ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

Detecting Internal Intrusion of the System Using Data Mining

Bhakyalakshmi.N[#1] , Durga.T[#2] , Gayathri.P[#3] , Jane sherin.B [#4,] Ramesh Kannan.M[5]

B.E. Student, Dept. of Computer Science and Engineering, Saranathan College of Engineering, Venkateshwara Nagar, Panjapur, Trichy, Tamilnadu, India[1, 2,3,4]

Assistant Professor, Dept. of Computer Science, Saranathan College of Engineering, Venkateshwara Nagar, Panjapur, Trichy, Tamilnadu, India[5]

ABSTRACT: The system proposes a security system, named the Internal Intrusion Detection and Protection System (IIDPS for short) at system call level, which creates personal profiles for users to keep track of their usage habits as the forensic features. The IIDPS uses a local computational grid to detect malicious behaviors in a real-time manner The proposed work is regarded with Digital forensics technique and intrusion detection mechanism. The number of hacking and intrusion incidents is increasing alarmingly each year as new technology rolls out. The system designed Intrusion Detection System (IDS) that implements predefined algorithms for identifying the attacks over a network. Therefore, in this project, a security system, named the Internal Intrusion Detection and Protection System (IIDPS), is proposed to detect insider attacks at SC level by using data mining and forensic techniques. The system can identify a user's forensic features by analyzing the corresponding SCs to enhance the accuracy of attack detection, and able to port the IIDPS to a parallel system to further shorten its detection response time.

KEYWORDS: Internal intrusion, user behavior, data mining

I. INTRODUCTION

The computer systems have been widely employed to provide users with easier and more convenient lives. However, when people exploit powerful capabilities and processing power of computer systems, security has been one of the serious problems in the computer domain since attackers very usually try to penetrate computer systems and behave maliciously, e.g., stealing critical data of a company, making the systems out of work or even destroying the systems. Generally, among all well-known attacks such as pharming attack, distributed denial-of-service (DDoS), eavesdropping attack, and spear-phishing attack , insider attack is one of the most difficult ones to be detected because firewalls and intrusion detection systems (IDSs) usually defend against outside attacks. To authenticate users, currently, most systems check user ID and password as a login pattern. However, attackers may install Trojans to pilfer victims' login patterns or issue a large scale of trials with the assistance of a dictionary to acquire users' passwords. When successful, they may then log in to the system, access users' private files, or modify or destroy system settings. However, it is very difficult to identify who the attacker is because attack packets are often issued with forged IPs or attackers may enter a system with valid login patterns.

II. EXISTING SYSTEM

In existing system, a model is proposed for such an attack based on network traffic flow. Specific network topology-based patterns are defined to model normal network traffic flow, and to facilitate differentiation between legitimate traffic packets and anomalous attack traffic packets. A novel approach for postmortem intrusion detection, which factors out repetitive behavior, thus speeding up the process of locating the execution of an exploit, if any. Central to our intrusion detection mechanism is a classifier, which separates abnormal behavior from normal one. This classifier is built upon a method that combines a hidden Markov model with k -means. Packet sniffer is not just a hacker's tool. It can be used for network traffic monitoring, traffic analysis, troubleshooting and other useful purposes. When computers communicate over networks, they normally just listen to the traffic specifically for them. The disadvantage is that they cannot easily authenticate remote-login users and detect specific types of intrusions.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

III. PROPOSED SYSTEM

The proposed system provide a security system, named Internal Intrusion Detection and Protection System (IIDPS), which detects malicious behaviors launched toward a system at SC level. The IIDPS uses data mining and forensic profiling techniques to mine system call patterns(SC patterns) defined as the longest system call sequence that has repeatedly appear several times in a user's log file for the user. The user's forensic features, defined as an SC pattern frequently appearing in a user's submitted SC sequence but rarely being used by other users ,are retrieved from the user's computer usage history. The system need to study the SCs generated and the SC-patterns produced by these commands so that the IIDPS can detect those malicious behaviors issued by them and then prevent the protected system from being attacked.

.

Scope: Our paper has a big scope to do. The security problem on system leads to find solution as privacy system.We have to find whether the user who logged into the syetem is authenticated user or not.

Applicability: This paper "Detecting internal intrusion of the system using data mining" is applicable for different classes of user's in various sector's like banking, multinational companies, etc.

Our project provides the modules like

- Mining of user habits
- User activities
- Attacker detection

A. *Mining of user habits:*

In this module, most commonly a data set corresponds to the contents of a single database table, or a single statistical data matrix, where every column of the table represents a particular variable, and each row corresponds to a given member of a dataset. Consider the user dataset with attributes as create, read, modify, delete, encrypt, decrypt. First the user dataset is loaded into the database. In preprocessing the missing values are identified. Range is fixed for each attributes. The attribute values which are beyond this range will be identified as noisy data and it is removed. After preprocessing the dataset the threshold value is calculated

B. *User activities:*

In this module, the user gets logged into the system. Now the user works on the file such as file create, file read, file modify, file delete, file encrypt, file decrypt. The user activities are then stored as user log file in the database.

.

