An Internal Intrusion Detection and Protection System by Victimization Data Mining and Forensic Techniques

Miss. Komal .S. Gaikwad[1] Miss. Harsha .S. Bhujbal[2]
[1,2]Student
[1,2]AISSM's Polytechnic Pune, India

Abstract— Currently, most laptop systems use user IDs and passwords because the login patterns to manifest users. However, many folks share their login patterns with co-workers and request these co-workers to help co-tasks, thereby making the pattern as one of the weakest points of computer security. Insider attackers, the valid users of a system who attack the system internally, are hard to detect since most intrusion detection systems and firewalls identify and isolate malicious behaviours launched from the skin world of the system solely. In addition, some studies claimed that analysing system calls (SCs) generated by commands can identify these commands, with which to accurately detect attacks, and attack patterns are the features of an attack. Therefore, in this paper, 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 IIDPS creates users' personal profiles to keep track of users' usage habits as their forensic features and determines whether a valid login user is the account holder or not by scrutiny his/her current laptop usage behaviours with the patterns collected within the account holder's personal profile. The experimental results demonstrate that the IIDPS's user identification accuracy is ninety four.29%, whereas the response time is less than 0.45 s, implying that it can prevent a protected system from insider attacks effectively and efficiently.

Key words: Data Mining, Internal Intrusion Detection and Protection System (IIDPS)

I. INTRODUCTION

In the past decades, computer systems have been widely employed to provide users with easier and more convenient lives. However, once folks exploit powerful capabilities and process power of laptop system, security has been one among the intense issues within the laptop domain since attackers terribly sometimes strive to penetrate laptop 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 like pharming attack, distributed denial-of-service (DDoS), eavesdropping attack, and spear-phishing attack business executive attack is one among the foremost difficult ones to be detected because firewalls and intrusion detection systems (IDSs) sometimes defend against outside attacks. To manifest users, currently, most systems check user ID and password as a login pattern.

However, attackers might install Trojans to lift victims' login patterns or issue an oversized scale of trials with the help of a lexicon to amass users' passwords.

When winning, they will then log in to the system, access users' personal files, or modify or destroy system settings. Fortunately, most current host-based security systems and network-based IDSs will discover a proverbial intrusion in an exceedingly time period manner. However, it's terribly difficult to spot UN agency the assailant is as a result of attack packets ar typically issued with cast IPs or attackers might enter a system with valid login patterns.

Although OS-level system calls (SCs) ar way more useful in sleuthing attackers and distinguishing users, process an oversized volume of SCs, mining malicious behaviours from them, and identifying possible attackers for an intrusion are still engineering challenges.

Therefore, in this paper, we propose a security system, named Internal Intrusion Detection and Protection System (IIDPS), which detects malicious behaviours 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 (SC-sequence) that has repeatedly appeared several times in a user's log file for the user.

The user's rhetorical options, defined as an SC-pattern frequently appearing in a user's submitted SC-sequences but rarely being used by other users, are retrieved from the user's computer usage history.

The contributions of this paper are: 1) determine a user's rhetorical options by analysing the corresponding SCs to boost the accuracy of attack detection; 2) able to port the

IIDPS to a parallel system to further shorten its detection response time; and 3) effectively resist insider attack.

The remainder of this paper is organized as follows.

Section II introduces the related work of this paper. Section III describes the framework and algorithms of the IIDPS.

Experimental results ar shown and mentioned in Sections IV and V, severally.

Section VI concludes this paper.

II. RELATED WORKS

Computer forensics science, which views computer systems as crime scenes, aims to identify, preserve, recover, analyse, and present facts and opinions on information collected for a security event. It analyses what attackers have done such as spreading computer viruses, malwares, and malicious codes and conducting DDoS attacks. Most intrusion detection techniques focus on how to find malicious network behaviours and acquire the characteristics of attack packets, i.e., attack patterns, based on the histories recorded in log files. Qadeer et used self-developed packet sniffer to collect network packets with which to discriminate network attacks with the help of network states and packet distribution. O' Shaughnessy and Gray acquired network intrusion and attack patterns from system log files. These files contain traces of computer misuse. It means that, from synthetically generated log files, these traces or patterns of misuse can be more

accurately reproduced. Wu and Banzhaf overviewed research progress of applying methods of computational intelligence, including artificial neural networks, fuzzy systems, evolutionary computation, artificial immune systems, and swarm intelligence, to detect malicious behaviours.

The authors consistently summarized and compared completely different intrusion detection strategies, thus allowing us to clearly view those existing research challenges. These said techniques and applications really contribute to network security.

