International Journal of Advanced Computer Technology
ISSN: 2319-7900 www.ijact.org Volume-IX, Issue-V, December 2020

Keystroke Dynamics Analysis to Enhance Password Security of Mobile Banking Applications

Ahmed Alsawwan[1], Maen Alrashdan[2]*, Qusay Al-Maatouk[3], Mohamed Abdulnabi[4], Veeramani Vijai Indrian[5], Mohammad Tubishat[6], Mosab Tayseer AlRashdan[7], Abdulaleem Z. Al-Othman[8]

[1,2,3,4,5,6] Asia Pacific University of Technology and Innovation, [7]Al-Madinah International University, [8]Universiti Kuala Lumpur, Kuala Lumpur, Malaysia

[1]Ahmadspa@hotmail.com , [2]dr.maen@staffemail.apu.edu.my, [3]mohamed.shabir@staffemail.apu.edu.my, [4]Qusay@staffemail.apu.edu.my, [5]Veeramani@staffemail.apu.edu.my, [6]Tubishat@staffemail.apu.edu.my, [7]mos3ab_83@yahoo.com, [6]Abdulaleem@unikl.edu.my

*Corresponding Author: email address: dr.maen@apu.edu.my

Abstract

Nowadays, there are many cases where users' accounts get hacked using their password. Such cases can vary depending on password strength and obvious passwords which are similar to the user's details such as usernames and emails. There are new ways of preventing such incidents from happening and strengthening the accounts' security. This paper studies the usage of keystroke analysis to enhance password security which includes biometrics and typing patterns. This paper will also discuss this method's previous research on many platforms, including touch screen devices. After that, this paper will look deeply into this technique's implementation process, followed by detailed experiments and analysis. They were using keystroke dynamics analysis to enhance password security on mobile devices proved to have a great chance of success and how it can affect the everyday users of banking applications.

Keywords

Cybersecurity, Keystroke, Mobile banking, Password protection, Experiment

Introduction

In this time, security breaches are common everywhere, which leads to many catastrophes happening, such as bank systems being hacked through getting into the employees' accounts using their usernames and passwords. For such cases, Keystroke analysis is proposed to prevent such incidents from happening as they can cause many businesses to be stolen or bankrupt. Keystroke analysis or keystroke dynamics is the analysis of users' keyboard interaction to draw the digital fingerprints of the users using Smart of Internet (SoT) [1].

Although many other authentication mechanisms are being used, such as fingerprint scanners, they only authenticate users' identity at the start of the process. An example of that is the laptop login process, which only authenticates the user at this stage. Still, it does not check user authenticity when making actions using the computer, which means that, if a computer account gets hacked, many actions can be done using that system without further authentication processes, including data breaches webcam usage.

Such incidents nowadays are very common and easy to happen. One of the victims of this case was Yahoo, which dominated the internet world in the past. In 2016, while it was in the process of offering itself to Verizon, it declared that it had been the casualty of the greatest information rupture in history where the assault traded off the clients' subtleties, for example, email locations and passwords[2].

This attacked affected 3 billion users which cost the internet dominator an estimation of $350 million of Yahoo's sale cost. This case is only one of the incidents where accounts' passwords got compromised; other cases include big companies and Marriott International and eBay.

Keystroke dynamics is a way of interaction with the keyboard, or the device used. It may be a physical keyboard or touch screen keyboard, which is considered a biometric factor. This type of biometrics works because it analyses the dwell time and flight time for changing keyboard actions [3]. Thus, this type of biometrics is used to authenticate users based on keyboard rhythm analysis. Biometrics is the most secure and convenient authentication tool as it is hard to be stolen, and it is impossible to figure out such details [4]. Biometrics has two main types: physical biometrics such as fingerprints and iris, and behavioural characteristics such as voice recognition and keystroke pattern.

Password protection is becoming easier to compromise, and it can cause lots of businesses to shut down. An example

of that is what happened to 7-Eleven in Japan as they had a mobile wallet system that allowed the users to pay with cash in-store when ordering from an online source. The issue with the system was that anyone could reset the user's password. This caused the business to lose ¥55 million, equivalent to about $510,000[5]. This led the business of 7-Eleven to bring down one of its assets, which was the system (7pay).

