enhAnced Mobile BiomEtRics

DELIVERABLE: D4.12
Academic paper 4.6:
Multi-modal Verification

Contract number:	675087
Project acronym:	AMBER
Project title:	Enhanced Mobile Biometrics
Project duration:	1 January 2017 – 31 December 2020
Coordinator:	Richard Guest, University of Kent, Canterbury, UK

Deliverable Number:	D4.12
Type:	Paper
Dissemination level	Public
Expected submission date	January 2021
Date submitted:	January 2021

Authors / contributors	Elakkiya Ellavarason
Contributing partners	University of Kent

RESEARCH & INNOVATION
Marie Skłodowska-Curie actions

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 675087

Multi-modal Verification

1. Introduction

Literature shows that a biometric authentication system using a single trait is susceptible to issues such as spoof-attacks and intra-class variability [17], [18], [125]. Owing to these existing challenges, the research direction shifted towards exploring methods to enhance the reliability and security of the touch-dynamics based biometric authentication. In view of this, a multi-modal authentication framework was developed that utilises characteristics from multiple touch modalities for performing the authentication. Unlike a uni-modal system, these systems depend on multiple sources of information to establish the identity of a person, hence enhancing security as it gets difficult for an attacker to spoof multiple traits simultaneously of a genuine user.

A multi-modal system can be built based on physiological or behavioural modalities, or a combination of both. However, literature reveals limited work on integrating exclusively touch-dynamics based behavioural biometric modalities in a multi-modal authentication system specifically performed on mobile devices. Based on this identified need, in this chapter an in-depth investigation of fusing multiple touch-dynamic based behavioural biometric modalities has been explored. Three modalities – swipe gestures, signature and keystroke dynamics have been integrated, multiple evaluations have been conducted, and the results are presented in this chapter. The novelty of this work is, in this chapter, the impact of multiple usage scenarios across the three simultaneous modalities has been evaluated in order to establish the reliability and accuracy.

In this chapter, performance assessment of different combinations of touch-dynamics based modalities has been reported. In order to conduct the assessment, a multi-modal framework has been developed using a feature-fusion method which combines two or more biometric modalities. This system makes use of multiple traits obtained from a single sensor of the mobile device. Only the touch screen sensor of the mobile device has been utilised to acquire the data as using multiple sensors can introduce additional noise. The acquired data was presented to multiple classifiers (SVM, k-NN and Naïve Bayes) for verification. Additionally, a score-fusion method was evaluated for signature and swipe gestures using the commercial signature verification system.
Based on the identified problems in the domain of multi-modality, various research questions aligning to the overall research objectives (described in **Error! Reference source not found.**) were developed. They are as follows:

- Does a multi-modal verification system using touch-dynamics based behavioural biometrics improve verification accuracy compared to a uni-modal solution?
- Does the impact from usage scenarios be seen using the multi-modal solutions?
- Combine different touch-dynamics based modalities and identify which combination performs the best.

The remainder of the chapter is organised as follows - Section 2 details the related work on this topic, Section 3 presents the description of the experimental framework. The detail of the dataset used, modes of operation, integration strategy and the multi-modal framework have been provided in this section along with the feature fusion and authentication phase information. Following this, Section 4 presents the results obtained based on different evaluations. Section 5 presents conclusions drawn based on the results.

2. Related Work

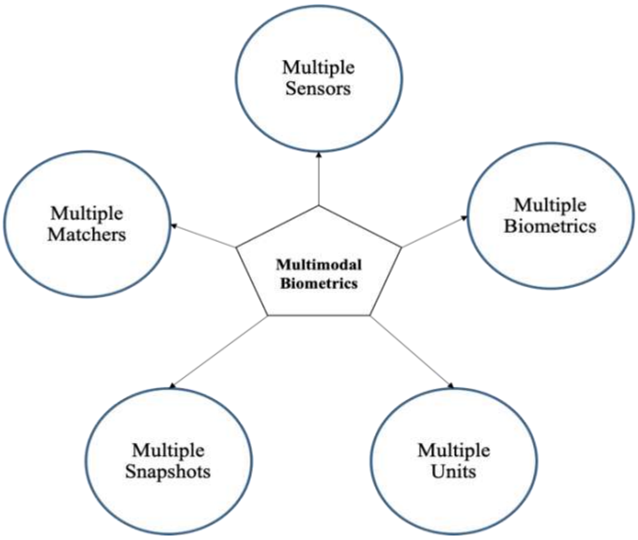

Figure 1. Multimodal biometric scenarios [128]

Typically, a multi-modal system adopts a specific fusion topology. The categories of the fusion topology taken from [126] are presented in Figure 1. Choosing the fusion topology depends on the number of modalities, feature sets and sensors. Based on these factors, the types of fusion topology are categorised by - multiple biometrics/modalities, units, snapshots, matcher and sensors. The multiple biometric/modalities method is used when different types of biometric modalities are combined. Multiple unit fusion involves utilising two or more units of the same biometric modality, for example, finger and stylus-based signatures. When two verification attempts or enrolment templates of the same biometric trait are used, it is known as multiple snapshot fusion. When multiple matchers are used on the same biometric trait, it is called as multiple matcher fusion. These classifiers can use different feature sets of the same modality or use the same feature set for processing. In case where a single biometric modality is captured through multiple sensors and the data from individual sensors are used for fusion, this method is known as the multiple sensors' fusion method.

Literature reveals that the multi-modal systems on mobile devices adopt one of the above-mentioned fusion topologies. The most common method of fusion is multiple biometric fusion. With respect to behavioural biometric modalities, two categories of biometric fusion are seen in the literature - fusing only behavioural traits [127], [128], [129] or fusing physiological and behavioural biometric modalities together [78]. **Error! Reference source not found.** presents the review of a number of studies using behavioural modalities for multi-modal based biometric systems.

Saevanee et al. [127] used behavioural profiling, linguistic profiling and keystroke dynamics fusion. Their experimental results showed that via text-entry method, the users can be authenticated with an average EER of 3.3%. They also report 91% reduction in the number of intrusive authentication requests using this system. Tanviruzzaman et al. [128] used location tracts and gait signals to generate a multi-modal system of 13 users. The data was captured using Google Nexus S device and acquired an average EER of 10%. Xu et al. [21]'s study combined touch-dynamics based modalities on mobile devices such as keystroke dynamics, handwriting, swipe, pinch and slide. Their experiment was conducted on the data collected of 30 users on a Samsung Galaxy SII device. They reported an average EER of 0% for slide and pinch when 3 to 5 consecutive operations are combined. However, keystroke and handwriting did not show a stable performance over time.

In conclusion, efforts are being taken to develop multi-modal systems with optimised accuracy and sensor data availability. This is an emerging domain, especially in terms of using behavioural biometric modalities in a multi-modal framework. Our work aims at exploring fusion of touch-dynamics based modalities acquired from a mobile device. The novelty of this experiment is that we have investigated the stability of fusing behavioural data in different usage scenarios and time-separated data, in context of a multi-modal application on mobile device. In the context of mobile biometrics, a feasible system utilising active-user touch interaction on a mobile device needs to be explored. Therefore, in this chapter, a *single sensor* based multi-modal system utilising swipe gestures, signature and keystroke dynamics has been developed and discussed. A number of evaluations were carried out based on this system and the results are presented in Section 4.

