JCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Volume-3, Issue-3 E-ISSN: 2347-2693

Feature Extraction Techniques in Keystroke Dynamics for Securing Personal Devices

N.Krishnaraj

Department of Information Technology,
Sree Sastha Institute of Engineering and Technology,Chennai, India

www.ijcseonline.org

Received: Mar/02/2015 Revised: Mar/08/2015 Accepted: Mar/22/2015 Published: Mar/31/2015

***Abstract*—** This paper presents a novel approach used to identify and analyze the features of keystroke dynamics. Keystroke dynamics is a authentication technique which aims to identify the person based on the behavioral characteristic (Typing Rhythms). The keystroke data like Duration, latency and digraph are measured using statistical techniques. The user keystroke patterns are collected and further it is analyzed to identify the quality features that will be given to feature subset selection for selecting the dominant features. The extracted features will be given to the feature subset selection for identifying dominant features.

***Keywords*— :** *Biometrics,Feature Extraction , Keystroke Dynamics , Latency, Digraph*

I. INTRODUCTION

Authentication[1,2,3,4] is an added security measure used to prove that someone or something is who or what they say they are before access is granted to personal or confidential information. User authentication[5] is the process of verifying the identity of a person. There are multiple authentication technologies that verify the identity of a user before granting access to system resources. However, these technologies provide different levels of security, and none can be said to be a secured system completely.

II. FEATURE EXTRACTION

Feature extraction[6] is used to characterize attributes common to all patterns belonging to a class. A complete set of dimensionality features for each pattern class can be found in feature extraction during the enrollment phase. There are many types of feature that can be gathered during a user's enrollment session. Biometric features[7] provide more security than traditional methods like identification cards, password. The performance of the whole system depends on how well each feature behaves. Extracting features in Keystroke dynamics need the user must type the system, to extract the features. In the enrollment phase, the user must type password/PIN of the user. The user must use the same Password/PIN for verification and Identification stages. After the Password/PIN entered , the feature vector has been created , stored in the user template.

III. DATA COLLECTION

In this keystroke experiment Duration , Latency and Digraph and its combinations are used for data analysis .The experiment was done with 100 users with different passwords/PIN whose length ranged from 3 to 10. During the data collection the user has advised to type the Password/PIN for the specified number of times. Keystroke duration, latency and digraph were measured for all the 20 samples typed by each user. Statistical measures mean, standard deviation and median are calculated. Figure 2.1 is an example of typing word "THE" and the measurement of duration, latency between keystrokes and digraph. There are clear differences in duration, latency and digraph and their mean, standard deviation and median.

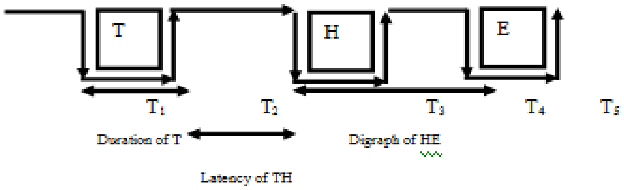


Fig2.1 Measurement of Duration, Latency and Digraph

IV. STATISTICAL MEASUREMENTS

The mean, standard deviation and median that is calculated for the features of the pattern set is in agreement with the following equations:

Mean (μ_i) = (1/N) Σ x (i)
Standard Deviation (σi) = √(1/N)Σ │x (i) - (μ_i)
Median (m) if m is odd = (N+1) /2
if m is even = (N/2) + (N/2) + ½

where I = 1..N, μ is the mean, σ is the standard deviation, m is the median, i is the feature and x is the feature set.

