International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 03 Issue: 07 | July -2016 www.irjet.net p-ISSN: 2395-0072

Internal Intrusion Detection Using Data Mining and Behaviometric Technique

Bini V. C.[1], Ms. Nimmy K.[2], Prof. P. Jayakumar[3]

[1]*MTech Cyber Security, Dept. of CSE, SNGCE, Kadayiruppu, Kerala, India*
[2]*Asst. Prof., Dept. of CSE, SNGCE, Kadayiruppu, Kerala, India*
[3]*Prof. & Head, Dept. of CSE, SNGCE, Kerala, India.*

Abstract - *Internal intrusions is one of the serious problem in the computer domain. Most of the computer system uses username and password as login pattern to enter in to the system. If any user shares their login pattern to their friends or co-workers then it will be the weakest point of security. Fang Yie Leu et al. proposed a security system named as Internal Intrusion Detection and Protection System (IIDPS) to detect the internal intrusion. In this system they analyze the System Calls (SCs) to identify the user computer usage habit. In this paper we present a security system named as IIDS. It is the combination of behaviometric technique and IIDPS. The behaviometric technique here we used is the Typing Speed of user for continuously authenticate the user's identity. Using typing speed we can uniquely identify users. Experimental analysis shows that typing speed has an accuracy of 96% shows that typing speed is a strong authentication mechanism. So the combinations of these two systems will increases the accuracy to detect the internal intruders.*

Key Words: System Calls (SC), Data Mining, Behavioral Biometrics.

1. INTRODUCTION

Computer security is one of the serious problems in the computer domain. Attackers are very usually trying to penetrate the computer security and behave maliciously. Intruders mainly grouped into two types; they are Internal Intruders and External Intruders. Internal intruders are the persons have some access privileges in the network and they are trying to penetrate the security system intentionally or unintentionally. Internal intruders are very difficult to detect in the network. The External intruders are the outsiders from the network.

The security systems like Intrusion Detection Systems (IDS) and firewalls usually block the attacks from outside network so the insider attack is one of the hardest attacks to be detected. The insiders have some access privileges in the network and using that privilege they are trying to penetrate the security in the network. To authenticate users computer systems use different type of authentication techniques. Authentication through username and password is the commonly used technique. If anyone shares their login pattern such as username and password to their friends or co-workers then it will be the weakest point of security. Authenticate user through the biometric techniques is one of the strongest authentication mechanism. Biometric authentication [1] can be divided into physical biometric authentication and behavioral biometric or behaviometric authentication. Physical biometric authentication uses the finger, retina, face, palm etc. for authenticate the user. Behavioral biometric includes the user behavior such as typing speed, typing sound on a keyboard or keypad. It is also known as keystroke dynamics. Because of uniquely identify user, it is more popular among strong authentication techniques.

In this paper we explain a security system called Internal Intrusion Detection System (IIDS) using data mining and bevaviometric technique to detect the internal intrusion. The basic idea of IIDS is IIDPS [2]. Along with IIDPS we use a continuous authentication mechanism using typing speed to authenticate the user.

IIDPS can detect the internal attacks at system call level. IIDPS store the user's computer usage habit by analyzing the system calls sequences that has stored in the user's log file. IIDPS can block internal intruders and can identify the intruders [3] in the network.

2.INTERNAL INTRUSION DETECTION SYSTEM

Internal Intrusion Detection system (IIDS) contains an authentication module and an IIDPS system. The behavioral authentication using typing speed is a strong authentication method it uniquely identify the user. The typing speed of a user can be calculated by the total time taken for typing. The typing speed of a person should not be same. Different users have different typing behavior so we can say that it is a strong authentication mechanism. Keystroke dynamics [4] along with IIDPS, it will increase the internal intruder detection accuracy.

Fig -1 gives the overview of IIDS. Before entering to the IIDPS system, authenticate the user using typing speed. If the typing speed is match to the original user we can entered into the IIDPS system. Otherwise the authentication system itself blocks the unauthorized user to login into the system.

