ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 3, March 2018

A Survey on Detection of Internal Intruder Using User's Habit File

Ameya N. Ghorpade, Prof. H. P. Channe

Student, PICT, Pune, India

Asst. Professor, PICT, Pune, India

ABSTRACT: Numerous computer systems utilize user IDs and passwords as the login patterns to approve users. In any case, the greater part of the general population shares their login designs with colleagues and demands them to help co-works. This makes the pattern as one of the weakest points of computer security. The principle issue of computer security is insider attack. Insider attack is one of the most troublesome attacks to be identified in the fact that firewalls and intrusion detection systems (IDSs) usually defend versus outside attacks. In an additional research, analyzing the OS-level System Calls (SCs) is helpful in detecting attackers and identifying users. Therefore, in this paper, a security framework, called Internal Intrusion Detection and Protection System (IIDPS), is intended to discover insider assaults at SC level by utilizing information mining and strategies. The IIDPS builds user's habit profiles to monitor user's habits and decides whether the login user is the account holder or not by contrasting the current computer usage behaviors of the user with the patterns gathered in the account holder's habit profile. The contribution work is performance analysis comparison between proposed algorithm and the machine learning classifier applied for intruder detection.

KEYWORDS: Data Mining, Insider Attack, System Calls (SCs), Habit File, User's Behaviors, Intrusion Detection, SVM

I. INTRODUCTION

Information mining (learning disclosure) is the way toward joining data from alternate points of view and studying it into important information.In other words the practice of examining large pre-existing databases in order to produce new information is known as data mining. Computer systems have been widely employed to provide users with simpler and more comfortable lives. However, when people make use of powerful capabilities and processing power of computers, security has been considered as one of the major and serious problems in the computer domain since attackers very usually try to break through computer systems and behave false heartedly, e.g., nicking critical data of a company, wrecking the systems or even making the systems out of work. Generally, among all well-known attacks such as 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 validate users, presently, most systems examine user ID and password as a login pattern. However, intruders may install Trojans to nick victims' login patterns or issue a large scale of trials with the aid of a dictionary to acquire user's passwords. When efficacious, they may then log in to the system, access user's private files, or modify or destroy system settings. Although OS-level System Calls (SCs) are helpful in detecting attackers and identifying users, processing a large amount of SCs, excavating malicious behaviors from them, and finding possible intruders for an intrusion are still implementing challenges. Therefore, in this paper, a security system, called Internal Intrusion Detection and Protection System (IIDPS), which finds malicious behaviors launched toward a system at SC level is designed.

Objectives:

1) Identify a user's forensic features by analyzing the corresponding SCs to enhance the accuracy of attack detection.
2) Able to port the IIDPS to a parallel system to further shorten its detection response time; and

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 3, March 2018

3) Effectively resist insider attack.
4) If attack detects then prevent it before attack.
5) When user remotely accessing the system, the no. of attacks should be detected the system and protect the system.

Motivation:

1. By using users' forensic features retrieved from their basic operations are helpful in detecting the users' malicious behaviors and tell us who the possible attackers are.
2. The SCs generated and the SC-patterns produced by using commands so that the IIDPS can detect those malicious behaviors issued by user and then prevent the protected system from being attacked.
3. Protect system from insider attack.

II. REVIEW OF LITERATURE

The paper [1] proposes a Malicious node Identification Scheme (MIS) that identifies and isolates malicious nodes, so that the pollution attack can cause harm to the network only for a short period of time and the subsequent streaming will no longer be influenced. MIS is block-based and can rapidly identify malicious nodes, ensuring that the system quickly recovers from pollution attacks. Advantages are: It has high computational efficiency, small space overhead, and the capability of handling a large number of corrupted blocks and malicious nodes, and does not require repeatedly pre-distributing verification information. MIS can rapidly identify malicious nodes. Disadvantages are: This scheme is applicable only when a limited number of blocks are corrupted.

