I J C T A, 9(11) 2016, pp. 5331-5334

USER AUTHENTICATION USING ONE-TIME PIN CODES AND KEYSTROKE DYNAMICS: A REVIEW

Amarjeet Badhani[*] and Baljit Singh Saini[**]

Abstract: Computer is the main source of information in today's world in every field of technology or study. With increase in the need of computer to store and process the sensitive information the threat of intruders also increases day by day. Keystroke dynamics is one of the methods to increase the security of your system as it is cheap and effective. Keystroke dynamics make use of typing pattern of every individual to identify him. One of the technique can be used is one time pin codes which changes randomly every time a user use it.

Key Words: Keystroke Dynamics, Latency, Authentication, One-Time Password.

1. INTRODUCTION

Keystroke dynamics is biometric ways which identify the user on the basis of its typing pattern [1]. In the 21st century everything is related to computers to store the data and forward the data. As the use increases the security risk of theft of data also increases, to protect data we use password but which are easily cracked by the hackers in today's world [2]. It is very essential to protect the computer network and data to build a new system [6]. There are two types of authentication in keystroke dynamics. Continuous authentication (structure and value are not fixed) and Static authentication (structure and value which is to be entered as in the text form are fixed) [10]. Two terms which are generally used in biometric are Authentication (here the user is verified whether it is who he poses himself) and Identification (a user having id is processed) [13]. As in current time people mostly use mouse to interact with computer so there is use of mouse dynamics (make use of pointing device (mouse) or touch pad) [15]. Keystroke dynamics is cheap solution to the problem of security threat in this user has not to worry about remembering of password or loss of password as it is different from one user to another and very hard to be used by someone else [7]. Keystroke dynamics can also help to find the malware in a user personal computer or laptop [4].

2. DATA ACQUISITION

Firstly a user is given a paragraph to see their timing pattern and key rhythm which is stored along with the paragraph of user, then during the authentication time when the user enters his id and password then it is typing rhythm is matched with stored one from the database if it verifies or

[*] Computer Science and Engineering, Lovely Professional University, Phagwara, India
Email: amarjeetbadhani@gmail.com

[**] Asst. Professor: Computer Science and Engineering, Lovely Professional University, Phagwara, India Research Scholar: Sri Guru Granth Sahib World University, Fatehgarh Sahib, Punjab, India Email: baljit.15359@lpu.co.in

match the user is allowed to proceed further and if not match then access is denied for that user [9], [13]. There are different ways to get data from the users.

(a) **Text entry:** Text can be classified into two types. Static or structured (which is fixed or common for all) and dynamic or free type (text according to the user choice or different for everyone). For example a static text can be like a pin code of length size of 6 digits varying values from 000000 to 999999 and dynamic can be of any size length and any value of user choice as it is not fixed [10]. In our experiment we will take about 50 users of different age and different profession like students, teachers etc. We will provide the same keyboard to the entire users and ask them to type the one time password and their typing pattern will be stored and will be matched during the authentication time.

(b) **Environment:** It is the important part of the experiment as it affects the typing pattern of the user. Environment can be categories in two types. Firstly controlled environment in which the machine, room light and other features are design or used according to the need of the researchers and secondly is the uncontrolled environment in which researchers have only partial control over the information entered by the users [13], [15].

(c) **Features:** Latency is the most common feature used in the keystroke dynamics. Latency can be categories in three different types that are press-to-press, release-to-release and release-to-press [3]. Flight time (release-to-press) can be defined as releasing of one key and pressing of another key. Dwell time is define as pressing of one key and releasing of same key [8].

(d) **Errors:** There are three types of error among which the first one is False acceptance rate(FAR) is define as the numbers of invalid user which do not have authentication are allowed in the system to use as the valid user in system [8]. Second is False rejection rate (FRR) is define as the number of valid user rejected to use the system posing them as invalid or impostor users. Thirdly is Equal error rate (EER) is define as the mean taken of both FAR and FRR. The less EER obtained in the experiment better the result is to be consider of the research. FRR is also said to be the type 1 error and FAR is also said to be type 2errors. . If we have high type 1 error then type 2 errors it is considered to be good result because there is less number of impostors in the system which will be helpful and beneficial for the system.

