Misuse Detection System Using Intelligent Agents for Online Transactions

Anuradha Sharma[1], Shama Parveen[2] and Puneet Misra[3]

Submitted in June, 2015; Accepted in January, 2016

Abstract-As E-Business is growing steadily with the increase in internet usage, it has increased the major security risk factors and as for novel attacks it becomes quite difficult to detect them. An Intrusion Detection System (IDS) is a device or software application that monitors network traffic and is able to analyse ongoing malicious activities and reports them to the administrator. Basically an IDS can be a Misuse Detection System or an Anomaly Detection System. There are several techniques for Misuse Detection System making use of Artificial Intelligence, Data mining etc. exploited by intelligent agents. This paper presents the model for Misuse Detection System exploiting data mining techniques with the help of intelligent agents which can be used by servers involved in E-Business. The agent makes use of pattern matching technique for the attack signatures already stored in the knowledge base and analyses audit log files. If it finds the one matching with those stored in knowledge base then the reporting agent immediately intimates the administrator about the malicious activity.

Index-Terms- Intrusion Detection System, Misuse detection, Multi-Agent system, intelligent agents, RETE pattern matching*.*

1.0 INTRODUCTION

Intrusion detection grew from the notion that computer misuse can be detected by analysing audit data in a computer system or network [1].Intrusion detection basically refers to the monitoring of network or system and the activities going on in them thus providing an important aid to the administrators about the various activities related to security. As networks are growing larger day by day, there must be some mechanism to detect intrusions. Organisations which are spreading their network across the world through internet that is E-Commerce need it the most. Nowadays, the intruders conceal their activities so as to extract maximum information and not only they are confined to information extraction but masquerading and misusing the extracted information for launching novel attacks on the target host for their benefit. Through intrusion detection it becomes possible to identify and recognise all such attacks and report them to the administrator.

An Intrusion Detection System (IDS)[15][16] is a combination of both hardware and software that makes it possible to detect malicious activities. It is able to monitor network traffic, analyse audit log files and match well known attack signatures already stored in knowledge base, etc. That is it gathers information from various known sources and is able to decide whether intrusion has taken place. The functional components of an IDS are- information source, analysis and response [2]. The most widely used IDS are Host Based IDS and Network Based IDS. Host Based IDS are able to detect intrusions on hosts and perform complete analysis on the host machines while the Network Based IDS are busy in sensing and monitoring network traffic that is intruders who attack on network traffic. The focus of this paper is on Host Based IDS.

Data Mining refers to a pattern finding technique in large data sets. Patterns related to normal behaviour and patterns related to malicious events can be computed using data mining [2]. The usage of various data mining techniques such as classification, clustering, association rule mining, outlier detection etc are being used for detection of intrusions. These techniques are also being used with other neural network techniques for more efficient anomaly detection. These techniques are applied on the incoming packets from the network to the server by intelligent agents. Moreover some of the intrusions can be detected by analysing log files and system call sequences. Any deviation from the normal behaviour is a key to detect vulnerabilities. Whenever the user logins for the first time a user profile is generated according to the information provided by the user and this user profile also helps in gaining information about suspicious activities.

The model presented in [3] is a distributed IDS and a similar model can be used for securing servers involved in E-Business. Active attacks are comparatively easier to detect than passive attacks as they involve certain kind of alterations whereas passive attacks are most of the time difficult to detect. Although the model is a combination of misuse and anomaly detection system we have proposed here misuse detection system. Misuse detection systems are good at detection of malicious activities if the patterns are known otherwise they are not able to detect novel attacks whereas anomaly detection systems can detect novel attacks. But anomaly detection systems generate more false alarms as compared to misuse detection agent. Earlier concepts of intrusion detection were based on central systems for monitoring but they had many disadvantages [8]. Although distributed IDS have drawbacks too some of them being [9]- the signature or the knowledge base update has to be done timely, most of the time it happens that the IDS and other security systems do not interoperate and thus it becomes quite inevitable to use any one of them for

