network. There are so many types of attacks which follow fixed patterns or frequently change their patterns. It is difficult to find the malicious attack which does not have any fixed patterns. The Distributed Denial of Service (DDoS) attacks like Botnets are used to slow down the system performance. To address such problems Collaborative Network Security Management System (CNSMS) is proposed along with the association mining rule. CNSMS system is consists of collaborative Unified Threat Management (UTM), cloud based security centre and traffic prober. The traffic prober captures the internet traffic and given to the collaborative UTM. Traffic is analysed by the Collaborative UTM, to determine whether it contains any malicious attack or not. If any security event occurs, it will reports to the cloud based security centre. The security centre generates security rules based on association mining rule and distributes to the network. The cloud based security centre is used to store the huge amount of tragic, their logs and the security rule generated. The feedback is evaluated and the invalid rules are eliminated to improve the system efficiency.

Keywords:

Collaborative UTM, Traffic Prober, Association Mining Rule, Anti-Phishing and Anti-Botnet

1. INTRODUCTION

Due to the increasing size of the network, the internet security has become one of the major challenges in current world. Attacks which have fixed patterns can be easily determined by matching the malicious attacks to known threats. But the pattern identification of polymorphic attacks and DDoS attacks are difficult. Botnet consists of a set of bots which is malicious software that can intrude on the computer. Botmaster infects victims with bot, then this bot connects to control and command server. Botnet also contains botmaster who sends commands to bot via the Command & Control server. This attacking process is repeated and soon the botmaster has an army of bots. The botmaster controls the whole army from a single point. The command and control(C & C) channel is backbone of botnet and it is responsible for setting up the botnet network, controlling bot activities, generating the commands, and finally achieving the goals. The life time of Botnet C & C exists until it gets detected. The existence of Command-and-Control (C & C) infrastructure is the main difference between Botnet and other kind of malwares. Botmaster communicates to the whole bot army through the C & C channel which passes the stolen information from infected machines to their master. [1] The victim is prevented from the botnet attack, by suppressing the C & C channel. For botnet suppression the cloud based security center will generate security rules and these rules are distributed into the networks.

Phishing is a kind of attack in which an attacker called phisher attempts to stolen the authorized client's confidential or sensitive information by mimicking the communication. For this the phisher initially sends an email which leads the victim to a dishonest webpage. When the victim sends the reply message for the email, the phisher forces the victim to provide their confidential data. Thus the phisher can use this collected confidential data of the victim for any unauthorized activities or money transfer. Initially, for preventing the victim from the phishing attack the collaborative network security management system will collect the internet traffic with the help of the traffic prober and evaluates the traffic. [4] If any security event is detected, the collaborative UTM will invoke the security centre to generate the security rule using the association mining rule. The generated rules are distributed to the network and the feedback is analysed for better performance.

CNSMS is used for the forensic analysis of huge open network traffic. Through the feedback evaluation and rules regeneration, the system will updated frequently. At any instant this will help to provide protection for the sensitive data. Thus the system becomes more efficient and reliable.

2. COLLABORATIVE NETWORK SECURITY MANAGEMENT SYSTEM

Collaborative Network Security Management is a system which provides a new collaborative system which integrates the UTM such as NetSecu. [2] A hierarchical architecture with three levels is implemented in CNSMS. In the first level of the hierarchy, the rule set library is stored in the central management system to have a big picture of the whole network. Domain NetSecu nodes is represented in the second level to manage the membership in corresponding sub-domains. Finally the third level of the hierarchy represents the basic NetSecu nodes.

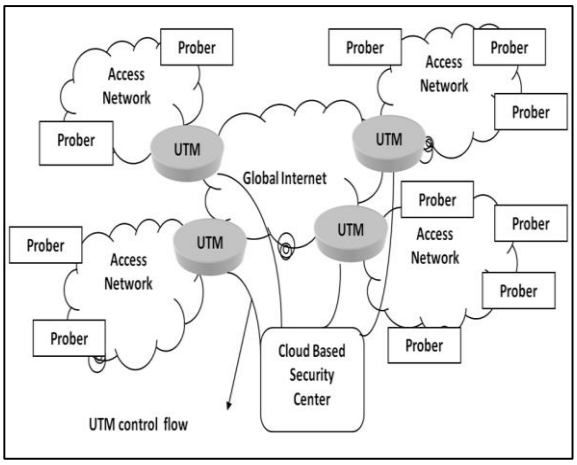

Fig.1. Work Principle of CNSMS in Cloud Computing

The Fig.1 shows the working principle of the collaborative network security management system. The traffic data is captured by the traffic prober and collaborative UTM and it is given to the security center. To exchange network events and security rules a peer-to-peer communication protocol is used in the collaborative UTM. During the systems operation, the CNSMS runs as defined and share the rulesets to the network. As instructed by the security center the new rulesets are distributed on demand. Operating reports are sends back to the cloud based security centre which are collected from each NetSecu node and Prober.

