A Novel Technique of Using Touch and Keystroke Dynamics for Sequential Authentication on Smartphones

Saravana Prakash Dhanabal

Department of Computer Applications,
Sri Krishna Arts and Science College,
Coimbatore, TN, India
dsp.prakash@gmail.ccom

Indian Journal of Advances in Science Engineering and Technology
Vol 1 No 1 2022: 28-38

IJASET

Original Article

A Novel Technique of Using Touch and Keystroke Dynamics for Sequential Authentication on Smartphones

Saravana Prakash Dhanabal
Department of Computer Applications, Sri Krishna Arts and Science College, Coimbatore, TN, India
dsp.prakash@gmail.ccom

Abstract

Smartphones are now more of a requirement than an accessory in the modern world. Thanks to all of the applications on the smartphone, it has evolved into our personal assistant. Every smartphone user is concerned about mobile security because they conduct a variety of transactions on their phones and store sensitive data there. Passwords can be stolen via finger oil or shoulder surfing, so password authentication is not really a more reliable authentication method. Sequential or continuous authentication on smart devices, touch and keystroke dynamics are behavioral biometric authentication techniques that use the user's typing and touching patterns to confirm their identity continuously. For all smart devices, the authors of this study suggested a TKDSmart system to improve authentication. The TKDSmart system's authenticity was confirmed by the EER 4.1 percent, FRR 6.73 percent, and FAR 1.66 percent, that were systematically proved.

Keywords: **Touch Dynamics, Keystroke Dynamics, TKDSmart system, EER, FAR, FRR.**

1. Introduction

Unquestionably, smartphones will continue to play a significant role in people's daily lives. This is due to the fact that smartphones are undergoing a massive and growing transformation but are no longer the common communication tools they once were. It has grown to be a major topic of attraction for both people and companies. 91 percent of those surveyed cite the importance of their smartphone, with 60 percent saying it would be even more essential than a cup of coffee [7]. This is due to the multiple amazing features and opportunities that smart phones through mobile services and applications offer. Everyone used to keep their personal data, conduct all of their monetary operations (banking, shopping, and payments of any kind), and preserve a password file in addition to using social media. Everyone is now completely occupied by and reliant on all these apps, and they find it difficult to fathom their lives without them. The proportion of people who use mobile applications on their smartphones is depicted in the Chart in Figure 1.

A true representation of the world today is shown in Chart 1. Keypad locks and passwords, which are less secure authentication techniques, are very popular solutions for protecting the data on smartphones. As username and password and keypad locks are simple for hackers to crack and everyone can grab them through finger oil and shoulder surfing, they are unable to protect the user's private information and files [10]. Password authentication is not, therefore, a more secure technique [12]. As a result, everyone is at risk because every smartphone contains sensitive personal information that can be obtained by unauthorized individuals in the event that it is stolen or lost accidentally. This data includes both personal and professional information. Chart in the Figure 2 displays the likelihood of unauthorized data access in the event that a smartphone is lost.

Continuous or unremitting authentication is an excellent solution, but it necessitates surveillance of the user all through the session. It is an advanced security mechanism for smartphones because it monitors users even after they have logged into an application in a smartphone. Touch Dynamics and Keystroke Dynamics are biometric security advancements driven adaptive behavior learned from user's touching and typing habits to confirm the identity of the smartphone user. This article proposes TKDSmart to enhance the authentication mechanism in smart devices.