[1,2] Department of Computer Science, Amity University
[3] Department of Computer Science, Lucknow University of Lucknow
[1]sharmaanuradhak@gmail.com
[2]naazparveen47@yahoo.in
[3]puneetmisra@gmail.com

security, IDS architecture needs improvement with time as new technologies come they have to be modified for adaptability. Moreover only these factors are not enough, an IDS has to be flexible, fault tolerant, adaptable and configurable [10]. When working on an agent oriented view, it soon becomes apparent that a single agent is not sufficient. Most problems require a single agent[14].

The next section gives an overview of various agents and multi-agent systems.

2.0 AGENTS AND MULTI-AGENT SYSTEMS

A software agent is a persistent, goal oriented piece of software that autonomously and continuously work in its environment to perform some specified function for the end user or another program. Agents are capable of doing the desired task without continuous direct supervision. The concept of software agents arise from Multi-Agent Systems which are again the concept somewhere used in Distributed Artificial Intelligence (DAI). Based on their mobility they can be classified as-

i. Static agents- They do not possess mobility hence they perform their jobs at the place where they are present.

ii. Mobile agents- They perform their goal at different hosts.

Agents can be classified on the basis of Nwana's primary attribute dimension as- smart agents, collaborative agents, collaborative learning agents and interface agents [12]. On the other hand agents are classified according to their task as mentioned by George F Luger in his book [13]–

- Rote Agents- They simply capture pieces of information and communicate them to others.
- Coordination Agents- These agents support interactions between tasks and other agents so as to prevent collision.
- Search Agents- These agents are used for analysing information and return the chosen ones.
- Learning Agents- They form concepts or generalize after analysing the information so far collected.
- Decision Agents- They are responsible for initiation of actions and come to conclusions from the information they receive.

Intelligent agent is an autonomous agent which senses its environment and accordingly acts upon it that is it is a software agent with intelligence. It can learn and based upon the knowledge it is capable of taking decisions.

Multi-Agent System consists of a network of autonomous software agents that interact with each other to solve the problem which cannot be solved by individual capability. Agents have the ability to act as mediators in solving problems. And such a concept of Artificial Intelligence is being used to tackle security related issues of a network. A Multi-Agent System always works in a dynamic environment on a set of objects with the help of autonomous agents. It can be viewed as a loosely coupled network of autonomous agents where each agent has its own goal to accomplish.

Some of the characteristics of Multi-Agent Systems are [11]-

i. As multiple agents are used, no agent gets complete information.

ii. There is no global control mechanism for any Multi-Agent System as they have less human intervention.

iii. Data is distributed that is there is no single source of data.

iv. Actions are not controlled as they are brought into action by autonomous software agents.

v. They are dynamic systems; they are not confined to work on a limited set of data but are designed to tackle new problems and new data.

vi. Each agent in a Multi-Agent System has a predefined goal and from the data they receive they work exactly on the relevant data of their use.

3.0. THE MISUSE DETECTION SYSTEM

The architecture of Misuse Detection System consists of a collection of agents which work cooperatively to collect and analyze the network data coming to the server. The various agents used in the proposed model are- sniffer agent, filter agent, CIA agent and a reporter agent. The process in brief is:

1. The Sniffer Agent gathers the incoming packet from the connected network for any online transaction.
2. The filter agent helps in filtering the packets being gathered by the Sniffer Agent and sorts them according to their packet format.
3. The filtered packets are then passed to CIA Agent (Confidentiality, Integrity and Availability Agent). This agent analyzes the packets and matches them with the predefined attack signatures stored in the Knowledge Base and is responsible for calculating the digests and comparing them with the already stored ones and accordingly passes the information to the Reporter Agent.
4. The Reporter Agent then analyzes the information passed by the CIA agent and if it is an attack it reports to the administrator.