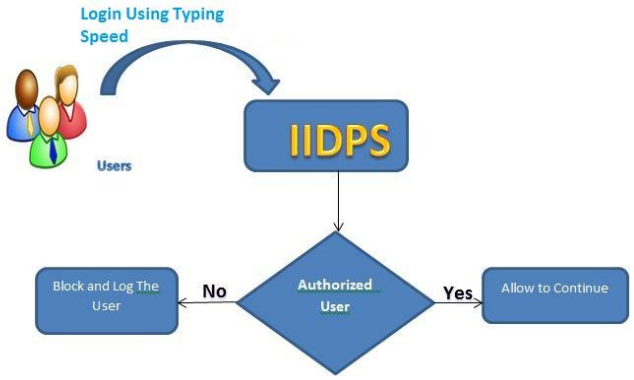

Fig -1: Overview of IIDS.

The overall typing speed of a user is based on the dwell time and flight time of a user. Dwell time is the total time taken for press and hold a key and the flight time is the time taken for find the right key. The Fig -2 shows the dwell time and flight time measurement [5].

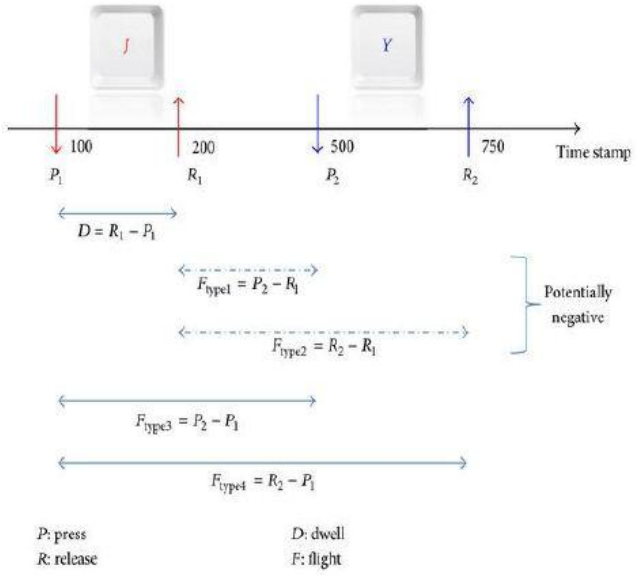

Fig -2: Flight and Dwell Time Measurement.

2.1 Internal Intrusion Detection and Protection System Framework

Fang-Yie Leu et al [2] proposed a security system named as Internal Intrusion Detection and Protection System (IIDPS). The main modules of this system are SC Monitor and Filter, Detection Server, Mining Server, Computational Grid and three Repositories – User Log File, User Profile and Attacker Profile. The Fig -3 shows the IIDPS system framework.

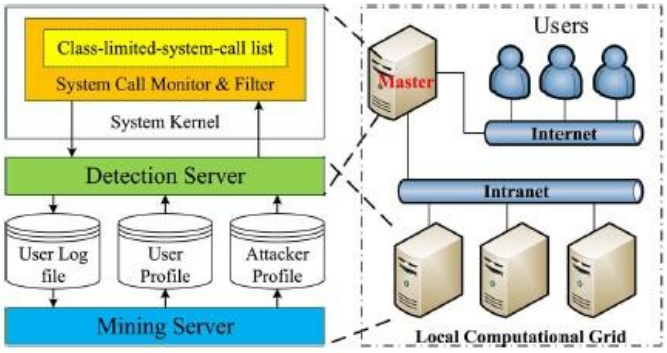

Fig -3: IIDPS System Framework

The main module of IIDPS is the SC Monitor and Filter which is a loadable module in kernel of the system. It collect all the SC submitted to the kernel and store in the user log file like <u_id,p_id,SC>. U_id is the user ID, p_id is the process ID and SC is the system call submitted by the underlying user. The Mining Server analyzes the user's log file using mining techniques. It identifies user's computer usage habit as user behavior pattern and saves to user profile. The detection server compares the user's behavior pattern with the SC pattern collected in the attacker profile. If any malicious pattern detected, it notifies the SC Monitor and Filter to isolate the user from the protected system. Using this IIDPS can detect the intruders in real time. The computational grid accelerates the IIDPS real time detection. Both detection sever and mining server run on the local computational grid. The generation of user profile can be explains using the control flow diagram. The Fig -4 shows the control flow diagram of generating user profile.

