Multimodal Behavioral Biometric Authentication in Smartphones for Covid-19 Pandemic

Original Scientific Paper

Amitabh Thapliyal
Department of Computer Science and Engineering, Delhi Technological University, Delhi, India
Samsung R&D Institute, Noida
amitabh.t@samsung.com

Om Prakash Verma
Department of Electronics and Communication, Delhi Technological University, Delhi, India
opverma@dce.ac.in

Amioy Kumar
Department of Data Science, Intel Corp, Bangalore, India
amioy.iitd@gmail.com

Abstract *– The usage of mobile phones has increased multi-fold in recent decades, mostly because of their utility in most aspects of daily life, such as communications, entertainment, and financial transactions. In use cases where users' information is at risk from imposter attacks, biometrics-based authentication systems such as fingerprint or facial recognition are considered the most trustworthy in comparison to PIN, password, or pattern-based authentication systems in smartphones. Biometrics need to be presented at the time of power-on, they cannot be guessed or attacked through brute force and eliminate the possibility of shoulder surfing. However, fingerprints or facial recognition-based systems in smartphones may not be applicable in a pandemic situation like Covid-19, where hand gloves or face masks are mandatory to protect against unwanted exposure of the body parts. This paper investigates the situations in which fingerprints cannot be utilized due to hand gloves and hence presents an alternative biometric system using the multimodal Touchscreen swipe and Keystroke dynamics pattern. We propose a HandGlove mode of authentication where the system will automatically be triggered to authenticate a user based on Touchscreen swipe and Keystroke dynamics patterns. Our experimental results suggest that the proposed multimodal biometric system can operate with high accuracy. We experiment with different classifiers like Isolation Forest Classifier, SVM, k-NN Classifier, and fuzzy logic classifier with SVM to obtain the best authentication accuracy of 99.55% with 197 users on the Samsung Galaxy S20. We further study the problem of untrained external factors which can impact the user experience of authentication system and propose a model based on fuzzy logic to extend the functionality of the system to improve under novel external effects. In this experiment, we considered the untrained external factor of 'sanitized hands' with which the user tries to authenticate and achieved 93.5% accuracy in this scenario. The proposed multimodal system could be one of the most sought approaches for biometrics-based authentication in smartphones in a COVID-19 pandemic situation.*

Keywords*: Fuzzy Logic, Keystroke, Multimodal Biometrics, Smartphone, Swipe*

1. INTRODUCTION

The last decade has seen many evolutions in smartphones with touch displays, bigger screens, large memory, and processors with high capability. The most powerful and advanced systems for smartphones in this decade are Android and IOS, developed by Google and Apple, respectively. In 2018, the mobile smartphone operating system market share worldwide from these platforms was 98% with Android (76%) and IOS (22%) [1]. With a report from counterpoint research, there were 1.43 billion smartphones sold in the year 2018. According to a report from Strategy Analytics, major players such as Samsung which sold 291.3 million smartphone units, and Apple sold 215 million smartphones worldwide. Smartphones have a huge impact on people's daily lives and are not limited to calls and messaging. Its utility has increased manifold with the availability of a huge number of utility applications available for the user, including social networking, entertainment, shopping, and financial transactions. Evolution and advancement in network technology with 5G have opened up several possibilities for streaming and Internet-based applications. While all these smartphones provide convenience and improved use cases, it also brings security and privacy issues for individuals as smartphones process a large amount of private and financial data, which can cause serious loss when it falls into the wrong hands. Therefore, a strong user authentication system that provides user access to smartphones is the most important requirement. Traditional authentication approaches in smartphones, such as swipe, PIN, password, and pattern, are prone to various attacks such as shoulder surfing, guessing attacks, brute force attacks, and dictionary attacks. Shoulder surfing is a very common attack in which the user's password is compromised by peeping into the password entry screen while the actual user types in the password [2]. Biometrics such as the face, fingerprints, voice, and iris are some of the authentication solutions that are the recent trend in Smartphones to provide user access. It utilizes the physiological property of the user that needs to be presented at the time of power-on; hence, it cannot be guessed or attacked through brute force and eliminates the possibility of shoulder surfing. However, fingerprints or facial recognition-based systems in smartphones may not be as applicable in pandemic situations like Covid-19, where hand gloves or face masks are mandatory to protect against unwanted exposure of the body parts. Fig. 1 depicts some of the cases where device operations are required to be performed using hand gloves.

This paper investigates the situations in which fingerprints cannot be utilized due to hand gloves and hence presents an alternative biometric system using the multimodal Touchscreen Swipe and Keystroke dynamics pattern. We propose a HandGlove mode of authentication where the system will automatically be triggered to authenticate a user based on Touchscreen Swipe and Keystroke dynamics patterns. The proposed system incorporates the user's touchscreen swipe and typing patterns as a security layer for authentication to ramp up the total security in the system. We propose the use of a fuzzy network classifier to learn the patterns in this multimodal system to reduce the effects of hand gloves and other external factors in user authentication. Our experimental results suggest that the proposed multimodal biometrics system can operate with high accuracy and the HandGlove mode of authentication has very little or no effect of hand gloves on the accuracy of the authentication system. The proposed multimodal system could be one of the most sought approaches for biometrics-based authentication in smartphones in a COVID-19 pandemic situation. The rest of the paper is organized as follows: Section 2 presents the related work on biometric authentication for smartphones and our proposal, Section 3 presents the proposed HandGlove mode, Section 4 presents the modules of the proposed multimodal system, and Section 5 depicts the experimental results, and finally, conclusions are discussed in Section 6.

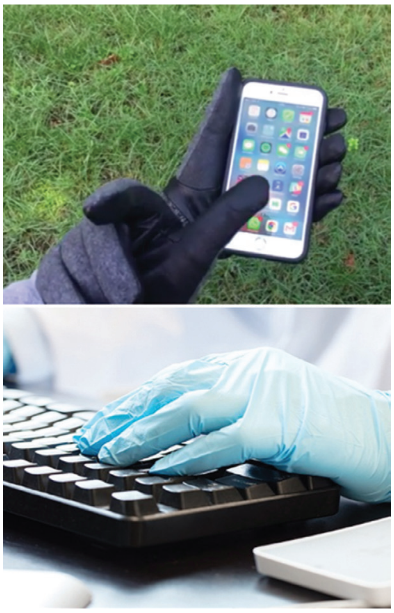


Fig. 1. Device operated using hand gloves

2. RELATED WORK AND OUR PROPOSAL

Several methods have been utilized for authentication purposes to grant users access to smartphones. Some of the popular authentication methods for smartphones are PIN, password, and pattern. However, these methods are not secure, and they have various shortcomings associated with them [3]. Owing to these shortcomings of PIN, password, and pattern-based methods, biometric-based solutions are the recent authentication trends in smartphones. Biometric-based authentication is based on the modalities and traits presented by the user, which can be physical or behavioral patterns of the user based on which they can be recognized by the system.

2.1. RELATED WORK IN SMARTPHONES

If the literature on personal authentication can be arranged chronologically, the biometric traits have been used for authentication for over a century [4], machine-based personal authentication is approximately forty years old [5], and the establishment of automatic biometric authentication as a specific area of research is more than a decade old [6]. In smartphones, the first attempt at bringing a fingerprint sensor was done by Toshiba for their G500 and G900 models in 2007, as shown in Fig. 2. Toshiba used Windows as an operating system in their smartphones, which became instantly popular among people in the days when the current mobile operating systems, Android and iOS, were still not in use. Another smartphone that attempted to implement a fingerprint sensor was the HTC P6500, which was available in the market after a few months of the G500 release. In 2013, Apple launched Touch ID where the fingerprint was used to unlock smartphones and made available to the iPhone 5S, iPhone 6, iPhone 6 Plus, iPad Air 2, and iPad Mini 3. One of the important security features in Apple Touch ID that makes it very difficult for external imposter attacks is that the fingerprint information is stored locally in a secure location on the Apple chip, instead of being stored remotely on Apple servers or iCloud. The popularity of fingerprint authentication in Apple has paved the way for almost all smartphones to add a fingerprint sensor to their flagships, for example, Galaxy S5/S6, iPhone 5S/6/6S, Huawei Mate S/Ascend, HTC M9+, Xperia Z5, One Plus Two, LG V10, etc.