However, they can't simply certify remote-login users and find specific styles of intrusions, e.g., when an unauthorized user logs in to a system with a valid user ID and password. In our previous work, a security system, which collects forensic features for users at command level rather than at SC level, by invoking data mining and forensic techniques, was developed.

Moreover, if attackers use several sessions to issue attacks, e.g., multistage attacks, or launch DDoS attacks, then it is not easy for that system to identify attack patterns.

Hu et al presented an intelligent lightweight IDS that utilizes a forensic technique to profile user behaviours and a data mining technique to carry out cooperative attacks.

The authors claimed that the system might find intrusions effectively and expeditiously in real time.

However, they did not mention the SC filter. Giffin et al. provided another example of integration pc forensics with a knowledge-based system.

The system adopts a predefined model, which, allowing SC-sequences to be normally executed, is employed by a detection system to restrict program execution to ensure the security of the protected system. This is helpful in detecting applications that issue a series of malicious SCs and identifying attack sequences having been collected in knowledge bases.

When Associate in nursing undetected attack is given, the system frequently finds the attack sequence in 2 s as its computation overhead.

Fiore et al. explored the effectiveness of a detection approach supported machine learning mistreatment the Discriminative Restricted physicist Machine to mix the communicative power of generative models with smart classification accuracy capabilities to infer a part of its data from incomplete coaching knowledge so the network Associate in Nursingomaly detection theme will give an adequate degree of protection from each external and internal menaces. Faisal et al. [20] analysed the possibility of using data stream mining to enhance the security of advanced metering infrastructure through an IDS.

The advanced metering infrastructure, that is one in all the foremost crucial parts of positive identification, serves as a bridge

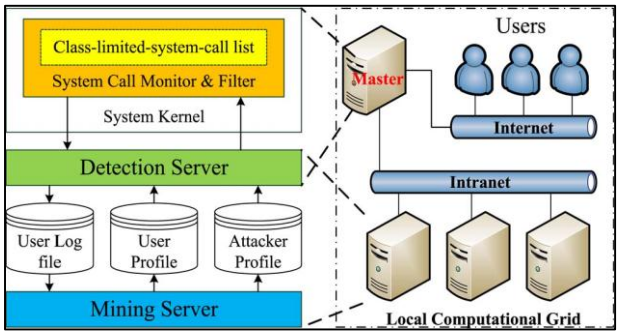

Fig. 1: IIDPS system framework.

For providing two-way info flow between the user domain and also the utility domain.

The authors treat IDS as a second-line security measure after the first line of primary advanced metering infrastructure security techniques such as encryption, authorization, and authentication.

III. IIDPS

In this section, we first introduce the IIDPS framework and describe components of the IIDPS in detail.
Two algorithms also are given for generating a user habit file and detective work an indoor trespasser.

A. System Framework

The IIDPS, as shown in Fig. 1, consists of an SC monitor and filter, a mining server, a detection server, a local computational grid, and three repositories, including user log files, user profiles, and an attacker profile.

The SC monitor and filter, as a loadable module embedded in the kernel of the system being considered, collects those SCs submitted to the kernel and stores these SCs in the format of in the protected system where uid, pid, and SC respectively represent the user ID, the process ID, and the SC c submitted by the underlying user, i.e., c ∈ SCs.

It conjointly stores the user inputs within the user's log file, which is a file keeping the SCs submitted by the user following their submitted sequence.

The mining server analyses the log data with data mining techniques to identify the user's computer usage habits as his/her behaviour patterns, which are then recorded in the user's user profile.

The detection server compares users' behaviour patterns with those SC-patterns collected within the assaulterprofile, called attack patterns, and those in user profiles to respectively detect malicious behaviours and identify who the attacker is in real time.

When Associate in nursing intrusion is discovered, the detection server notifies the SC monitor and filter to isolate the user from the protected system.

The purpose is to forestall him/her from unendingly assaultive the system.

Both the detection server and also the mining server area unit run on the native procedure grid to accelerate the IIDPS's on-line detection and mining speeds and enhance its detection and mining capability.

If a user logs in to the system by mistreatment another person's login pattern, the IIDPS identifies who the underlying user is by computing the similarity scores between the user's current inputs, i.e., SCs,

and the behaviour patterns stored in different users' user profiles.

In the IIDPS, the SCs collected in the class-limited-SC list, as a key component of the SC monitor and filter, are the SCs prohibited to be used by different groups/classes of users within the underlying system, e.g., a secretary cannot submit some specific privileged SCs.

Therefore, commands that generate these SCs are prohibited to be utilized by all secretaries.