3. Experimental Framework

This section describes in detail the dataset used, modes of operation, integration strategy, classifiers and the framework design.

3.1. Dataset

The dataset used for this experiment is the multi-modal dataset described in **Error! Reference source not found.**. A total of 50 participants donated three different touch-dynamics based behavioural biometric modalities - swipe gestures, signature and keystroke dynamics in two sessions, separated by a week. The data was acquired under various usage scenarios as well. Different UI contexts were used to capture these data. For example, for a keystroke task, the UI context was an alphabetical key entry task using a soft keyboard on the mobile device and for swipe gesture capturing task, it was image navigation for vertical and horizontal swipes. As the participants used fingers to perform the swipe gesture and keystroke dynamics tasks, therefore, only the finger-based signature data were used in this multi-modal experiment. The stylus-based signature data were excluded.

3.2. Mode of Operation

A multi-modal system can work in different operational modes – serial or parallel. In the serial mode of operation, the outcome of one modality is used to verify the identity before using the next modality. When the outcome of multiple modalities is used at the same time in the authentication process, it is known as parallel mode of operation.

With regards to the data capturing method, the modalities were captured in serial mode, one after another. In this study, only the touch sensor of the mobile device has been utilised for data acquisition, therefore, the data inconsistency from other sensors was avoided. However, in the data collection setup, the touch operation of all three modalities were independent of each other. That is, only one specific modality was captured at a given time and therefore the modalities could not be captured in parallel.

With regards to the data processing method, parallel mode of operation was chosen. The experimental evaluation was performed off-device. Therefore, one of the hypotheses of this analysis was availability of data from all three modalities. Considering the practical implementation, in order to conduct a feature-level fusion of two or three modalities, it is important to have data from all the modalities available in real-time.

3.3. Integration Strategy

The integration strategy adopted in this study was feature-level fusion. One of the objectives of the current analysis is to show the advantages of using feature-fusion techniques on behavioural biometric modalities and the subsequent authentication accuracy improvement. Compared to the information available when using a match score outcome, the feature set from a biometric modality contains richer information from the raw data. Hence, fusion at this level can enhance the authentication accuracy.

However, feature-level integration has inherent issues such as incompatibility of scaled feature sets. This is due to an increased dimensionality of a single feature vector due to concatenation of features from different modalities. In this experiment, we use a feature selection method to deal with the *'curse of dimensionality'* problem. All three modalities generate fixed length temporal feature sets and these feature sets were normalised, hence achieving feature compatibility.

3.4. Multi-Modal Framework

Based on the above-mentioned mode of operation and feature fusion strategy, a multi-modal framework was developed. This multi-modal system worked on two different phases – feature fusion phase and authentication phase. The feature fusion phase is depicted in Figure 2 and the authentication phase is depicted in Figure 3, which is further divided into enrolment and verification steps.

3.4.1. Feature Fusion Phase

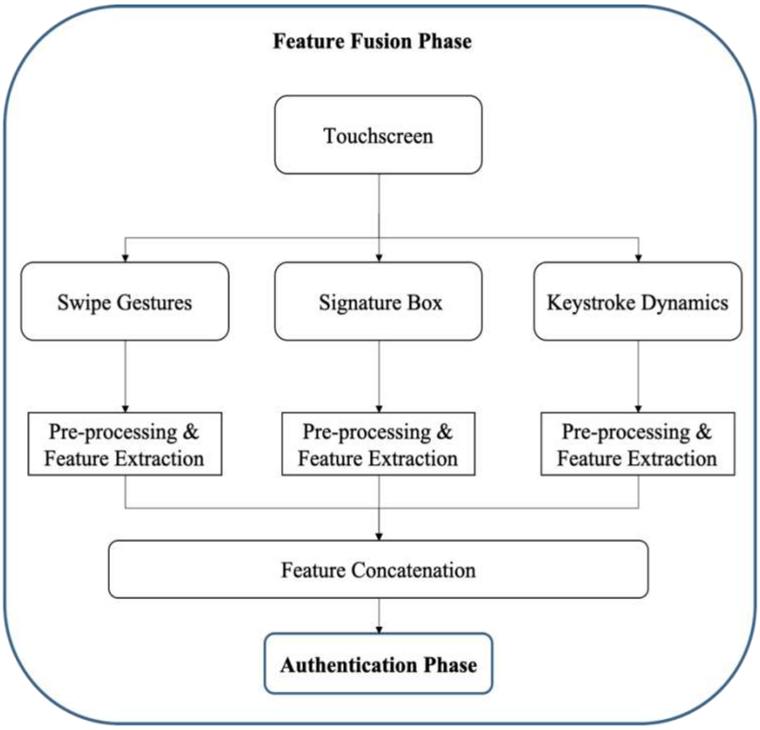

Figure 2. Feature fusion phase

This phase consisted of pre-processing, feature extraction and feature concatenation of the touch input.

- *Step 1. Pre-processing* – For swipe gestures, the pre-processing steps consisted of first separating the horizontal and vertical swipes. The next step was to identify outliers in terms of low number of data points and invalid swipe inputs. Swipes containing less than three data points were discarded. The swipes which had ACTION_UP value missing from the TOUCH ACTION parameter were also considered invalid. The pre-processing of signatures involved removal of incomplete signatures, normalisation to avoid inconsistency due to screen location and generation of fixed length signatures. For keystroke dynamics, the pre-processing step involved removal of incomplete data entry samples, such as incomplete phone number entry.

- *Step 2. Feature Extraction* – The feature extraction step for all the three modalities were performed individually. The number of features and the type of features extracted for each modality were different. The list of features along with the description of each feature is provided below.
  - Swipe gestures - For every swipe stroke, a set of 28 global features were computed (listed and detailed in **Error! Reference source not found.**).
  - Signature - The pre-processed inputs for a signature were used to extract global features listed in **Error! Reference source not found.**.
  - Keystroke Dynamics - The input sample of a keystroke dynamics data consisted of the entire sentence/phone number entry typed by the user. Based on this input sample, the global features extracted are described in Table 1.

Feature	**Description**
Total time	Total time spent on typing the input sample
Number of errors	Total number of errors committed while typing the input sample
Average flight time	Average flight time of different digraphs extracted from the input sample

Table 1. Keystroke Dynamics feature set

- *Step 3. Feature Concatenation* - In this step, all the feature sets from individual modalities were concatenated together to form a single feature vector. 28 features from swipe gestures, 22 features from signatures and 3 features from keystroke dynamics were concatenated. A total of 53 features were used. This feature vector was fed into the authentication phase.