Fig. 2. Popular Smartphones which initiated using fingerprint sensor Toshiba 6500, HTC P6500, iPhone 6s

To the best of our knowledge, Apple introduced Face-ID using face recognition in iPhone X for the first time in 2017. The popularity of Face-ID has led to various other Android-based smartphones introducing face recognition for user access. However, face recognition also has several limitations, such as low light accuracy, spoofing attacks using photographs, and user inconveniences [7]. Consequently, behavioral biometrics-based methods have also been utilized for user authentication. Researchers have attempted to understand and learn user behavior patterns and how they interact with systems such as keystrokes, touches, and tapping patterns on the device. These behavioral biometric methods provide several benefits over physiological methods, such as behavioral patterns that can be collected continuously and without user knowledge; they do not require any additional hardware sensors to support them. Some of the key works that researchers have attempted on behavioral biometric traits and their accuracy are listed in Table 1.

Keystroke typing pattern-based biometric authentication is based on the fact that each user's typing pattern is unique and consistent. Many approaches to authenticate a device by keystroke biometrics have been utilized in the literature. Clarke and Furnell [8] studied user authentication using keystroke dynamics on mobile devices. In their work, they have used the key typing pattern of 11-digit telephone numbers and 4-digit security PINs to distinguish users. Their models were based on the generalized regression networks with an accuracy of EERs ranging from 9% to 16%. Sunghoon Park et al., in their paper "Keystroke dynamics-based authentication for mobile devices", achieved an EER of 13% when applying the "Arthematics rhythms with Cues" [9].

Table 1. Behavioral biometric keystroke and touch dynamics

Study	Work Description	Modality	EER
N. L. Clarke et al. [8]	Authentication using keystroke dynamics	keystroke	9% to 16%
Hwang et. al [9]	Arthematics rhythms with Cues	keystroke	13%
Nan Zheng [10]	Tapping patterns	Touch	3.65%
Wang Y. et al [11]	Support Vector Machine	keystroke	8.70%
Meng et al. [12]	Neural Network with PSO	Touch gestures	2.92%
Pin Shen Teh et al [13]	Gaussian, Z-Score, Standard deviation	Touch	8.50%
Ka-Wing Tse et al [14]	RNN	Touch, keystroke	Accuracy 83.9%

Nan Zheng et al. used the union of four features that is pressure, acceleration, time, and size pulled out from smartphone sensors. Experimental results have shown that their verification system achieves accuracy with averaged equal error rates of 3.65% [10]. Meng et al. [12] leveraged touch behavioral patterns from touch gesture data collected from 20 Android phone users for training several classifiers including neural networks. In their work, they also performed optimization of neural networks by using Particle Swarm Optimization (PSO) and achieved an equal error rate of 2.92%. Pin Shen Teh et al. [13] performed an experiment in which data is collected from 150 subjects, and this dataset is shared in three packages of 50 each. In this process, subjects have to enter the same string 10 times, resulting in 20 samples per subject. The timing data and finger touch size features were captured during subject interaction. Three matching functions were used to compute the likelihood of a test sample. These three functions are Gaussian estimation (GE), Z-score (ZS), and standard deviation (SD) drift. FAR and FRR are measured to estimate the accuracy of a biometric authentication system. The Gaussian estimator (GE) gives the lowest EER value in both cases, that is, 8.55% EER when the input string is 4-digit and 5.49% EER when the input string is 16-digit. Ka-Wing Tse [14] evaluated their approach and formulated a dataset of 31 subjects where each subject had to enter a password 50 times. Temporal features, spatial dynamics features, and swipe features were extracted from the dataset. They used the RNN method, and three unique RNNs were implemented and trained separately. The results from each model were fused to obtain the final results. The results indicate that late fusion yields better results than early fusion, and the best result is achieved by spatial features, which were 83.91%.

2.2. OUR PROPOSAL

The increasing popularity of biometrics in smartphones has attracted considerable research work; thus, the literature has shown the number of potential attempts made in this area. However, our literature survey shows the following areas, which are less explored:

- Most of the biometrics utilized in smartphones are physiological, such as fingerprints, iris, face, etc. Some attempts have been made to use behavioral biometrics such as voice and signature, gait, and keystroke. However, these attempts are very few and are currently not well industrialized in smartphones.
- Most of the available work explores a single-modal biometric approach for user authentication in smartphones. Multimodal systems are mostly not considered because of the complexity of the fusion of two different biometric traits in real-time in smartphones. Multimodal refers to systems that can process and relate information from multiple modalities, in our case touchscreen swipe and keystroke typing patterns.
- Most of the available biometrics in smartphones do not consider a pandemic situation like Covid-19, where hand gloves or face masks are mandatory to protect against unwanted exposure of the body parts. In such situations, the acquisition of biometrics from the user itself is a difficult task, which further limits the use of biometric authentication.

In this research, we propose a multimodal-based behavioral biometric system that uses touchscreen swipe and keystroke dynamics patterns to uniquely identify the user and distinguish them from imposters. The highlights of the proposed work are as follows: We propose a behavioral multimodal biometric system with the fusion of the Touchscreen swipe and Keystroke dynamics. The acquisition of these two biometrics is easy and user-friendly, as both of these modalities can be acquired in one action of the hand. Another important highlight of this work is that it investigates the proposed multimodal for situations where hand gloves can be present at hand. We propose a HandGlove mode of authentication where the system will automatically be triggered to authenticate a user based on Touchscreen swipe and Keystroke dynamic patterns. The proposed HandGlove mode will be triggered by the user, and the system will incorporate the user's touchscreen swipe and typing patterns as a security means to authenticate the user.

We develop a fuzzy network classifier to learn the patterns in this multimodal system to reduce the effects of hand gloves and untrained samples in user authentication. Our experimental results suggest that the proposed multimodal biometrics system operates well with high accuracy and the HandGlove mode of authentication has very little or no effect of hand gloves on the accuracy of the authentication system. We experiment with different classifiers to obtain the best authentication accuracy of 99.55% with 197 users on the Samsung Galaxy S20. The proposed multimodal system could be one of the most sought approaches for biometrics-based authentication in smartphones in a COVID-19 pandemic situation.

A block diagram of the proposed system is presented in Fig. 3. Data collection is the first major module of our proposed system, which is responsible for extracting the keystroke and touch dynamics data from the user's input sample. Details of how the input samples are acquired are explained in Section 4.1. The feature extraction module extracts feature data from the collected sample. In the proposed work, we have used a multimodal approach, and the user sample has features for Touchscreen Swipe and keystroke dynamics. In the training module, a combined feature vector is generated with the touch-swipe and keystroke dynamics and is passed to the feature classifier after being normalized. The fuzzy logic controller unit in the fuzzy classifier is configured to convert a crisp input into a fuzzy value termed fuzzification (explained in Section 4.4). The authentication unit then makes final decisions on accepting genuine users or rejecting imposters.