IV. DISCUSSION

In this paper, an IIDPS is developed to detect insider attacks at SC level by using data mining and forensic techniques. The experimental results show that the IIDPS can effectively resist several aforementioned attacks. The outcome extends the features of, confirming that data mining and forensic techniques used for intrusion detection provide effective attack resistance.

The second experiment indicates that the average detection accuracy is 94.29%. However, in Table VI, the accuracy of user backup is 89.97% since backup's log file has more common SCs than the other users'.

It conjointly shows that the IIDPS could observe inaccurately once user's habit suddenly changes.

Nevertheless, in most cases, the IIDPS can still identify the legality of a login user. When a user inputs a command, hundreds or thousands of SCs will be generated. Analyzing a huge number of SCs often takes a long time. As shown in Table VII, the IIDPS spends 0.45 s to identify a user. Although alternative systems consume longer time for knowledge analysis than the IIDPS will, how to mine SCs in an efficient method should be addressed. Employing an area procedure grid will accelerate the process speed of the miming server and detection server.

Generally, users' forensic features retrieved from their basic operations are helpful in detecting the users' malicious behaviours and tell us who the possible attackers are. This can also detect malicious behaviours for systems employing GUI interfaces. However, many third party shell commands have been developed, including those used in Oracle Database, Oracle Web Logic, IBM Web Sphere MQ, and some user-developed applications. We need to study the SCs generated and the SC-patterns produced by these commands so that the IIDPS can detect those malicious behaviours issued by them and then prevent the protected system from being attacked. Additionally, mining user profiles by using an unsupervised cluster approach can also improve the performance of the mining process because processing big data is indeed an engineering challenge.

Moreover, to observe AN attack and scale back the corresponding latent period, we need a cluster workload monitor, a faster filter, an efficient detection algorithm, and a fault-tolerant environment provided by a computational grid. Furthermore, a mathematical analysis on the IIDPS's behaviours is helpful in deriving its formal performance and cost models, with which users can predict performance and cost of the IIDPS before using it. The model proposed in can be further used to increase detection accuracy and improve the decisive rate. Furthermore, one may ask how a behaviour record is created for a new user and how the IIDPS updates a user profile. The answer of the first question is that, when there is a new user k, the IIDPS creates k's log file, habit file, and user profile on his/her first login and then follows the procedure shown in Fig. 3 to generate k's user profile. The answer of the second question is that, each time when k logs in to the system, in addition to the SC k that directly submits, the IIDPS also identifies those SC-sequences generated by submitted commands. These SC sequences are then used to detect whether k is an attacker or not by invoking algorithm 2. When k is recognized as a legal user, the IIDPS continues collecting k's SCs and updating his/her profile according to the procedure shown in Fig. 3 on his/her logout.

Once k's user profile is updated, it can be used for next time detection.

V. CONCLUSION

In this paper, we have proposed an approach that employs data mining and forensic techniques to identify the representative SC-patterns for a user.

The time that a habitual SC pattern seems within the user's log file is counted, the most commonly used SC-patterns are filtered out, and then a user's profile is established.

By characteristic a user's SC-patterns as his/her laptop usage habits from the user's current input SCs, the IIDPS resists suspected attackers.

The experimental results demonstrate that the typical detection accuracy is above ninety four once the decisive rate threshold is zero.9, indicating that the IIDPS can assist system administrators to point out a business executive or AN offender during a closed setting.

The any study is done by up IIDPS's performance and investigation third-party shell commands.

REFERENCES & BIBLIOGRAPHY

[1] U. Fiore, F. Palmieri, A. Castiglione, and A. D. Santis, "Network anomaly detection with the restricted physicist machine," Neurocomputing, vol.
[2] 122, pp. 13–23, Dec. 2013.
2 B. Sayed, I. Traore, I. Woungang, and M. S. Obaidat, "Biometric authentication victimisation mouse gesture dynamics," IEEE Syst. J., vol. 7, no. 2, pp. 262–274, Jun. 2013.
[3] H. Lu, B. Zhao, X. Wang, and J. Su, "DiffSig: Resource differentiation-based malware behavioral brief signature generation," Inf. Commun. Technol., vol. 7804, pp. 271–284, 2013.
[4] S. C. Arseni, E. C. Popovici, L. A. Stancu, O. G. Guta, and S. V. Halunga, "Securing AN alerting system for a keystroke-based user identification system," in Proc. Int. Conf. Commun., Bucharest, Romania, 2014, pp. 1–4.
[5] G. M. Amdahl, "Validity of the one processor approach to achieving giant scale computing capabilities," in Proc. AFIPS Spring Joint Comput. Conf., New Brunswick, NJ, USA, 1967, pp. 1–4.