Although there can be many apps that help store passwords, some of them contain some flaws that expose their users' information, an example of that is following the report of Independent Security evaluators (ISE). The research involved situations where the master password was residing in the memory of the system as a readable text file [6].

Literature Review

Using Keystroke Analysis for Verification on Touch Screen Devices.

As days go by, the new mobile devices included touch screens that users would interact with the phone with. For that, many applications nowadays require verification for several processes; i.e., financial transactions, Identity verification and many others. The next Article by [7] talks about applying keystroke dynamics analysis and using it as a secondary authentication factor. It also discusses suggesting a prototype for a keyboard application to gather information from the way the keys are being stroked. They also suggest using a neural network as a way of enhancing the password security level that is based on multilayer perception.

This paper uses authentication based on the user's behaviours when typing or entering information such as rhythm and timing. This makes it less costly to apply as it does not require any complicated hardware devices as only the keyboard or the keypad is necessary for their experiment. The behaviour-based authentication depends on the two-factor authentication as passwords are not enough to securely verify users remotely specifically when it comes to personal data such as banking applications or mobile network systems.

To collect data for this experiment, the static-based and dynamics-based approaches are also used in this article. There will be two primary methods to examine the data collected: "statistical model and learning model. Statistical Models compare reference typing characteristic with user typing profile characteristics using statistical measures" [8]. There are multiple learning algorithms utilized by the learning models such as Genetic algorithms, ANN [10], the Artificial Neural Network, and the Machine Learning algorithm (ML). To develop the profile which the user can refer to, including the way they type, those algorithms are used, and then the input of that user is compared with the profile. As it was measured in the previous article, the current one also uses error rates which are FAR, FRR and EER to analyze the keystroke dynamics.

I. Keystroke Dynamics (KSD) Proposed Focal Points

Many focal points helped to give the keystroke rhythms specific characteristics. They were obtained and expressed through the authentication system interfaces. The interfaces would be used through several devices with different inputs starting from standard keyboards to touch screen devices that use the software as a keyboard. This article experiments using certain processes that would enable extraction of specific focal points from the handheld platforms, such points are timing and non-timing characteristics.

As mentioned, the KSD has two major types that are timing and non-timing. Timing characteristics are described as pressing time, release time and holding time. On the other hand, non-timing qualities include characteristics that include pressure, size and finger positioning.

Authentication using behavioural typing is to gather typing rhythms and to examine them so that the outcome can be used to authenticate the user. The often-used feature of typing behaviour is information about timing characteristics. Those are collected through a unique timer on handsets, either by a plugin inside the device's OS or by utilizing event handlers that are triggered at the time of pressing and releasing a certain button. There are two major event parameters:

- Pressing Time: the time when the button is pressed.
- Release time: the time of key release.

TABLE 1: Main timing and non-timing characteristics

Feature	Notation	Type	Android function
Duration	DU = $\{DU_1,...,DU_n\}$, where each value $DU_i = U_i - D_i$.	Timing	getEventTime(),
Pressure	P (pressure) = $\{P_1,..,$ Pn$\}$. For n successive characters.	Non-timing	getPressure()
Position	L (position) = $\{X = \{X_1,...,X_n\}$, Y $=\{Y_1,...,Y_n\}\}$	Non-timing	getX(), getY().
Size	S (size) = $\{S_1,...,S_n\}$.	Non-timing	getSize(),

In this article's proposed model," the signature of user typing rhythm is analysed by catching timing and non-timing parameters of each event of typing n characters, where n is the total number of characters"[7] in the sequence as shown in the following sequence.

$K = \{K_1, K_2, \ldots, K_n\}$, is the a set of consecutive keystrokes (K) .

$D = \{D_1, D_2, \ldots, D_n\}$, is the set of timestamps when keys are pressed.

$U = \{U_1, U_2, \ldots, U_n\}$, a set of timestamps for when keys are released.

The main timing and non-timing characteristics were distinguished after being extracted from the unprocessed data, explained in the table below. The vectors of eight entries represented each of the features. The Android operating system had provided the system calls.