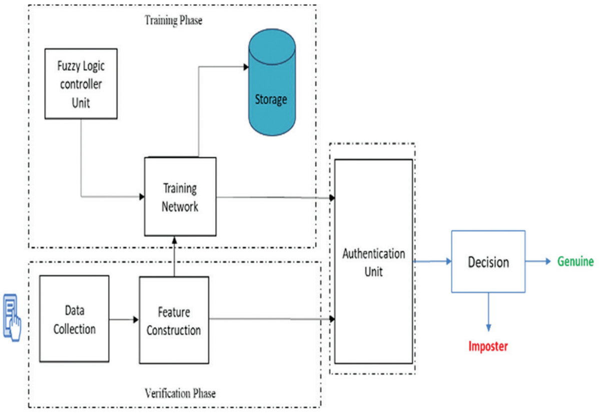


Fig. 3. Multi-Modal Behavioral Authentication Systems

3. HANDGLOVE MODE

The HandGlove mode is used to ease the user. This mode will trigger the multimodal behavioral authentication system and allow device access based on user acceptance by the proposed multimodal system using user swipe and keystroke dynamics. A depiction of the HandGlove mode in mobile devices is shown in Fig. 4.

To detect hand gloves or other external factors on the surface of mobile phones, three main techniques are considered based on the popularity and usability of touch panel devices. Fig. 5a-5c illustrate various detection mechanisms for the detection of hand gloves and other external factors such as wet hands.

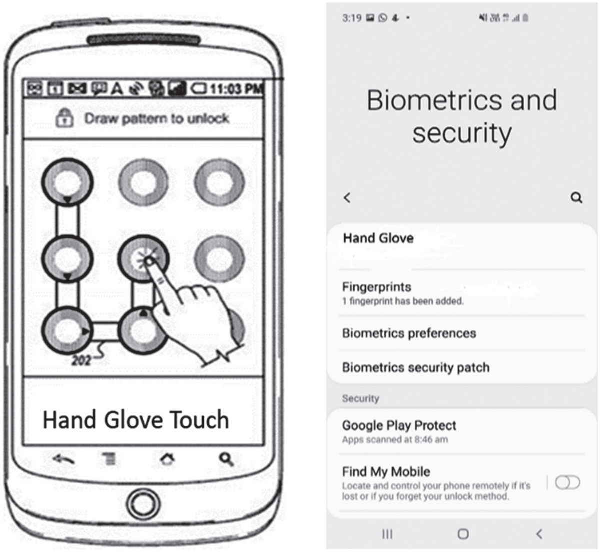


Fig. 4. Hand Glove Mode - Multimodal Behavioral Biometric

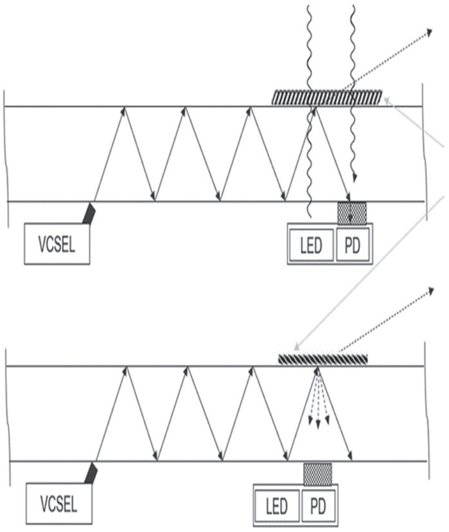


Fig. 5(a)

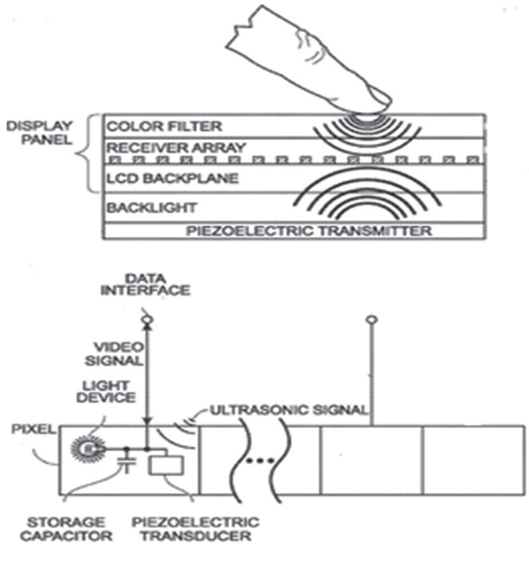


Fig. 5(b)

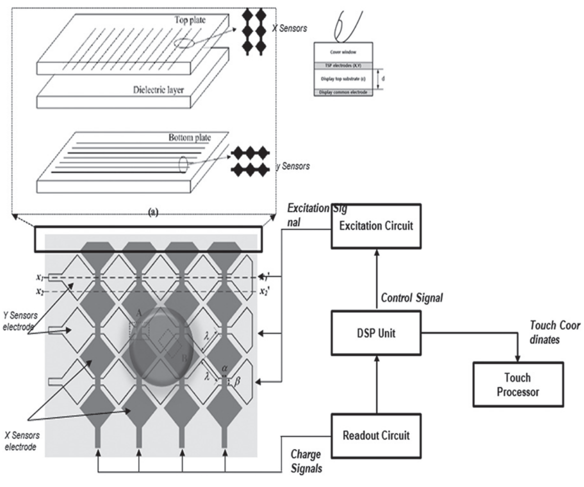


Fig. 5. Detection Mechanism of External Factors

Fig. 5(c)

A detection mechanism based on total internal reflection (TIR) within the display technique is illustrated in Fig. 5a. In physics, TIR is a phenomenon in which the complete reflection of a ray of light within a medium such as water or glass from the surrounding surfaces is reflected into the medium [15]. This phenomenon occurs if the angle of incidence is greater than a certain limiting angle, called the critical angle. Using this principle of total internal reflection, an object along with an external agent is identified. Referring to Fig. 5a, a vertical-cavity surface-emitting laser (VCSEL) or other types of light-emitting diodes capable of producing a controlled beam of infrared light via a lens are provided. When a film/layer of foreign object/contamination/external agent (e.g., finger, gloves, grease, facial oil, water, or other viscous contaminants that may prevent functionality) is present over a proximity sensor (LED), infrared light is reflected into the light detector. As long as the surface is not touched, the light remains inside the screen. However, when an object touches the screen based on the total internal reflection, the light is shattered, so the light escapes from the exact point where the pressure is applied; thus, the position of an object is accurately determined by the sensor registering the light loss. This diversion of the light is also utilized to detect the presence of external agents, such as water, on the surface of the touch screen. The detection mechanism of ultrasonic sensor-based reflection within the display technique is shown in Fig. 5b, which has several advantages over existing technologies for touch screen applications [16].

Finally, the detection mechanism was based on a capacitance-based false positive detection mechanism. Unlike resistive-based touch screens, capacitive screens do not use the pressure of an object to create a change in the flow of electricity. Instead, it works with anything that holds an electrical charge similar to that of human skin. The basic principle of the capacitance-based false-positive detection mechanism is explained below. As already known, the simplest form of a capacitor consists of two conductors, for example, two metal plates separated by an insulator. The following formula shows the parameters that influence the capacitance:

$$C = \varepsilon \times \frac{A}{d}$$
$$\varepsilon = \varepsilon_0 \times \varepsilon_r \tag{1}$$

Where, C is the capacitance, ε_r is the relative permittivity (also called the dielectric constant) of the insulating material between the plates, ε_0 and is the permittivity of free space (8.854 × 10^{-12} F/m). A is the area of the plates, and d is the distance between the plates. As shown in Fig. 5c, a flexible and thin display for smart devices having a large coupling capacitance between the sensor electrode of the touch screen panel (TSP) and the display electrode is provided to detect the external agents by utilizing a varying capacitance value that occurs due to the presence of external factors on the touch screen.