3.4.2. Authentication Phase

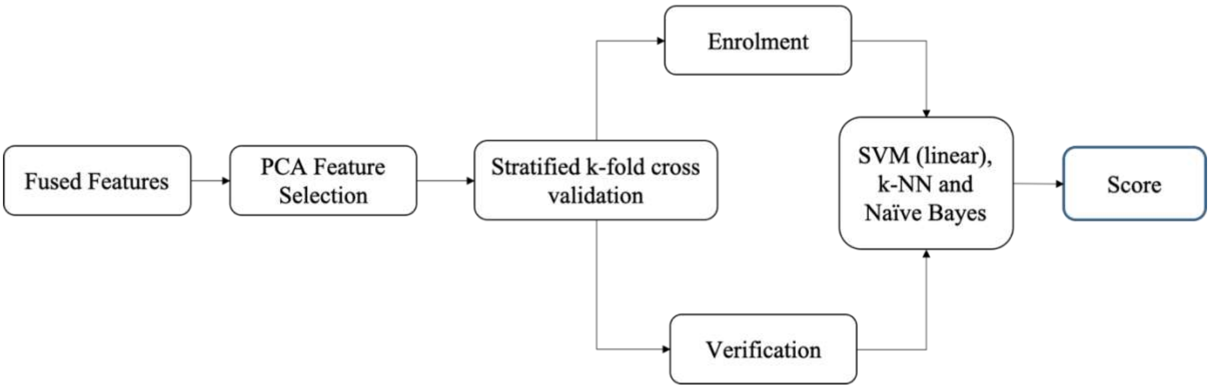

Figure 3. Authentication phase

After the completion of the feature fusion phase, the authentication phase received the fused feature vector as an input. This fused feature vector underwent feature normalisation and selection process. In order to have a dimensionally-reduced feature set and to optimise the computational time and prediction performance, Principal Component Analysis [130] feature selection technique has been applied. The dimensionally reduced features were then split into enrolment and verification sets from the input data. The verification was performed based on three different classifiers – SVM, k-NN and Naïve Bayes and a final match score was generated for the incoming input sample.

Step 1. Feature Normalisation - Individual feature values of two feature vectors X and Y may have different range and distribution. Feature normalisation is performed to adjust the mean and variance of each individual feature value to normalise and compare the contribution of each feature to the final match score. A min-max normalisation technique was adopted in this phase. The formula used to find the normalised feature *x*' of every individual feature (F_x) is provided below:

$$x' = \frac{x - \min{(F_x)}}{max(F_x) - min(F_x)} \tag{1}$$

Each feature in the feature set was normalised based on this formula. After this, feature selection was performed on the normalised feature set.

Step 2 - Feature Selection - Concatenating two feature vectors - x' and y', results in a new vector – z' = {$x'_1, x'_2, \ldots x'_n, y'_1, y'_2, \ldots y'_m$}, where n and m are the total number of features in x' and y' respectively. The idea behind the feature selection process is to choose a minimal feature set of size k, where k < (n+m) contains maximum characteristics, hence improving the classification performance. PCA has been used to perform the feature selection. The process in PCA involves the calculation of a matrix that defines the relation of different variables with each other. This matrix is then divided into two components – direction and magnitude. Finally, transformation of the original feature set data to align with the direction is obtained by deriving the principal components 1 and 2, which contains the maximum explained variance ratio. Figure 4 shows that first two principal components (PC1 and PC2) accounted for maximum variance ratio of 30% and 27% respectively. Therefore, these two principal components were considered, and the discriminative features were selected based on this.

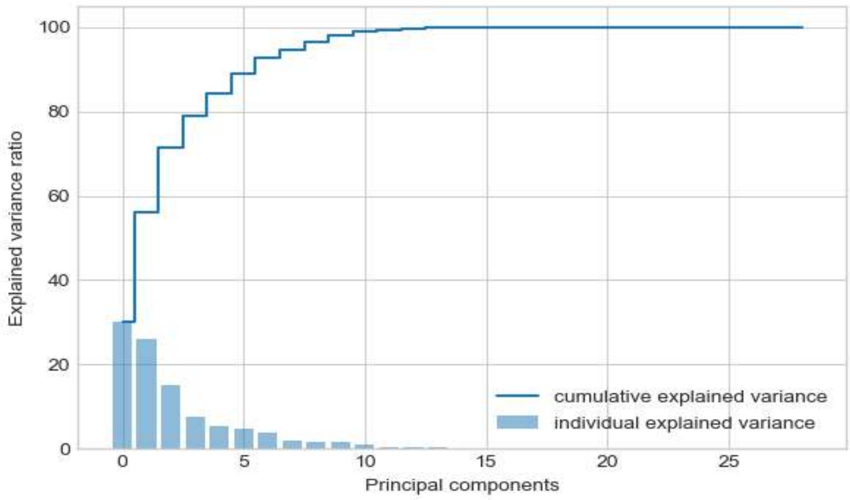

Figure 4 Principal components explained variance

Step 3. Stratified K-fold Cross validation - Once the normalised and dimensionally reduced feature set were acquired, input samples were split into enrolment and verification sets. In order to do this, stratified k-fold cross validation method was adopted.

Cross-validation is a resampling procedure used for evaluating the models on limited data samples such as the one used in our experiment. Parameter '*k*' refers to the number of groups that a given data sample is to be split into. These are the steps carried out using the stratified k-fold cross validation:

- Feature set samples were shuffled randomly.
- Feature sets were split into 5 folds.
- For each unique group:
  - One group was set aside as test data set
  - The remaining groups were taken as a training data set
  - The model was fit to the training set and evaluation was performed on the test set
  - An evaluation score was obtained
- Finally, the evaluation scores were accumulated and average EER was obtained.

The number of folds was chosen as five owing to the limited number of samples. The shuffle parameter was set to true, for shuffling the indices of each class samples before the class split.

Step 4. Enrolment - The authentication configuration for individual user present in the dataset was built. The owner of the mobile device was considered as the genuine user and the imposter user was chosen using the random forgery method. The authentication model was enrolled with the concatenated features of the genuine user.

Modality	Session 1	Session 2
Swipe gestures	175	155
Signature	18 Finger-based	18 Finger-based
Keystroke dynamics	13 alphabetical, 15 numerical	13 alphabetical, 15 numerical

Table 2. Total number of samples from each modality per individual user

The samples were divided based on the usage scenarios (Scenario 1, 2, 3 and 4) and sessions (Session 1 and Session 2). The swipe gesture data were divided as horizontal and vertical swipes. A total of 175 horizontal swipe and 155 vertical swipe samples were used from Session 1 and Session 2 respectively. Each user had 18 finger-based signature samples and the keystroke dynamics had 13 alphabetical samples and 15 numerical samples were used.

As the number of samples acquired from each of the modality varied, a different combination of the enrolment samples set was formed with the existing number of samples from each modality. For swipe gestures, each user had around 60 samples from each usage scenario of Session 1. These 60 samples were split into 50% in training and 50% in testing class. The finger signatures were limited to 10 samples per scenario; therefore, the same samples were repeated on different swipe data / keystroke data to form the input sample. Similarly, with the keystroke dynamics data, same data were used with different swipe gestures to increase the number of input samples used.

The enrolment strategy differed based on the research question under investigation. For example, horizontal or vertical swipe with a combination of keystroke dynamic data and signature data from the baseline scenario were used in enrolment. Additionally, the enrolment was performed for each combination of fused modality as listed below:

- Enrolment combination 1 – Swipe gesture features + Signature features
- Enrolment combination 2 – Keystroke Dynamics features + Swipe gesture features
- Enrolment combination 3 – Keystroke Dynamics features + Signature features
- Enrolment combination 4 – Swipe Gesture features + Keystroke Dynamics features + Signature features

All the enrolment samples belonged to Sitting Indoors scenario (Scenario 1) of Session 1 when Session 2 evaluation was performed and Sitting Indoors scenario (Scenario 1) of Session 2 when Session 2 evaluation was performed.

