International Journal of Multidisciplinary Research and Development

Volume: 2, Issue: 5, 526-528
May 2015
www.allsubjectjournal.com
e-ISSN: 2349-4182
p-ISSN: 2349-5979
Impact Factor: 3.762

Swapnil Dargude
Department of Information Technology, ISB&M School of Technology, Nande, Pune, India

Anurath Shinde
Department of Information Technology, ISB&M School of Technology, Nande, Pune, India

Dipak S U
Department of Information Technology, ISB&M School of Technology, Nande, Pune, India

Milind R Hegade
Department of Information Technology, ISB&M School of Technology, Nande, Pune, India

Correspondence:
Swapnil Dargude
Department of Information Technology, ISB&M School of Technology,Nande, Pune, India

Keystroke dynamic biometrics for user authentication

Swapnil Dargude, Anurath Shinde, Prof Dipak S U, Prof.Milind R Hegade

Abstract
Traditionally, the validation is based on the mixture of user name and PIN, password. Though this technique is frequently used, there are convinced flaws that makes it risky. Biometric access method connected with security in computer is gaining popularity in today's cosmos. Biometrics is measuring physical uniqueness of an human being for identification purpose. Keystroke biometrics is a fresh variety of biometric identification which is inexpensive and unobtrusive method that provides enhanced security as compared to traditional login and password. It does not necessitate any hardware as it uses the simple keyboard. The phrase keystroke dynamics refers to the latency, key push durations and the typing pattern which is fairly unique to an individual.
This project aims at analyzing the typing pattern of an individual for strengthening the security in a computer system. In this project we are desighning tool that accepts the password and based on specific calculations a user is legitimate. For the standard of taking the security to the next phase, we are tolerating the user to include special characters (shift,ctrl,shift etc) in their password.

Keywords: Mean, standard Deviation, Weighted latancy

Introduction
Keystroke Biometrics for user verification uses the behavioral typing prototype of a user to validate [1] Computers have become an omnipresent component of the current culture. Computer is repeatedly used simultaneously with the gap force we are all recognizable with that is the internet. every person uses internet for special reason or proficient reason or may be for both the purposes. Today's People over independent on internet. With the enlarge raise of internet, there comes a thing which is valueable and perhaps the riskiest term in the field of internet that is Security. It can be compromised in definite conditions. In early 2011, there was an online attack that happened on multiple companies total network shutdown and all the vital data and passwords of the workers were lost. This is the level of attacks that can occur if there is not a tremendous security available, so too much dependency on internet so there is a need of caring users and protecting their data from the hackers. We all require a easy, low cost yet inconspicuous scheme for security principle. All this reasons led to the Keystroke Biometrics for user validation coming into the picture.The science of measuring the physical uniqueness of human being is known as biometrics. There are two types of Biometric properties 1) Physiological properties: It includes retina, hand, palm scan [4], face detection [3], iris detection [5], 2)Behavioral properties: It includes typing pattern, speech,voice[6], handwriting[8] etc.
biometric system concern with two modes [2]. 1) Authentication: scrutiny if the user is whom he claims to be. Here the authentication device used a user Identity number, smart card number and identification: Giving an uniqueness to an unknown user.
Keystroke dynamics based verification system needs users to enroll their information for model development and classifier preparation and testing purposes for confirmation. The standard of register is to extract typing pattern from which a representation of the typing method can be generated. To capture keystroke dynamics, it is necessary for users to type their own password a number of times during enrollment can begin and that is before registration. 2) dynamic mothod of registration, the users register after they have granted access to the system [9]-[11]. This is the dynamic version of enrollment.

Related Works
Ahmed A. Ahmed and Issa Traore [13], projected "Enhancing password based authentication scheme with keystroke".on the study of base research papers, the integral performance parameters such as False Acceptance Rate (FAR) and False Rejection Rate (FRR).

FAR is refers to the percentage of an fraud was accepted by the system, FRR refers to the percentage of authorized users was rejected from the system. This method provides enhanced security as compared to traditional login and password which has least FAR and FRR.

Ying Zhang and Lin Liu [14], mentioned main factor is "Multi-factor authentication" combination to reap benefits-in terms of security and convenience. This security mechanism in order to reinforce user authentication. A string of password complemented with its matching typing pattern which represents something the user must be.

Zahid Syed [15], proposed work is "No additional hardware required".Using by default a simple keyboard and a software and normally keyboard is by default hardware associated with computer system. So that no any extra hardware cost arises.

Marcus Karnan and M. Akila [12], "Typing of password is unique to an individual and complicated to copy". the system deals with user biometric it become difficult to intruder to hack the system.