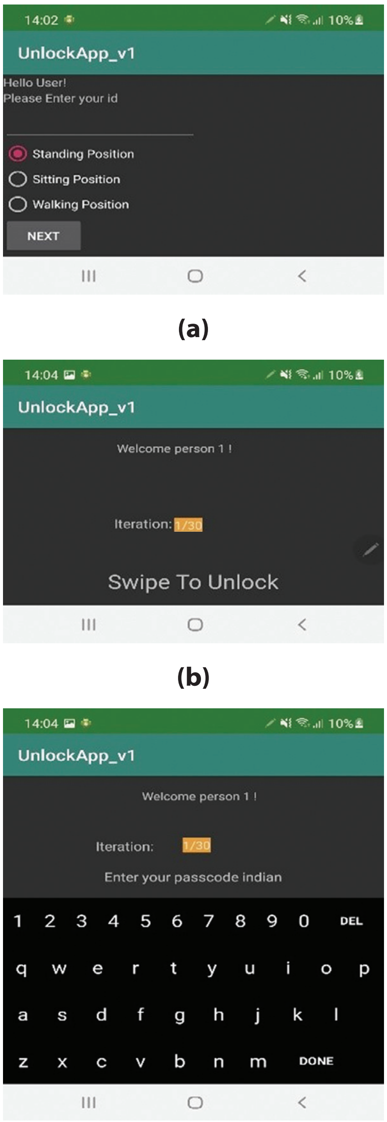
4. MODULES OF MULTIMODAL SYSTEM

In this section, we discuss in detail the various modules of our system which work together to authenticate the user.

4.1. DATA COLLECTION AND ENROLLMENT

The data collection for the proposed multimodal system can be performed once the android applications trigger the physical sensors to read touchscreen to swipe touch patterns and keystroke password-key input by users. For touchscreen swipe, accelerometer and gyroscope are the sensors used to acquire the user inputs. It captures the touch speed and distance of swipe features corresponding to each enrolled user. For keystroke, we captured the hold-time and inter-key time as a feature for each enrolled user. In contrast to the enrolment module of other biometric systems, the input to the enrolment system in the proposed multimodal system may work in continuous enrolment mode. It can read the above-mentioned features whenever a user swipes and types in a smartphone for better learning of the authentication system. The enrolment system works in the background and reads the swipe pattern and keystroke inputs when the user logs into the system. The application was developed on a Samsung Galaxy S20 device using Google Android OS, 11. We collected data from 197 volunteers (124 men, 73 women) aged between 25 and 40 years. The data collection was done in three different postures: standing, sitting, and walking. The users who participated in the data collection process are presented with a mobile application to collect sensor measurements required to calculate feature values encompassing the behavioral patterns in touch-screen swipe and keystroke dynamics. The users are required to swipe on the application and then type the password appearing on the screen. Each user recorded the data 30 times for each posture and with three different scenarios of external factors namely dry hands, wet hands, and hands with gloves, making a total of 270 data samples for each user, or 53190 data samples for 197 users. The schematic of the data collection application is presented in Fig. 6.

For the experiments reported in this paper, we collected 30 patterns from each individual in each posture. We also asked users to provide inputs with dry hands, wet hands, with gloves as part of data collection, to handle such scenarios to better train the model in HandGlove mode. In total, we collected 53190 samples from 197 users under the three mentioned postures and three external factor cases. Data collection was performed in two separate sessions for each user. The entire enrolment process took 2 weeks period to collect sample data from all 197 users. Data collection and all experiments were performed at the Samsung Research Institute, India R&D, where one of the authors is working. A multimodal spectrogram with data collected considering three scenarios- normal dry hands, gloves, and wet hands–is shown in Fig. 7 for 8 users for better representation.



14:02 10%

UnlockApp_v1

Hello User!
Please Enter your id

● Standing Position
○ Sitting Position
○ Walking Position

NEXT

(a)

14:04 10%

UnlockApp_v1

Welcome person 1 !

Iteration: 1/30

Swipe To Unlock

(b)

14:04 10%

UnlockApp_v1

Welcome person 1 !

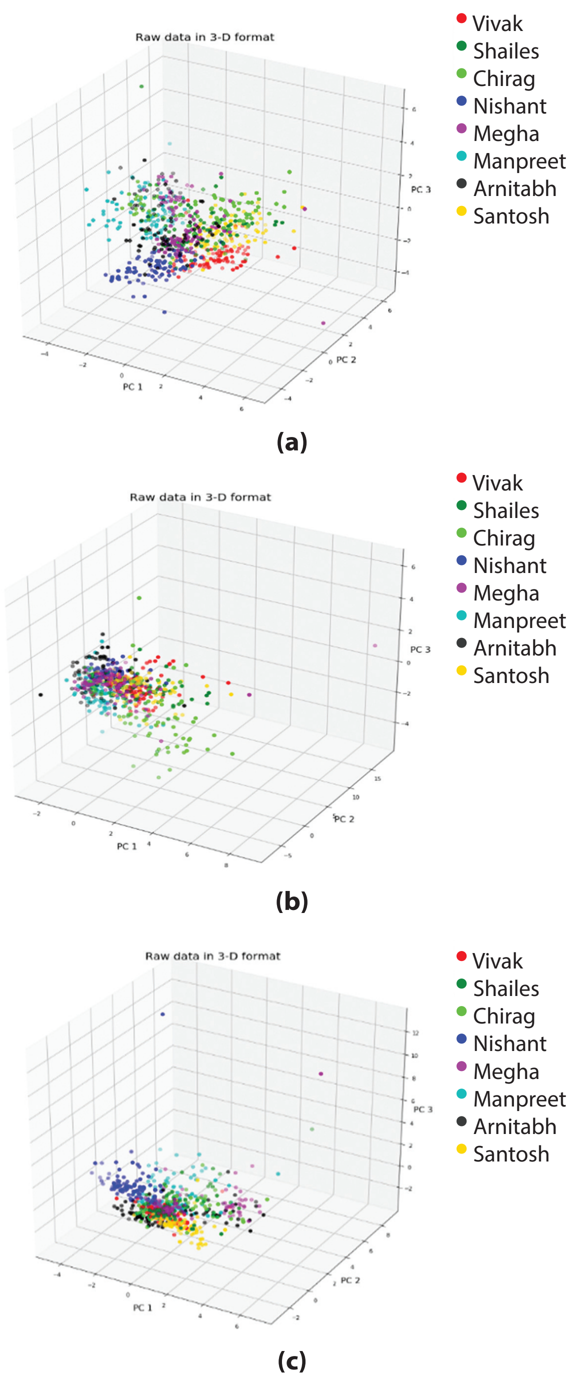
Iteration: 1/30

Enter your passcode indian

(c)

Fig. 6. Schematic of the Keystroke and Touch-Swipe behavioral data collection application from users. **(a)** Application home-screen where users set their current position. **(b)** Swipe layout where participants are asked to swipe on the screen to capture touch-swipe related feature values. **(c)** Password layout where users type the displayed password on the keyboard to capture keystroke dynamics.

In a preliminary analysis of the data set collected, it was observed that the touch-swipe and keystroke typing patterns of the users when collected have a unique pattern to distinguish the users considering the typing speed (key hold time and key switch time), touch swiping speed, and distance to unlock the device we can identify the smartphone user uniquely.