Step 5. Verification- Every incoming sample was verified against the existing template generated in the enrolment phase. The classification algorithms used for this analysis were SVM, k-NN and Naïve Bayes (described in **Error! Reference source not found.**). The verification sample was compared against the enrolment samples and a probability score was generated based on the classifier's output. This score was used to generate the false acceptance rate and false rejection rate. Further, the equal error rate was generated using these parameters to analyse the performance of the system. Different thresholds on the match scores were used to classify between genuine and imposter samples and the analysed results based on these thresholds are provided in the results section.

3.5. Score Fusion

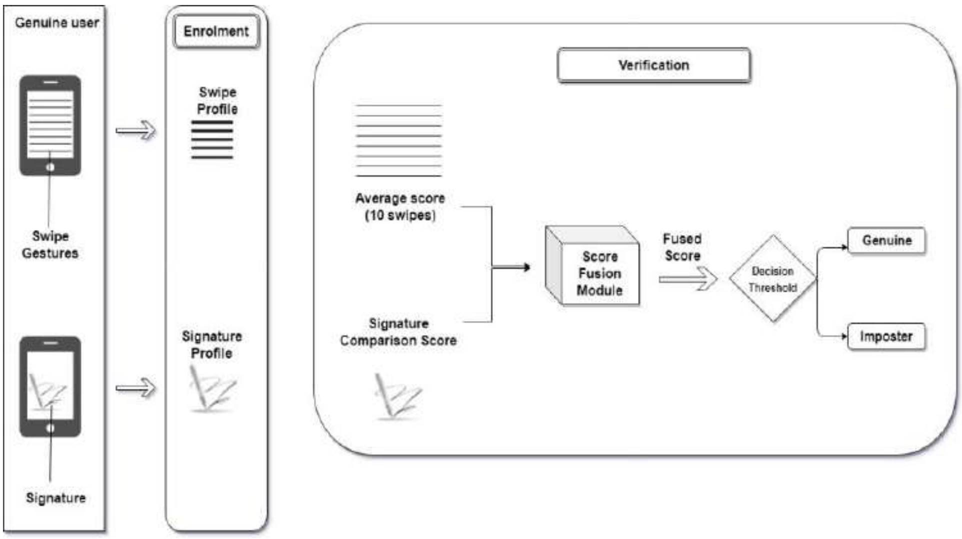

Figure 5. Score fusion method adopted using the commercial signature verification system

Score fusion method was applied for combining signature and swipe gesture modalities using the commercial signature verification system. The context of the application that was considered for using such a method was a contract signing scenario, where the user first browses through the document browsing (hence producing swipe gestures during this process) and finally, signs contract.

The method adopted to perform the score fusion has been presented in Figure 5. The swipe gestures and signatures of the individual user from the dataset were enrolled separately under two different profiles of the same user. Only finger-based signatures were used for this analysis. It was assumed that the user performed a number of swipe gestures while reading the document on the device before signing at the end of the document. Therefore, instead of combining individual swipe gestures to the signature, the match scores generated from ten swipe gestures were combined to generate an average score of swipe gestures. This score was combined with the match score from the signature.

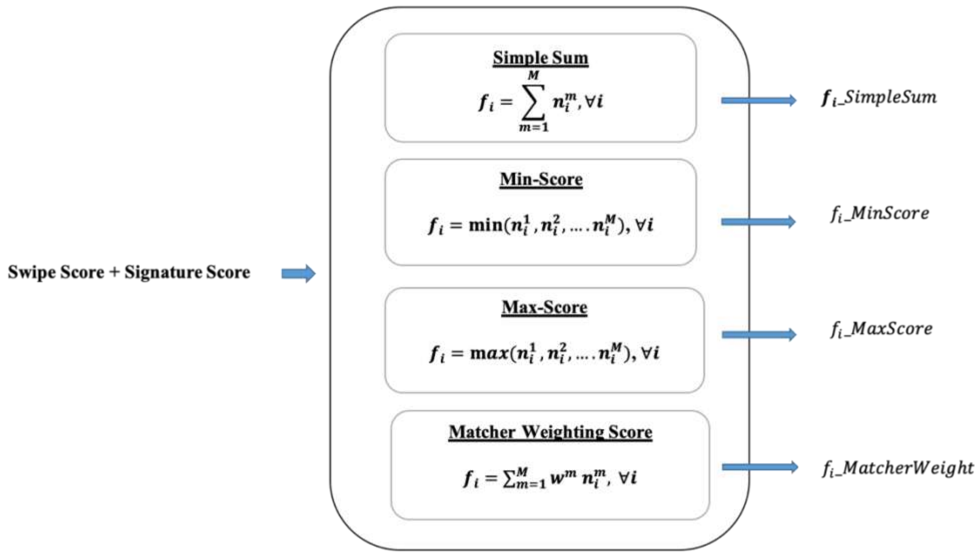

Figure 6. Score fusion module (*f*-fused score, n-score, m-matcher, M - total number of matchers)

The score fusion module combined the individual scores from signature and swipe gestures using four different methods – simple sum, min-score, max-score and matcher weighted score. The simple sum method simply added the scores from both the modalities. The min-score and max-score methods chose the minimum and maximum scores respectively, obtained from both the modalities. Finally, the matcher weighting method assigned weight to each modality. The weights given for each modality were calculated using the formula

$$w^m = \frac{\frac{1}{\sum_{m=1}^{M} \frac{1}{r^m}}}{r^m} \tag{2}$$

Where w^m is the weight of a matcher and r^m is the equal error rate of a matcher. Based on this, the signature matcher was given 0.8 weight and swipe gestures matcher were given 0.2 weight.

4. Results

In this section, the results obtained for different evaluations are presented. The metric used for this study is EER. Different combinations of modalities: a) swipe gestures and signature, b) swipe gestures and keystroke, c) signature and keystroke dynamics, and d) all three modalities are analysed. The results acquired from each of the individual combinations are explained in detail in the following sub-sections.

The results are presented as a mean EER percentage, acquired from 50 verification configurations having varied thresholds. The verification configurations belonged to the individual users present in the dataset. All the modalities were combined using the feature fusion method except for the swipe gesture and signature combination, which has been analysed using both -feature and score fusion methods.

4.1. Swipe gestures and Signature

Two techniques were used for combining the swipe gestures and the signature – a) feature-fusion and b) score-fusion method. The score-fusion method was applied on the commercial signature verification system. Here, the match scores generated from the swipe gesture verification and signature verification of the same user were combined. Verification using both the modalities were performed on the commercial signature verification system. The match scores were fused using different score fusion techniques and the results obtained using all these techniques are presented in this section.

