IOP Conference Series:
Materials Science and
Engineering

PAPER • **OPEN ACCESS**

Multiple Fusions Approach for Keystroke Dynamics Verification System with Soft Biometrics

To cite this article: Mohd Noorulfakhri Yaacob *et al* 2020 *IOP Conf. Ser.: Mater. Sci. Eng.* **917** 012075

IOP Conf. Series: Materials Science and Engineering **917** (2020) 012075 doi:10.1088/1757-899X/917/1/012075

Multiple Fusions Approach for Keystroke Dynamics Verification System with Soft Biometrics

Mohd Noorulfakhri Yaacob[1], Syed Zulkarnain Syed Idrus[1], Wan Azani Mustafa[2], Mohd Aminudin Jamlos[2], Mohd Helmy Abd Wahab[3]

[1]Faculty of Applied and Human Sciences, Universiti Malaysia Perlis, 01000 Kangar, Perlis, Malaysia.
[2]Faculty of Engineering Technology, Universiti Malaysia Perlis, UniCITI ALAM Campus, Sungai Chuchuh, 02100 Padang Besar, Perlis, Malaysia.
[3]Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia, Johor, Malaysia.

Abstract. Computer security is a process that controls the entire information system, including network, system and hardware. Important information that must be controlled in a system is the data or information contained in a system. Various methods have been used to ensure that only users with legitimate access to data can use a system. Usernames and passwords have been a common practice by many systems as the first requirement to be fulfilled to access the system, but some systems use the secondary verification for additional confirmation. In this article, Keystroke Dynamics has been used as the user's second level authentication for the systems that use the keyboard to login into a system. A common problem of system intrusions is that the system fails to identify the user who signs in using the keyboard when the login is correct. There is a possibility that someone else tries to break into the system. To ensure and improve users' recognition who use the keyboard to enter their logins into the system, Keystroke Dynamics is used as a next-level verification if the login is correct. Soft biometrics is used in the user authentication process using KD method in this study. The soft biometric elements used in this study are culture, gender, educational level (CGPA - Cumulative Grade Point Average) and region of birth (ROB). All of these four soft biometric elements are expected to enhance capabilities in the user authentication process.

1. Introduction

Information security is one of the most important things for every system. It should cover several important aspects such as data privacy [1], data access and data integrity [2]. Various methods of data security such as Smart Card usage, Token, Key and biometric have been implemented to ensure the safety of the data in a system [3, 4]. This article will explain the data security management using biometric methods. Authentication using biometric methods has been popular nowadays for systems that require high levels of security. However, most people's mindset with regard to biometrics is the use of the fingerprint and retinal scan. This is because these two biometric methods are commonly used in a system. This study will use one of the biometric elements that is keystroke dynamic to help

improve the security of a system. Keystroke Dynamics (KD) can be categorized as a biometric behavior because it studies the behavioral differences of a person using a keyboard.

2. Keystroke Dynamic System

The There are a few things that system developers or system researchers should do if they want to use KD in their study or system. The things are data collection, feature extraction, classification, and decision-making process. The first thing to do is the keystroke data collection process. This process involves the hardware such as a keyboard or a touch screen. A variety of keyboards can be used for this purpose such as QWERTY Keyboard [5], AZERTY Keyboard [6], pressure keyboard [7], virtual keyboard [8] and numpad keyboard. Recorded data is the interval in milliseconds for each letter pressed and released.

The second stage of the KD system is the identification of feature extraction. Feature extraction is a process or a method required in order to do classification on gathered data and conduct data normalization [9]. This is a crucial process for researchers to analyze raw data obtained. There are multiple data extraction methods implemented by researchers to analyze obtained KD data. Method normally used for this process is by using time and pressure measurement. Time-based measurements can be divided into two parts, namely dwell time and flight time [10]. Dwell time is calculated based on the time range of a single letter pressed and released [11, 12]. Flight time is a time range between two consecutive letters pressed or released [13-15].

In addition to the current measurement, there are other features used by previous researchers in the study of KD such as pressure, finger movement [16], finger size used to touch screen [17, 18], finger used to press a letter and the position of the finger or hand when using the keyboard [19]. Other measurement methods are useful for devices that do not have a physical keyboard such as devices that use the touch screen. The next stage is the classification phase. At this stage the dynamic keystroke raw data obtained will be categorized according to the similarities that exist for each category [20]. Various classification techniques can be used to perform this classification such as statistical approach, neural network, fusion [21] and hybrid techniques [22].