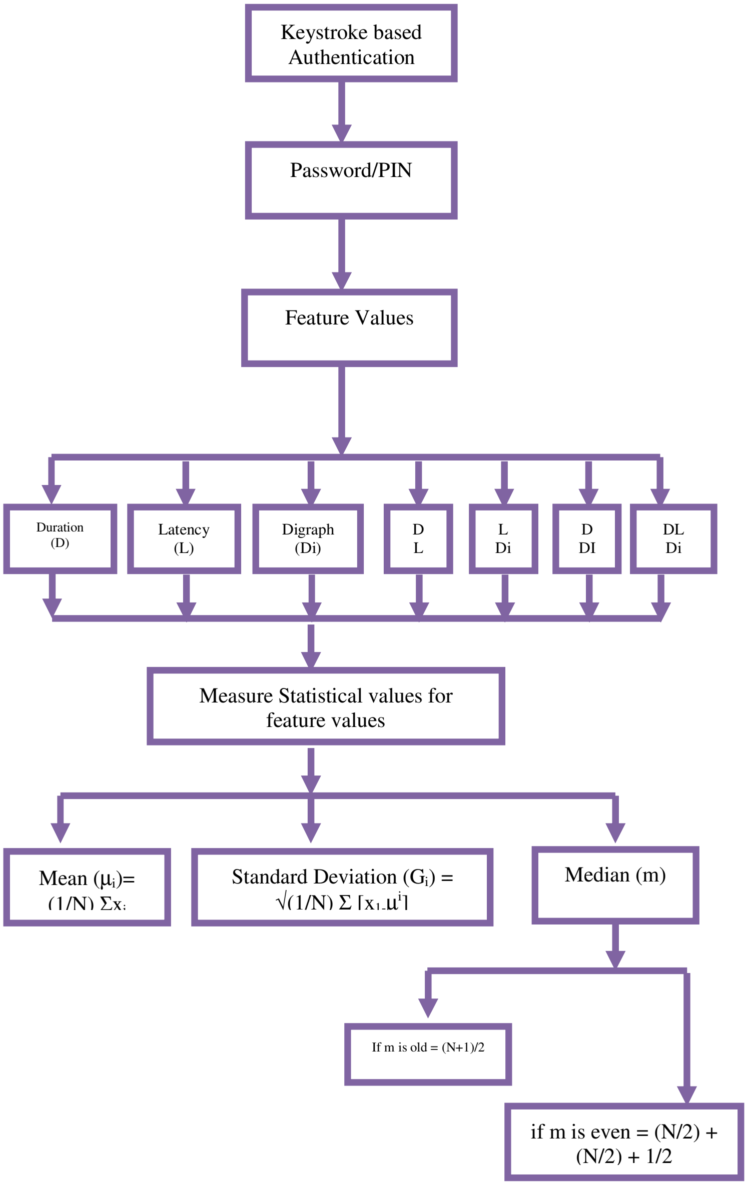

Keystroke based Authentication
Password/PIN
Feature Values
Duration (D)
Latency (L)
Digraph (Di)
D L
L Di
D DI
DL Di
Measure Statistical values for feature values
Mean (μi)= (1/N) Σxi
Standard Deviation (Gi) = √(1/N) Σ [x1-μi]
Median (m)
If m is old = (N+1)/2
if m is even = (N/2) + (N/2) + 1/2

Figure 4.1 Schematic Diagram of Feature Extraction

4.1 *Experiment and Results*

Using equations shown in section III, the mean, standard deviation and median values of duration, latency, digraph and their combinations are computed. For instance, the password "computer" results in a timing vector are as follows:

[30, 28, 8, 27.5, 25.5, 30.21, 25.9, 22.01, 36]

This serves as the reference signature or template or feature string which is used for further investigation. The extracted feature results are displayed in Fig Fig.4.2 to 4.8.