C. *Attacker detection:*

In this module, the user log file in the database is the compared with the user habit (i.e) user profile, to identify whether the user is authenticated user or not. If the user log file does not exceeds user habit, the user is regular user. Otherwise the system will detect that attacking is done in the file system due to some intruder. It is shown in Fig.4.

IV. SYSTEM ARCHITECTURE

Loading the dataset for the current user into the database. The dataset contains the users behavioural data and do the preprocessing. The preprocessed data is stored into database. The threshold value is calculated from the preprocessed dataset from which the the calculations for the nearest values are made. The users behaviour pattern is found by the nearest value which is stored as User's profile.

When the authenticated user is log into the system his/her current system usage activities are updated in the data base. The Admin updates the user login details into the database. The logged user's behavioural values are stored as the User's log file. The comparison is made between the current user's profile which contains his/her behavioural pattern and user's log file. If there is any deviations occur while the comparison the it is notified that the currently logged user is attacking the system. Thus the Insider attack is identified.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

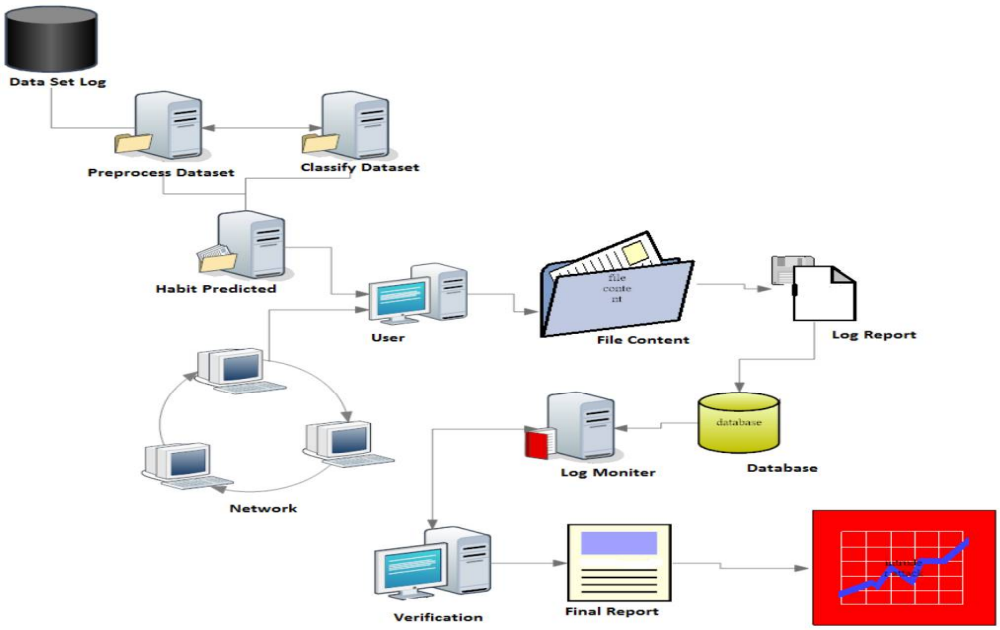

Fig 1. System architecture

V. RESULTS AND OUTCOME

The user profile is generated from the dataset. The dataset contains the user's usage history. The log file for the current logged user is generated. By comparing the user profile and log file the intruder is detected.