4.1.1. Feature-Fusion Method

In order to combine the swipe gesture and signature modalities, the features extracted from each modality were concatenated separately for the genuine and the imposter user. First, the swipe gesture data for the genuine user was categorised into horizontal and vertical swipes. Once they were separated, feature extraction was performed, and swipe gesture feature set was obtained. The finger-based signatures belonging to the same user was used for generating signature feature set. The swipe gesture and signature features of the same user were then concatenated together to form the input to the multi-modal framework. Here a combination of fused data was formed as following:

- Horizontal swipes and Finger signatures (Scenario 1, 2, 3 of Session 1 and Scenario 1,3 & 4 of Session 2)
- Vertical swipes and Finger signatures (Scenario 1, 2, 3 of Session 1 and Scenario 1,3 & 4 of Session 2)

Each of these combinations were used to evaluate the performance of the multi-modal system. First, horizontal swipes and finger signatures obtained in the baseline scenario (Sitting Indoors - Scenario 1) were combined. Both the enrolment and the verification samples belonged to this scenario.

Next, the enrolment samples belonging to Sitting Indoors (Scenario 1) were compared with samples from Treadmill (Scenario 2) and Walking Outdoors (Scenario 3) of Session 1 for both horizontal and vertical swipes respectively. Similarly, the Session 2 enrolled samples from Sitting Indoors (Scenario

1) were compared with Walking Outdoors (Scenario 3) and Travelling on a Moving Bus (Scenario 4). The mean EER% obtained using different classification algorithms are presented in Figure 7 and Figure 8.

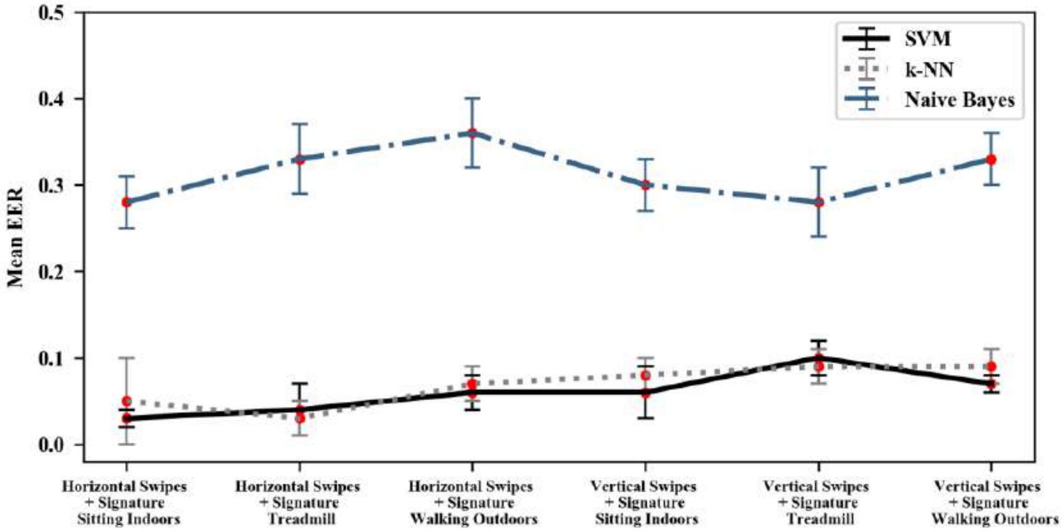

Figure 7. Mean EER attained from intra-session comparison of Session 1 for swipe gestures and finger-based signature fusion

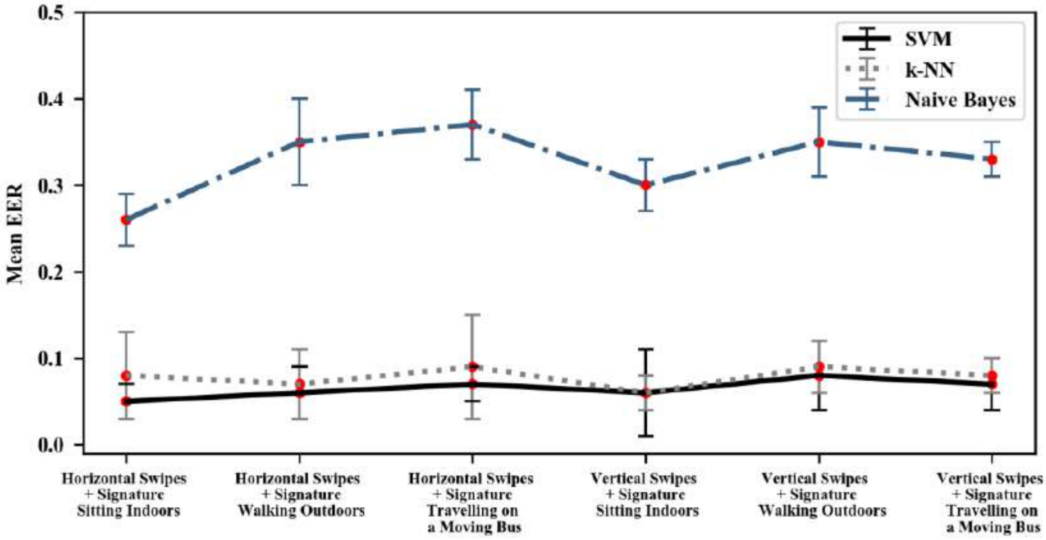

Figure 8. Mean EER attained for intra-session comparison of Session 2 for swipe gestures and finger-based signature fusion

As shown in the above figures, the acquired mean EER % for Sitting Scenario of horizontal swipes with finger-based signature showed that SVM attained best performance with the lowest EER of 3% for Session 1 and 5% for Session 2. It can be observed that the performance using the Naïve Bayes algorithm is the worst compared to SVM and k-NN for all the evaluations in Session 1 and Session 2 as the mean EER% obtained were significantly higher for each evaluation. The SVM algorithm performs the best with lowest mean EER% in case of both horizontal and vertical swipe gesture combinations with signature.

On comparing between horizontal and vertical swipe combinations, the combination of vertical swipes with the signature obtained relatively higher mean EER % using SVM and k-NN for Session 1. However, for Session 2, vertical and horizontal swipe gesture combinations with signature in walking Outdoors and Travelling on Moving Bus attained performances in similar ranges. This may suggest that for both the swipe category (horizontal and vertical swipes), the combination with signature show similar performances. The results also show that the intra-session EERs, even for usage scenarios involving movement, are in acceptable ranges for SVM or k-NN algorithms. A comparison of results acquired using the uni-modal approach of swipe gestures and signatures has been presented in Table 3.

