

Behavioral biometric authentication on smartphone using keystroke dynamics

Pragya Vaishnav[§]

Amity Institute of Information Technology

Amity University Rajasthan

Jaipur 303007

Rajasthan

India

and

Department of Computer Applications

Manipal University Jaipur

Jaipur 303007

Rajasthan

India

Manju Kaushik[†]

Amity Institute of Information Technology

Amity University Rajasthan

Jaipur 303007

Rajasthan

India

Linesh Raja^{*}

Department of Computer Applications

Manipal University Jaipur

Jaipur 303007

Rajasthan

India

[§] E-mail: pragya.vaishnav23@gmail.com

[†] E-mail: mkaushik@jpr.amity.edu

^{*} E-mail: lineshreja@gmail.com (Corresponding Author)

Abstract

Over the past few years, smartphones have revolutionized our life. Authentication plays important role as all private and business data are stored in smartphone. Password authentication is not secure technique as password can be easily stolen by hackers. Hence, need to use advance technology to increase the security of authentication. Keystroke Dynamics is behavioral based authentication technique. It identifies the user identity through their typing behavior to check user is genuine or imposter. In this paper author is presenting a KDSP system for Smart devices to increase the security of authentication at login time. On the basis of timing feature set, typing speed, flight time and error rate of keystroke dynamics author is checking the identity of the user. Proposed system works on all the Android smartphones and tablets. All experimental work was performed on 268 subjects to capture keystroke data on server. In Statistical analysis, Equal-error-rate (EER) is measured 7.415% of the 60 subjects for False acceptance rate (FAR) and 208 subjects for False rejection rate (FRR).

Subject Classification: 4A62.

Keywords: KDSP, Pass-Value, EER, FRR, FAR, SD, CV.

1. Introduction

Presently, for most of the people smartphones and tablets has become main tool for communication and computation for their personal and business works, because of functionality and ease to use, smartphones and tablets has beaten PCs and laptops. It must be required to secure the personal and professional data saved on smartphone [1], as security risk is higher in smart devices, it can be lost or stolen, and are easier to attack if used in social interactions. User authentication is the key technology on smart devices to control the access for data security. Traditionally user authentication was dependent upon passwords, this approach is not reliable because using shoulder surfing or through keyloggers, or network sniffing passwords can be found by hackers. Two-way authentication has been proposed as a secure solution, like to send a one-time-password (OTP) to another communication device, but this is not a safe solution if the OTP is captured by an intruder who supervises to access that another device [2-5]. Keystroke Dynamics is a technique based on behavioral biometrics that checks the typing patterns of the user by monitoring the keystrokes timings and the latency between key presses [6-20].

2. Background

Keystroke dynamics studies are increased in the last few years in a scientific discipline where systems and models are evaluated using collected samples, and examination results are accepted by everyone. The several authors research aspects are pointed in Table 1.

Table 1
Comparative Study of Existing Work

Refer-ences	Classification/ Proposed work	Outcome	Limitation
[5]	Naïve Bayes, Random Forest, and K Nearest Neighbours (KNN)	KNN performance was more accurate	Accuracy Rate was high and did not use Dynamic Authentication
[17]	Naïve Bayes, Random Forest, and k Nearest Neighbours (KNN)	KNN provided the good percentage of accuracy for KD and TD sensors	User identification authentication was of 60% to 70% which is bit higher
[18]	Proposed continuous biometrics authentication using keystroke dynamic and touch dynamics	98.2% of accuracy of proposed work	Dynamic storage was not done
[8]	Micro behaviours on smartphone	FAR is 14% and FRR is 2% value	FAR rate was too high
[12]	SVM classification with RBF kernel	Using misclassification analyses result found by Digraphs and Trigraphs features.	Keystroke Dynamics analysis (FAR, FRR, EER) not done
[19]	Comparison of training vector and testing vector has been applied on classifier	Anomaly detection performed, Hit Rate is 92%, EER is 0.08 less than 99.999% the criteria of industry.	Continuous Authentication was not done
[7]	KSD method with NN learning algorithm applied on structure of Behavioural authentication	FAR value received 2.2, FRR 8.67 and EER 5.43	Dynamic authentication was not done