To apply keystroke dynamics authentication on touch screen phones, two main phases need to be analysed: "Enrolment" and "Authentication". This article mainly focuses on those major stages. The handheld devices which this experiment uses operate based on Android system. The two phases function simultaneously to accomplish the following functions:

1- Identification: coming up with a usable format after taking samples and processing them, then see how they match in opposition to a group of template profiles. At the end of this phase, the samples are compared to the best possibilities. That happens after generating a list that displays how similar are the samples to each other.

2- Verification: In this function, at least one sample is caught, prepared into a configuration that could be used, and then coordinated facing an information layout profile, where the correlation results are determined.

Enrolment: The behavioural enrolment systems function within a structure that contains four levels:

1- Capture: At first, the behavioural samples that belong to the user using the system for the first time are gathered by the behavioural platform, then registered into the system. Students and employees volunteered to enter their true information, which then gathered the authors' data set.

2- Extract: To extract the data, a different way was followed, which differed from the way the samples and templates were made. The system then extracts the Distinct characteristics and transformed into a behavioural identity belonging to that specific user.

3- Pre-processing extracted characteristics: this happens because not storing collected characteristics directly in the individual's template. Part of the pre-processing was completed to get the data ready for the authentication stage.

4- Store: The final step of the enrolment as the user template is created and stored.

Authentication: The behavioural authentication system includes three-layered architecture.

1- Enroll: The device collects a sample for one time, transformed into a structure that made a layout and then returned to the application. The learning procedure is depending on enrolment.

2- Compare: the existing template is compared with the newly introduced sample. Next, a behavioural template will represent the behavioural data for the user.

3- Match/Non-match: After the comparison stage, the newly extracted characteristics stands against the profile layout to determine if they were identical or not. When required to verify the identity, the user has an interaction with the Keystroke Dynamics platform. It then takes the behavioural sample from that interaction and compares it with the existing template. Unless they turn out to be identical, the user's identity will be unknown, and it will declare a non-match. If they are identical, the identity of the individual will be confirmed.

II. Keystroke Dynamics Proposed Keyboard Prototype

Collecting the user's input rhythm can be done by developing a specialized virtual keyboard done by the authors of the article, making it easier to install it on the selected smartphone. The smartphone selected in this article is operating on Android OS. Data gathering was done on Sony Xperia tablet Z, operating on Android version (5.1.1). In the case of Android, the application programming interface MotionEvent is used to collect input data using getDownTime() and getEventTime() methods. The OS also calculates the value of the pressure using getPressure() method[9]. The table mentioned earlier talks about the main android features which are utilized to develop the virtual keyboard. The authors chose DU duration time only to represent timing information in this paper[7].

III. KSD Neural Network Model: Multilayer Perception

In this research, the authors chose the Multilayer Perception (MLP) Neural Network Model, which has the following characteristics:

1- Layers

Many layers make the MLP: input layers and output layers that contain many un-seeable layers. It is a completely integrated network that every node inside of it stays in a specific layer linked to the nodes in the layer after it, after putting specific weight (W) on node input.

2- Weights

Knowing the weight comes from the set of practices that move to the output layer from the network's input layer. The weights are used repeatedly by utilizing various parameters such as "gradient, momentum and learning rate" to reduce faults. To calculate the weight change, the authors multiply the gradient with the learning rate value and then add the former weight change, which is then multiplied with the momentum value, which is explained in the equations below.

Certain parameter configurations set up the MLP like:

A- Learning Rate: this parameter is carefully chosen to make sure the values of the weights change quickly to responses, with no turbulence in the output of the network (the outcome should range from zero to one, without changing the default value which = 0.3 always).

B- Rate of Momentum: this parameter includes the applied weight deviation to each resulting weight deviations, accomplishing quicker learning rate. It is frequently utilized in the neural network models (the outcome should range from zero to one, without changing the default value which = 0.2 always).

