A Keystroke Dynamics Based Authentication System

Moitrei Bharadwaj[1], Amreesha Kakati[2], Ankita Bharali[3] and Aniruddha Deka[4]

[1,2,3]Royal School of Engineering and Technology
[4]Department of Computer Science & Engineering Royal School of Engineering and Technology
Email: [1]moitreibharadwaj1871, [2]amreesha35, [3]ankitabharali @gmail.com, [4]dekaaniruddha@gmail.com

Abstract—*In this paper we briefly discuss about the various authentication techniques and we primarily give our importance to one specific technique that is "Keystroke Dynamics". It is a type of biometrics which takes into accounts not only the user id and password as a measure of security for authentication but also how an individual types a password. We did a literary survey based on this technique and a brief study about the research works done on this topic both nationally and internationally .We have proposed an authentication system based on Keystroke Dynamics with five modules namely registration module, authentication module, password verification module, identification module and decision module. Any user accessing the system is asked to type a few words in conjunction with his/her user name and password. Access is granted if his/her typing pattern matches the pattern which has been already stored in the database. This safeguard is effective as there is usually no remote access allowed to the system and the only entry point is via console login.*

Keywords*: Authentication, Keystroke Dynamics, Biometric.*

1. INTRODUCTION

In today's era, with the advancement of technology, the rate of crimes such as credit card frauds, security breaches etc. have also increased. With such high crime rates, network security and authentication has become a major concern of today's IT departments.

In this paper, we focus on a technology which can help us build a strong authentication system that can overcome most of the hurdles faced by the traditional user id-password based authentication system. This technology, known as keystroke dynamics, is based on a technique called biometrics. The word biometric is derived from the Greek word 'bio' means life and 'metrics' means to measure. It is a technique which takes into account the biological properties of an individual for authentication, which is difficult to fake. Some of the technologies based on biometrics are- keystroke dynamics, face recognition, iris recognition, dynamic signature analysis, etc.

Keystroke dynamics is a type of biometrics which takes into accounts not only the user id and password as a measure of security for authentication but also how an individual types a password. Thus, keystroke dynamics is based on the habitual rhythm patterns in the way a person types.

2. OBJECTIVE OF THE STUDY

Password based authentication system is a traditional and efficient way to authenticate a user. The user mainly has two information with him, that is, the username which is known by almost everyone and his/her password which is known only by the user. The user generally accesses the account by entering his/her username and password. The system aims in developing a secure and effective security system for protecting the computer applications and data based on typing pattern biometric.

The proposed system which is based on typing biometric will be an efficient, low cost and secured security system. The authentication technique proposed here works on two modes, registration mode and verification mode. In the registration mode, the username and password is being typed and the typing pattern values are recorded. The values of the typing pattern are captured and stored. In the verification mode, the user is asked to enter the username and password about 10 times and is matched with the stored values.

3. LITERATURE SURVEY

Keystroke dynamics is a technology based on biometrics which helps to minimize the security problems faced by traditional user id-password based on authentication systems, as it depends on not only the password but also on how a person types a password i.e. dwell time (amount of time a key is pressed) and flight time (amount of time between releasing one key and pressing the next). Keystroke dynamic system can measure one's input up to 1000 times per second. Like any other biometric technique, it also requires a reference template, which involves several sessions of a person using keystroke dynamic system so that the system can construct the reference template by detecting one's typing rhythms. Thus keystroke dynamics is based on the habitual rhythm patterns in the way a person types, and takes into account the following-

a) Latencies between successive keystrokes
b) Duration of each keystroke
c) Finger placement
d) Pressure applied on keys
e) Overall typing speed

Some of the classifiers used in this method for recognition are Euclidean Distance Measure, Non-Weighted Probability Measure, and Manhattan Distance Measure etc.