4.1.2. Score Fusion

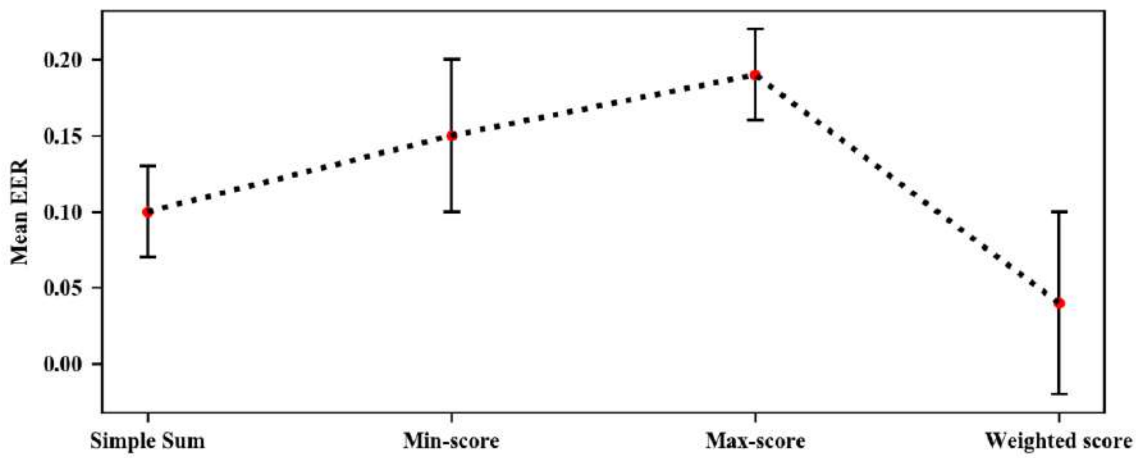

Figure 9. Mean EER from score fusion methods for swipe gestures and signature combination using the commercial signature verification system from scenario 1 of Session 1

Using the score fusion method, the mean EERs obtained from different score fusion techniques are shown in Figure 9. The enrolment and verification samples were horizontal swipes and finger-based signatures belonging to Scenario 1 of Session 1. Based on the obtained results, the weighted score method obtained the lowest mean EER of 5% and the max-score method obtained the highest mean EER of 18%. Therefore, it was concluded that the weighted score method performed the best compared to all the other score fusion methods.

4.2. Keystroke Dynamics and Swipe Gestures

Similar to the fusion method chosen for combining the swipe gestures and signature, the keystroke and swipe combinations were also divided into,

- horizontal swipes and keystroke dynamics features
- the vertical swipes and the keystroke dynamics features.

Keystroke dynamics consisted of both alphabetical and numerical inputs. The results obtained for performing intra-session comparison of samples belonging to Session 1 has been presented in Figure 10.

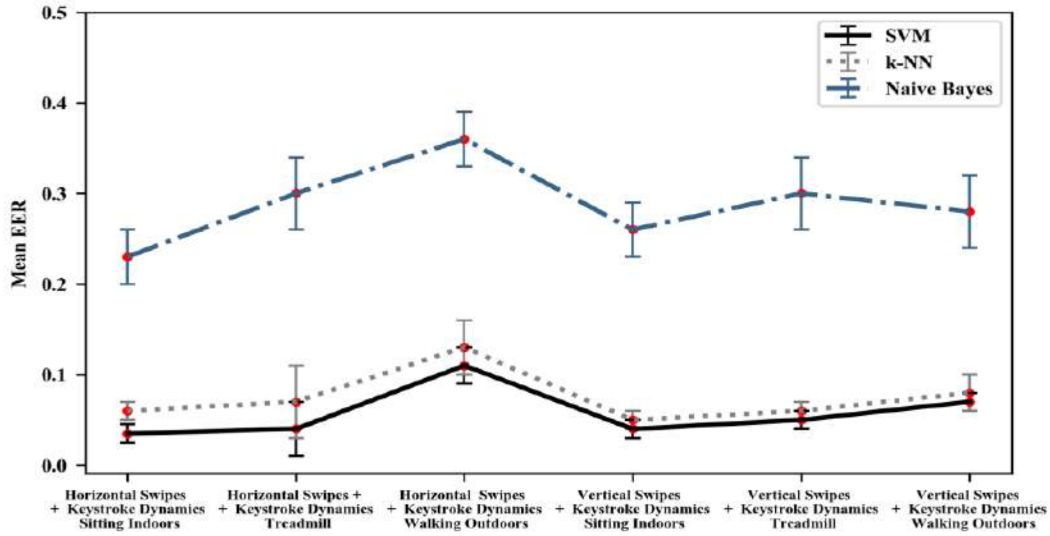

Figure 10. Mean EER attained for combining keystroke dynamics and swipe gestures belonging to Session 1

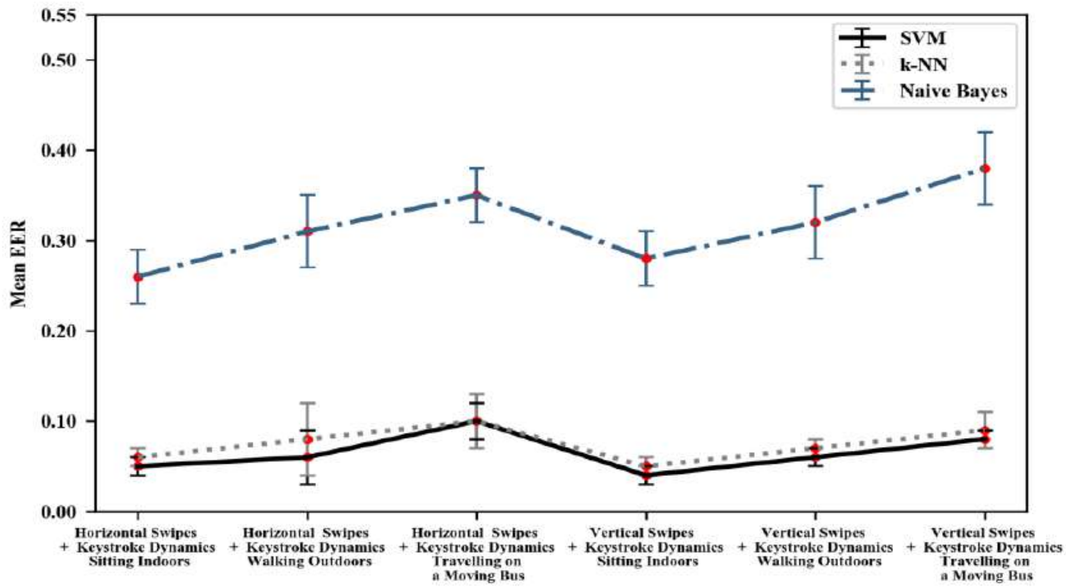

Figure 11. Mean EER obtained for combining keystroke dynamics and swipe gestures belonging to Session 2

It can be observed from Figure 10 and Figure 11 that SVM algorithm performed the best compared to k-NN and Naïve Bayes for both sessions as it obtained the lowest mean EER across all scenarios and combinations. Across all the comparisons, the performance obtained using the Naïve Bayes algorithm is the worst. Even with the best performing algorithm (SVM), the results from Session 1 and Session 2 show that horizontal swipes and keystroke dynamics combination belonging to the Walking Outdoors (Scenario 3) scenario performed the worse with a mean EER of 11% (Session 1) and 10% (Session 2). The combination of horizontal or vertical swipes with keystroke dynamics belonging to Sitting Indoors (Scenario 1) scenario yielded the lowest mean EERs of 5.2% (Session 1) and 5% (Session 2).