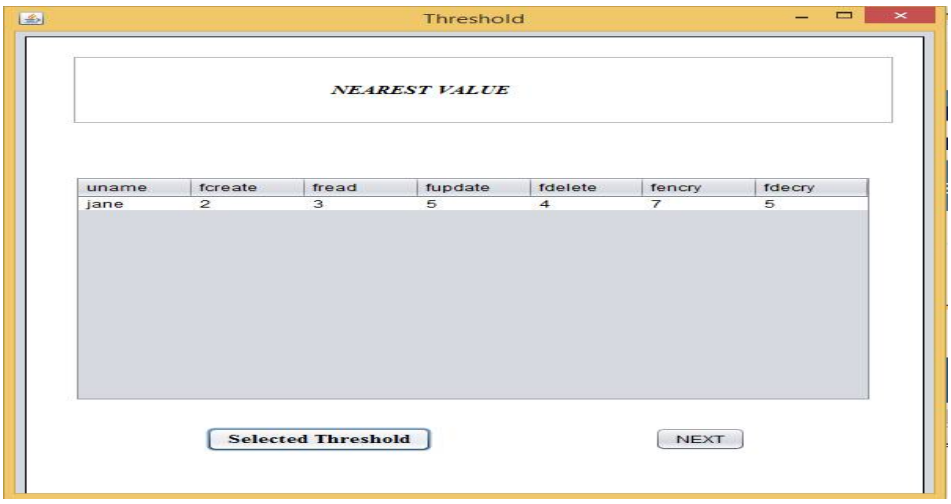

Fig 2: Mining of user habits

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

Fig 3: User profile

Fig 4: User login form

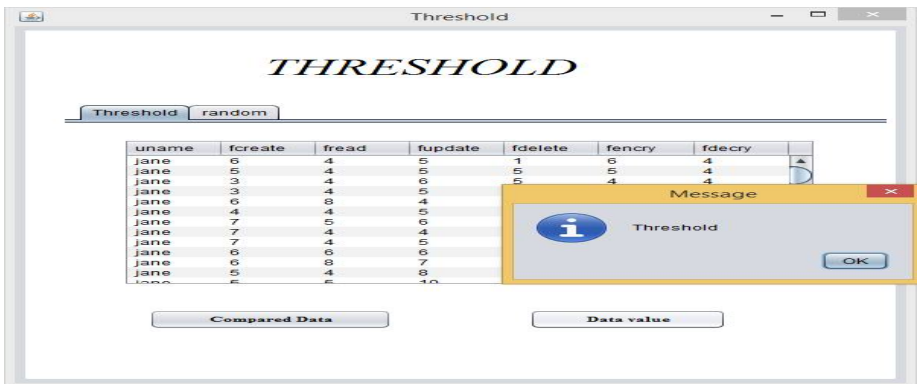

Fig 5: User log file

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

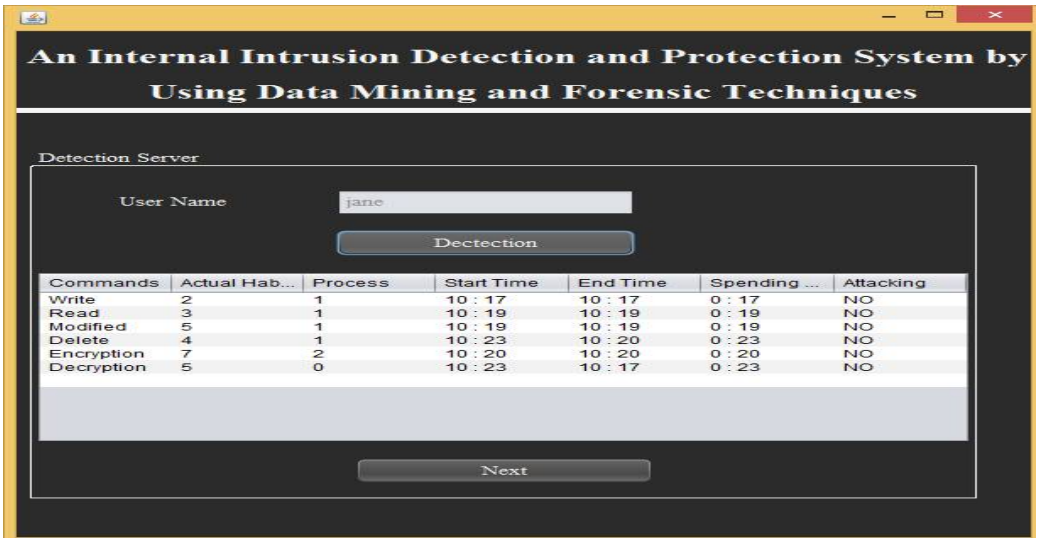

Fig 6: Attacker detection

VI. CONCLUSION AND FUTURE WORK

The IIDPS (Internal Intrusion Detection and Protection System) employs data mining and forensic techniques to identify the user behavioral patterns for a user. The time that a habitual behavior pattern appears in the user's log file is counted, the most commonly used patterns are filtered out, and then a user's profile is established. By identifying a user's behavior patterns as his/her computer usage habits from the user's current input, the IIDPS resists suspected attackers. The future work of insider attack detection research will be about collecting the real data in order to study general solutions and models. It is hard to collect data from normal users in many different environments. It is especially hard to acquire real data from a masquerader or traitor while performing their malicious actions. Even if such data were available, it is more likely to be out of reach and controlled under the rules of evidence, rather than being a source of valuable information for research purposes.

REFERENCES

1. F. Y. Leu, M. C. Li, J. C. Lin, and C. T. Yang, "Detection workload in a dynamic grid-based intrusion detection environment," *J. Parallel Distrib. Comput.*, vol. 68, no. 4, pp. 427–442, Apr. 2008.
2. J. T. Giffin, S. Jha, and B. P. Miller, "Automated discovery of mimicry attacks," *Recent Adv. Intrusion Detection*, vol. 4219, pp. 41–60, Sep. 2006.
3. S. X. Wu and W. Banzhaf, "The use of computational intelligence in intrusion detection systems: A review," *Appl. Soft Comput.*, vol. 10, no. 1, pp. 1–35, Jan. 2010.
4. F. Y. Leu, K.W. Hu, and F. C. Jiang "Intrusion detection and identification system using data mining and forensic techniques," *Adv. Inf. Comput. Security*, vol. 4752, pp. 137–152, 2007.
5. M. A. Faisal, Z. Aung, J. R. Williams, and A. Sanchez, "Data-stream based intrusion detection system for advanced metering infrastructure in smart grid: A feasibility study," *IEEE Syst. J.*, vol. 9, no. 1, pp. 1–14, Jan. 2014.
6. Z. A. Baig, "Pattern recognition for detecting distributed node exhaustion attacks in wireless sensor networks," *Comput. Commun.*, vol. 34, no. 3, pp. 468–484, Mar. 2011.