The use of keystroke dynamics for verification and identification purposes was first proposed and investigated back in the 1970s. Spillane, in the year 1976, suggested in an IBM technical bulletin that typing rhythms might be used for identifying the user at the keyboard. In the year 1977, Forsen et al. conducted preliminary tests of whether keystroke dynamics could be used to distinguish typists and it was found that a person typing his or her name is distinguishable from another person typing the same name. The research conducted by Gaines et al. in 1980 showed that the keystroke dynamics field was effectively initiated during the initial manual phase of telegraphy, where operators had been observed to have a unique "fist" (tapping style) by which their colleague could often identify them. [1]

There have been a number of researches conducted by different individuals. We will summarize some of them below-

In the study of keystroke dynamics user authentication based on Gaussian Mixture Model (GMM) and Deep Belief Nets (DBN) by Yunbin Deng and Yu Zhong, it was found that the performances, measured in Mean and Standard deviation of equal error rate, in case of GMM-UBM was 0.055 (0.052) and that in case of DBN was 0.035 (0.027). [2]

In another paper proposed by FabianMonrose and Avial D. Rubin, it was found that the correct identification rate using the Weighted Probabilistic classifier was approximately 87.18% on a data set of 63 users whereas the performance using Euclidean distance was 83.22% and that by the Non-Weighted scoring approach was 85.63%. [5]

Shivshankar Rajput and PriyankaVijayawargiya proposed an implementation of keystroke dynamics for identifying emotional state and it was found that, for a data size 5, the overall accuracy using Euclidean distance function and Manhattan distance function was 80% and 70% respectively and mean absolute error was 0.8 and 0.3 respectively. Again for a data size 10, the accuracy using Euclidean distance function and Manhattan distance function was 84% and 70% respectively and the mean absolute error was 0.4 and 0.5 respectively. [4]

In the paper titled-"Keystroke dynamics as a biometric for authentication" by Fabian Monrose and Aviel D. Rubin published in 1999, they examined an emerging non-static biometric technique that was aimed to identify users based on analyzed habitual rhythm, patterns in the way they type. They argued that the use of keystroke rhythm in a natural choice for computer security. When a person types, the latencies between successive keystrokes, keystrokes durations, finger placement and applied pressure on the keys can be used to construct a unique signature for that individual. [3]

In 2013, Yunbin Deng and Yu Zhong in their paper titled-"Keystroke Dynamics User Authentication based on Gaussian Mixture Model and Deep Belief Nets" introduced two new algorithms to the domain: the Gaussian mixture model with the universal background model(GMM-UBM) and the deep belief nets(DBN).[2]

In 2015, Shivshankar Rajput, PriyankaVijayawargiyan in their paper titled-"Implementation of Keystroke Dynamic applications for Identifying Emotional State" described the concept based on using standard input devices as sources of data recognition of user's emotional states. In this paper, they investigated emotional states via Keystroke Dynamics. [4]

The three methods that typically describe the performance of a biometric authentication system are as follows:

- False rejection rate (FRR) - the percentage of valid user attempts identified as imposters.
- False acceptance rate (FAR) - the percentage of imposter access attempts identified as valid users.
- Equal error rate (ERR) **-** the point at which FRR is equal to FAR.

Many commercial systems are being developed based on Keystroke dynamics to authenticate a user. Some of the systems are as follows:

a) Coursera: it is an online education site which uses keystroke recognition in order to authenticate and identify a user specifically enrolled in course through the site's "Signature Track" feature.
b) Admit one security: it is a patented commercial system that uses keystroke dynamics in addition to other biometrics methods. It links the user to their digital identity. The Admit one security system can also detect online fraud.
c) DelFig. o security: it is a system that provides multifactor risk based authentication. The system uses the keystroke biometric to create a unique identity of an individual.
d) ID control: this systems provides a keystroke ID which delivers a low FRR and FAR for identification. Keystroke ID is easy to enrol through their fully centralized identity management solution server.

Thus, we can see that keystroke dynamics is quickly emerging biometric technique and its accuracy and efficiency is increasing day to day. Keystroke dynamics is also preferred because no special equipment is required for installation and thus it minimizes the cost.

4. WORKING PRINCIPLE

In the system, initially the user is directed to an authentication page. If the password entered by the user matches the password typing pattern stored in the database, then the user can have an access to the account. Otherwise, he/she will be rejected and asked to re-enter the password.

If the person does not have a login ID, then he/she will be directed to the registration page. In the page, he/she is asked to enter the password 10 times and the press time of the keystrokes is calculated and stored in the database. Again when the person tries to log into the account, the press time of the keystrokes of his/her password is compared with the previously stored press times. If they match, he/she can have an access to the account, otherwise will be rejected.