3. APPROACHES

Approaches are used for the classification of users of different algorithms used by the researchers in their research. Different type of algorithm are used for the same experiment as to compare the result of the experiments and find out which one gives us the better result[10], [13], [7].

(a) **Statistical Algorithms:** These types of algorithms are used for computing the mean and standard deviation in the experiment for the different values taken by the researcher. It is used for the comparison using assumption testing, t-tests and distance events such as absolute distance, weighted absolute distance, Euclidean distance, Manhattan distance etc. Statistical algorithms has disadvantage of lack of training which is very essential part of identifying the patterns of user in the keystroke dynamics which result into the poor result of the experiment [10], [13].

(b) **Neural Networks:** It is non-linear statistical data modeling tools which make the use of neurons interconnection for the result of the experiment. In this we have two types of learning that are supervised learning and unsupervised learning. Supervised learning makes the use of back propagation technique of its other different techniques while unsupervised learning makes the use of Hopfield neural networks [13]. Neural networks have the advantage of using multiple parameters in their algorithms but it make them slowly processing in the training and

application phase of research. It also have the disadvantage of selecting features for experiment in the classification as it is cannot be seen from inside. It can make problem when it is use in real time for continuous keystroke authentication.

(c) **Pattern Recognition and Learning based algorithms:** The way of using the pattern or objects in different algorithms to classifying them in different categories. It makes the use of simple machine learning algorithms such as nearest neighbor algorithms and clustering [5]. Support vector machine (SVM) one of the supervised learning technique which can be used for both identification and authentication purpose in the research [9]. Probabilistic learning provides us confidence value by ignoring the outputs of low confidence values in the experiment whereas unsupervised learning techniques identify the patterns in the data automatically [7].

4. FACTORS AFFECTING THE PERFORMANCE

There can be multiple factors affecting the performance such as the type of text used by the researcher in his research is structured text or unstructured text. Another factor can be overloading of the system due to more amount of data to be processed which can affect the timing inaccuracy problem for this the system should be made in robustness form. The length of the text can be one of the factors as longer text make easy to identify the User while short text make difficult for identifying the user. Number of samples taken for the experiment will also result in the error rate of the system which can be recognized one more factor for affecting the performance. One of the main factors can be the state of a user, algorithm used and the position of the user while he or she is typing the text [11].

4.1 Proposed theory

In my research I am going to use the one time pin codes (OTP) of 8 digits where the length size is fixed but value will vary from 00000000 to 99999999. As it is was there in paper of one time pin code that with increase in length the equal error rate decreases [10]. There will be about 50 participants from which the data will be collected for the research in which there will be some common OTPs to all participant and some personal OTPs to all participants. Distance metrics algorithms will be used such as Scaled Manhattan Distance (SMD), Scaled Euclidean Distance (SED) and Adapted Scaled Euclidean Distance (ASED). Data collection tool software will be built for the collection of data from the users which will include an external keyboard for the typing of numbers from (0, 0) to (9, 9) from 10 to 20 times on both sides from number pad and number row which will be taken as the training data and the timing of user for pressing each key and releasing each key every time will be recorded. In phase two of this tool a testing phase will be taken where the 20 fix eight digits one time pin code will be ask from user to enter and it will be matched from the training data of that user. Later the weka toll will be used for analyzing the data of data collection software to find the error percentage.

4.2 Security issues which can be resolved by keystroke dynamics

(a) **Shoulder surfing and user mimicking:** It is an attack in which the invalid user tries to use typing pattern of a valid user behavior [3].

(b) **Spyware:** It is software which is use to record the accurate timing pattern of user when he pressed the key and when he release it [3].