3. Proposed KDSP System

3.1 Background

The proposed KDSP (Keystroke Dynamics Smartphone) system enhances the security of all the applications where people store their private data. This system provides two phases: Registration phase and

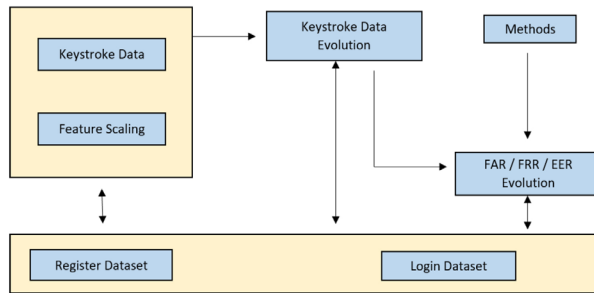


Figure 1

A Keystroke Dynamics authentication framework for System Login

Login phase. Registration Phase collects all the keystrokes information and saved that information in MySQL database on server, base64 sha1 encryption used by author to encrypt the unique id and password. Login keystroke data compares with registered keystroke data in login phase and if compared data are less than with pass-value then user will be able to login the application. Author collected 200 subject's keystroke data on server hence collected data are secured on the server, due to high data encryption user cannot do any modifications. Figure 1 explains working of KDSP system for login the application. This system confirms the identity of the user to check if user is genuine or fake through keystrokes data at login time.

3.2 Feature of KDSP

All previous studies on keystroke dynamics have been done on desktop and laptop keyboard, all that study was only comparative study. Smartphone has advance features that can be evaluated, such as dwell time, flight time, finger size, finger pressure and touch area using sensor readings. In this research study, authors used timing feature set and total typing speed and typing error rate has also been included. Authors also used encryption. This feature includes typing speed, flight time, error rate and encryption.

3.3 Registration and Login Phase

The KDSP system provides a registration phase module. Author has taken 200 keystroke data for registration phase. All received sample data are saved on server in database. Unique id and password are encrypted on server. In this phase user's name, email-id and password has taken and typing speed, flight time are measured. Author gathered current

time of every pressed key to measure Flight time and calculated the time difference between two pressed keys and saved this time difference.

In the login phase, using user’s email-id and password registered data of the user is retrieved from the database. This developed system is used further to analyses statistical error metrics (ERR, FAR, FRR).

3.4 Design of Algorithm

The design of algorithm to check the login data are explained in the Table 2 and 3.

Table 2
Registration

Step 1: Capture user Name, Email-Id and Password
Step 2: If typing speed length is equals to 1 set start time goto step 3
Step 3: If typing speed equals to sample text set end time goto step 4
Step 4: Calculate typing speed
Step 5: Capture input text
Step 6: Index is not equals to sample text length set current time set array index as current time initialize index set first number in array at 0 index
Step 7: calculate result of first number to N number
Step 8: calculate Flight Time

Table 3
Login

Step 1: Capture Email-Id and Password
Step 2: If typing speed length is equals to 1 set start time goto step 3
Step 3: If typing speed equals to sample text set end time go to step 4
Step 4: Calculate typing speed
Step 5: Capture input text set key info as input text
Step 6: index is not equal to key info set current time set array index as current time initialize index set first number in array of 0 index
Step 7: calculate Result first number
Step 8: calculate Flight Time

On the basis of Table 2 and 3 successful login or failed cases are compared for the user's identification.

4. Result and Analysis

In this research study, 268 subject's keystroke data gathered by author on server in database. The gathered data contains 3 features subsets (Typing speed, Flight Time, Error Rate).