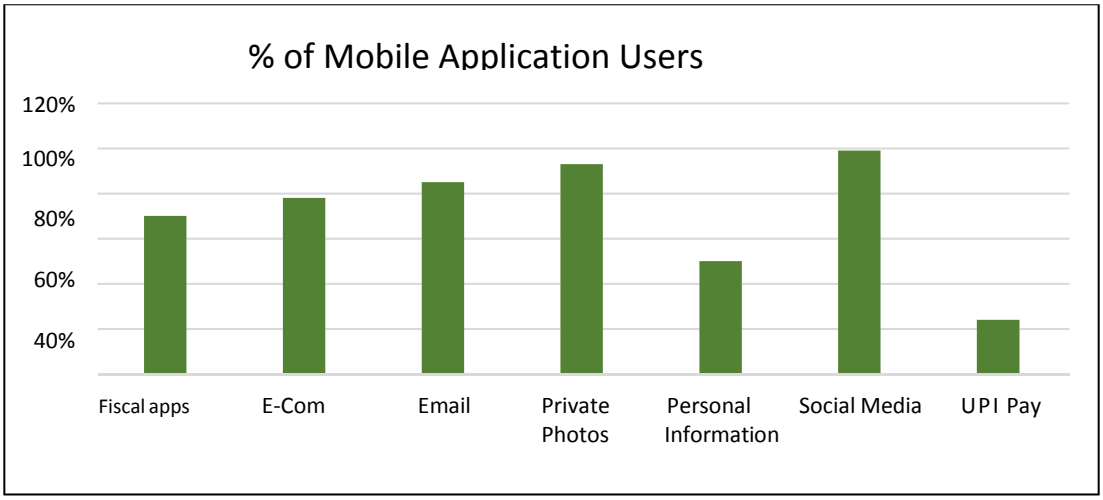

Figure 1: % of Mobile Application Users

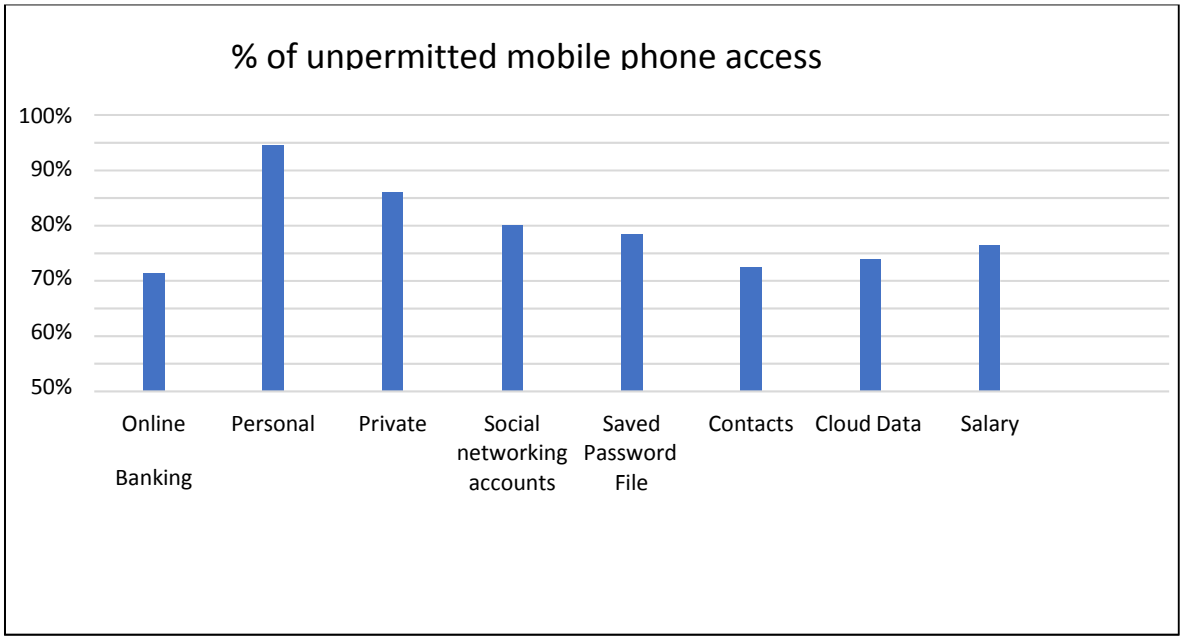

Figure 2: % of unpermitted mobile phone access

In the TKDSmart system, the keystroke dynamics technique is used for application login and continuous authentication, while the touch dynamics technique is used to monitor the user's behavior. In this study, the authors collected EER, FAR, and analysis data from 208 participants. To verify the authenticity of the device narrowed t-test, 200 individuals provided 600 data points, each individual provide three for in depth analysis of results in Section 3. Using the keystroke dynamics and touch dynamics described, this article proposes a frame work to improve the security features of smartphone authentication.