4.3. Signature and Keystroke Dynamics

The enrolment samples for this evaluation belonged to the Sitting Indoors (Scenario 1) of Session 1 and Sitting Indoors (Scenario 1) of Session 2 separately. The intra-session results obtained for Session 1 and Session 2 have been presented in Figure 12 and Figure 13. It can be observed that for Sitting Indoors (Scenario 1) from Session 1, the SVM algorithm attained the lowest mean EER of 8%, followed by k-NN with 12%. k-NN performed worse in Walking Outdoors (Scenario 3) of Session 1 and Travelling on a moving bus (Scenario 4) of Session 2. On comparing with the results of other combinations of modalities, the lowest EER attained by signature and keystroke combination was 8%, even on comparing the samples from the same scenario, Sitting Indoors (Scenario 1). Therefore, the performance of signature and keystroke dynamics combination the worse out of all the three combinations.

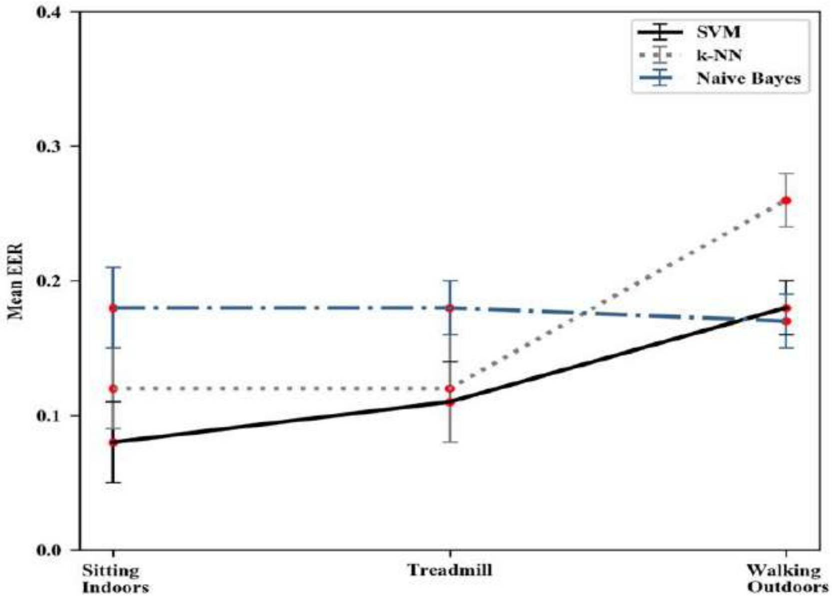

Figure 12. Mean EER attained for combining signature and keystroke dynamics data of Session 1

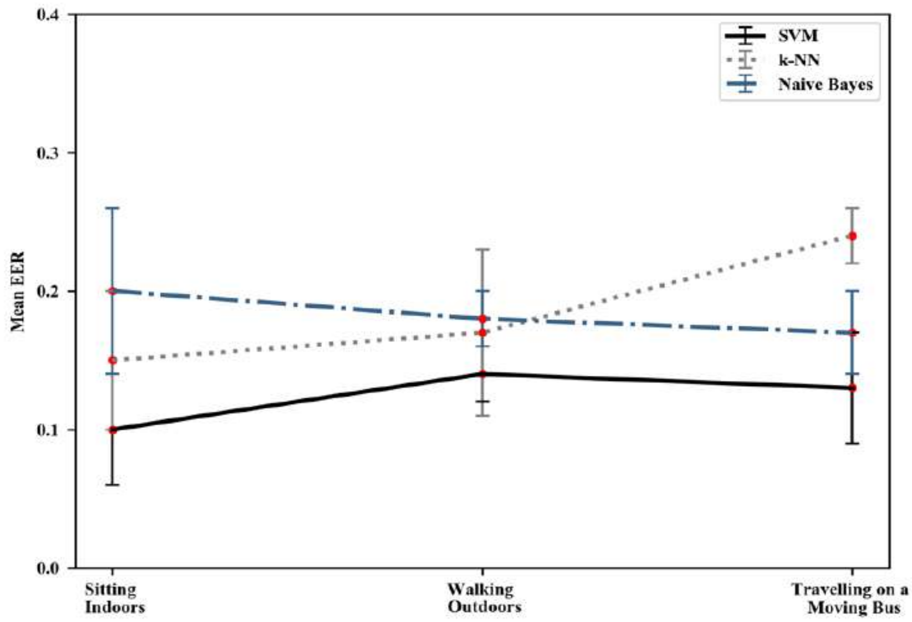

Figure 13. Mean EER attained for combining signature and keystroke dynamics data of Session 2

4.4. Swipe Gesture, Signature and Keystroke Dynamics

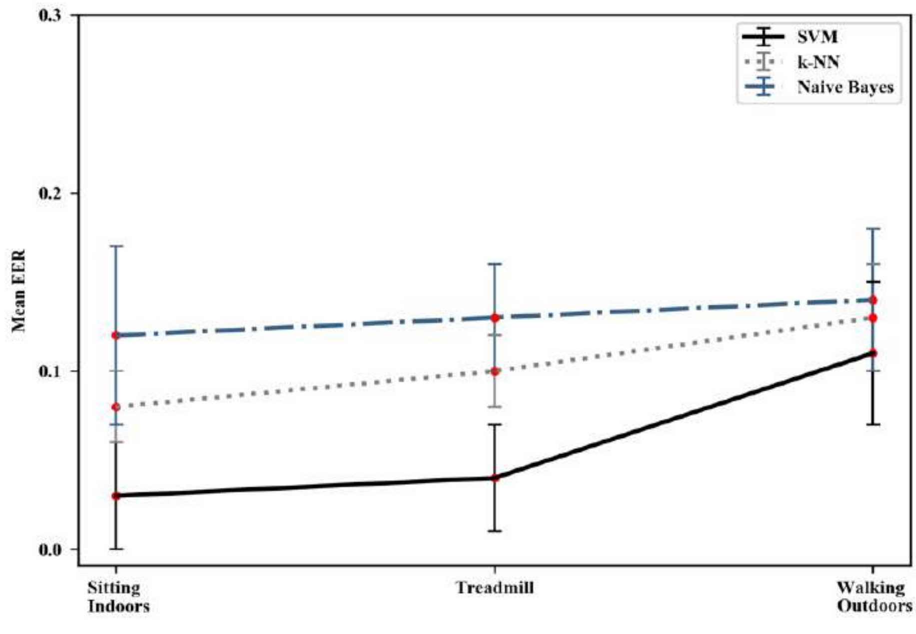

Figure 14. Mean EER attained on combining swipe gestures, signature and keystroke dynamics for Session 1.