2.1 COLLABORATIVE UTM

Collaborative Network Security Management is a system which provides a new collaborative system which integrates the Unified Threat Management (UTM) such as NetSecu. NetSecu is a network security device which can be used as an attack detector and traffic controller. It is the building block of collaborative network security system. It will visualize the security event and react to resist attacks in a consistent and unified way. [5]

Traffic prober in the NetSecu node records the raw network traffic and the traffic controller controls the internet traffic based on the QoS requirements. The NetSecu node also contains Collaborator Element to manage other security elements based on Security Centre's command, for collecting logs and threat events it contains Reporting Element and generates human-readable report and for local maintenance a local manager also in the NetSecu node. [6] The main two tasks of the collaborative UTM are: Internal Protection and Server Protection.

In internal protection the UTM filtered and monitored the network flows between internet and enterprise network by bringing the advantage of its ability of firewalling and intrusion detection. However, in server protection, to provide protection for the resources, the UTM should disallow any access to server network from outside and restrict access to the network from unauthorised internal users.

2.2 SECURITY CENTER

The major function of the security center is the evaluation of the collected traffic and security rule generation. Large amount of traffic data from different origins are stored in the security center and forensic analysis is carried out for generating security rule. The close loop control is one of the important features of this system. In this system, the invalid rules are removed to make the system more efficient and reliable and the rules are redistributed. The rules are generated according to the association mining rule. The frequently occurred attacks are evaluated from the logs and rules are generated based on the support and confidence values.

2.3 TRAFFIC PROBER

Traffic Prober is another important functional part of the CNSMS which is used for the frequent traffic collection. Traffic prober has the ability to focus on particular traffic occasioned by certain events whenever necessary. It can concentrate on specific traffic which is incurred by the security centre.

3. ASSOCIATION MINING RULE

Association mining rule is one of the important techniques in the data mining which is used to extract the frequent patterns and their correlations from the items stored in the data repositories. The association rules are used in several areas like inventory control and telecommunication networks. Association mining rule derives the association mining rule based on the predefined minimum support and confidence value stored in the database. [7] The problem is divided into two sub problems. Initially, the first problem finds the itemsets whose occurrence exceeds the predefined support and confidence value stored in the database. Those itemsets are called frequent itemsets. The second problem will generate the association mining rule for the frequent itemsets based on the constraints of minimum support and confidence values. Consider the large itemsets L_k, $L_k = \{I_1, I_2 \ldots I_k\}$, association rule for this itemsets are generated as $\{I_1, I_2 \ldots I_{k-1}\} \Rightarrow \{I_k\}$, the rest of the rules are generated by removing the last items in the antecedent and inserting it in the consequent, then the confidence and support thresholds are re-evaluated.

The first sub-problem can be again sub-divided into two sub-problems: candidate itemsets generation process and frequent itemsets generation process. The itemsets whose support threshold exceeds the minimum value are called frequent itemsets and those itemsets which are expected to be frequent are called candidate itemsets. In several cases the association mining algorithms are used to generate large number of association rules. In the case of large complex association rules, it is impossible to validate the rules; thereby limiting the number of association rule will speed up the validation of rules. Several algorithms are proposed for limiting the number of rules. Apriori algorithm is one of the well known algorithms for association rule.

Several strategies have been proposed to reduce the number of association rules, such as generating only "interesting" rules, generating only "non-redundant" rules, or generating only those rules satisfying certain other criteria such as coverage, leverage, lift or strength. The most well-known algorithm for producing association rules is Apriori algorithm.

Let $I = I_1, I_2 \ldots\ldots, I_m$ be a set of m distinct attributes, T be transaction, D be a database with several transaction records T_s. Association rule is represented as implication in the form of $X \Rightarrow Y$, where $X, Y \subset I$ and $X \cap Y = \phi$, where X represents the antecedent and Y indicates the consequent. The important basic measures of association rules are support(s) and confidence(c). [8] The users are always bothered about the frequently accessed items, thus the minimal support and minimal confidence are predefined. The additional constraints for association mining rules are also can be specified by the users.