The knowledge base consists of the rules and signatures of already known attacks as well as some vulnerable ipaddresses, message digests so that the calculated message digest can be compared with the valid one. The figure. 1 shows the proposed model of the Misuse Detection System.

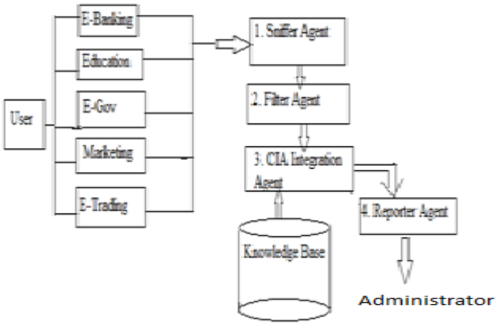


Figure 1. Misuse Detection System

Various components have been described below-

3.1 Sniffer Agent

Basically sniffer is a device that captures all the traffic passing through a network. It helps in detection of intrusions by analysing network traffic as it gathers all the packets passing through it. Sniffer agents read packets from the machine and cache them in memory[3]. Thus they are being used for capturing packets and the output of this agent is being sent as an input to the Filter Agent.

3.2 Filter Agent

The Filter Agent is responsible for filtering out specific data as in a network there are various data sources so while considering intrusion as a matter of concern [3]. It differentiates the various fields of the received packets and sorts them according to their category concerning about the intrusions.

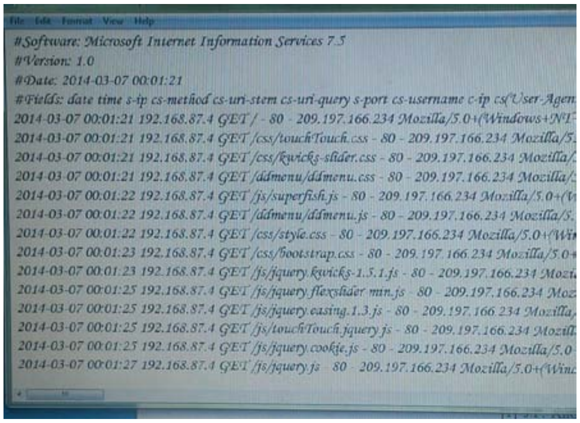
3.3. CIA IntegrationAgent

The CIA Integration Agent is the intelligent agent playing the most crucial role for detection of malicious activities going on in the server. It focuses completely on the three parameters namely *Confidentiality, Integrity* and *Availability.*The name being used as the agent proposed here is responsible for checking the requisite constraints for these three issues. The agent being proposed in this paper has multiple functions to perform so that malicious activities can be detected effectively. The CIA Integration Agent after receiving packets from the Filter Agent checks whether they are from an authentic source and for this it checks the signatures of the packets and compares with the known attack signatures already stored in the knowledge base then it consequently raises an alert to the Reporter Agent and removes those malicious packets.

As the attack signatures are already stored in the knowledge base it requires a rule set and a pattern matching technique so as to detect intrusions. For this RETE algorithm is proposed [5]. The Knowledge Base consists of attack patterns for known attacks as provided by experts of this field. The reasons for using RETE algorithm being its time saving capability and that it does not leads to exhaustive matching of thousands of conditions [4]. As we are focusing on online transactions we assume that the messages are encrypted and before initiating the transaction for the first time public keys are exchanged and these public keys are stored in the knowledge base thus the next time transaction occurs decryption is performed by using the combination of the saved public key. We assume that for the first time if a user is requesting for a service, the details of the user are saved and keys are exchanged and for maintaining integrity of the user credentials a strong hash function is being used. Now the hash values are computed and stored in the knowledge base and the same hash function is used for the entire transaction and thus the CIA Integration agent computes the hash value and matches with the one already stored for each incoming packet from the same source. Apart from this the CIA Integration agent being proposed here is also responsible for analysing user profile because user profile gives an idea of the sort of behaviour expected from the user involved in the transaction, it is like monitoring the user's activities. Moreover as the intrusion detection is host based it also keeps track of the sequences of the system calls being executed. Any behaviour deviating from the normal is considered to be anomalous. The various log files are analysed by one of the modules of CIA Integration agent known to be Log File Analyzer. The contents of the log files are compared to the patterns present in the knowledge base[6]. The system log files are categorised as- system, application and security logs corresponding to their events [7]. Analysis of host machine's log files provides extra security to the host as time to time the security log files can be analysed. And since the rules are stored in the Knowledge Base it also makes use of rule mining to accomplish its goal.