A summary of these previous studies is listed in table 1. The results of this previous study found that the accuracy in user classification or user validation using keystroke dynamics is still low when the number of volunteers involved is high. Therefore, this study combines the keystroke dynamic with four additional soft biometric features with the expectation of enhanced recognition and user authentication in KD.

2.1. Performance Evaluation

Biometric performance measurement is essential to assess the accuracy of a biometric system [23]. In order for this performance measurement to be completed, more authentic user data needs to be recorded and more impostor data needs to be generated [24]. The similarities and differences in scores between the genuine user and the imposter were recorded [25]. The result obtained will be used for False Acceptance Rate (FAR) and False Rejection Rate (FRR) calculations. Generally, biometric systems with the lowest FAR and FRR readings will make them better[26]. This study involved user behavior by studying how an individual uses a keyboard. Similar behaviors may exist from one individual to another [27]. For example, a man who resembles women's tendencies will have some of the characteristics possessed by women.

2.2. Score Level Fusion

Most biometric systems perform data integration at the score level because, at this stage, the results obtained are easy to be analyzed and have complete information [28]. Additionally, merging at this score level is easy as it involves only two different scores merging from two different entities. At this stage, the scores obtained from each entity will be combined in various methods and calculations to produce a new score that will be used at the decision stage. There are two approaches to perform a score fusion: classification and combination [29]. In a classification approach, a vector score

calculation will be created to allow data to be classified as accepted or not whereas in a combination approach, every individual score obtained will be combined and a new score is used at the decision level. The two steps needed to perform merging at this level are normalization and merger. For example, the score for entity A may be between 1 to 200, while the score for entity B may be between 1 to 5000. Both of these scores are difficult to combine because the range is different. Therefore, normalization is necessary.

Table 1: Summary of previous research

Researcher	**Year**	**Research Focus**	**Method**	**Results**
Rumelhart and Norman [16]	1982	Research on typing speed and errors	Activation Triggered Schema system	Found inter-keystroke, interval times
Yu and Cho [12]	2003	Password verification using Keystroke Volunteer: 21	Auto associative, multilayer perceptron, SVM	FAR: 0 % FRR: 0.814%
Nonaka and Kurihara [7]	2004	Estimating the real moment of keystroke	Pressure Sensor	Different Latency for Operating System compared to Sensor
Ng, Oh [5]	2007	Studies on the use of QWERTY keyboard	Design suitability	
Saevanee and Bhattarakosol [17]	2009	Keystroke Dynamic on Mobile. Volunteer: 10	Probabilistic Neural Network	Accuracy 99%
Trojahn, Ortmeier [8]	2012	Authentication on Mobile Device		12-Key : FAR 8.31 % ; FRR 5.26 % QWERTZ : FAR 9.53% ;FRR: 5.88%
Idrus, *et. al.* [6,15,20, 21,37] Idrus [36]	2013-2016	Keystroke Dynamics and Soft Biometric (Gender and Use of Hand) Volunteer : 110	Support Vector Machine (SVM)	Recognition rate: 89% -96%
Monaco [14]	2016	Detect Anomaly in Public Keystroke. Volunteer: 300	Neural Networks	EER: 5.32%
Zhao, Yang [18]	2019	Volunteer : 104	deep belief network.	Accuracy 97.58%
Lu, Yu [19]	2019	Keystroke attack on Touch Screen Volunteer: 24	binary tree-based	Accuracy : 96.2%

2.2.1. Score Level Fusion

The process of score normalization has been widely discussed in previous studies. The following are the most commonly used normalization methods, namely Min-Max, Z-Score and Hyperbolic Tangent (TanH). Table 2 shows a brief description of the normalization methods. The normalization technique used in this study for the user authentication process is TanH as it performs well and manages data outliers. Besides, TanH can distinguish small differences in the data analyzed.

2.2.2. Score Combination

Score combinations involve combining multiple scores into a single score [28]. Before a combination can be performed, the obtained score needs to be normalized as described in section 2.2.1. There are several methods for combining scores such as Maximum Score, Minimum Score, Simple Sum [29] and Weighted Sum [30]. This study will use a method similar to Simple Sum which is by way of

summation and multiplication. The combination formulas used in this study will be described in detailed in the methodology section below.