2. Related Works

In the past ten years, touch dynamics and keystroke dynamics research have evolved into scientific disciplines which are using community sample data to process experimental modes. Tanapat et al. suggested three categorization techniques for dynamics data: Naive Bayes, Random Forest, and k Nearest Neighbor. They utilized 10-fold cross verification to validate the data set and evaluate the EER accuracy percentage during the comparison categorization methods [14]. The worst performance is offered by Naive Bayes, while Random Forest categorization offers the best performance for time-based features [10]. The accuracy rate is higher than with simple passwords. As a result, kNN offers touch sensor accuracy, keystroke dynamics, and touch that is best in class and very competitive with Random Forest. Keystroke count for a digraph (2G) on two consecutive keys, a tri-graph (3G) on three consecutive keys, hold time, and typing time. The minimal FAR (roughly 7.0 percent) of a prototype indicates it is a well performing model at preventing attempts to intentionally access another person's account. As a result, identifying is just a little bit greater than the standard range of 60 percentage points to 70%. This does not imply that the system is 100% accurate in identifying smartphone users [10].

It is a suggested fusion strategy that combines two modes. The first is keystroke dynamics, which includes typing behavior; the other one is touch gestures, which includes tapping, swiping, and pinching. Different machine learning classifications enable both authentication methods. This authentication system is always active on Android devices [2]. Continuous integration authorization and performance tests are more accurate than FAR in the FRR modality experiment.

Draffin et al. provide a structure for micro-behavior but no information about the text that was entered. It determines the precise location where each key is touched, the length of time the key is depressed, the pressure used to press the key, or even the location where the user of a figure press is located. The FRR and FAR of five key downs with 4.6 and 32.3 percent, respectively, followed by FRR and FAR of fifteen key downs with 2 and 14 percent, respectively, were needed by the author [5] to identify an unauthorized user.

When Jong et al. combined the various features, many investigations were conducted on the posture of each user to look at how these feature combinations, like table, hand, and walk, were used when performing postures. Without preprocessing, the distance algorithm also yielded better results by utilizing mean deviations or standard deviation [8]. Preprocessing achieved excellent results with scalability and standardization.

Marlie offers a new continuous biometrics authentication technique that combines the dynamics of keystrokes and touch gestures as two authentication features. Because they used it in the Android-compatible mobile banking application, this evaluates the approach's viability. The accuracy rate for this assessment, which gathers data from 25 users, is 98.2 percent [11].

A method for smartphone authentication that uses a finger print of the user, information on login, and password hashing based on biometric rules was proposed by Chandrasekar et al. Three phases and two stages were used by the author. The first stage is fingerprints, keystroke dynamics, and username and password information for login time for enrollment (training) and time for verification. Based on this methodology, they are providing an additional layer of security [6].

Sung et al. examined sensor-based attributes, specifically for mobiles that are equipped with sophisticated sensors, to discover precise findings for smartphone devices. They acquired the "Up Up" (UU) characteristic to organize each user in turn. To obtain an accurate result, they made use of the timing features set, statistical features, and sensor-based features. In comparison to timing-based features, this indicates a higher accuracy rate. Although they do not achieve good accuracy, regular and shifting postures achieve greater accuracy when using statistical features as opposed to attributes of sensors [13].

Every user was arranged by Arwa according to the timing features using SVMs and DTs classification. For user authentication, it is free-text. They have taken this into consideration. Their output was precise. For user authentication, they offer complete knowledge. Acceptable results were obtained using the FAR and FRR rates, but the FAR rate performed marginally better. This suggested technique for user authentication in Arabic has worked well [3].