4.1 Analysis of Standard Deviation and Coefficient of Variation

The Standard deviation (SD) and coefficient of variation (CV) are the statistical analysis of the KDSP dataset, which is the ratio of all feature set element. Table 4 represents the coefficient of variation and standard deviation of registration phase for feature categories (Typing Speed and Flight Time) of the 200 timing features subset. Table 5 represents the standard deviation and coefficient of variation of login phase for feature categories (Typing Speed, Flight Time and Error Rate) for the 208 timing features subset.

Table 4

Represents the Standard deviation and coefficient of variation of Registration phase

Analysis	User	Typing Speed	Flight Time
Standard Deviation	200	2.187727	0.582395
Coefficient of variation	200	0.169979	0.255612
Mean Value	200	12.87059	2.278435

Table 5

Shows the Standard Deviation and Coefficient of Variation of Login Phase

Analysis	User	Typing Speed	Flight Time	Error Rate
Standard Deviation	208	1.784253	0.542824	0.919664
Coefficient of variation	208	0.160584	0.245808	1.406545
Mean Value	208	12.58162	2.208322	0.653846

Statistical analysis was done on every feature of both registration phase and login phase. Value of SD and CV in both the phases are closed to each other. Therefore, users typing behavior were almost same in both the phases, and not changed the typing pattern in same position of using smart phone at login time.

4.2 Analysis of FAR and FRR

FAR and FRR value achieved by using the pass-value for keystroke data of all the users to check the performance of the KDSP system at login time. Table 6 is representing the analysis of FRR for keystroke data of 208 subjects and Table 7 is representing the analysis FAR for keystroke data of 60 subjects.

Table 6
FRR Analysis

No Of User	FRR
208	8.1730%

Table 7
FAR Analysis

No Of User	FAR
60	6.66%

4.3 Analysis of EER

The dataset has been measured through the FAR and FRR value that calculates the EER, in which the pass-value is a variable, therefore it is determined independently for each subject (Typing Speed, Flight Time and Error Rate). This analyzes keystroke feature data only at login time. The results of EER in Table 8 represents the EER value.

Table 8
EER Analysis

FRR	FAR	EER
208 Users	60 Users	7.415

4.4 Comparative Study of FAR, FRR and EER

On the basis of existing literature and proposed KDSP system, Table 9 is showing comparative values of FAR, FRR and ERR values computed at the time of system login.

Table 9
Comparative Analysis of FAR, FRR and ERR

Study	User	FAR%	FRR %	EER %
[17]	6	7.0	40	-
[12]	24	4	3	3.93
[5]	20	-	-	5.1
[11]	100	14.54	14.6	14.6
[3]	100	9.78	19	10
[16]	50	15.8	9.1	12.2
Proposed KDSP	208/60	6.66	8.1730	7.415

5. Conclusion and Future Work

In previous study where number of users are greater than 50 then ERR value is above 10% and FAR value counted almost 10% or more than 10%. In [16] FAR value is higher and greater than FRR as users are only 50. In [11] FAR, FRR, and EER value are almost same. In [12] FAR value is higher than FRR, which makes system dangerous. In this proposed KDSP author measured FAR value 6.66% of 60 subjects, FRR 8.1730% of 208 subjects and EER value 7.415 of the KDSP system. FAR is less than FRR, it makes system reliable and authentic. KDSP system is secure and working good in large amount of people.

For future work, using keystroke dynamics possible improvement of the authentication process on smart phone will be as: add more timing feature such as dwell time, add touch feature data and apply continuous authentication throughout the session.

References

- [1] Teh, Pin Shen, et al. "A survey on touch dynamics authentication in mobile devices." *Computers & Security* 59 (2016): 210-235.
- [2] Mondal, Soumik, and Patrick Bours. "Combining keystroke and mouse dynamics for continuous user authentication and identification." 2016 IEEE International Conference on Identity, Security and Behavior Analysis (ISBA). IEEE, 2016.