Table 2: Score normalization method from previous study

No	Normalization Technique	Description
1	Min-Max [30]	This normalization strategy converts the value of X to Y. The score will be matched within the range [0,1]. The Equation for this normalization is ; y = (x – min(x))/(max(x) – min(x))
2	Z-Score[31]	The technique will use standard and standard deviation. Therefore, the average value of the score and the variance of the score should be first calculated. This technique will be less accurate when there are outliers in the data collected. The Z-Score Equation is: y = (x – mean(X))/std(X)
3	Hyperbolic Tangent (TanH) [32].	TanH is a long-standing mathematical formula. It was first used by Sauri [33]. This technique is an effective technique, performs well in normalization and intelligently handles data outliers.
4	Decimal scaling [34]	Used when the score of two entities to be merged has a logarithmic factor difference; for example, entity A has a score of [0,1], whereas entity B has a range of [0,100000]

3. Data Collection

The data used in this study were obtained from data collection among university student volunteers in Malaysia. A total of 250 students volunteered in this study. Each volunteer is required to type 5 sentences correctly 10 times using the computer provided. Software GREYC is used for the data recording. For this research, soft biometric information consisting of CGPA, gender, region of birth of 250 students was collected and obtained during the enrollment process for the purpose of this study. All this information is recorded into the computer to perform analysis for this study. Each user typing pattern involved in this research will be matched with the personal information provided during the enrollment process. In real situations, this biometric soft information can be obtained from the system registration form for example student information and account opening forms.

3.1. Data Analysis

The data obtained were separated into four categories of soft biometric, namely culture, region of birth (ROB), gender and education level. The culture category was divided into four sections, of which three were made up of the three main cultures in Malaysia namely Malay, Chinese and Indian, while the other section consisted of minority cultures in Malaysia namely Bajau, Murut, Siam, Suluk, Iban, Kadazan, Bisaya, Kedayan, Iranum, Tidong and others. Minor culture is consolidated and labelled as 'Óthers'. The ROB category is also divided into four parts, north of the peninsular, east of the peninsular, center of the peninsular and south of peninsular Malaysia. In addition, data is also segregated into education level categories. CGPA is used as a separation between these categories. Students with CGPA 3.0 and above are categorized as one group, while students with CGPA 3.0 and below belong to another group. The classification of gender categories is divided into two, male and female. The statistical breakdown of the data is described in Figure 1 to Figure 4.

Several approaches were used to integrate soft biometric elements with the user authentication process in KD. The approach is divided into two categories. The first approach is to create a benchmark score for user authentication for each sentence used. A detailed description of the steps taken to get this benchmark score is described in section 3.1 below.

The second method is to incorporate soft biometric scores obtained from the classification process. To analyze and classify the raw data collected, Support Vector Machine (SVM) was the first choice technique opted. Math Lab is the software used to perform SVM. SVM is used for soft biometric classification purposes only. Please refer to figure 5 to see where this SVM is used.

There are three ways of combining soft biometric scores with benchmark scores performed using this second method:

i. Benchmark scores are combined with soft biometric scores one by one. Table 3 shows the merger elements.

ii. The overall score for the combination of culture element and region of birth by category. This merger is called second fusion and produces a new score that will be used in the next step. The new score is called the culture score and the Region of Birth Score. Table 4 shows a clear picture of how the merger was conducted.

iii. All elements of soft biometric score elements are incorporated with benchmark scores. Table 5 shows a clear scenario of how this merger is completed.

iv. All elements of soft biometric score elements are incorporated with benchmark scores. Table 5 shows a clear scenario of how this merger is completed.

The process of the approach used for this second method is expected to perform better than the first approach in terms of recognition. The measurement used during this user authentication process is to obtain Equal Error Rate (EER) readings. EER is a reading derived from the FAR and FRR readings, which is to measure the number of incorrect confirmations and the number of validations. Figure 5 shows an overview of the methodology to be used in this study.

v.

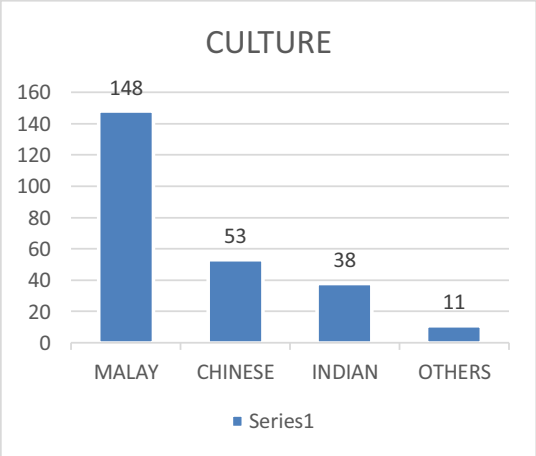

Figure 1: Shows the breakdown number of volunteers participated according to culture fragments