Data Collection Procedure

The dataset used in this study was gathered by Shanthi Bhatt and T.Santhanam [1], over the Internet. Shanthi and santhanam are created a client-side frontend using Java applets in the NetBeans integrated developing environment (IDE). This is executing on web browsers, that is Sun microsystem Java Console installed. The information collected at the client-side was stored on the MySQL Database. The client established a connection with the server and then entered the keystroke records. The keystroke records was transmitted to the server through the data connection.The data set series consisted of early register and subsequent information entry. the time of enrollment phase each user was given a record through the user interface. all data comprised of a username and a password, the user firstly register all its personal information, for example name,email ID,contact number,even password. after that the next step user enter its five times passwords then the system extract the biometric feature, then next formulate the biometrics from it and store it in a database, In Verification Phase. It is initiated when a previously registered user tries to login into the system. All the three steps included in verification phase is similar to the First phase, only variation is that after receiving the biometric feature of the user that is trying to log in, a pattern matcher is called. pattern matcher compares the biometric data received at the time of login for the current user with the previously stored biometric data of the same user in the database and if the timing difference is in the permissible limit, access is approved otherwise the access is discarded and the user is prompted to attempt once more. The data capture for each user was stored using files, each file containing Keystroke records for the username, passwords typed by the user.If any entries the user made a inaccuracy in typing a passwords and pressed backspace then the user can retype its passwords.

Table 1: Notations used in the algorithm

Notations	Description
x'	Mean
S	Standard Deviation
W	Weighted Latency.

Proposed Scheme

The algorithm will develop a signature profile for each of the three input strings entered by the user. The algorithm uses two variables, a test signature "T" which is required at the login period and a mean reference signature "M" where

$M= \{M_{username}, M_{password}, M_{phrase}\}$.

Verification is performed by comparing the test signature T with M and determining the magnitude of the difference between the two profiles.

$M=\{m1,m2,.....m_n\}$ and $T= \{t1,t2,t3,t4....t_n\}$

n=total number of latencies in the signature

The algorithm will compute the amount of the variation and positive identification is declared when the difference is within the threshold variability if the reference signature. When more than 80% of all the possible latencies passed this test then input for that string would be considered valid.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x \qquad (1)$$

Here mean is calculated as shown in equation (1).

$$S = \sqrt{\frac{1}{n-1}}\sum_{i=1}^{n}(x-\bar{x})^2 \qquad (2)$$

Where, x= value for each latency, x' is the mean, n is the number of logins and S is the standard deviation.

The usual way of calculating standard deviations is shown below in the equation, 2) to calculating standard deviation, mean is required. This scheme requires that all of the login diagraph times need to be stored the population standard deviation can then be calculated from all of these diagraphs.

This process means that every previous login needs to be stored just so that the standard deviation can be calculated. Manipulating the above standard deviation formula, a method can be used that does not require all the previous latency values for the standard deviation to be calculated.

$$S = \sqrt{(((\sum X)\,/n - ((\sum X)/n)^2)} \qquad (3)$$

Where, S is the standard deviation, n is the number of logins, X=x-x', x' is the mean.

Equation (3) written below will only use the value for the sum of the values for each latency and the squared sum for each latency. This enables the standard deviation to be counted more quickly and reduces the amount of storage space required for the user's profile.

The main problem with using the standard deviation directly is that a diagraph may be approved even though it has a high variability and consequently allows many more diagraphs to be approved because the allowed error is high[Joyce 1990]. So that even if more than 80% of latencies fit within one standard deviation of the reference profile, the profile generated may not be significant.

A way to remedy this problem is to use weighted latencies by assigning a weight to each of the deviations with regards to how the standard deviation compares with the mean as a percentage. In such a case, if the standard deviation is high in comparison to the mean, then the approval for the latency will have a low weighting. Equation (4) takes care of this problem by calculating weighted values for the latencies as shown below. If the sum of the computed weights using the following equation is at least 70%, the generated contour is approved.

$$W = [(1/(\sum(\bar{X}/S)) * (\bar{X}/S)] * 100\ \% \qquad \mathbf{(4)}$$

Where W is weighted latency.

For a biometric profile to be approved, more than 80% of the latencies must fit within one standard deviation of the

reference profile and for each valid latency that fits within the standard deviation of the sum of the weights must add up to at least 70%.

Encryption

MD 5(Message-Digest)

The **MD5** message-digest algorithm is a generally used cryptographic hash function creating a 128-bit (16-byte) hash significance, usually uttered in text design as a 32 digit hexadecimal number. MD5 has been exertion in a huge collection of cryptographic applications, and is also frequently used to authenticate data integrity.