4.0 LOG FILE ANALYSIS

As log files are the archives of various events taking place in a system and in windows are generally stored in EVTX format [7]. And every event has some attributes associated with it. The log files after*Denial of service attack* as provided by Centre for Cyber Forensics & Information Security and moreover we analysed that there were no considerable changes in log files of Windows XP when it was exploited with the help of BackTrack5.



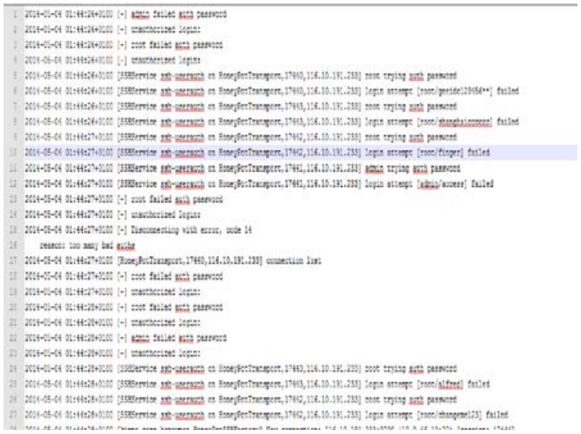
```
#Software: Microsoft Internet Information Services 7.5
#Version: 1.0
#Date: 2014-03-07 00:01:21
#Fields: date time s-ip cs-method cs-uri-stem cs-uri-query s-port cs-username c-ip cs(User-Agen
2014-03-07 00:01:21 192.168.87.4 GET / - 80 - 209.197.166.234 Mozilla/5.0+(Windows+NT
2014-03-07 00:01:21 192.168.87.4 GET /css/touchTouch.css - 80 - 209.197.166.234 Mozilla/5.
2014-03-07 00:01:21 192.168.87.4 GET /css/kwicks-slider.css - 80 - 209.197.166.234 Mozilla/
2014-03-07 00:01:21 192.168.87.4 GET /ddmenu/ddmenu.css - 80 - 209.197.166.234 Mozilla/
2014-03-07 00:01:22 192.168.87.4 GET /js/superfish.js - 80 - 209.197.166.234 Mozilla/5.0+(W
2014-03-07 00:01:22 192.168.87.4 GET /ddmenu/ddmenu.js - 80 - 209.197.166.234 Mozilla/5.
2014-03-07 00:01:22 192.168.87.4 GET /css/style.css - 80 - 209.197.166.234 Mozilla/5.0+(Wi
2014-03-07 00:01:23 192.168.87.4 GET /css/bootstrap.css - 80 - 209.197.166.234 Mozilla/5.0+
2014-03-07 00:01:23 192.168.87.4 GET /js/jquery.kwicks-1.5.1.js - 80 - 209.197.166.234 Mozi
2014-03-07 00:01:25 192.168.87.4 GET /js/jquery.flexslider min.js - 80 - 209.197.166.234 Moz
2014-03-07 00:01:25 192.168.87.4 GET /js/jquery.easing.1.3.js - 80 - 209.197.166.234 Mozilla,
2014-03-07 00:01:25 192.168.87.4 GET /js/touchTouch.jquery.js - 80 - 209.197.166.234 Mozill
2014-03-07 00:01:25 192.168.87.4 GET /js/jquery.cookie.js - 80 - 209.197.166.234 Mozilla/5.0
2014-03-07 00:01:27 192.168.87.4 GET /js/jquery.js - 80 - 209.197.166.234 Mozilla/5.0+(Winc
```

Figure 2. Log Files after Dos attack

After performing DHCP snooping through Man In The Middle (MITM) attack the system log file showed some errors and warnings which has to be analysed by the CIA Integration agent and simultaneously the CIA Integration agent has to make an observation on each user's profile so as to detect any malicious event which can be harmful to the server involved in online transactions.