C- Number of nodes/neurons: Every layer is made of other layers, the MLP notation (x,y,z) represents how many neurons are in every layer in the MLP structure, made up of a three-layer architecture. The authors selected the notation of (10,10,20). The equation above describes the parameters aforementioned with weight change relationship where:

3- Operation of Activation:

There are two major characteristics of activation for MLP: linear and non-linear. Nonlinear activation has two major functions that are illustrated in the equations below. Those two functions help with data which is non-linearly distinguishable. Its behaviour tends to seem similar to the human brain's biological neurons behaviour. The MLP functions flow starts by "entering inputs to the neurons by utilizing a linear combination of attributes and their weights as described in the first equation. V represents the outcome vector of the neurons, X is the input vector, W is the weight vector". Then adding V vector to the function of activation, which is shown in the second equation. In this article, MLP was selected to represent KSD for adaptability, ease of use, robustness, nonlinearity, and popularity.

The toolkit for the NN training in this research is WEKA, a data mining tool built at the University of Waikato in New Zealand. It employs a group of Machine Learning (ML) algorithms for data mining tasks. The algorithms are used directly in a dataset in various formats (.txt, .xls and .csv). WEKA applies algorithms for data pre-processing, classification, and clustering; it also involves visualization tools and GUI. The version used is version (3.6), which is based on java used in various development areas. The training rate was set to 0.3, while momentum set to 0.2, with a 3-layer structure made of (10,20,20) neurons for each layer, respectively.

Experiment with KSD

In the following part of this paper, experiments that were made in the article will be discussed.

A. Experiment 1: Classification capability

Proving that KSD authentication measures are understandable and measurable features that can be utilized to verify, and label users are the following experiment's mission. The result is emphasized by examining the model's capability to categorize individuals based on how they type the password. To tune the experiment, the authors used the toolkit of WEKA based on method and classifier.

To analyse the KSD, FAR and FRR were used. These measures could be processed from experiments results provided by WEKA. In the framework of the author's approach, there are four possible measures for classification:

1- True Positive (TP) represents the instances classified correctly, as instances, in this case, are the samples that belong to five individuals.

2- True Negative (TN) represents the fake users that the system that was accepted.

3- False Positive (FP) represents the users that were genuinely rejected from the system.

4- False Negative (FN) represents the fake users that were rejected from the system.

5-

TABLE 2: Experiment 1 parameters

Experiment parameters	Description
Number of volunteers	Five users. Each user is associated with his own user ID {1,2,3,4,5}
Number of samples	50 samples were obtained. In addition average values are added for each user. (50 +5) 55 samples
Password template	"P@ssw0rd", users were asked to type the same password template.
Number of trials	Ten trials for each user.
Device	Sony Xperia tablet Z, with Android version 5.1.1. It was used for training and testing.
Local machine specification	Toshiba Laptop with core i7 processor, and 6 GB RAM.
Toolkit for NN	WEKA 3.6 windows version.
WEKA-MLP parameters tuning	- Learning rate = 0.3 - Momentum = 0.2 - Number of Hidden layers =1 - Total number of layers =3 , MLP (10,10,20) - Test Option =Cross-Validation, Folds =10.

TABLE 3: Experiment 1 results

Results	TP Rate	FAR	FRR
	0.855	0.036	0.145

Correctly Classified Instances	47	85.4545 %
Incorrectly Classified Instances	8	14.5455 %

B. Experiment 2: Implementation

In the second experiment of this article, the model suggested by the authors will go through analysis with other classifiers than the MLP to see how the model reliability stands against classifications under more than one classifier. The previous experiment's configuration is repeated with changes in classifiers' tuning and changes of iterations of trials. For example, using a standard "P@ssw0rd" and asking five different individuals to enter the same password for a total of thirty times.

Keeping in mind that MLP was preferred to represent the KSD mode, other classifiers will be analyzed for comparison. Other classifiers help with modelling KSD such as:

1-Naïve Bayes: are a group of algorithms used for classification based on the "Bayes Theorem". This is an algorithms family where each shares a common value. They are often utilized to analyze sentiments and to filter spam messages/mail[23].

2-J48: Provided by WEKA software, J48 algorithm is a well-known machine learning algorithm based on J.R. Quilan C4.5 algorithm. By categorization, it assists in making decision trees that give binary trees. It is made and applied as a classifier that overlooks the dataset's unavailable values[11].