MD5 was praposed by Ron Rivest in 1991 to change an previous hash function, MD4. The RFC source code 1321 contains a "by attribution" RSA certificate. In 1996 fault was found in the intend of MD5. While it was not a terminal error at the time, cryptographers started recommending the use of additional algorithms, for example SHA-1—which has been found to be risky as well. In 2004 shown that MD5 is not collision unbounding.MD5 is not right for applications similar to SSL certificates or digital signatures that rely on this possessions for digital security. Also in 2004 more solemn mistakes were exposed in MD5, use of the algorithm for security motive problematic; particularly, group of researchers described how to make a pair of files that distribute the matching MD5 checksum.

more advances were made in flouting MD5 in 2005, 2006, and 2007.[1] In December 2008, a group of researchers used this method to bogus SSL certificate ability, and CMU Software Engineering Institute at the moment says that MD5 "should be measured cryptographically busted and incompatible for additional use", and the government of U S applications now want the SHA-2 family of hash functions. In 2012, the Flame malware demoralized the weaknesses in MD5 to bogus a Microsoft digital signature.

Conclusion

The project presents a novel approach to harden the passwords by incorporating a biometric authentication method into the system. The system designed is not complicated or fancy. It is a simple implementation to achieve high efficiency. The biometric authentication tool is able to do its objectives by logging data and verifying login by using statistical analysis methods. This system increases the time indefinitely that would take for an attacker to intrude into the system. Overall, the project builds a successful system to establish that typing dynamics can be used to build a secure system for user authentication.

References

1. Shanthi Bhatt, T.Santhanam, ”Keystroke Dynamics for Biometric Authentication –A Survey”in International Conference on Pattern Recognition, Informatics and Mobile Engineering (PRIME) February 21-22,2013
2. Anil K.Jain, Arun Ross, Salil Prabhakar, "An Introduction to Biometric Recognition," in IEEE Transactions On Circuits and Systems for Video Technology, vol 14, no. 1, 2004.
3. Voth. D (2003), ‘Face Recognition Technology’, IEEE IntelligentSystems, Vol. 18, No. 3, pp. 4-7.
4. Shu W. and Zhang. D (1998), ‘Automated Personal Identificationby Palmprint’, Optical Engineering, Vol. 37, No. 8, pp. 2359-2362.
5. Li Ma, Tieniu Tan, Yunhong Wang and Dexin Zhang (2003), ‘Personal Identification Based on Iris Texture Analysis’, IEEETransactions on Pattern Analysis and Machine Intelligence, Vol.25, No. 12, pp. 1519-1533.
6. Shaughnessy. D (1986), ‘Speaker Recognition’, IEEE ASSP Magazine, Vol. 3, No. 4, pp. 4-17.
7. [7] Germain R.S., Bolle R., Califano A., Colville S., Pankanti S. andRatha N. (1997),‘Issues in large scale automatic biometric identification’, Proceedings of IEEE Workshop on Automatic Identification Advanced Technologies, Stony Brook, NY, pp. 43-46.
8. [8] Tappert. C.C (2009), ‘Rationale for adaptive online handwriting recognition’, Proceedings of International Workshop on Frontiers in Handwriting Recognition, Montreal, Canada, pp. 13-22
9. [9]Leggett. J and Williams. G (1998), ‘Verifying Identity via Keystroke Characteristics’, International Journal of Man-Machine Studies, Vol. 28, No. 1, pp. 67-76.
10. Rajkumar Janakiraman and Terence Sim (2007), ‘Keystroke Dynamics in a General Setting’, Advances in Biometrics, Springer Berlin / Heidelberg, Vol. 4642, pp. 584–593.
11. Marcus Karnan, M. Akila,“Personal Authentication based on Keystroke Dynamics using Soft Computing Techniques”,in IEEE Second International Conference on Communication Software and Networks, 2010.
12. Ahmed A.Ahmed and Issa Traore, "Biometric Recognition Based on FreeText keystroke Dynamics ", in IEEE Transactions on Cybernetics, 2013.
13. Ying ZHANG, Guiran CHANG, Lin Liu and Jie JIA "Authenticating User’s Keystroke Based on Statistical Models", in IEEE Fourth International Conference on Genetic and Evolutionary Computing, 2010.
14. Zahid Syed, Sean Banerjee, Qi Cheng, Bojan Cukic, "*Effects of user habituation in keystroke dynamics on password security policy*", in IEEE 13th International Symposium on High-Assurance Systems Engineering, 2011.