Asma Salem postulated a behavioural authentication architecture for touch-screen Android smartphones that is based on the KSD method and uses a NN learning algorithm. They are recording duration and non-timing features for each letter, such as its size, finger pressure, and area, using a virtual keyboard for each letter. KSD provides an allowable level of performance measures for second factor authentication; they received 5.43, 8.67, and 2.2 of EER, FRR, and FAR, respectively [4].

The passwords are now limited and pose serious security risks to the authentication systems that are currently in use. Since the 1970s, Forsen et al. have studied the habitual patterns of people's typing rhythms over the previous two decades. They found that people were differentiated in the way they typed names [9].

Time, acceleration, size, and other features were used by Zheng et al. to quantify the user's finger habits on a smartphone touch screen. They computed the system's outcome using those experimental studies using data from 4 and 8-digit PINs [14].

The first of its kind was the PIN that Buchoux et al. used, and the keystroke measurement during the logging process was the second. For further analysis, the authors recorded important occurrences and latencies of inter-key. Twenty subject groups were used to measure the deployment. The quantitative classification method is suitable for smartphones, but the author claims that a 4-digit PIN is insufficient to produce the desired results [1].

In the past ten years, studies on keystroke dynamics have been conducted that relied on the rhythm and touch capabilities of smartphones. Using a smartphone to define various positions, the authors computed EER, FRR, and FAR based on those positions. It is suggested that a pared t-test be used in this study to distinguish between real and fake users. After logging into the application, the researcher applied sequential authentication for smartphones to verify the login credentials. For an accurate analysis, the author collected 600 datasets from 200 users three times each.

3. Proposed TKDSmart System

3.1. Background

The TKDSmart System that has been proposed is compatible with all devices that run the Android operating system. This system offers security to all mobile applications where users conduct financial transactions, store personal information, and perform other sensitive tasks. The registration phase, the login phase, and the concluding testing phase are the three phases of this system.

All keystrokes and touch data are recorded during the registration phase and stored in a MYSQL database on the server. The unique ID and password were encrypted using base64 sha1 by the author. In the login phase, the user can log into the application if the data matches the pass-value when compared to the registered keystroke data. However, after successfully logging into the application, the touch data is compared with the enrolled touch data during the last testing phase of the session. The present display will turn off and show the home screen if the touch data of the user is inconsistent with the enrolled data. As a result, the user won't be able to use the application in any way.

The 200 subjects provided the author with keystroke and touch data on a server. Data is therefore protected on the server. Data is encrypted; therefore, there is no way for a user to alter the application in any way. The operation of the TKDSmart system is shown in Figure 1. Keystrokes and touch data are used by the system to determine the user's identity all throughout the session, even after logging into the application, in order to determine whether the user is a genuine user or an impostor.

3.2. Feature Selection in TKDSmart System

Since continuous authentication (during the session) has not yet been studied, all preceding keystroke dynamics research has been conducted using hard keyboards (desktop and laptop keyboards). The additional features of a smartphone, such as finger size, finger pressure, typing error rate, dwell time, and flying time, could be read by sensors and measured. The timing feature, touch feature, typing speed, and error rate are all used in this study by the author. Encryption is being used by the author to protect the data. Here is an explanation of these feature set components:

- Typing Rate: Overall speed of typing.
- Flight Time: The amount of time between two consecutive keys being touched.
- Dwell Time: The interval of time between pressing a key and releasing it while holding it.
- Error Rate: Count of incorrect key presses, backspaces and delete key presses.
- Finger Length: The impact of finger while touching.
- Encryption: Password and unique ID are stored with Base64 sha1 encryption.