3-Sequential Minimal Optimization (SMO): SMO algorithm was brought up by John C. Platt, an optimization technique used for training Support Vector Machines[12]. When there are empty values, the SMO replaces them and converts nominal attributes to values written in binary. Be default; all the attributes are normalized with this algorithm. Although it is repetitive, it is considered more complicated, and it needs expensive third-party solutions.

TABLE 4: Experiment 2 results

Classifier \ Results	TP %	FAR %	FRR %	EER %
MLP	91.33 %	2.2 %	8.67 %	5.43%
SMO	86 %	3.5 %	14 %	8.75%
Naïve Bayes	83.33 %	4.2 %	16.67 %	10.43%
J48	81.33 %	4.7 %	18.67 %	11.68%

The experiment's results of the classifiers were compared. There were four attempts, each implementing one of the classifiers above, including the MLP, with the same group of data for the examination, the same set of features for measurements and the same parameters. As it is shown below, the best classifier according to the KSD performance measures, FRR and FAR was the MLP classifier, with having the minimal errors in the FRR (8.667%), FAR (2.2%). On the other hand, the classifier with the largest error results was the J48, getting FRR (18.667%) and FAR (4.7%).

Comparative Analysis

In this section, the authors compare the experiment results with already existing ones. Several other kinds of research kept going regarding KSD usage for smartphones, specifically with touchscreens, focusing on timing and non-timing features. This part had many experiments aiming to determine how the non-timing features affect the KSD authentication system's accuracy. There were several

attempts made in the experiment, focusing on time-only, time and size, time, pressure and size, and time pressure, size, and position features. The results are shown in the table below.

Table 5: Comparative analysis results

Experiment	# Features	FAR %	FRR %	EER %
MLP – time only	8	9.1 %	36.4 %	22.75 %
MLP – time and pressure	16	5.3 %	21.33 %	13.31 %
MLP – time and size	16	4.5 %	18 %	11.25 %
MLP – time , pressure, and size	24	3.8 %	15.33%	9.56 %
MLP- time, pressure, size, and position	**40**	**2.2 %**	**8.67 %**	**5.43%**

This article proposed behavioural authentication model based on KSD features with Neural Network Learning algorithm for Android smartphones. To accomplish that, a keyboard software was developed to gather non-timing and timing characteristics; size, pressure, position and duration time of pressing one letter. To build the MLP model, the authors used WEKA toolkit depending on their group of data. After many experiments, the results below were discovered:

- KSD is a substantial component that can be utilized to enhance authentication since the KSD behaviour is special.
- KSD delivers satisfactory level in execution quantifies as a secondary authentication factor (5.43 for EER, 8.67 for FRR, and 2.2 for FAR). Enhancements can be done to the execution by reducing FAR as much as possible by adapting the KSD model using algorithms with good fault tolerance techniques.
- The security of KSD authentication systems can be enhanced by using non-timing characteristics.
- Since it does not require further hardware, this model is suggested to be the cheapest 2nd-factor.