The paper [2] proposes a novel, effective and feasible method to traceback to individual mobiles who involve in attacks. Designed a practical traceback scheme to identify active mobile bots with single IP packet. Advantages are: The proposed method takes the advantage of IMEI number of mobile phones, and employed the marking on demand philosophy. The feasibility analysis indicates that the proposed method is feasible. Disadvantages are: The traceback architecture is centralized. Scalability and detection at WAP gateways these are the issues.

The paper [3] describes a security system, named the Intrusion Detection and Identification System (IDIS), which mines log data to identify commands and their sequences (together named command sequences (C-sequences in short)) that a user habitually submits and follows respectively as the user's forensic features. By comparing a user's current input commands with all others' profiles, the IDIS can identify who the user is. Advantages are: The accuracy is high. Accurately and completely collecting user behaviors on much more basic operations, such as system calls. Disadvantages are: Need a fast algorithm and a distributed computing environment to speed up data processing.

In [4] paper, proposes a security network framework for enterprise security management. The traffic data is analyzed and stored by data mining process which is used to find various traceback paths easily. The logging-based traceback method is used for summarizes and stores only main information from large-scale data by using data mining. Advantages are: Reduces an enormous amount of calculation used for traceback. Reduce the storage overhead.

Paper [5] focus on the security of advanced metering infrastructure (AMI), which is one of the most crucial components of Smart Grid (SG). Proposes a new AMI IDS architecture based on the AMI architecture presented by OPENMeter, which is a project deployed by several European countries to reduce gap between the state-of-the-art technologies and AMI's requirements. Advantages are: AMI's main functionalities encompass power measurement facilities, assisting adaptive power pricing and demand side management, providing self-healing ability, and interfaces for other systems. Data speed is very high as it has to handle a huge amount of meter data, event data, commands, etc. Disadvantages are: The main problem is the limitation of computing resources in the current smart meters.

The paper [6] elaborates a method of integration between HTTP GET flooding among DDOS attacks and MapReduce processing for fast attack detection in cloud computing environment. This method is possible to ensure the availability of the target system for accurate and reliable detection based on HTTP GET flooding. Advantages are: Accurate and reliable detection based on HTTP GET flooding method. Processing time of proposed method is shorter with increasing congestion. Proposed method is better than Snort detection method. Disadvantages are: Needs the various pattern recognitions for DDoS attack detection in cloud computing environment.

The [7] paper proposes new approach; sensors collect system and network parameters and send the data to the forecasters and the intrusion detection systems (IDSes). A multi-objective controller selects the optimal protection

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 3, March 2018

method to recover the system based on the signature of attacks. Advantages are: Proposed approach works properly for a wide range of exploits with low overheads. The accuracy of the ARIMA forecasters is high enough to meet the demands of system early detection as well. Disadvantages are: This detection attack algorithm is time consuming.
In [8] paper a robust adversary model for a distributed node exhaustion attack is formulated. In addition, a distributed pattern recognition scheme is defined to efficiently detect such an attack. A distributed, pattern recognition scheme for distributed node exhaustion attack detection in wireless sensor networks is defined. Advantages are: Attack detection, without incurring significant overhead on the limited energy resources of the sensor nodes. Reduce the network overhead.
The paper [9] develop a VM commitment system called Secom to automatically eliminate malicious state changes when merging the contents of an OS-level VM to the host. Secom consists of three steps: grouping state changes into clusters, distinguishing between benign and malicious clusters, and committing benign clusters. Advantages are: A lower false positive rate of behavior-based malware detection approach. VM is to identify compromised objects thoroughly and lightly.

III. SYSTEM OVERVIEW

The IIDPS includes an SC monitor and filter, Intrusion detection, three repositories containing user log files, user habit, an attacker habit, user profiles, an attacker profile, and a Data Extraction from log File.

System call monitoring and filtering
The system call monitor & filter, collects the System Calls (SCs) i.e. user‘s actions and stores the SCs in the protected system. The user inputs are stored in the user‘s log file i.e., a file used to keep the SCs submitted by the user. F or example, when a user changes his/her password, up to 2916 SCs will be generated by capitulating a password command to a Linux operating system, including open(), close(), read(), write(), update(), delete(), etc. So, it is hard for a system to keep track of all SCs at the same time. As a result, we need to sifter out some frequently used safe SCs.