Module Description

In our proposed system we have use the following modules:

a) Registration
b) Password Verification
c) Identification
d) Verification
e) Decision

a. REGISTRATION

The data collector will collect the data with User Name, password and 10 Reference Samples and stores it into the database. The raw data will undergo cluster analysis and verification units training to create Clusters and Matrix. The cluster and matrix with the user name will be stored in the database.

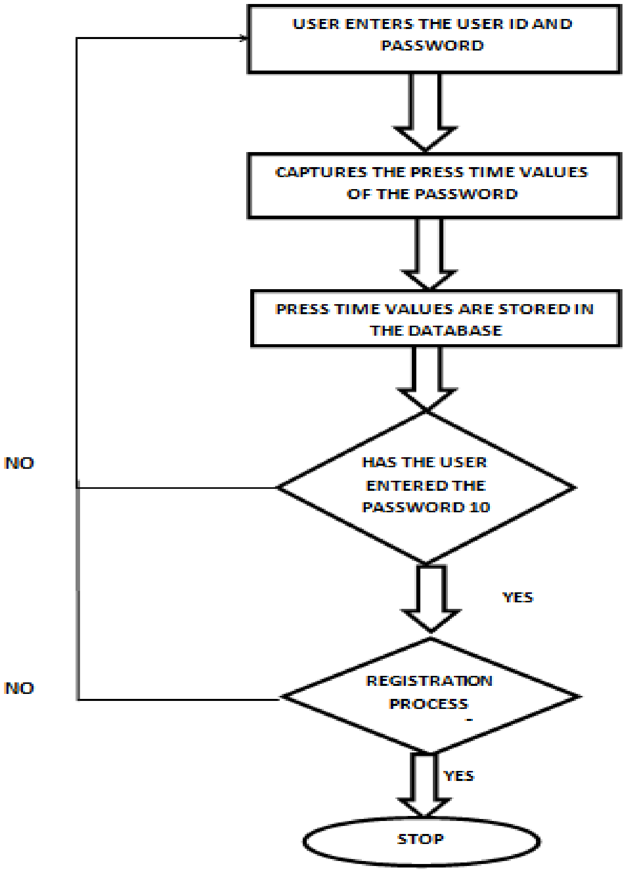
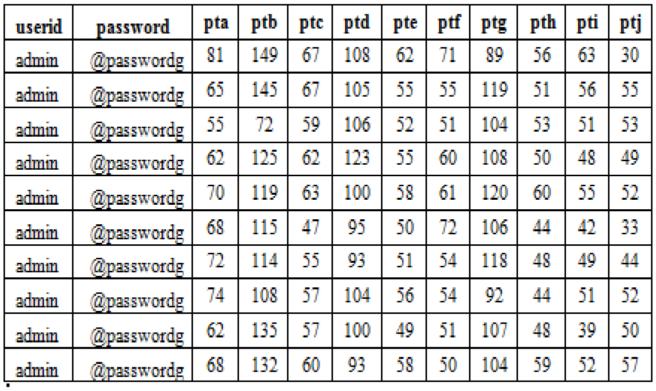

USER ENTERS THE USER ID AND PASSWORD
CAPTURES THE PRESS TIME VALUES OF THE PASSWORD
PRESS TIME VALUES ARE STORED IN THE DATABASE
NO
HAS THE USER ENTERED THE PASSWORD 10
YES
NO
REGISTRATION PROCESS
YES
STOP

Fig. . 1: Flow chart of registration module

In this module, the registration process starts. At first, the user is asked to enter the user id and password. For the training session, the user needs to enter the password for 10 times. After the first password is entered, the press time values of the characters are captured and stored in the database. The user is again directed to enter the password. Again, the press time values are captured and sent to the database. In this manner, the user requires to enter the password for 10 times. This completes the training session.