(a)

(b)

(c)

Fig. 7. 3D plot spectrogram of data collected from users with (a) Dry hands, (b) Hand Gloves, and (c) Wet hands

4.2. FEATURE EXTRACTION

The next step is to extract features from the data collected from the user. In the proposed work, we used a multimodal approach, and the user sample has features for Touchscreen Swipe and keystroke dynamics. The feature set captured for Touchscreen Swipe includes the speed of swipe, duration of swipe, and orientation of the touch area with axis along the x-axis when touched on the screen, the orientation of the touch area along the y-axis, accelerometer, and gyroscope. The feature set captured for keystroke dynamics for 6-digit passcode entry-key hold time, key switch time-the time interval to switch from one key to another, also called flight time. A combined vector comprising both modality inputs is the final feature set to be trained with the model. A description of the Touchscreen swipe and keystroke dynamics features is presented in Table 2. To use these features in multimodal authentication, we need to fuse the information extracted from them. Fusion of these features can occur at various levels, such as feature level [17-18], match score level [19], rank level [20], and decision level [21]. The literature work in the areas of biometric authentication has shown that the data fusion at the feature level incorporates the best performance. Hence, in the proposed work, we have utilized the feature level data fusion of the two behavior modalities that is keystroke and Swipe touch features. We combined the feature vectors of the two modalities and generated a combined feature vector with a total of 18 features. However, features extracted from different modalities have different value ranges; therefore, these values should be normalized to represent them in the common range of values.

Table 2. Feature set of the proposed system

Event		Features	Description
Swipe	Touch	MajorAxis	Orientation of touch area with axis along x-axis when touched on a screen
		MinorAxis	Orientation of touch area with axis along y-axis when touched on a screen
		SwipeTime	Duration of swipe
		Speed	distance covered by swipe in touch duration
	Accelerometer	A_axisMean	mean value of the list of accelerometer values
	Gyroscope	G_Jitter	the difference in the ideal signal we type or touch from the gyroscope value
		G_axisMean	axis standard deviation from the list of gyroscope values

Event	Features	Description
Keypad	Key1_Latency	Hold Time Key1
	Key2_Latency	Hold Time Key2
	Key3_Latency	Hold Time Key3
	Key4_Latency	Hold Time Key4
	Key5_Latency	Hold Time Key5
	Key6_Latency	Hold Time Key6
	Key1_2_Latency	key switch time K1->K2
	Key2_3_Latency	key switch time K2->K3
	Key3_4_Latency	key switch time K3->K4
	Key4_5_Latency	key switch time K4->K5
	Key5_6_Latency	key switch time K5->K6

In our work, we utilize the min-max normalization, which maps the minimum of a feature to zero, the maximum to one, and everything else to a decimal between 0 and 1 [22]. Given a set of *N* feature vectors x_1, $x_2, \ldots x_N$, we normalize them as

$$x_{ij} = \frac{x_{ij} - x_{min,j}}{x_{max,j} - x_{min,j}} \quad (3)$$

Where, x_{min} and x_{max} are calculated as

$$x_{min,j} = \min_{i=1\ to\ N} x_{ij}$$

&

$$x_{max,j} = \max_{i=1\ to\ N} x_{ij}$$

4.3. CLASSIFIER

After the feature extraction step, we experiment with three different classifiers namely

- Isolation Forest (IF)
- k-NN Classifier
- Radial SVM.

We partitioned the dataset into training and test sets and trained these classifiers on the training set. Each model was trained on the combined dataset of the presence of different external factors. The external factors considered in our experiments are dry hands (normal), wet hands (water), and hands with gloves. Each classifier was trained on the combined dataset collected under these three external factors present from each volunteer. The dataset also constitutes keystroke and swipe dynamics collected under three different positions of the subject, while: Standing, Sitting, and Walking. Evaluation of the model network involves computation of the false acceptance rate (FAR), false rejection rate (FRR), and equal error rate (EER), as discussed in more detail in Section 5.

4.3.1. Isolation Forest

Isolation Forest works on the principle of the decision tree algorithm and is an unsupervised technique mostly utilized for anomaly detection. This algorithm recursively generates partitions on the datasets by randomly selecting a feature and then randomly selecting a split value for the feature. Anomalies are patterns that have features that are dissimilar to the usual cases. It exploits the fact that anomalous feature observations are few and significantly different from normal observations. Let *S* be an anomaly score at an instance *t*.

Then,

$$s(t, m) = 2^{-\frac{E(p(t))}{k(m)}} \quad (4)$$

Where, *p(t)*: length of a point *t* is computed by the number of edges *t* covered in the tree until the traversal is terminated.

k(m) is the average of *p(t)* for specified *m*

$$k(m) = 2p(m-1) - \frac{2(m-1)}{m}$$

Here, E(p(t)) is the mean of p(t) from a group of isolation trees. Using the anomaly score, we can make the following assessments:

- Values close to 1 are considered an anomaly
- smaller than 0.5 considered as normal instances

We split the dataset of samples from each individual into test and training sets and train an individual isolation forest model for each individual. The samples from each person are divided into training and test sets at 85:15 proportion. For the model trained on each individual, we use the rest of the individuals' samples as test samples to evaluate the accuracy of that model.

4.3.2. k - Nearest Neighbor

k–Nearest Neighbor (k-NN) is a simple supervised classification algorithm that can be applied to both classification and regression problems. For each query sample, it finds the k number of nearest samples from the train set in the feature space according to a distance metric. We train a k-NN classifier model on our dataset as a multi-class classification model assigning a label of target identity for the test sample. We divide the entire dataset into training and test sets randomly at 85:15 proportion and classify the test samples and record the FAR, FRR, and EER of the model for evaluation. By tuning the hyper-parameters using the validation set, we used k=5 in all our experiments with k-NN. For the distance metric, we used the Minkowski distance metric which is computed as follows.

Let $X=(x_1, x_2, \ldots, x_n)$ and $Y=(y_1, y_2, \ldots, y_n)$ be the two points in the feature space. Then the Minkowski distance of order p between those two points is given by

$$D(X, Y) = \left(\sum_{i=1}^{n} |x_i - y_i|^p\right)^{1/p} \quad (5)$$

4.3.3. Radial Support Vector Machine

Support Vector Machines are primarily used for binary classification problems. They simply generate the hyperplanes to separate/classify data in some feature space into different regions. The nonlinearity is added into SVM to work well on high dimensional and linearly inseparable data using a mechanism called Kernel Trick. The Kernel function is of the form

$$K(X,Y) = (1 + \sum_{j=1}^{p} x_{ij} y_{ij})^{d} \tag{6}$$

Here *d* is the degree of the polynomial. In our experiments, we use the Radial Kernel function, which is of the form,

$$K(X,Y) = exp(-\gamma \sum_{j=1}^{p} (x_{ij} - y_{ij})^{2}) \tag{7}$$

Where *γ* is the hyper-parameter that controls the smoothness of the decision boundary and in turn regularizes the model. The regularization strength of the model is inversely proportional to *γ*. The SVMs can be used for multi-class classification problems in many different ways. We train the *N* number of SVM classifiers, where *N* is the number of identities/classes in the dataset. Each classifier learns the decision boundary between its specific class and the rest of the classes. For a new test sample, we compute the score on each classifier and decide the target class by combining all the scores.

4.4. FUZZY NETWORK

In the proposed work for HandGlove mode, we trained the network for external factors such as dry hands, wet hands, and hand gloves. However, there can be external factors other than hand gloves that could impact user input. For example, the user's hand could be affected by sanitizers, dust, oil or grease, cloth gloves, and so on. This may bring vagueness to the input presented from the user during the authentication phase with HandGlove mode, and we observed high false-positive cases. In such scenarios, the conventional machine learning-based classifiers may not be decisive and fail to handle the test input because their network is not trained for all external factors. To handle such a situation, we train a fuzzy logic classifier with SVM to incorporate fuzziness to minimize the effect of external factors on verification accuracy.