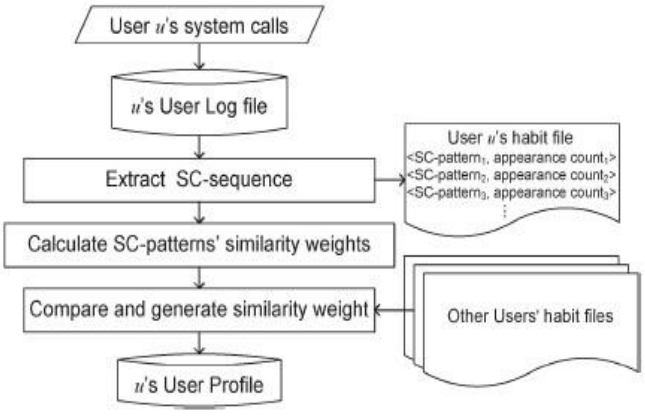

Fig -4: Control Flow of the Generation of a User Profile.

The Mining server creates the user profile using the SC's collected in the user log file. Mining server extract specific SC patterns using data mining techniques [2] from the user log file and store in the user's habit file. After this SC pattern's similarity weights are calculated and compare it with the

other users habit file and make sure that none of them have same SC patterns.

The Fig -5 shows the algorihm1 [2] for generating the user's habit file. In this u is the user of the system. The algorithm1 compare k-grams and k'-grams. The consecutive system calls up to the sliding window is the k-grams k=1,2,3,….|sliding window|. K'-grams is also the system calls for comparing with the k-grams.

Algorithm 1: The algorithm for generating a user habit file
Input: u's log file where u is a user of the underlying system
Output: u's habit file

1. G = |log file| - |Sliding window| ;
   /* |Sliding windows|=|L-window|=|C-window| */
2. for (i=0 ; $i \le G-1$; i++) {
3. for (j=i+1 ; $j \le G$; j++) {
4. for (each of $\sum_{k=2}^{|\text{Sliding window}|}(|\text{Sliding window}| - k + 1)$ k-grams in current L-window){
5. for (each of $\sum_{k'=2}^{|\text{Sliding window}|}(|\text{Sliding window}| - k' + 1)$ k'-grams in C-window){
6. Compare the k-grams and k'-grams with the longest common subsequence algorithm;
7. if (the identified SC-pattern already exists in the habit file)
8. Increase the count of the SC-pattern by one;
9. else
10. Insert the SC-pattern into the habit file with count=1; }}}}

Fig -5: Algorithm1 for Generating a User's Habit File.

Detection server detects the internal intrusion using the algorithm 2 [2]. The Fig -5 shows the algorithm 2 for detecting the internal intruder or an attacker. In this algorithm u is the user. Detection server tries to identify the underlying user is an account holder or not by calculating the similarity score between the newly generated SCs, denoted by NSCu, in the u's current input and usage habit stored in the in user's user profile to verify u.

There are three types of attacks [2] are blocked by IIDPS. They are 1. A user of specific groups submits an SC, which the group members are prohibited to use. 2. An attack that launches a sensitive SC, which is defined as one that may erase or modify sensitive data or system settings, to change the environmental settings of the system or attack the system. 3. SC level attack patterns that are an attacker mixing specific SC can sometimes penetrate a security system.

Algorithm 2: Detecting an internal intruder or an attacker
Input: user u's current input SCs, i.e., NSC_u, (each time only one SC is input), and all users' user profiles
Output: u is suspected as an internal intruder or a known attacker

1. $NCS_u = \emptyset$;
2. while (receiving u's input SC, denoted by h) {
3. $NCS_u = NCS_u \cup \{h\}$;
4. if ($|NCS_u| >$ |Sliding window|) {
5. L-window = Right(NCS_u, | Sliding window|); /* Right(x, y) retrieves the last L-window of y from x */
6. for (j= $|NCS_u|$ - |Sliding window| ; j >0 ; j--) {
7. C-window = Mid (NCS_u, j, |Sliding window|);/* Mid (x, y, z) retrieves a sliding window of size z beginning at the position of y from x */
8. Compare k-grams and k'-grams by using the comparison logic employed in Algorithm 1 to generate NHF_u;}
9. for (each user g, $1 \le g \le N$)
10. Calculate the similarity score $Sim(u, j)$ between NCS_u and g's user profile by invoking Eq. (8);
11. if (($|NCS_u|$ mod paragraph size) == 0) { /* paragraph size = 30, meaning we judge whether u is an attacker or the account holder for every 30 input SCs */
12. Sort similarity scores for all users;
13. if ((the decisive rate of u's user profile < threshold_1) or (the decisive rate of attacker profile > threshold_2)){
    /* threshold_1 is the predefined lower bound of average decisive rate of user u's user profile, while threshold_2 is the predefined upper bound of average decisive rate of attacker profile*/
14. Alert system manager that u is a suspected attacker, rather than u himself/herself; }}}}

Fig -6: Algorithm2 for Detecting Internal Intruder or an Attacker.