Data Extraction from log File
A mining server extracts the system calls from the user‘s log file and counts the number of times that a SC-pattern appears in this file and stores it in the user‘s habit file. A habit file is an SC-pattern‘s appearance count that is substituted by its corresponding similarity weight. Then the SC-pattern‘s similarity weights are considered to sifter out those SC-patterns usually used by all or most users. An attack pattern or a signature may be an attacker-specific pattern or a pattern usually used by attackers.

Intrusion Detection
It detects an internal intruder or an attacker. The detection server records the SCs submitted by the underlying user u and stores the SCs in the u‘s log file. Then, the server tries to identify whether u is the valid account holder or not by computing the similarity score between the newly generated SCs, denoted by NSC, in the u‘s current inputs (in u‘s log file) and the usage habits, i.e., forensic signatures (also known as behavior patterns), stored in u‘s user profile to verify u. The current habit file NHF is established by |SC Su|-|sliding window| pair-wised. The Detection accuracy‖ is the accuracy of finding out the user‘s identity.

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 3, March 2018

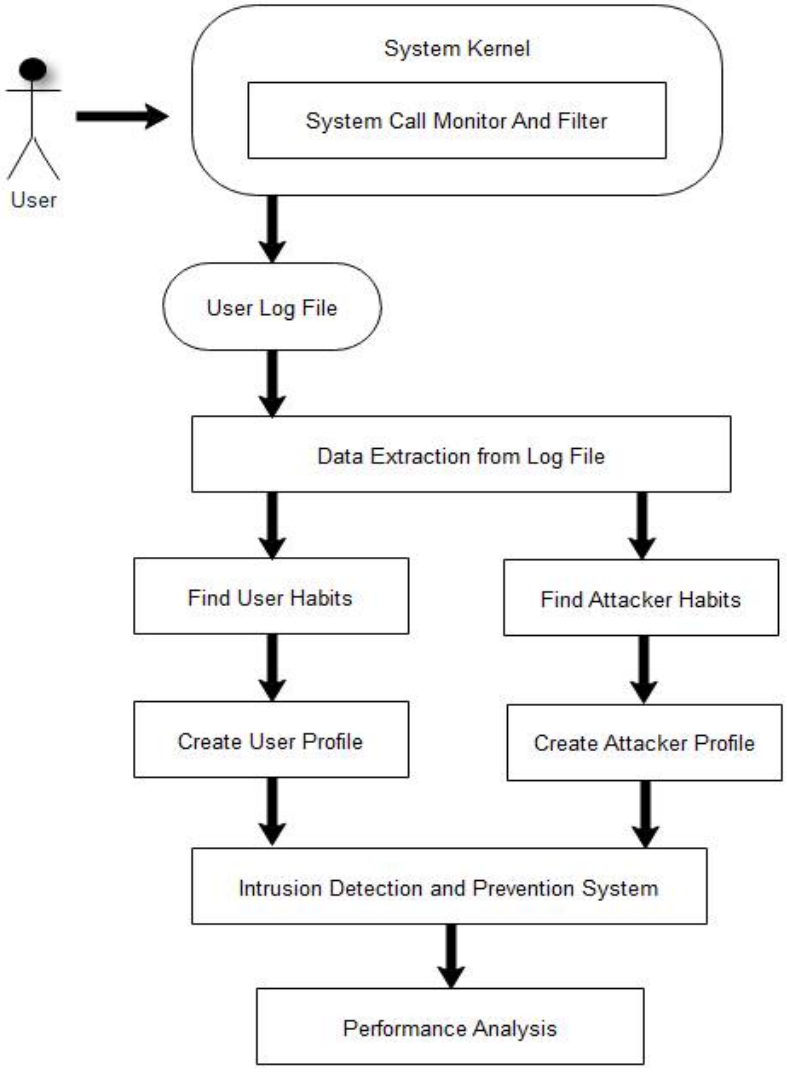

Fig.1. System Architecture

Algorithm:

Input

Input given to the system is: -User Log File (User log data).