IV. KSD ON ANDROID DEVICES: GYROSCOPE DATA AND DYNAMIC TEXT TYPING

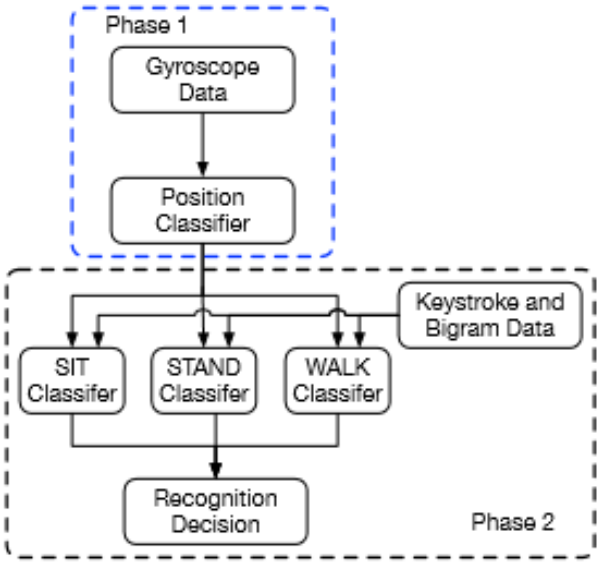

Figure 1: Gyroscope Data and Dynamic Text Typing

The next article by [13] talks about doing experiments about KSD authentication using gyroscope data and dynamic text typing data. Gyroscope data is collected by the gyroscope sensors built in the new smartphones such as motion, orientation, environmental conditions and humidity. That type of data is used for activity recognition to address typing inaccuracies[14] and to create key loggers and to determine on-device entry mistakes. The authentication research started to consider if accelerometer and gyroscope data might be utilized as a unique identifier. KSD was used in the UNAGI system made by[15]. In that article, there were experiments done using gyroscopic data and accelerometer when twenty participants keyed in a fixed group of passwords and discovered that they could reach less than 1 percent ERR values. The use of fixed passwords as a basis shows that their experiment's goal was to make it difficult to write passwords than aiming for a dynamic text experience. Another way of analysing data is encoding using Matlab to show the perceptual quality score speech signals [16].

The members in the Crawford explore, in any case, utilized an exceptionally manufactured Android application to type sentences given to them that fluctuated both their position and the gadget direction while composing. To be explicit, members were advised to hold their telephones in a specific direction (representation or scene) and to type while either situated, standing or strolling. The members were

likewise incited to type regularly; explicitly, their composing speed as there were no speed restrictions.

This article didn't give any direction about how to do their developments. For instance, numerous members decided to continue standing while inclined toward a divider or sit with their arms on a table. During the trial, the main guidelines were to continue strolling if the member halted while composing in a mobile condition. They rounded out a short segment study before beginning to type, and they were given the authorization to rest between conditions on the off chance they needed. All of them were allowed to compose the beginning of the principal condition; the preparation information was disposed of before beginning the examination.

39 members (6 females, 33 male) were enrolled through comfort testing methods like individual greeting, messages to contacts inside their surroundings and verbally expressed words. The information from three members was prohibited from the investigation because of procedural blunders, leaving information from 36 members (5 ladies and 31 men). The remainder of the members alludes to the investigation of information. The normal period of members was 28.3 years (SD = 11.3). They were not required to have any earlier information composing on contact screen gadgets, albeit every one of them professed to realize how to utilize contact screen gadgets. 2 members were left-given, and 34 were correct given. Members experience utilizing their telephones were different: 14 utilized an Android-based gadget, 18 utilized iOS gadgets, 2 utilized different brands and 2 utilized an element (non-PDAs. 2 members had utilization of their telephones for once per week, 3 as normal use (more than once yet few out of every odd day) and 31 as specialists (consistently and a few times in a day). The majority of them were understudies, personnel or staff at the college of the creators; They all had probably some post-auxiliary instruction, extending from undergrad experience to graduate levels[13].

Each participant was given an LG Nexus 5 device to use during the experiment. Every phone was operating on version 4.4.4 of Android with stock android apps only. Typing was made simpler due to using two custom Android apps. In the beginning, they displayed the phrase to be typed which was not changeable, an empty box in which the same phrase was keyed in, and a counter showing how many phrases they wrote in the state of the experiment. The application selects a phrase randomly from an edited version of the default set of phases given by[17]. Secure Data Storage System, which protects data leakage, may use this method [24].

The second app utilized in this paper was a custom-designed keyboard, made to visually mirror the default keyboard of Android for a detailed simulation of the regular typing environment; all of the participants used the same keyboard. It was this app that collected the needed bigram metrics and keystrokes. Whenever the participant hits a key, the app records they pressed key, holding time, inter-key delay, device orientation, user position and instantaneous gyroscope data. Holding time is described as the duration of holding down a certain character. Inter-key delay is described as the duration between releasing a key and pressing another or the same key.

Timing of such composing occasions is debatable in the KSD field as wrong planning can affect the deliberate composing example of a member, which affects the detailed examination results [19]. In the examination, the creators relieved such significant assets of a mistake by using a gathering of four gadgets of a similar kind with the equivalent working framework construct, which had been reset to default settings preceding the start of the trial. Moreover, the creators used a similar Android application on each gadget and prohibited the previous members' information and restarted the application between members. With these precautionary measures, the creators put forth all attempts to decrease the effects of clock errors on the consequences of the examination.