3.3. Stages of TKDSmart System

A. Stage of Registration

To register the data, the TKDSmart system offers a registration module. A total of 200 keystrokes and touch data were gathered for the registration phase. A unique ID is generated for each user based on the records of the collected information that is kept on the server in a MySQL database. The password and unique ID are encrypted using base64 sha1. While the password, email address and the user name are entered during this phase, the flight time and typing speed are also computed. The author gathered the current time of the device at the starting and end points of typing from user in order to calculate typing speed. The author recorded the current times of each key that was pressed, then added up the times between the two keys that were pressed to determine the flight time. The author gathered the difference in time of each key that was pressed as a result. K stands for the key, N for the key's position within the word, and counts the time interval between the two keys being pressed. Touch data is gathered after keystroke data has been registered. In this study, the TKDSmart system uses touch data for continuous authentication. 200 participants' finger length and dwell time are gathered.

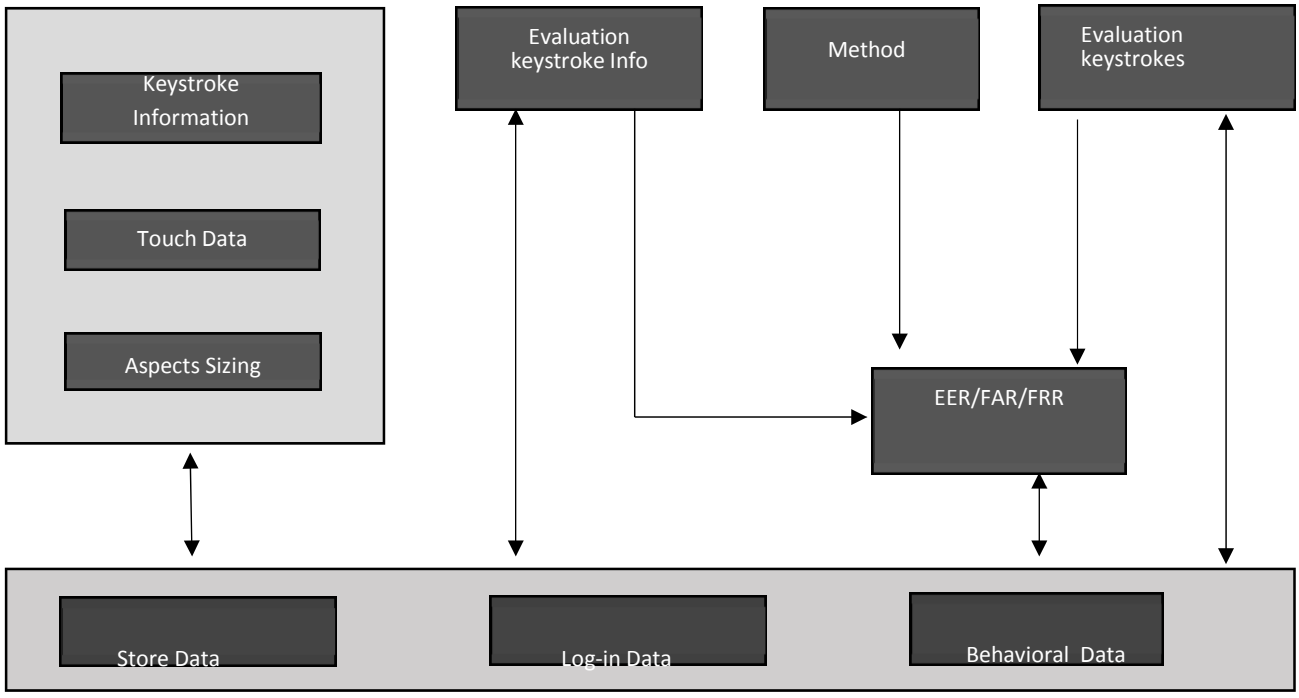

Figure. 1 : Architecture of TKDSmart system

B. *Login Phase*

The user's saved and stored data is extracted from the database during the login phase using the email address and password of the user. The server-registered keystroke information and the keystroke information used to log in are compared. The person is allowed to login to the system if the keystroke login data matches the pass-value. 208 users repeatedly tried the login phase module in this study. This comparison identifies whether a user is a fraud or the real deal. The error rate, flight time, and typing frequency are computed and verified with the stored data on the server. Their typing errors are used to calculate the error rate. Backspace keystrokes, deleted keystrokes, and incorrect key presses reduce the value by one. The statistical error metric analysis is another application for this TKDSmart system (FRR, FAR, ERR).