A membership function for a fuzzy set A on the universe of discourse *X* is defined as $\mu_A : X \longrightarrow [0,1]$, where each element of *X* is mapped to a value between 0 and 1. This value, called membership value or degree of membership, quantifies the grade of the membership of the element in X to the fuzzy set A. Membership functions allow us to graphically represent a fuzzy set. The input to the membership functions are the feature values and the output is the degree of membership in the [0, 1] interval for each fuzzy set. In our experiment, we implement a triangular membership function as shown in Fig. 7.

It contains a lower limit '*a*', upper limit '*b*', and '*m*', where $a < m < b$. In our case, as the feature vector $X=(x_1, x_2, \ldots, x_M) \in \mathrm{R}^M$ is the input to the membership function, the parameters *a*, *b*, and *m* are also M-dimensional i.e., $a, b, m \in \mathrm{R}^M$. The membership function is as follows.

$$\mu_A(x_i) = \left\{ \begin{matrix} 0, & x_i \le a_i \\ \frac{x-a}{m-a}, & a_i < x_i < m_i \\ \frac{b-X}{b-m}, & m_i < x_i < b_i \\ 0, & x_i \ge b_i \end{matrix} \right\} \text{for all } i = 1, 2, \ldots, M \tag{8}$$

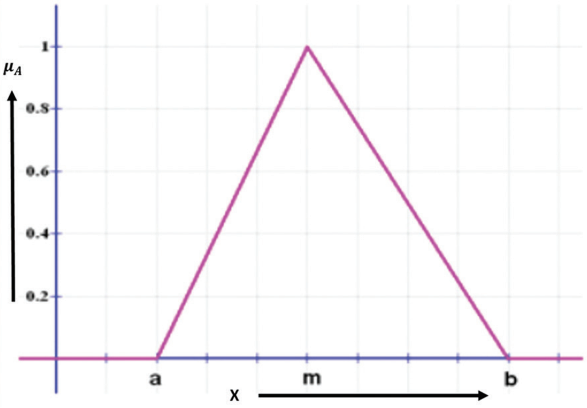


Fig. 8. Triangular membership function

We tune the values of vectors *a*, *m*, and *b* based on the training set to minimize training recognition error. The authentication module makes the authentication decision that the claimant sample matches with the owner of the device. In this case, the claimant user sample features are matched against the stored model, and the degree of membership for each fuzzy set is computed. In the matching process, the degree of membership is compared to the threshold value; if the membership degree is higher than the threshold value, the sample is classified as genuine otherwise imposter. During HandGlove mode in the case of input from the user impacted by the external factor, which is not trained example hands with sanitizer, oil, grease, etc., the proposed fuzzy logic classifier helps to get consistent performance on untrained cases.

5. EXPERIMENTAL RESULTS AND ANALYSIS

This section describes the evaluation method for the proposed multimodal behavioral biometric system on the test data set. The following sub-sections cover the evaluation metrics, methodology, results, and analysis.

5.1. EVALUATION METRICS

The accuracy of the proposed multimodal behavioral biometric system was measured using the following metrics

- False rejection rate (FRR): It is defined as the probability of a genuine user being rejected as an impostor. It is measured as the fraction of the genuine user's score below the predefined threshold.
- False acceptance rate (FAR): FAR is defined as the probability of an impostor being accepted as a genuine user. It is measured as the fraction of the impostor score (a matching score that involves comparing two biometric samples originating from different users) exceeding the predefined threshold.
- Equal error rate (EER): This is used to determine the accuracy of the proposed biometric system.

When both FAR and FRR rates are equal, the intersection point is the EER. The lower the value of EER, the higher the precision of the biometric system.

The relationship between FRR, FAR, and EER is shown in Fig. 9.

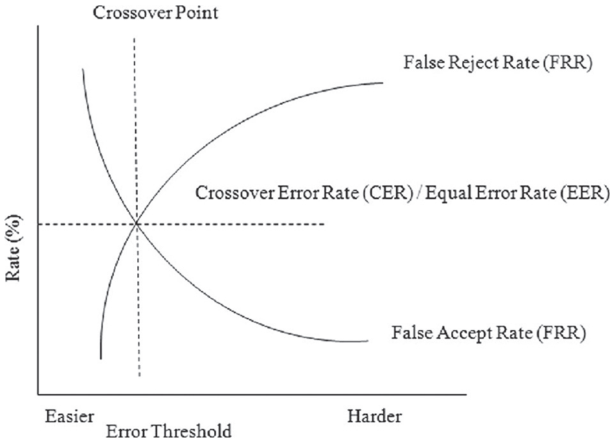


Fig. 9. Relation between FAR and FRR

5.3. EVALUATION METHODOLOGY

To evaluate the accuracy of the proposed multimodal behavioral biometric system based on touchscreen Swipe and keystroke dynamics [23-32], we performed the following task in our experiments to train with binary classifiers such as Isolation Forest, k-NN, SVM, and Fuzzy with SVM Classifier. First, we divided the subjects into two parts: one was treated as the genuine subject and the other as the imposter subject. In our experiment, a total of 197 users participated; for every mobile device, one user is the owner of the device, and his/her samples are labeled as genuine and the remaining 196 users are labeled as imposters. We partitioned the collected dataset into training and test sets in a ratio of 85:15 and trained these classifiers on the training set. We generated four models using four different training sets for different postures: sitting, standing, walking, and all postures. Both the training and the test sets contained all the variations in the external factors (dry hands, wet hands, and hands with gloves). Finally, based on the decision, the evaluation metric values were computed on testing data.

We have four sets of data samples: a genuine training set, a genuine testing set, an imposter training set, and an imposter testing set. Once we have acquired the sample sets, they are used to evaluate the above metrics of the proposed multimodal behavioral biometric system. In the experiments with fuzzy classifier with SVM, users also presented the inputs with non-trained external inputs such as hands with a sanitizer.

For training the k-NN classifier, we simply consider the identities of the users as class labels and train the model as a multi-class classification model. Generally, SVM doesn't support multi-class classification in its normal form. For multi-class classification, the basic SVM principle is utilized after breaking down the multi-class classification problem into smaller sub-problems, all of which are binary classification problems.

5.3. EVALUATION RESULTS

In experimental results, the EER value was computed for the Isolation Forest Classifier from FAR and FRR values while controlling the 'ease of acceptance' of the isolation forest by varying the contamination factor, and the intersection point in the graph between FAR and FRR for varied contamination level gives us the EER value as shown in the Receiver Operating Characteristic (ROC) Curve plot in Fig. 10.

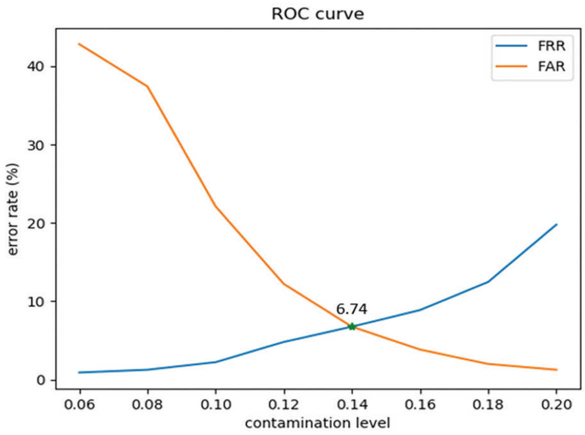


Fig. 10. ROC Curve plot of the proposed system

The equal error rate with isolation forest is obtained at around 6.74% for authentication as shown in Fig. 10. These results are obtained on the combined dataset with and without the presence of external factors such as hand gloves, wet hands, etc. for both training and validation. We conducted experiments by including individual positions in the dataset separately as well as the complete dataset with all three positions standing, sitting and walking. We also experiment with other classifiers such as k-NN and SVM as well and summarize our results in Table 3.