User	No.of Samples	Duration & Latency & Digraph		
		Mean	Standard Deviation	Digraph
USER1	10	1.911682	0.695573	1.717000
		1.763480	0.624071	1.340000
		1.934955	0.709995	1.903000
		2.082760	0.723969	1.890000
		1.69884.	0.578580	1.333000
		1.787727	0.713956	1.670000
		1.727520	0.644846	1.260000
		1.465818	0.534560	1.114500
		1.791182	0.655496	1.612500
		1.840682	0.713252	1.610000

Table 3.3.Statistical Measurement of Duration & Latency & Digraph

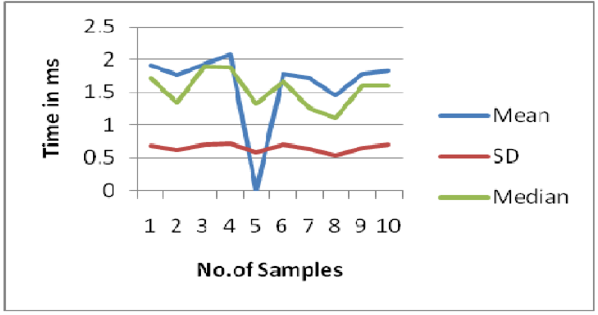


Fig4.2 (a) Duration using Mean, Standard Deviation and Median

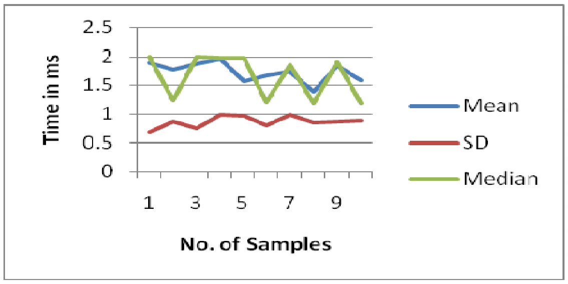


Fig4.3 (a) Latency Using Mean, Standard Deviation and Median

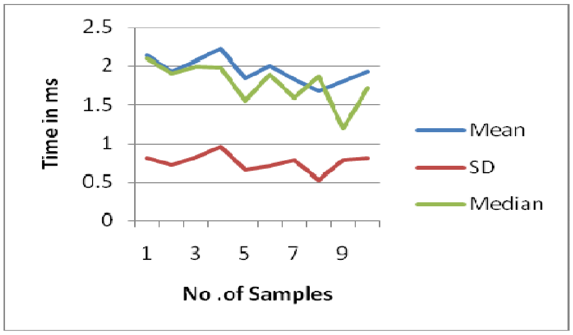


Fig 4.4 Digraph using Mean, Standard Deviation and Median

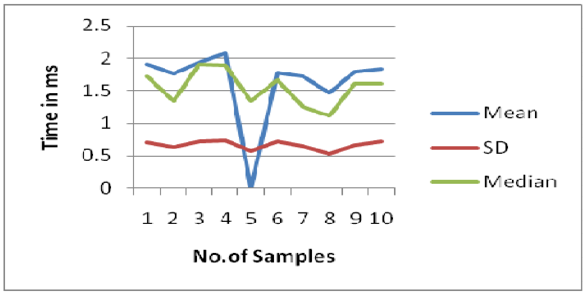
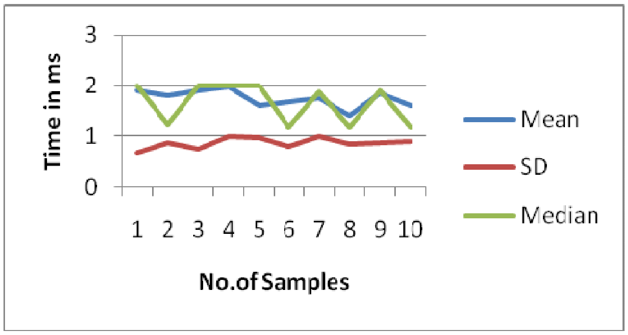


Fig4.5 Duration & Latency using Mean ,Standard Deviation and Median



.