In the experiment process, every individual had to answer a short demographic survey, to begin with. They were then introduced to the app and the spoken soft keyboard they will be using to key in the phrase. They were provided with the option to practice using the regular OS keyboard if they did not know about it. Many disagreed as they believed they already had ample typing practice on the regular keyboard. They had the possibility of taking short breaks after every condition of the experiment. The authors informed the participants to type as they normally do, and their precision or quickness was not monitored. They were instructed not to use auto-correction and auto-capitalization, and that it was up to them if they wanted to correct or not. They were also instructed to keep the device orientation and to stay in the specified movement (sitting, standing, walking) they were put into by the researcher. However, they were not instructed about their walking speed, nor were they instructed about typing (leaning against the wall, sitting down).

Each individual was put into each of six experimental conditions by the researcher, and they were instructed to return to him after writing a minimum of 22 phrases. Showing the number of phrases was to collect appropriate for the study and provide an estimated data amount for each of them. When every condition has ended, there were instructions to get back to the researcher, which would put the handheld in the next state and tell the participants depending on their phone orientation and mode ("please type the next set of phrases with the device in portrait while you're seated"). After completing all the conditions, the participants were thanked for their time, and they were given a choice to go.

Development and Review of the experiments

The laboratory-based research used a "within-subjects, repeated measures design, in which the participants were assigned to one of six experimental groups that differed only in the order in which the participant completed each of the six study conditions" [13][18]. The participants finished all six conditions. The participants were allocated to each study state using a design of 6*6 Latin squares to minimize fatigues and learning effects. Each of the session went on for one hour.

V. Limitations And Recommendations

Like many kinds of research go, there will be many things that can limit how an idea is applied or how it was conducted. The concept of applying KSD is to make sure that the data is readable and available. Unlike Graphical password strategy to remember the password only [20]. There should be a soft enhancement, whether on the keyboard or the gyroscopic sensor, to collect data from the user's movements or key presses. Another limitation that is found in the process that there should be more experiments using standard physical keyboards as it would bring more successful research results for those keyboards, especially when it is done using other experiment configurations and topologies, such as the FF MLP, which can give it a higher chance of success and more accurate results. Based on the importance of text classification in different applications, the butterfly optimisation algorithm was improved and used to improve classification accuracy [21]. Besides, it should be noted that using keystroke analysis should not bring any inconvenience as there are ways that exist, which, if added to the keystroke analysis, can cause trouble for the users.

Furthermore, the KSD, as of now, is best to be considered as a second authentication factor for the platforms. Recommendations can include ways of making the KSD more acceptable between users and more applicable to systems. Also, can use secure blockchain database management system to protect valuable data [22].

Conclusion

To conclude this report, using keystroke dynamics analysis to enhance password security on mobile devices proved to have a great chance of success and how it can affect the everyday users of banking applications. This can open the door to many other fascinating kinds of research to improve the password security level. It can use even further biometrics or mixes between types of authentication and methods that can make it even more convenient and ensure users feel safe using mobile banking applications. However, there have been many limitations to the research, and the biggest part of that can be that there were no correct tools to deal with this sort of experiments. The research questionnaire to get a good insight into how the people feel about this type of addition to the system responded to warm welcomes and huge acceptance among mobile banking applications users.

References

[1]. Yaseein Soubhi Hussein, Ahmed Saeed Alabed, Mustafa Al Mafrachi, Maen Alrshdan, Qusay Al-Maatouk, Li-Fi Technology for Smart Cities, Solid State Technology, p 2391-2399, 2020

[2]. Norton Setup Blog, Top 10 most infamous digital assault in history, 2019. online: https://norton.comsetup-activate.com/blog/top-10-biggest-cyber-attacks-in-history/.

[3]. S. T. Prof P. D. Thakare, "Graphical-Based Password Keystroke Dynamic Authentication System " irjet, vol. 5, no. 2, pp. 2395-0072, 2018.

[4]. S. K. Swarna Bajaj, "Typing Speed Analysis of Human for Password Protection (Based On Keystrokes Dynamics)," International Journal of Innovative Technology and Exploring Engineering (IJITEE), vol. 3, no. 2, pp. 2278-3075, 2013.