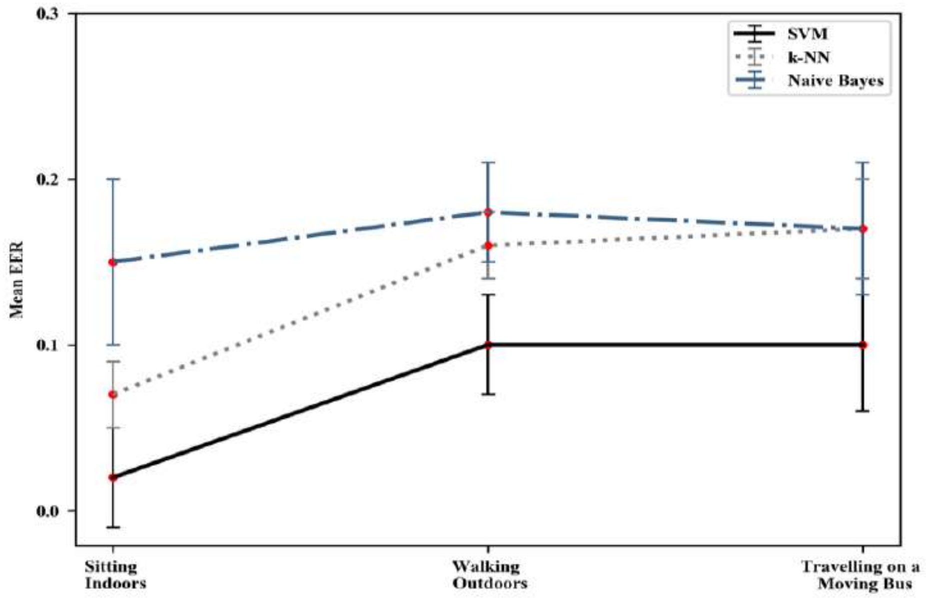

Figure 15. Mean EER attained on combining swipe gestures, signature and keystroke dynamics for Session 2

For this evaluation, the enrolment samples belonged to Sitting Indoors (Scenario 1) of Session 1 for Session 1 evaluations and Sitting Indoors (Scenario 1) of Session 2 for Session 2 evaluations. The verification samples belonged to Sitting Indoors, Treadmill and Walking Outdoors for Session 1 and Sitting Indoors, Walking Outdoors and Travelling on a Moving Bus in Session 2. Figure 14 (Session 1) and Figure 15 (Session 2) show the mean EERs obtained using the SVM, k-NN and Naïve Bayes algorithms when all the three modalities were combined. It can be observed that the SVM algorithm performed the best with the lowest mean EER of 2% (Session 2) and 3% (Session 1) in Sitting Indoors scenario of Session 2. Compared to the results obtained in combination of two modalities, Naïve Bayes algorithm performed better when all three modalities were combined. However, the mean EER % for k-NN are better than Naïve Bayes algorithm in all the scenarios for Session 1 and Session 2. Walking Outdoors (Scenario 3) of Session 1 as well as Session 2 obtained the highest mean EERs, followed by Travelling in a Moving Bus of Session 2.

4.5. Comparison of Uni-modal and Multi-modal Verification

Table 3 shows the best mean EER percentage values attained using different classifiers for both - uni-modal and multi-modal verification systems used for the evaluations in this thesis. The uni-modal based verification attained best mean EERs when the enrolment and the verification samples belonged to the same usage scenario of the data collection. However, there was performance deterioration when the verification samples belonged to a different usage scenario than the enrolled ones. For uni-modal systems, out of all the usage scenarios, the baseline scenario (Sitting Indoors) with no variation in environment or user's body motion acquired best mean EER.

Scenario	Uni-modal (Mean EER (%))			Multi-modal (Mean EER (%))			
	Swipe gestures	**Signature**	**Keystroke Dynamics**	**Swipe + Finger-based signatures (SVM)**	**Swipe + Keystroke Dynamics**	**Keystroke + Finger-based signatures**	**Swipe + Keystroke Dynamics + Signatures**
Scenario 1	Horizontal - 1%	13%	2%	3%	3.5%	8%	3%
	Vertical –2 %	-	-	6%	4%	-	-
Scenario 2	Horizontal - 23	15%	52%	4%	4%	11%	4%
	Vertical –28%	-	-	10%	5%	-	-
Scenario 3	Horizontal –27%	28%	53%	6%	11%	18%	11%
	Vertical -31%	-	-	7%	7%	-	-
Scenario 4	Horizontal –30%	17%	45%	8%	11%	13%	10%
	Vertical -26%	-	-	10%	8%	-	-

Table 3. Mean EER percentages of uni-modal and multi-modal verification

On the contrary, a multi-modal verification showed better performance even when the enrolment and the verification samples do not belong to the same usage scenario. When all the three modalities were combined, the best mean EER attained was for the scenarios that were performed indoors – 3% and 4%. The scenarios having body movement either whilst the user was walking or due to the transport, the performance deteriorated attaining a mean EER of 11% (Scenario 3) and 10% (Scenario 4). However, the error rates are in acceptable range compared to the uni-modal results.

On comparison, the multi-modal solution certainly obtained better performance compared to the uni-modal methods. However, the need for combining such modalities together depends on the requirement of the biometric application.

5. Conclusion

In this chapter, the effectiveness of a multi-modal framework using solely the behavioural biometric modalities has been demonstrated. Multiple evaluations with different combinations of modalities have been carried out and the results have been presented. The multi-modal dataset of 50 participants, acquired using a Samsung Galaxy Note 5 smartphone over two sessions separated by one week with multiple usage scenarios, has been used in the experiment. Feature fusion and score-fusion methods had been applied on combining the swipe gesture, signature and keystroke dynamics data. The classification algorithms such as SVM, k-NN, Naïve Bayes and commercial signature verification system were used for the evaluation.

In terms of compatibility of combination of behavioural biometric modalities, swipe gesture and signature feature fusion method attained an average EER of 3% with SVM and swipe gesture and keystroke dynamics obtained 3.5% and signature and keystroke dynamics feature fusion obtained 8% average EER using SVM classifier. This shows that fusing signature and swipe features gives best results in terms of the average EER. However, fusing signature and keystroke modalities showed comparatively poor performance. The best combination using the feature-fusion method is swipe gestures with signature. The score-fusion method using the commercial signature verification system showed that the weighted score method performed the best for combining swipe gestures and signature modalities.

One of the reasons for fusing different modalities was to enhance the mobile device security. The experimental results obtained reveal a boost in the performance with 3% mean EER using SVM classifier when all three modalities were fused. This is because different modalities possess distinct characteristics corresponding to that data source. When these features are integrated together, a valuable and distinctive feature set is constructed that aids in performance enhancement. The results of the multi-modal verification system yielded better results compared to the uni-modal verification.

One of the challenges faced while conducting this study was the limitation of the number of samples from individual modality. There were unequal number of samples from every modality per scenario. For example, Session 1 contained 175 swipe gesture samples, 18 signature and 28 keystroke dynamics samples. Therefore, 18 unique samples of swipe, signature and keystroke dynamic data could be formed. However, in order to generate more samples to be used for enrolment and verification, samples from signature and keystrokes were repeatedly combined with the remaining swipe samples. Hence, although the swipe gesture samples were new, the signature and keystroke data were repeated to conduct the analysis.

A challenge regarding the practical implementation of such a multi-modal system is that the system would work only when data from all the three modalities are available. Hence, introducing a wait in the authentication score calculation process until data from all three modalities were acquired. The system is designed to combine the data acquired from one sensor – the touchscreen, hence, it is not possible to attain all the three-modality data at the same time.

In conclusion, this chapter details a multi-modal framework that is built on behavioural biometric modalities such as swipe gesture, signature and keystroke dynamics. Using this framework, an evaluation of different combinations of modalities has been explored. Additionally, intra-session evaluation results are also presented that reflect the stability of this framework across different usage scenarios of a mobile device.