Support is defined as the fraction of records that contain $X \cup Y$ to the total number of records. During the scanning process, the count for each item is incremented by one whenever the item is encountered in different transactions. That means the support count does not consider the quantity of the items. For example, if a customer buys three bottles of beer, the support count is increased by one. It means that the transaction contains an item then the support count of that item is increased by one. Support is computed by the following formula:

$$Support(XY) = \frac{Support\ count\ of\ XY}{Total\ number\ of\ transaction\ in\ D}$$

Suppose support value of an item is 0.2%, that means 0.2% of the transaction contain the accessed data item. The provider will not pay much attention to the items which are not bought so frequently, high support value is desired for the good association rules. [13] User can specify the minimum support as a particular threshold that means certain association rules that are generated from those itemsets whose supports exceed that threshold.

Confidence is defined as the fraction of the number of transactions that contain $X \cup Y$ with the total number of the records which contain X, thus the association rule generated will be $X \Rightarrow Y$. According to the association rules, confidence is a measure of strength, suppose the confidence of X, Y is 80% which means that 80% of the transaction contains will also contains that much of Y.

$$Confident(X|Y) = \frac{Support(XY)}{Support(X)}$$

Association mining rule is for generating association rules which satisfy the pre-defined minimum support and confidence value stored in the database. Association rule mining is used in the CNSMS to generate the security rules based on the support and confidence level. These rules are given to the network to prevent from the malicious attacks. Based on the feedback evaluated, when new attacks are detected in the network the rules are regenerated and distributed to the network.

4. SUPPRESSION OF BOTNET

Botnet is a collection of interconnected computers whose have been controlled by the malicious attacks. The group of affected computers are controlled by one or more set of attacker called Botmaster. Botmaster will send command to perform in the victim machine. [3] The botmaster will hide behind the C & C server and this C & C server is the backbone of the botnet attacks. The main functions of the C & C server are controlling the activities of the bots, botnet setup and achieving the final goal. The life time of Botnet C & C exists until it gets detected. The existence of Command-and-Control (C & C) infrastructure is the main difference between Botnet and other kind of malwares. Botmaster communicates to the whole bot army through the C & C channel which passes the stolen information from infected machines to their master. The victim is prevented from the botnet attack, by suppressing the C & C channel. For botnet suppression the cloud based security center will generate security rules and these rules are distributed into the networks.

To spread the botnet in an easy way, botmaster will keep the bots as smaller as possible, thus the suppression of botnet becomes more difficult. In order to hide and rescue the C & C server of the botnet can automatically change their server structure. For the distributed botnet suppression, the CNSMS analyze the automatically collected network traffic from the traffic prober. Then the collected traffic is processed by the cloud based security centre. When the collaborative UTM detects the botnet attack, the security centre generates the security rule for the suppression of the attack by using the association mining rule.

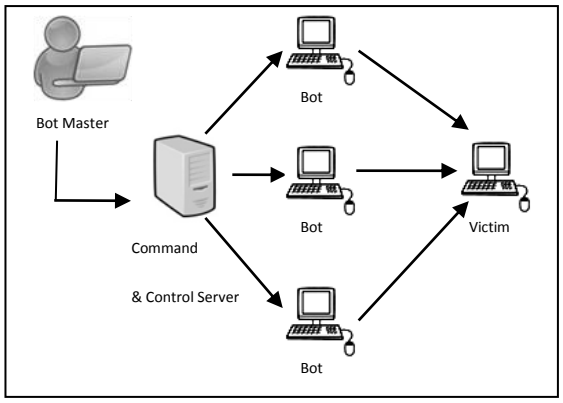

Fig.3. Botnet Structure

5. PHISHING ATTACK PREVENTION

Phishing is a type of online identity theft which is done to steal sensitive information such as credit card information, net banking passwords etc. A phishing attack has mainly three roles of phishers. Initially the mailer sends a large number of dishonest emails to direct the user to a dishonest webpage. [9 - 11] Secondly, the dishonest websites will prompt the users to provide their confidential data. Finally, the casher will use the confidential data for money transfer or for any unauthorized purpose. The information flow is shown in Fig.4.