[5]. R. Chan, "7-Eleven Japan shut down a mobile payments app after only two days because hackers exploited a simple security flaw and customers lost over $500,000," ed, 2019.

[6]. K. O'Flaherty, "Password Managers Have A Security Flaw -- Here's How to Avoid It," ed, 2019.

[7]. A. Salem, D. Zaidan, A. Swidan, and R. Saifan, "Analysis of Strong Password Using Keystroke Dynamics Authentication in Touch Screen Devices," Amman, Jordan, 2016: IEEE.

[8]. S. Mondal, "Context Independent Continuous Authentication using Behavioural Biometrics," 2015: IEEE.

[9]. T.-Y. C. Cheng Jung Tasia, Pei Cheng Cheng, Jyun Hao Lin, "Two novel biometric features in keystroke dynamics authentication systems for touch screen devices," Security and Communication Networks, vol. 7, no. 4, pp. 750-758, 2014.

[10]. Yadav, Amrendra Singh, et al. "Increasing Efficiency of Sensor Nodes by Clustering in Section Based Hybrid Routing Protocol with Artificial Bee Colony." Procedia Computer Science 171 (2020): 887-896.

[11]. R. Arora, "Comparative Analysis of Classification Algorithms on Different Datasets using WEKA," International Journal of Computer Applications vol. 54, no. 13, pp. 0975–8887, 2012.

[12]. J. Jayan, "Sequential Minimal Optimization for Support Vector Machines," in towardsdatascience, ed, 2020.

[13]. H. Crawford, "Authentication on the Go: Assessing the Effect of Movement on Mobile Device Keystroke Dynamics," Santa Clara, 2017: Usenix.

[14]. M. Goel, "WalkType: Using accelerometer data to accommodate situational impairments in mobile touch screen text entry," Seattle, 2017: Research Gate.

[15]. C. Giuffrida, "I Sensed It Was You: Authenticating Mobile Users with Sensor-Enhanced Keystroke Dynamics," 2014: Springer.

[16]. A. Z. Al-Othmani, A. A. Manaf, A. M. Zeki, Q. Almaatouk, A. Aborujilah and M. T. Al-Rashdan, "Correlation Between Speaker Gender and Perceptual Quality of Mobile Speech Signal," 2020 14th International Conference on Ubiquitous Information Management and Communication (IMCOM), Taichung, Taiwan, 2020, pp. 1-6, DOI: 10.1109/IMCOM48794.2020.9001793.

[17]. R. W. S. Scott MacKenzie, "Phrase Sets for Evaluating Text Entry Techniques," New York, 2003: York University.

[18]. Mewada, Arvind, et al. "Network intrusion detection using multiclass support vector machine." Special Issue of IJCCT 1.2-4 (2010): 172-175.

[19]. Syed Zulkarnain Syed Idrus. Soft Biometrics for Keystroke Dynamics. Computer Vision andPattern Recognition. Universit´e de Caen Basse-Normandie, 2014. English. <tel-01108638>

[20]. Yap Sing Chuen, Maen Al-Rashdan, Qusay Al-Maatouk, “Graphical Password Strategy”, Journal of Critical Reviews, Vol 7, Issue 3, 2020

[21]. M. Tubishat, M. Alswaitti, S. Mirjalili, M. A. Al-Garadi, M. T. Alrashdan and T. A. Rana, "Dynamic Butterfly Optimization Algorithm for Feature Selection," in IEEE Access, vol. 8, pp. 194303-194314, 2020, doi: 10.1109/ACCESS.2020.3033757

[22]. Teo Min Xuan, Maen T. Alrashdan, Qusay Al-Maatouk, Mosab Tayseer Alrashdan, “Blockchain Technology in E-Commerce Platform”, International Journal of Management, vol. 11, issue 10, pp. 1688-1697, 2020, doi: 10.34218/IJM.11.10.2020.154.

[23]. R. Gandhi, "Naive Bayes Classifier," in towardsdatascience, ed, 2018.

[24]. Derrick Chan Jianli, Maen Al-Rashdan, Qusay Al-Maatouk, Secure Data Storage System, Journal of Critical Reviews, Vol 7, Issue 3, 2020.