3. RESULT AND ANALYSIS

In this section we examine the efficiency of the methods that we added to the IIDPS system to improve the internal intrusion detection.

Table -1: The total time taken by typing same username =hello and password=save.

User	Total Time(Mille Seconds)
1	893
2	973
3	820
4	908
5	865

Using the Table -1 shows the total time taken by different users for typing the same username=hello and password =save.

By analyzing the time in the table it is clearly shows the typing speed is unique for each user. We can draw a chart using this.

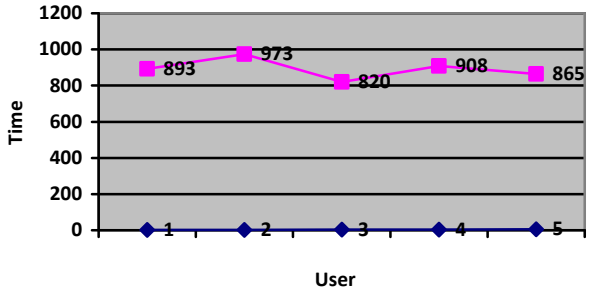

Chart -1: Typing Time comparison Chart.

Char -2 shows the ROC for typing speed. ROC plots the True Positive rate of detection as the false positive rate increases.

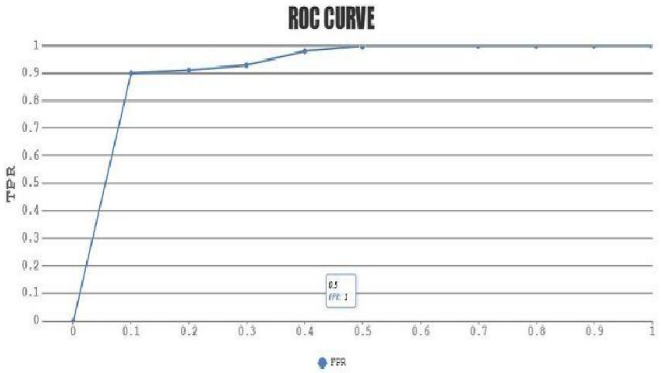

Chart -1: ROC for Typing Speed.

To calculate the accuracy of Typing speed, the experiment is conducted with 25 users. By analyzing the Table -2 we can calculate the accuracy. The True Positive Rate is 0.92 and False Positive Rate is 0 so the accuracy of the system is 96.

Table -2: Analysis of Typing Speed

Total Population=25	Positive	Negative	Total Row
Genuine User	23	2	25
Imposter	0	25	25
Total Column	23	27	50

Analysis of Typing Speed has an accuracy of 96% shows that it is a strong method to authenticate the users.

4. CONCLUSION

The Behavioral Biometric Authentication such as Typing Speed is a strong authentication mechanism to authenticate the user's identity. The existing system, IIDPS is proposed by Fang Yie Leu et al, has an accuracy of detection of internal intruder is above 94%. The Typing Speed method has an accuracy 96% .The combination of these two systems named as IIDS, will increases the accuracy of Internal Intrusion Detection.

REFERENCES

[1] Sandhya Avasthi and Tanushree Sanwal, "Biometric Authentication Techniques: A Study on Keystroke Dynamics," International Journal of Scientific Engineering and Applied Science (IJSEAS), vol.1, ISSN: 2395-3470, Jan. 2016.

[2] Fang Yie Leu and Kun Lin Tsai, "An Internal Intrusion Detection and Protection System by Using Data Mining and Forensic Techniques ", IEEE Systems Journal,2015.

[3] 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, pp. 111–120, 2011.

[4] S. C. Arseni, E. C. Popovici, L. A. Stancu, O. G. Guta, and S. V. Halunga, "Securing an alerting subsystem for a keystroke-based user identification system," in Proc. Int. Conf. Commun., Bucharest, Romania, pp. 1–4,2014.

[5] Ankur Kumar, Abhijeet Patwari and Sagar Sabale, "User Authentication by Typing Pattern for Computer and Computer based devices," International Journal of Advanced Research in Computer and Communication Engineering, Vol. 3, Issue 10, October 2014.