Output

Results demonstrate that the intrusion detected or not.

Process

1. Start
2. First we identify the representative SC-patterns for a user.
3. After find the habitual SC pattern from user log files. We used specific algorithm for filtering most commonly used SC-patterns.
4. Create the user profile using habitual data

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 3, March 2018

5. Compare the user and attacker profile (By identifying a user's SC-patterns as his/her computer usage habits from the user's current input SCs, the IIDPS resists suspected attackers).
6. Finally results demonstrate that the intrusion detected or not.
7. Stop

SVM algorithm:
Algorithm Steps:
Step1: SVMs maximize the margin around the separating hyperplane.

- Assume linear separability for now:
  - in 2 dimensions, can separate by a line
  - in higher dimensions, need hyperplanes
- Can find separating hyperplane by linear programming (e.g. perceptron):
  - separator can be expressed as ax + by = c

Step2: The decision function is fully specified by a subset of training samples, the support vectors.
Step3: Quadratic programming problem
Step4: Text classification method

V. CONCLUSION

In this paper, an approach that utilizes data mining and forensic techniques to identify the representative SC-patterns for a user is proposed. The time that a frequent SC- pattern appears in the user's log file is counted, the most frequently used SC-patterns are filtered out, and then a user's profile is established. By identifying a user's SC-patterns as his/her computer usage habits from the user's current input SCs, the IIDPS resists suspected intruders. The performance analyzing between proposed algorithm with SVM machine learning classifier for intrusion detection system.

REFERENCES

[1] Fang-YieLeu, Kun-Lin Tsai "An Internal Intrusion Detection and Protection System by Using Data Mining and Forensic Techniques" IEEE SYSTEMS.
[2] Q. Wang, L. Vu, K. Nahrstedt, and H. Khurana, "MIS: Malicious nodes identification scheme in network-coding-based peer-to-peer streaming," in Proc. IEEE INFOCOM, San Diego, CA, USA, 2010, pp. 1–5.
[3] S. Yu, K. Sood, and Y. Xiang, "An effective and feasible trace back scheme in mobile internet environment," IEEE Commun. Lett., vol. 18, no. 11, pp. 1911–1914, Nov. 2014.
[4] F. Y. Leu, K.W. Hu, and F. C. Jiang "Intrusion detection and identification system using data mining and forensic techniques," Adv. Info. Computer Security, vol.4752, pp. 137–152, 2007.
[5] H. S. Kang and S. R. Kim, "A new logging-based IP trace back approach using data mining techniques," J. Internet Serv. Inf. Security, vol. 3, no. 3/4, pp. 72–80, Nov. 2013.
[6] M. A. Faisal, Z. Aung, J. R. Williams, and A. Sanchez, "Data-streambased 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
[7] J. Choi, C. Choi, B. Ko, D. Choi, and P. Kim, "Detecting web based DDoS attack using Map Reduce operations in cloud computing environment", J. Internet Serv. Inf. Security, vol. 3, no. 3/4, pp. 28–37, Nov. 2013.
[8] Q. Chen, S. Abdel wahed, and A. Erradi, "A model-based approach to self-protection in computing system", in Proc. ACM Cloud Autonomic Comput. Conf., Miami, FL, USA, 2013, pp. 1–10.
[9] Z. A. Baig, "Pattern recognition for detecting distributed node exhaustion attacks in wireless sensor networks", Comput. Commun., vol. 34, no. 3, Mar. 2011.
[10] Z. Shan, X. Wang, T. Chiueh, and X. Meng, "Safe side effects commitment for OS-level virtualization", in Proc. ACM Int. Conf. Autonomic Comput., Karlsruhe, Germany, 2011, pp. 111–120
[11]S. Gajek, A. Sadeghi, C. Stuble, and M. Winandy, "Compartmented security for browsers—Or how to thwart a phisher with trusted computing," in *Proc. IEEE Int. Conf. Avail., Rel. Security*, Vienna, Austria, Apr. 2007, pp. 120–127.