C. *Final Testing Phase*

This is the last stage of the TKDSmart system's continuous authentication process. As long as the user is genuine and has successfully logged into the application, this module begins to monitor them and keeps doing so until the session is over. It continuously compares received touch data with registered touch data while monitoring every touch activity by the user. The session will end immediately and the user will be unable to perform any operations on the app because they will remain treated as imposters.

4. Results and Discussions

4.1. Gathering Data

The 200 subjects provided keystroke and touch data, which the author entered into a database on a server. There were 208 subsets in the data that was gathered (Error Rate, Finger Size, Dwell Time, Flight Time, Typing Speed). Each participant's touch and keystroke data were gathered in different stages. The registration stage required 200 typing attempts; the login phase required 200 typing attempts for true user testing; and the third phase required a total of 600 typing efforts by 200 users to implement a pared t-test. The author collected 180 attempts from 60 subjects for ERR, FRR, and FAR (three attempts from each subject). To familiarize readers with the TKDSmart system before using it, the author provided a demo, which lowers the likelihood that they will type something incorrectly.

A. *Pared t – test*

For the same subject, the difference between these two variables is computed using a paired t-test. The two variables are typically separated by time. When our statistical measures are paired measurements, we can apply this test. Consequently, we might have measurements taken before and after for a group of people. The null hypothesis as well as the alternative hypothesis are the two categories of theories for a parsed t-test on samples. The following is a definition of the substitute theory used in this study.

- The true mean of the sample is greater than the comparison value (m0), according to the upper-tailed alternative theory (H1).
- The true mean of the sample is lower than the comparison value (m0) , according to lower-tailed alternate theory (H1).

Below is a definition of the void as well as alternative hypotheses' mathematical representations:

- Upper-tailed (H1): $m0 < \mu$
- Lower-tailed (H1): $m0 > \mu$

This study compared the parameters (registered and login data), which involves calculating the variation between both the variables, to see if there is a difference in the users' touch and typing patterns between the times of registration and login, to determine whether the TKDSmart system is functioning accurately, and to verify their identity. To obtain accurate results and determine an average of three attempts, the researcher

collects primary data users' three times at login time. Using SPSS software's pared t-test, this estimate was compared to the registered data. Below are the findings from these pared t-tests on dwell time, flight time, and typing speed.

B. Authorized Users' Pared T - test

Typing Speed Test

Parameter 1	Parameter 2	Void Hypothesis	Received P Value
Key_Speed	Normal	Key speed and average don't differ much from one another.	0.226

As P is greater than 0.05, the conclusion from the above is that the variation between mean value and key speed is null.

Test on Time of Flight

Variable Name 1	Variable Name 2	Void Hypothesis	Received P Value
Key_Flight	Normal	Key Flight and Average do not significantly differ from one another.	0.312

As P is greater than 0.05, the conclusion from the above is that the variation between mean value and Key Flight is null.

Test on Time of Dwell

Variable Name 1	Variable Name 2	Void Hypothesis	Received P Value
R_Touch_size	Normal	R Touch size and average do not significantly differ from one another.	0.586

As P is greater than 0.05, the conclusion from the above is that the variation between mean value and Dwell Time is null.

c. Unauthorized users' paired T – test.

The purpose of the test is to validate the TKDSmart system's reliability. To test whether the fake user could log into the application, the author gave them a valid email address and password. Consequently, 60 users provided 180 samples (each user gave three attempts). Tests on typing speed and flight time were conducted. A significant difference between the registered and logged-in data was discovered by the test.Test on Speed typing of fake users

Variable Name 1	Variable Name 2	Void Hypothesis	Received P Value
Registration_Speed	Normal	The difference between average and registration speed is substantial.	0.025

It can be deduced that the variation between mean value and typing speed significant as P value s is less than 0.05.