5.0 CONCLUSION

In this paper a novel Misuse Detection System is proposed for the security of servers involved for online transactions. The basic goal of this system is to intimate the administrator about the various malicious activities occurring in the host machine so that appropriate actions are taken immediately. The Misuse Detection System being proposed here is basically a multi-agent system making use of intelligent agents having the capability to exploit data mining techniques to detect intrusions. The CIA agent not only checks availability but also checks for confidentiality and integrity by making use of public and private keys and calculating message digests. The system being proposed in this paper also makes use of log files analysis and system call analysis since changes in log files and deviations from the normal system call usage depict vulnerabilities.



```
2016-05-04 01:44:26+0100 [-] admin failed auth password
2016-05-04 01:44:26+0100 [-] unauthorized login:
2016-05-04 01:44:26+0100 [-] root failed auth password
2016-05-04 01:44:26+0100 [-] unauthorized login:
2016-05-04 01:44:26+0100 [SSHService ssh-userauth on HoneyPotTransport,17460,116.10.191.233] root trying auth password
2016-05-04 01:44:26+0100 [SSHService ssh-userauth on HoneyPotTransport,17460,116.10.191.233] login attempt [root/genie123456**] failed
2016-05-04 01:44:26+0100 [SSHService ssh-userauth on HoneyPotTransport,17463,116.10.191.233] root trying auth password
2016-05-04 01:44:26+0100 [SSHService ssh-userauth on HoneyPotTransport,17463,116.10.191.233] login attempt [root/shanghaiccccc] failed
2016-05-04 01:44:27+0100 [SSHService ssh-userauth on HoneyPotTransport,17462,116.10.191.233] root trying auth password
2016-05-04 01:44:27+0100 [SSHService ssh-userauth on HoneyPotTransport,17462,116.10.191.233] login attempt [root/finger] failed
2016-05-04 01:44:27+0100 [SSHService ssh-userauth on HoneyPotTransport,17461,116.10.191.233] admin trying auth password
2016-05-04 01:44:27+0100 [SSHService ssh-userauth on HoneyPotTransport,17461,116.10.191.233] login attempt [admin/access] failed
2016-05-04 01:44:27+0100 [-] root failed auth password
2016-05-04 01:44:27+0100 [-] unauthorized login:
2016-05-04 01:44:27+0100 [-] Disconnecting with error, code 14
    reason: too many bad auths
2016-05-04 01:44:27+0100 [HoneyPotTransport,17460,116.10.191.233] connection lost
2016-05-04 01:44:27+0100 [-] root failed auth password
2016-05-04 01:44:27+0100 [-] unauthorized login:
2016-01-04 01:44:28+0100 [-] root failed auth password
2016-01-04 01:44:28+0100 [-] unauthorized login:
2016-01-04 01:44:28+0100 [-] admin failed auth password
2016-01-04 01:44:28+0100 [-] unauthorized login:
2016-01-04 01:44:28+0100 [SSHService ssh-userauth on HoneyPotTransport,17463,116.10.191.233] root trying auth password
2016-05-04 01:44:28+0100 [SSHService ssh-userauth on HoneyPotTransport,17463,116.10.191.233] login attempt [root/alfred] failed
2016-05-04 01:44:28+0100 [SSHService ssh-userauth on HoneyPotTransport,17462,116.10.191.233] root trying auth password
2016-05-04 01:44:28+0100 [SSHService ssh-userauth on HoneyPotTransport,17462,116.10.191.233] login attempt [root/changeme123] failed
```

Figure 3. Log Files After DHCP Snooping

The system can be extended by adding new task agents and can be used for E-Business where the involved servers store user credentials. The future work includes the additional capabilities to detect novel attacks with less false positive rate without compromising the resources available and thus creating a proactive shield for future attacks.