userid	password	pta	ptb	ptc	ptd	pte	ptf	ptg	pth	pti	ptj
admin	@passwordg	81	149	67	108	62	71	89	56	63	30
admin	@passwordg	65	145	67	105	55	55	119	51	56	55
admin	@passwordg	55	72	59	106	52	51	104	53	51	53
admin	@passwordg	62	125	62	123	55	60	108	50	48	49
admin	@passwordg	70	119	63	100	58	61	120	60	55	52
admin	@passwordg	68	115	47	95	50	72	106	44	42	33
admin	@passwordg	72	114	55	93	51	54	118	48	49	44
admin	@passwordg	74	108	57	104	56	54	92	44	51	52
admin	@passwordg	62	135	57	100	49	51	107	48	39	50
admin	@passwordg	68	132	60	93	58	50	104	59	52	57

Fig. . 2: A screen shot of stored pressing time value

b. PASSWORD VERIFICATION:

For security purposes the password that is taken from the user interface is stored in the database.

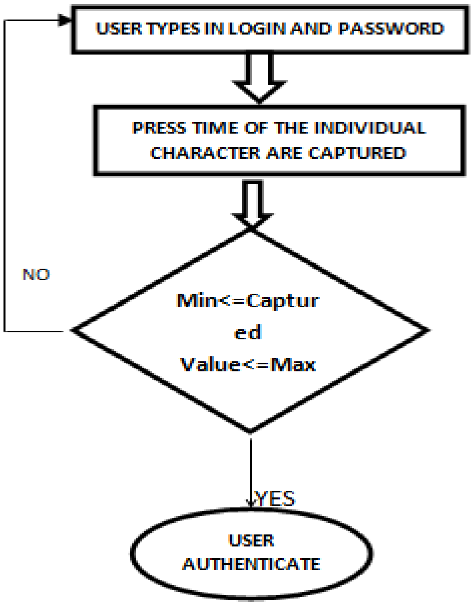

USER TYPES IN LOGIN AND PASSWORD
PRESS TIME OF THE INDIVIDUAL CHARACTER ARE CAPTURED
NO
Min<=Captur ed Value<=Max
YES
USER AUTHENTICATE

Fig. 3: Flow chart of verification module

c. IDENTIFICATION

In this sub system the user input is checked with the test sample that the user entered. If it is verified then the results are sent to the verification.

d. VERIFICATION

This subsystem will verify all the entities with clusters and taking into account the presstime, interkeytime and total time, the results are sent to the decision.

e. DECISION

The decision will take care about the tolerance and identification of fraud user from the legitimate user and then it will take care about granting or denying the access.

5. DISCUSSIONS

Combined with traditional authentication system, the user typing information can be of some help to identify users more precisely. Considering behavioral typing information to authenticate users can be very convenient because no extra hardware is necessary. All the behavioral information can be obtained by software systems, which generally implies lower cost than hardware development.

The authentication mechanism proposed here can operate in two modes: new user registration and user authentication. In the first case, the mechanism will record the user username, password and typing profile. The typing profile is then analyzed and stored so that it can be used during the authentication phase. In this mode, the user will be asked to type his user name and password about 10 times.

6. CONCLUSION

Here in this report we propose a method through which we can access the speed of typing the password and identify the user. When the user gets registered into the system then the system accesses the speed and the user name and user id. When the user need to get login again, user need to enter his/her user id and password so that the system will access the speed to typing and then authenticates. After the authentication process the system verifies the user id and password given by the user. If the user id and the password are correct and the authentication is done successfully.

In future the proposed system can be tested and applied in real time application. For more security we can store the value of key pressing time, password and user id in encrypted form. In future we are going to incorporate those modules and analyse this hybrid module to give better performance.

REFERENCES

[1] Kevin S. Killoushy, “A Scientific Understanding of Keystroke Dynamics”, CMU-CS-12-100, January 2012.

[2] Yunbin Deng and Yu Zhong, “Keystroke Dynamics User Authentication Based on Gaussian Mixture Model and Deep Belief Nets”, Hindawi Publishing Corporation, ISRN Signal Processing Volume 2013, Article ID 565183.

[3] Fabian Monrose, Aviel D. Rubin, “Keystroke Dynamics as a Biometric for Authentication”.

[4] Shivshankar Rajput, Priyanka Vijayawargiya, “Implementation of Keystroke Dynamics Application for Identifying Emotional state”, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 5, Issue 5, May 2015.

[5] Aniel D. Rubin, Fabrian Monrose, “Keystroke Dynamics as a Biometric for Authentication”.