Test on Time of Flight of fake users

Variable 1	Variable 2	Void Hypothesis	Received P Value
Registered_Flight_Time	Normal	The difference between Registered Flight Time and the average is substantial.	0.001

As p value is less than 0.05, the variation between mean value and flight time is significant.

D. Pared T – test Results: -

The TKDSmart system, which was created to improve the security of smartphones, was subjected to the Pared t-test to ascertain its validity. This test, which was conducted on both real and fake users alike, demonstrated that the system is perfectly capable of identifying both legitimate users and intruders.

There was no discernible variation between login and registered data when the tests were run on actual user data. Thus, it demonstrates that the information gathered at login time comes from actual users. In order to verify the system's validity, fake users who are included in the test are also given a real user's email ID and password. The test revealed discrepancies between logged-in and registered data. As a result, the system becomes more precise and a fake user cannot log in if registered and keystroke data of them differs. Additionally, it establishes the individuality of touch and keystroke patterns. The users' touch and typing patterns are protected and cannot be copied or stolen. These tests demonstrate that the proposed is fully functional.

4.2. FRR and FAR Analysis

The TKDSmart system determines FAR and FRR values using pass-value for each user in order to determine system performance. Both the login phase and the final phase are subject to FAR and FRR. The FRR analysis of login attempt of 208 valid users was performed 6.73% percent of users denied entry in to application. The FAR analysis of login attempt of 60 invalid users was performed and only 1.66% of them got access.

4.3. Analysis of EER

The FRR and FAR are used to analyze the dataset, which determines the EER value and a variable pass-value i.e., each subject's value is calculated independently (touch size, dwell time, flight time, typing speed, and error rate). Only data with timing features is used for the analysis. The EER rate is shown in figure 3.

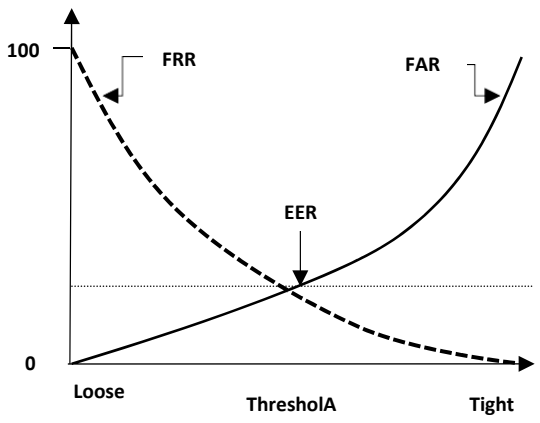

Figure 3: Error Rate

4.4. FAR, FRR and EER Analysis

6.73 percent is the FRR result, and 1.66 percent is the FAR. The system's performance is determined by the ratio of FRR to FRR. It can be taken into consideration if the FRR percentage is high, but it cannot be considered if the FAR percentage is high. It cannot be taken into consideration since it could be harmful to the system. The accuracy of the system is demonstrated by the EER, which is 4.1 percent. As a result, it demonstrates that the TKDSmart system's precision is far above and that it is an appropriate authentication system.

5. Conclusions

The TKDSmart system was designed with sequential authentication through keystroke and touch behavioral data for smart phones. The architecture was created to offer a reliable authentication process. The TKDSmart system uses touch and keystroke dynamics technology that is very reliable, effective, and safe. An algorithm has been developed to enhance the authentication procedure. Any Android-based application can use this algorithm. In order to verify the validity of the TKDSmart system, the FAR, FRR, and EER methods were also used, along with the Pared T test for result analysis. The designed system is effective, according to test results, with an extraordinary ability to distinguish between the identities of legitimate users and intruders.