REFERENCES

[1]. J.P. Anderson, "Computer Security Threat Monitoring and Surveillance,"James P. Anderson Co., Tech. Rep., 26 Feb 1980.

[2]. Chang-Tien Lu, Arnold P. Boedihardjo, PrajwalManalwar, "Exploiting Efficient Data Mining Techniques to Enhance Intrusion Detection Systems." in *Proceedings of the 2005 IEEE International Conference on Information Reuse and Integration,* pp. 512-517.

[3]. ImenBrahmi, Sadok Ben Yahia, HamedAouadi and Pascal Poncelet, "Towards A Multi agent Based Distributed Intrusion Detection System Using Data Mining Approaches ", 2011, ADMI'11 *Proceedings of 7th international conference on Agents and Data Mining interaction, Springer- Verlag*, pp. 173-194.

[4]. Dan W. Patterson,"Introduction to Artificial Intelligence and Expert Systems", *Prentice Hall of India*, first edition.

[5]. C. Forgy. RETE: "A Fast Algorithm for the many pattern/ many object pattern match problem Artificial Intelligence," 19(1):17-37, 1982.

[6]. Firkhan Ali Bin Hamid Ali, Yee Yong Len, "Development of Host Based Intrusion Detection System for Log Files" , in *proceedings of 2011 IEEE symposium on Business, Engineering & Industrial Application*, pp.281-285.

[7]. PavitraChauhan, Nikita Singh, Nidhi Chandra, "Deportment of Logs for securing the Host System"in *2013 5th International Conference on Computational Intelligence and Communication Networks* , pp. 355-359.

[8]. R. Gopalakrishna and E.H. Spafford, "A Framework for Distributed Intrusion Detection using Interest Driven Cooperating Agents," *in Proceedings of the 4th International Symposium on Recent Advances in Intrusion Detection*, Davis, CA, USA, 2001.

[9]. W. Huang, Y. An and W. Du, " A Multi-Agent Based Distributed Intrusion Detection System," in Proceedings of the 3rd International Conference on Advanced Computer Theory and Engineering, Chengdu, Sichuan province, China, pp: 141-143, 2010.

[10]. E. Mosqueira-Rey, B. Guijarro-Berdias, A. Alonso-Betanzos, D. Alonso-Ros and J. Lago-Pieiro, " A Snort-based Agent for a JADE Multi-agent Intrusion Detection System". International Journal of Intelligent Information and Database System, 3(1) : 107-121, 2009.

[11]. Jennings, N.R., Sycara, K., Wooldridge, M., " A Roadmap of Agent Research and Development ," Autonomous Agents and Multi-Agent Systems1(1),7-38, 1998.

[12]. AleksanderPivk, Matjaz Gams, "Intelligent Agents in E-Commerce".Available at http://citeseerx.ist.psu.edu .

[13]. George F. Luger, "Artificial Intelligence Structures & Strategies for Complex Problem Solving", Addison-Wesley 2008, sixth edition.

[14]. S. Ajitha, T.V. Suresh Kumar, K. Rajanikanth, "Framework for multi-agent systems performance prediction process model: MASP^3", BIJIT-2014, July - December, 2014; Vol. 6 No. 2; ISSN 0973 – 5658 774.

[15]. Bilal MaqboolBeigh, "Framework for choosing best intrusion detection system", BIJIT – 2015; January - June, 2015; Vol. 7 No. 1; ISSN 0973 – 5658 821.

[16]. Mitra, Sulata, and ArkadeepGoswami. "Load Balancingin Integrated MANET, WLAN and Cellular Network."BIJIT – BVICAM's International Journal ofInformation Technology, (2011): 304.