Table 3. Results of proposed Multimodal Behavioral Biometric System with Isolation Forest, k-NN, and SVM classifiers with the test data

Classifier	Posture	Average EER (%)
Isolation Forest	Standing	8.65
	Sitting	6.55
	Walking	8.92
	All	6.74
k-NN	Standing	4.54
	Sitting	4.08
	Walking	4.76
	All	1.58
SVM	Standing	2.04
	Sitting	0.68
	Walking	2.70
	All	0.45

5. 4. ANALYSIS

As per the results mentioned in table 3, we observed that SVM gave the best result of 0.45% equal error rate when including all the positions (Sitting, walking, and standing). SVM is closely followed by k-NN at 1.58% and then isolation forest at 6.74% ERR. The error rates are shown for each posture setting as shown in Fig. 11. Classifiers gave the best results when all the positions are included except for the isolation forest which gave the best result with the 'Sitting' position. This shows that the presence of samples of each identity in diverse positions helps to form precise decision boundaries for that identity which further increases the identification accuracy. We note that both touch swipe and keystroke dynamics for all the subjects were considered in the dataset to achieve the results. Further, we observe that the results obtained in the 'Sitting' position are better than other positions for all the classifiers as expected because the users are generally more stable while in the sitting position and the variance among the different samples obtained will be minimum. On contrary, the users will be most unstable while walking and so the variance of the samples would be considerably high, and thus walking position accuracy is the lowest.

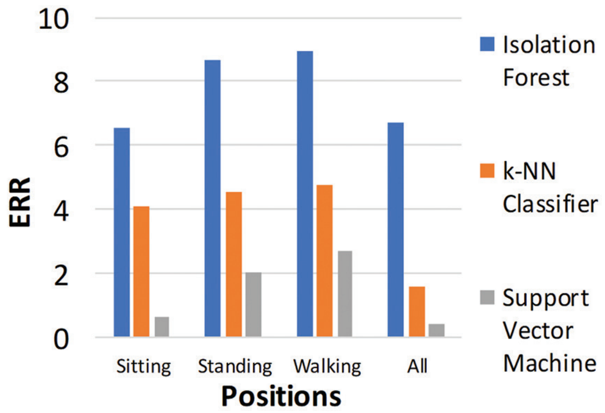


Fig. 11. Chart showing performance of classifiers for each position and all positions together

We also quantitatively compared our work with the recent existing methods utilizing touch-swipe and/or keystroke dynamics behavioral patterns for authentication/verification [8-14] in Table 4. We observed from Table 4, that our approach achieves the Equal Error Rate of 0.45% with the SVM classifier.

However, in a real case scenario, users can try to access the authentication system in presence of varied types of external factors. For example, in the current situation of the pandemic, the user may likely try to authenticate his/her mobile by swiping/typing a passcode with hands containing sanitizer, dirt or dust, etc. In such cases, the typing or swiping behavioral characteristics may vary slightly due to the presence of such external factors. So, there is a need for a system that can recognize the true owner/ imposter even when the behavioral patterns are slightly varied because of external factors. Since training the model on the dataset under the influence of all possible external factors is infeasible and impractical, we aim to explore the neighborhood similarities in feature space in an unsupervised manner to solve this problem. We argue that the behavioral features influenced by an unknown external factor 'a', will be in near neighborhood space to the behavioral features of 'closely related' external factor 'b'. For example, the behavioral patterns influenced by sanitized hands will be in the near neighborhood to the patterns influenced by wet hands in the feature space because of the closely related physical properties of sanitizer and water. We utilize this contextual neighborhood to train the model to classify into fuzzy sets instead of sharp binary sets such that it incorporates relations between the 'closely related' external factors.

Table 4. Behavioral Biometric Keystroke and Touch Dynamics

Study	Work Description	Modality	Average ERR
N. L. Clarke et al. [8]	Authentication using keystroke dynamics	keystroke	9% to 16%
Hwang et. al [9]	Arthematics rhythms with Cues	keystroke	13%
Nan Zheng [10]	Tapping patterns	Touch	3.65%
Wang Y. et al [11]	Support Vector Machine	keystroke	8.70%
Meng et al. [12]	Neural Network with PSO	Touch gestures	2.92%
Pin Shen Teh et al [13]	Gaussian, Z-Score, Standard deviation	Touch	8.50%
Ka-Wing Tse et al [14]	RNN	Touch, keystroke	Accuracy 83.9%
Proposed work	SVM	Touch, Keystroke	0.45%
	k-Nearest Neighbor		1.58%
	Isolation Forest		6.74%
	Fuzzy Classifier		6.50%

For this reason, we train a fuzzy logic classifier on the collected dataset with samples affected by only two external factors namely wet hands and gloves. We considered the triangular membership function as explained in Section 4.4 to decide positive/negative class for a sample and tuned the value of a, b, and m based on the training set. We then utilized the trained fuzzy logic classifier to classify samples of the same individuals affected by an untrained external factor like hands with sanitizer as positive/negative. The results on the untrained external factor are summarized in Table 5. We observe that the error rates of the traditional machine learning-based classifiers increased when tried to evaluate an untrained external factor case. The fuzzy classifier gave the best evaluation results on untrained cases with a 6.46% error rate. This shows that our approach can minimize the effect of external factors like sanitizer, gloves, etc. which are common during the pandemic times like COVID-19 by making use of fuzzy logic.

Table 5. Validation results of the authentication system in the presence of untrained external factor: Hands with sanitizer

Classifier	Average EER (%)
Isolation Forest	22.4
k-NN Classifier	18.25
SVM	16.5
Fuzzy Logic Classifier with SVM	6.46

6. CONCLUSION

This paper investigates the situations in which fingerprints cannot be utilized due to hand gloves and hence presents an alternative biometric system using the multimodal Touchscreen swipe and Keystroke dynamics pattern. We propose a HandGlove mode of authentication where the system will automatically be triggered to authenticate a user based on Touchscreen swipe and Keystroke dynamics patterns. The proposed system incorporates Touchscreen swipe and typing patterns as a security layer for authentication to increase the total security in the system. We propose the use of a fuzzy classification network to incorporate fuzziness in the authentication system with SVM, thereby reducing the effects of unknown external factors such as dust or sanitized hands in user authentication. Our experimental results suggest that the proposed multimodal biometrics system can operate with high accuracy and that the HandGlove mode of authentication has very little or no effect of hand gloves on the accuracy of the authentication system. We experimented with multiple commonly used machine learning-based classification algorithms to obtain the best authentication accuracy of 99.55% with 197 users on the Samsung Galaxy S20. With the developed work accuracy of 99.55% with 197 users with a Samsung Galaxy S20 device and Android R OS, 11.0. The importance of this work is mainly due to the following reasons. First, most of the biometrics utilized in smartphones are physiological, such as fingerprints, iris, face, etc. Some attempts have been made to use behavioral biometrics such as voice and signature, gait, and keystroke. However, these attempts are very few and are currently not well industrialized in smartphones. This work provides a framework for the implementation of a multimodal approach for user authentication in smartphones using touch swipe and keystroke patterns of users. It also provides extensive experimentation on a dataset created using a smartphone (Samsung Galaxy S20). The experimental results established the usability and importance of the presented work for smartphones. We use a fuzzy network to learn the patterns in this multimodal system to reduce the effects of hands with sanitizer in user authentication and achieved 93.5% accuracy on novel external factor case with SVM. The results are shown for 197 users; however, it is sufficient to conclude the potential of the presented work for user authentication in smartphones. More extensive experiments on large smartphone datasets with more variations in acquisition could be a future scope. To further increase the scope of this work, other modalities such as application usage patterns, battery charging patterns, and walking patterns of an individual can be explored as future research work for smartphone security under a behavioral biometric research scope.