(c) **Social engineering:** It is the way to get the private information of the user by manipulating and tricking him or her [3].

(d) **Guessing:** It is not very easy to guess anyone typing especially in the free text or the dynamic text used for authentication [3].

5. CONCLUSION

As keystroke dynamics does not depend on what the user is writing but how he or she is writing so it make more easy to find the invalid users in the system trying to authenticate as the valid users. It is one of advance feature or a way to increase the security in the modern world to prevent the hackers from their ways to enter in a system. From the earlier research papers we can see that some of the features are more important and useful to us in comparison of other features. Future work with the combination of features can increase the accuracies of or research and provide us better result than we have seen from the existing literature.

References

[1] A. Ahmed and I. Traore, "Biometric Recognition Based on Free-Text Keystroke Dynamics", IEEE TRANSACTIONS ON CYBERNETICS, VOL. 44, NO. 4, APRIL 2014.

[2] A. Alsultan and K. Warwick, "User-Friendly Free-text Keystroke Dynamics Authentication for Practical Applications", IEEE International Conference on Systems, Man, and Cybernetics 2013.

[3] D. Shanmugapriya and Dr. G. Padmavathi, "A Survey of Biometric keystroke Dynamics: Approaches, Security and Challenges", (IJCSIS) International Journal of Computer Science and Information Security, Vol. 5, No. 1, 2009.

[4] D. Stefan and D. Yao, "Keystroke-Dynamics Authentication against Synthetic Forgeries", International Conference on Collaborative Computing 2010.

[5] F. Monrose and A. Rubin "Keystroke dynamics as a biometric for authentication", Elsevier Science 2000.

[6]. F. Bazrafshan, A. Javanbakht and H. Mojallali, "Keystroke Identification with a Genetic Fuzzy Classifier", IEEE 2nd International Conference on Computer Engineering and Technology 2010.

[7] G. Azevedo, G. Cavalcanti and E. Filho "Hybrid Solution for the Feature Selection in Personal Identification Problems through Keystroke Dynamics", IEEE Proceedings of International Joint Conference on Neural Networks, Orlando, Florida, USA, August 12-17, 2007.

[8] L. Araújo, L. Sucupira ,G. Lizárraga, L. Ling, and J. Yabu-Uti, "User Authentication through Typing Biometrics Features", IEEE TRANSACTIONS ON SIGNAL PROCESSING, VOL. 53, NO. 2, FEBRUARY 2005.

[9] M. Antal, L. Szabó and I. László, "Keystroke Dynamics on Android Platform", Science Direct 8th International Conference Interdisciplinary in Engineering, INTER-ENG, Tirgu Mures, Romania, 9-10 October 2014.

[10] P. Bours and E. Masoudian, "Applying Keystroke Dynamics on One-Time Pin Codes", IEEE International Workshop on Biometrics and Forensics (IWBF) 27- 28 march 2014.

[11] R. Giot, M. El-Abed and C. Rosenberger, "Dynamics Authentication for Collaborative Systems", IEEE International Symposium on Collaborative Technologies and System (CTS) 18-22 may 2009.

[12] S. Bleha and M. Obaidat, "Computer Users Verification Using the Perceptron Algorithm", IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS, VOL. 23, NO. 3, MAY/JUNE 1993

[13] S. Banerjee and D. Woodard, "Biometric Authentication and Identification using Keystroke Dynamics: A Survey", Journal of Pattern Recognition Research 7 (2012) 116-139. Received April 28, 2012. Revised July 10, 2012. Accepted July 1, 2012.

[14] S. Bhatt and T. Santhanam, "Keystroke Dynamics for Biometric Authentication –A Survey", International Conference on Pattern Recognition, Informatics and Mobile Engineering (PRIME) 2013.

[15] Z. Jorgensen and T.Yu,"On Mouse Dynamics as a Behavioral Biometric for Authentication", ACM New York, USA 2011