Fig 4.6 Latency & digraph using Mean, Standard Deviation and Median

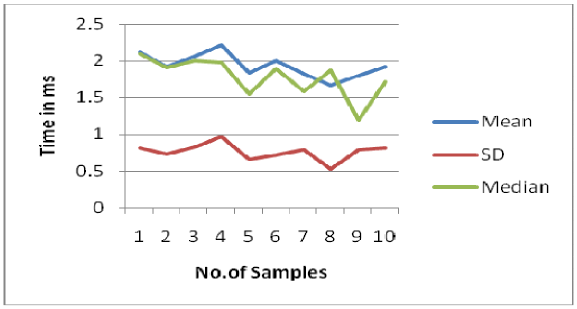


Fig 4.7 Duration &Digraph using Mean, Standard Deviation and Median

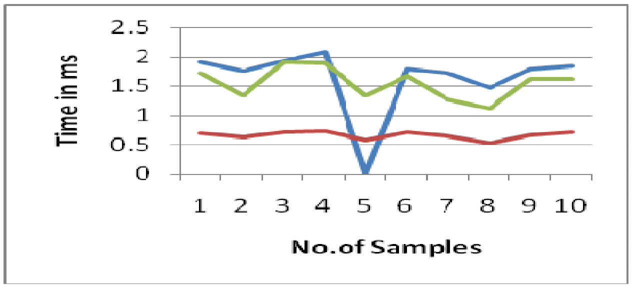


Fig 4.8 Duration , Latency and Digraph using Mean, Standard Deviation and Median

V CONLUSION

The features extracted in keystroke dynamics , the time measurement and statistical measurement of the features are explained. Feature vectors representing keystroke characteristics are derived from keystroke times, key release times, and information on which keys are being pressed. In keystroke mean, duration, latency and digraph were calculated. The feature extraction is the basic stage in any personal authentication techniques inorder to improve the efficiency of the system. The proposed techniques measures various feature and its combination in keystroke dynamics for additional security.

VI REFERENCES

1. Marcus Karnan, N.Krishnaraj , " Biopassword – A Keystroke Dynamics Approach to Secure Mobile Devices" , in IEEE International Conference on computational Intelligence and Computing Research (ICCCIC), pp.1-4,2010.
2. Marcus Karnan,M. Akila," Personal Authentication based on Keystroke Dynamics using Soft Computing Techniques The 2010 International Conference on Communication Software and Networks (ICCSN 2010) 26 - 28, February 2010.
3. Marcus Karnan,M. Akila,N.Krishnaraj , " Biometric Personal Authentication using Keystroke dynamics – A Review" , in International Journal of Applied soft Computing ,Vol 11, Isssue 2, pp.1565 – 1573 , 2011.
4. Marcus Karnan, N.Krishnaraj , "A Model to Secure Mobile Devices Using Keystroke Dynamics through Soft Computing Techniques" in International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-2, Issue-3, July 2012.
5. Seong-Seob Hwang, Hyoung-joo Lee and Sungzoon Cho, " Improving Authentication Accuracy Using Artificial rhythms and cues for Keystroke Dynamics based Autentication" , Expert system with Applications : An International Journal , Vol.36, Issue.7, pp.10649-10656 , 2009 .
6. Duane Blackburn, Chris Miles, Brad Wing, Kim Shepard, Biometrics Overview, National Science and Technology Council (NSTC) Committee on Technology Committee on Homeland and National Security, 2007.
7. Brown, M., Rogers, J. , "User Identification via Keystroke Characteristics of Typed Names using Neural Networks". International Journal of Man-Machine Studies, vol. 39, pp. 999-1014, 1993.
8. Leggett J, Williams G, Usnick M, Longnecker M. Dynamic identity verification via keystroke characteristics. Int Journal in Man Machine Stud, 1991.

Dr. N. KRISHNARAJ received the B.Tech in Information Technology in 2005 from Anna University, Chennai, Tamilnadu, India. He obtained the Master of Engineering Degree in Software Engineering in 2007 from Anna University, Chennai, Tamilnadu, India. He Completed his Ph.D in Computer Science and Engineering from M.S.University, Tirunelveli. Currently he is working as Assistant Professor & Head of Department of Information Technology, Sree Sastha Institute of Engineering and Technology,, Chennai, Tamilnadu, India. Her area of interest is Personal Security, Image Processing , Data Mining , Wireless networks.