6. References

1. Buchoux A, Clarke NL. Deployment of keystroke analysis on a smartphone. Australian Information Security Management Conference. 2008. Available from: https://ro.ecu.edu.au/ism/48/
2. Alsultan A, Warwick K. Keystroke dynamics authentication: a survey of free-text methods. International Journal of Computer Science Issues (IJCSI). 2013 Jul 1;10(4):1. Available from: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.412.2833&rep=rep1&type=pdf
3. Alsultan A, Warwick K, Wei H. Free-text keystroke dynamics authentication for Arabic language. IET Biometrics. 2016 Sep;5(3):164-9. Available from: https://ietresearch.onlinelibrary.wiley.com/doi/full/10.1049/iet-bmt.2015.0101
4. Salem A, Zaidan D, Swidan A, Saifan R. Analysis of strong password using keystroke dynamics authentication in touch screen devices. In2016 Cybersecurity and Cyberforensics Conference (CCC) 2016 Aug 2 (pp. 15-21). IEEE. Available from: Available from: https://ieeexplore.ieee.org/abstract/document/7600204
5. Draffin B, Zhu J, Zhang J. Keysens: Passive user authentication through micro-behavior modeling of soft keyboard interaction. InInternational Conference on Mobile Computing, Applications, and Services 2013 Nov 7 (pp. 184-201). Springer, Cham. Available from: https://link.springer.com/chapter/10.1007/978-3-319-05452-0_14
6. Vaishnav P, Kaushik M, Raja L. Design An Algorithm For Continuous Authentication On Smartphone Through Keystroke Dynamics And Touch Dynamics. Available from: https://www.ijcse.com/docs/INDJCSE22-13-02-111.pdf
7. https://saucelabs.com/blog/how-smartphones-and-mobile-internet-have-changed-our-lives
8. Roh JH, Lee SH, Kim S. Keystroke dynamics for authentication in smartphone. In2016 international conference on information and communication technology convergence (ICTC) 2016 Oct 19 (pp. 1155-1159). IEEE. Available from: https://ieeexplore.ieee.org/abstract/document/7763394/
9. Kaushik M, Kumar G, Preeti, Sharma R. Availability analysis for embedded system with N-version programming using fuzzy approach. International Journal of Software Engineering, Technology and Applications. 2015 Jan 1;1(1):90-101. Available from: https://www.inderscienceonline.com/doi/abs/10.1504/IJSETA.2015.067533
10. Corpus KR, Gonzales RJ, Morada AS, Vea LA. Mobile user identification through authentication using keystroke dynamics and accelerometer biometrics. InProceedings of the International Conference on

Mobile Software Engineering and Systems 2016 May 14 (pp. 11-12). Available from: https://dl.acm.org/doi/abs/10.1145/2897073.2897111

11. Temper M, Tjoa S. The applicability of fuzzy rough classifier for continuous person authentication. In2016 International Conference on Software Security and Assurance (ICSSA) 2016 Aug 24 (pp. 17-23). IEEE. Available from: https://ieeexplore.ieee.org/abstract/document/7861645
12. Vaishnav P, Kaushik M, Raja L. Survey on Smartphone Securities. InIOP Conference Series: Materials Science and Engineering 2021 Mar 1 (Vol. 1099, No. 1, p. 012067). IOP Publishing. Available from: https://iopscience.iop.org/article/10.1088/1757-899X/1099/1/012067/meta
13. Lee SH, Roh JH, Kim S, Jin SH. A study on feature of keystroke dynamics for improving accuracy in mobile environment. InInternational Workshop on Information Security Applications 2016 Aug25 (pp. 366-375). Springer, Cham. Available from: https://link.springer.com/chapter/10.1007/978-3-319-56549-1_31
14. Anusas-Amornkul T. Strengthening password authentication using keystroke dynamics and smartphone sensors. InProceedings of the 9th International Conference on Information Communication and Management 2019 Aug 23 (pp. 70-74). Available from: https://dl.acm.org/doi/abs/10.1145/3357419.3357425