- [3] Tasia, Cheng-Jung, et al. "Two novel biometric features in keystroke dynamics authentication systems for touch screen devices." *Security and Communication Networks* 7.4 (2014): 750-758.
- [4] Mathur, Bhawana, and Manju Kaushik. "Data analysis utilizing principal component analysis." *Int J Eng Res Technol* 11.2 (2018): 333-348.
- [5] Anusas-Amornkul, Tanapat. "Strengthening password authentication using keystroke dynamics and smartphone sensors." *Proceedings of the 9th International Conference on Information Communication and Management*. 2019.
- [6] Al-Obaidi, Noor Mahmood, and Mudhafar M. Al-Jarrah. "Statistical keystroke dynamics system on mobile devices for experimental data collection and user authentication." *2016 9th International Conference on Developments in eSystems Engineering (DeSE)*. IEEE, 2016.
- [7] Salem, Asma, et al. "Analysis of strong password using keystroke dynamics authentication in touch screen devices." *2016 Cybersecurity and Cyberforensics Conference (CCC)*. IEEE, 2016.
- [8] Draffin, Benjamin, Jiang Zhu, and Joy Zhang. "Keysens: Passive user authentication through micro-behavior modeling of soft keyboard interaction." *International Conference on Mobile Computing, Applications, and Services*. Springer, Cham, 2013.
- [9] Mathur, Bhawana, and Manju Kaushik. "Empirical Analysis of Metrics Using UML Class Diagram." *International Journal of Advanced Computer Science and Applications* 7.5 (2016).
- [10] Jia, Weichen, et al. "High Security User Authentication Based on Piezoelectric Keystroke Dynamics Applying to Multiple Emotional Responses." *IEEE Sensors Journal* 22.3 (2021): 2814-2822.
- [11] Chang, Ting-Yi, Cheng-Jung Tsai, and Jyun-Hao Lin. "A graphical-based password keystroke dynamic authentication system for touch screen handheld mobile devices." *Journal of Systems and Software* 85.5 (2012): 1157-1165.
- [12] Cilia, Darren, and Frankie Inguanez. "Multi-model authentication using keystroke dynamics for smartphones." *2018 IEEE 8th International Conference on Consumer Electronics-Berlin (ICCE-Berlin)*. IEEE, 2018.

- [13] Mathur, Bhawana, and Manju Kaushik. "Comparing Different Classification Techniques Using Data Mining Tools." Journal Homepage: <http://www.ijmra.us> 8.9 (2018).
- [14] Baljit Singh Saini, Navdeep Kaur, Kamaljit Singh Bhatia & Ashish Kumar Luhach (2019) "Analyzing user typing behaviour in different positions using keystroke dynamics for mobile phones", *Journal of Discrete Mathematical Sciences and Cryptography*, 22:4, 591-603, DOI: 10.1080/09720529.2019.1638580.
- [15] Vishnu Shankar & Karan Singh (2019) "An improved user authentication scheme on smartphone using dominating attribute of touch data", *Journal of Discrete Mathematical Sciences and Cryptography*, 22:8, 1549-1561, DOI: 10.1080/09720529.2019.1695903.
- [16] Karatzouni, Sevasti, and Nathan Clarke. "Keystroke analysis for thumb-based keyboards on mobile devices." IFIP International Information Security Conference. Springer, Boston, MA, 2007.
- [17] Corpus, Kyle R., et al. "Mobile user identification through authentication using keystroke dynamics and accelerometer biometrics." Proceedings of the International Conference on Mobile Software Engineering and Systems. 2016.
- [18] Temper, Marlies, and Simon Tjoa. "The applicability of fuzzy rough classifier for continuous person authentication." 2016 International Conference on Software Security and Assurance (ICSSA). IEEE, 2016.

Received June, 2021

Revised July, 2022