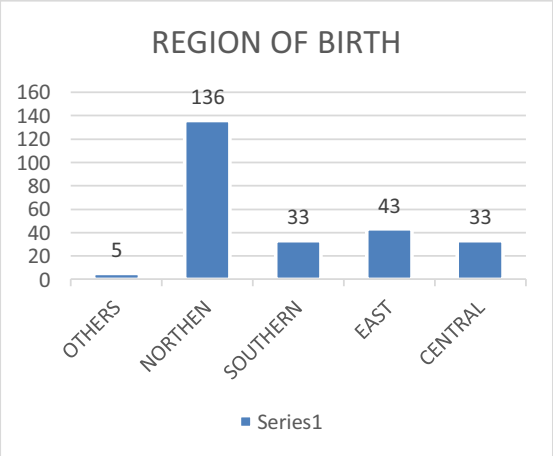

Figure 2: Shows the breakdown number of volunteers participated by region of birth fragments

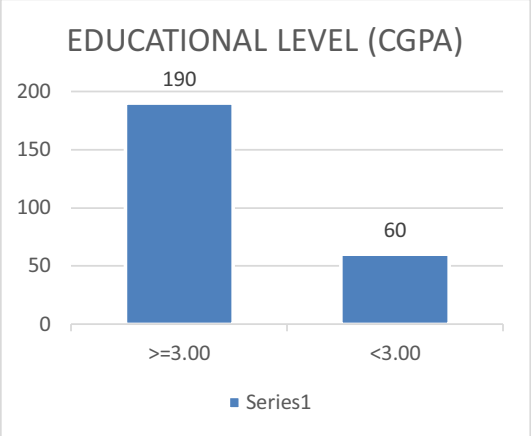

Figure 3: Shows the breakdown number of volunteers participated according to the CGPA fragments

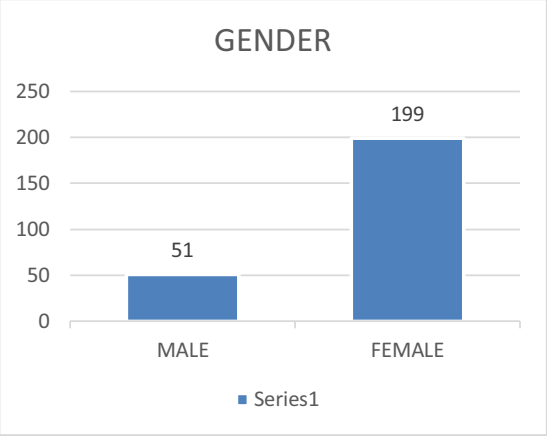

Figure 4: Shows the breakdown number of volunteers participated according to the Gender fragments

Table 3: Combining elements of benchmark scores and soft biometric scores respectively

No	First Fusion Level	Soft Biometric Category
1	Benchmark + (Chinese VS Indian)	Culture
2	Benchmark + (Others VS Chinese)	
3	Benchmark + (Others VS Indian)	
4	Benchmark + (Malay VS Chinese)	
5	Benchmark + (Malay VS Indian)	
6	Benchmark + (Malay VS Others)	
7	Benchmark + (Central VS South)	Region of Birth
8	Benchmark + (Central VS East)	
9	Benchmark + (East VS South)	
10	Benchmark + (North VS South)	
11	Benchmark + (North VS Central)	
12	Benchmark + (North VS East)	
13	Benchmark + (Female VS Male)	Gender
14	Benchmark + (3.0 Above VS 3.0 Below)	Educational Level

Table 4: Combined soft biometric score for culture and region of birth.

No	Second Fusion Level	Result name for Soft Biometric
1	(Chinese VS Indian) + (Others VS Chinese) + (Others VS Indian) + (Malay VS Chinese) + (Malay VS Indian) + (Malay VS Others)	Culture Score
2	(Central VS South) + (Central VS East) + (East VS South) + (North VS South) + (North VS Central) + (North VS East)	Region of Birth Score

Table 5: The overall combination of soft biometric scores with benchmark scores

No	Combined Fusion Score
1	Benchmark + (Culture Score)
2	Benchmark + (Region of Birth Score)
3	Benchmark + (Region of Birth Score) + (Culture Score)
4	Benchmark + (Region of Birth Score) + (Culture Score) + (Female VS Male)
5	Benchmark + (Region of Birth Score) + (Culture Score) + (Female VS Male) + (CGPA 3.0 Above VS CGPA 3.0 Below)