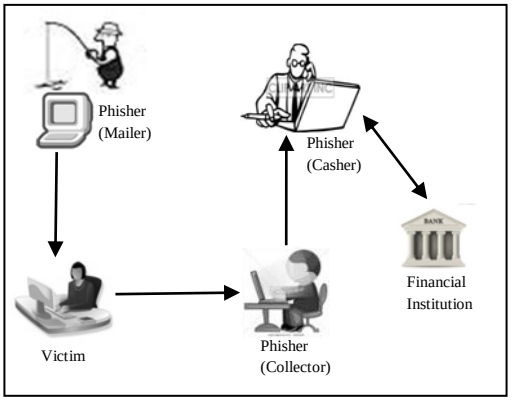

Fig.4. Information flow of Phishing attack

Cloud computing has the ability to provide offline phishing attack analysis. The traffic prober traces the internet traffic and when the collaborative UTM detects the botnet attack, the security centre generates the security rule for the suppression of the attack by using the association mining rule. The security centre will distributes the generated security rule to each node in the network.

6. CONCLUSION

The Collaborative Network Security Management is a system which aims to provide security in an open network by providing security rules to each node in the network. The security rules are generated by the association mining rule in the data mining concept. So in cloud computing platform association mining rules can used for providing security in the open network. This paper also deals with the forensic analysis of the phishing attack and the botnet attack suppression.

7. FUTURE WORK

CNSMS will analyse huge amount of traffic data within less time and executing the attack detection techniques for CNSMS simultaneously by cloud computing, attack detection and hence resolution becomes more efficient. Cloud storage can be used to store huge network traffic and in future parallel processing can also use for data classification. Thus cloud computing and parallel processing can incorporate for high speed vulnerability analysis. High speed classification of vulnerabilities using collaborative network security management can also be used for the detection of other network attacks in the future.

REFERENCES

[1] Jignesh Vania, Arvind Meniya and H. B. Jethva, "A Review on Botnet and Detection Technique", *International Journal of Computer Trends and Technology*, Vol. 4, No. 1, pp. 23-29, 2013.

[2] B. Mu, X. Chen and Z. Chen, "A collaborative network security management system in metropolitan area network", *Proceedings of the 3rd International Conference on Communications and Mobile Computing*, pp. 45-50, 2011.

[3] F. Han, Z. Chen, H. Xu and Y. Liang, "Garlic: A distributed botnets suppression system", *32nd International Conference on Distributed Computing Systems Workshops*, pp. 634-639, 2012.

[4] T. Li, F. Han, S. Ding and Z. Chen, "LARX: Largescale anti-phishing by retrospective data-exploring based on a cloud computing platform", *Proceedings of the 20th International Conference on Computer Communications and Networks*, pp. 1-5, 2011.

[5] Zhen Chen, Fuye Han, Junwei Cao, Xin Jiang and Shuo Chen, "Cloud Computing-Based Forensic Analysis for Collaborative Network Security Management System", *Tsinghua Science And Technology* Vol. 18, No. 1, pp. 40-50, 2013.

[6] X. Chen, B. Mu and Z. Chen, "Netsecu: A collaborative network security platform for in-network security", *3rd International Conference on Communication and Mobile Computing*, pp. 59-64, 2011.

[7] Qiankun Zhao and Sourav S. Bhowmick, "Association Rule Mining: A Survey", Technical Report, Nanyang Technological University, Singapore, No. 2003116, 2003.

[8] Bing Liu, Wynne Hsu and Yiming Ma, "Integrating Classification and Association Rule Mining", *KDD-98, American Association for Artificial Intelligence*, 1998.

[9] Radha Damodaram and M. L. Valarmathi, "Phishing Website Detection and Optimization Using Particle Swarm Optimization Technique", *International Journal of Computer Science and Security*, Vol. 5, No. 5, pp. 477-490, 2011.

[10] S. Sheng, B. Wardman, G. Warner, L. F. Cranor, J. Hang and C. Zhang, "An empirical analysis of phishing blacklists", *Sixth Conference on Email and Anti-Spam*, pp. 1-10, 2009.

[11] P. Knickerbocker, D. Yu and J. Li, "Humboldt: A distributed phishing disruption system", *IEEE eCrime Researchers Summit*, pp. 1-12, 2009.

[12] F. Deng, A. Luo, Y. Zhang, Z. Chen, X. Peng, X. Jiang and D. Peng, "TNC-UTM: A holistic solution to secure enterprise networks", *9th IEEE International Conference for Young Computer Scientists*, pp. 2240-2245, 2008.

[13] Sotiris Kotsiantis and Dimitris Kanellopoulos, "Association Rules Mining: A Recent Overview", *GESTS International Transactions on Computer Science and Engineering*, Vol. 32, No. 1, pp. 71-82, 2006.