7. REFERENCES

[1] Counterpoint, "Global Smartphone Market Share: By Quarter", https://www.counterpointresearch.com/global-smartphone-share/ (accessed: 2022)

[2] T. Zhao, G. Zhang, L. Zhang, "An Overview of Mobile Devices Security Issues and Countermeasures", Proceedings of the International Conference on Wireless Communication and Sensor Network, Wuhan, China,13-14 December 2014, pp. 439-443.

[3] M. Raza, M. Iqbal, M. Sharif, W. Haider, "A survey of password attacks and comparative analysis on methods of secure authentication", World applied sciences Journal, Vol. 19, No. 4, 2012, pp. 439-444.

[4] T. Sabhanayagam, V. P. Venkatesan, K. Senthamaraikannan, "A Comprehensive Survey on Various Biometric Systems", International Journal of Applied Engineering Research, Vol. 13, No. 5, 2018, pp. 2276-2297.

[5] N. Ortiz, R. Beleno, R. Moreno, Mauledeoux, O. Sanchez, "Survey of Biometric Pattern Recognition via Machine Learning Techniques", Contemporary Engineering Sciences, Vol. 11, No. 34,2018, pp. 1677-1694.

[6] A. Kataria, D. Adhyaru, A. K. Sharma, T. H. Zaveri, "A survey of automated biometric authentication techniques", Proceedings of the Nirma University International Conference on Engineering, Ahemadabad, India, 28-30 November 2013, pp. 1-6.

[7] S. Ohlyan, S. Sangwan, T. Ahuja, "A Survey On Various Problems & Challenges In Face Recognition", International Journal of Engineering Research & Technology, Vol. 2, No. 6, 2013.

[8] N. L. Clarke, S. M. Furnell, "Authenticating mobile phone users using keystroke analysis", International Journal of Information Security, 2007, pp. 1-14.

[9] S. Hwang, S. Cho, S. Park., "Keystroke dynamics-based authentication for mobile devices", Computers & Security, Vol. 28, No. 1-2, 2009, pp. 85-93.

[10] N. Zheng, K. Bai, H. Huang, H. Wang, “You Are How You Touch: User Verification on smartphones via tapping behaviors”, Proceedings of the IEEE 22nd International Conference on Network Protocols, Raleigh, NC, USA, 21-24 October 2014, pp. 221-232.

[11] Y. Wang, C. Wu, K. Zheng, X. Wang, “Improving Reliability: User authentication on smartphones using keystroke biometrics”, IEEE Access, Vol. 7, 2019, pp. 26218-26228.

[12] Y. Meng, D. S. Wong, R. Schlegel, L. Kwok, “Touch Gestures Based Biometric Authentication Scheme for Touchscreen Mobile Phones”, Proceedings of Information Security and Cryptology, Lecture Notes in Computer Science, Vol 7763. Springer, Berlin, Heidelberg.

[13] P. S. Teh, N. Zhang, A. B. J. Teoh, K. Chen, ”Recognizing your touch: Towards strengthening mobile device authentication via touch dynamics integration”, Proceedings of the 13th International Conference on Advances in Mobile Computing and Multimedia, December 2015, pp. 108-116.

[14] K. Tse, K. Hung., “User behavioral biometrics identification on a mobile platform using multimodal fusion of keystroke and swipe dynamics and recurrent neural network”, Proceedings of the IEEE 10th Symposium on Computer Applications & Industrial Electronics, 2020, pp. 262-267.

[15] L. Whitehead, M. Mossman, "Reflections on Total Internal Reflection", Optics and Photonics News, Vol. 20, No. 2, 2009, pp. 28-34.

[16] J. Dolcourt, “Galaxy S10 has an ultrasonic fingerprint scanner. Here’s why you should care”, https://www.cnet.com/tech/mobile/galaxy-s10-has-ultrasonic-fingerprint-scanner-heres-why-you-should-care-explainer (accessed: 2022)

[17] A. Ross, R. Govindarajan, “Feature level fusion using hand and face biometrics”, Proceedings of the SPIE 2nd Conference Biometric Technology Human Identification, Orlando, FL, USA, 2005, pp. 196-204.

[18] K. Chang, K. W. Bower, S. Sarkar, B. Victor, “Comparison and combination of ear and face images in appearance-based biometrics”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 25, No. 9, 2003, pp. 1160-1165.

[19] A. Ross, A. K. Jain, “Information fusion in biometrics”, Pattern Recognition Letters, Vol. 24, No. 13, 2003, pp. 2115-2125.

[20] A. Ross, K. Nandakumar, A. K. Jain, “Handbook of Multibiometrics”, Springer, Boston, MA, 2006, pp. 59-90.

[21] T. Kinnunen, V. Hautamäki, P. Fränti, “Fusion of spectral feature sets for accurate speaker identification”, Proceedings of the 9th Conference Speech and Computer, St. Petersburg, Russia, 2004, pp. 361-365.

[22] S. Gopal, K. Sahu, “Normalization: A Preprocessing Stage”, International Advance Research Journal in Science, Engineering and Technology, Vol. 2, No. 3, 2015, pp. 20-22.

[23] J. Kim, H. Kim, P. Kang, “Keystroke dynamics-based user authentication using freely typed text based on user-adaptive feature extraction and novelty detection”, Applied Soft Computing, Vol. 62, 2018, pp. 1077-1087.

[24] D. Stefan, X. Shu, D. D. Yao, “Robustness of keystroke-dynamics based biometrics against synthetic forgeries”, Computers & Security, Vol. 31, No. 1, 2012, pp. 109-121.

[25] A. Motwani, R. Jain, J. Sondhi, “A Multimodal Behavioral Biometric Technique for User Identification using Mouse and Keystroke Dynamics”, International Journal of Computer Applications, Vol. 111, No. 8, 2015, pp. 15-20.

[26] V. Stanciu, R. Spolaor, M. Conti, C. Giuffrida, “On the Effectiveness of Sensor-enhanced Keystroke Dynamics Against Statistical Attacks”, Proceedings of the Sixth ACM Conference on Data and Application Security and Privacy, New Orleans, USA, 9-11 March 2016, pp. 105-112.

[27] C. Giuffrida, K. Majdanik, M. Conti, H. Bos., “I sensed it was you: authenticating mobile users with sensor-enhanced keystroke dynamics”, Lecture Notes on Computer Science, Springer, Vol. 8550, 2014, pp. 92-11.

[28] X. Huang, G. Lund, A. Sapeluk, “Development of a typing behavior recognition mechanism on android”, Proceedings of the IEEE 11th International Conference on Trust, Security and Privacy in Com-

puting and Communications, Liverpool, UK, 25-27 June 2012, pp. 1342-1347.

[29] C. J. Tasia, T. Chang, P. C. Cheng, J. H. Lin, "Two novel biometric features in keystroke dynamics authentication systems for touch screen devices", Security and Communication Networks, Vol. 7, No. 4, 2014, pp. 750-758.

[30] D. Umphress, G. Williams, "Identity verification through keyboard characteristics", International Journal of Man-Machine Studies, Vol. 23, No. 3, 1985, pp. 263–273.

[31] R. V. Yampolskiy, V. Govindaraju. "Behavioral Biometrics: A survey and classification", International Journal of Biometrics, Vol. 1, No. 1, 2008, pp. 81-113.

[32] A. Masri, "Active Authentication Using Behavioral Biometrics and Machine Learning," George Mason University, Fairfax, VA, USA, Ph.D. Thesis, 2016.