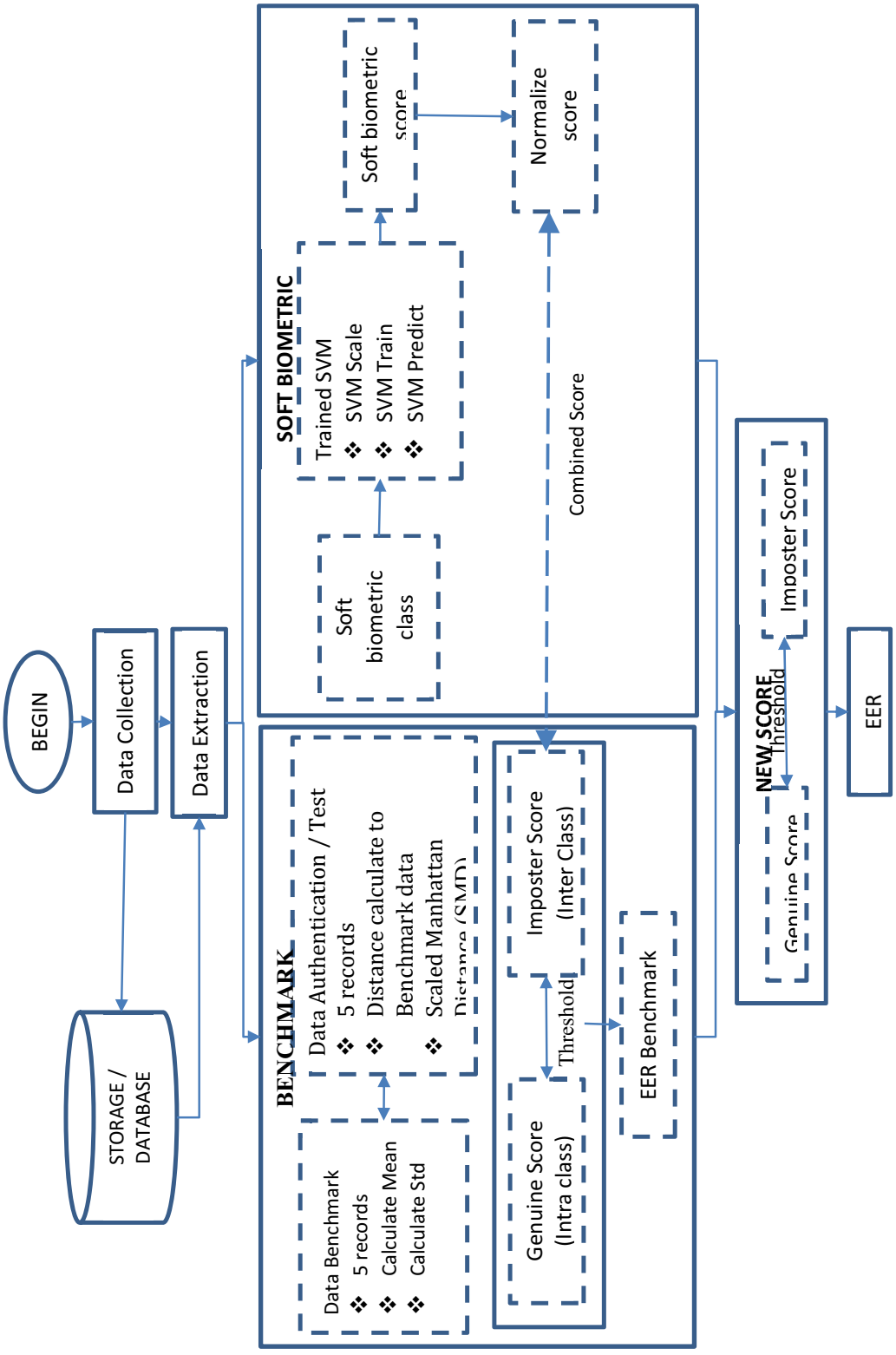

Figure 5: Overview of the methodology to be used in this study.

3.2. Benchmarking Score Calculation

Firstly, is to identify the steps and factors that need to be calculated. All of the keystroke data that had been compiled were the time recorded during Press Release (PR), Release Release (RR), Release Press (RP) and Press Press (PP) for each sentence to be the benchmark for each individual. PR is the time taken for an individual to press and release the respective letter. RR is the amount of time each individual needs to release two consecutive letter sequences. The time taken between two consecutive letter sequences, the first letter release and the second letter press are known as RP. Finally, PP is the recorded time for individuals to press two consecutive letters. The entire time is recorded in milliseconds. Figure 6 below shows a clear picture of the time range recorded

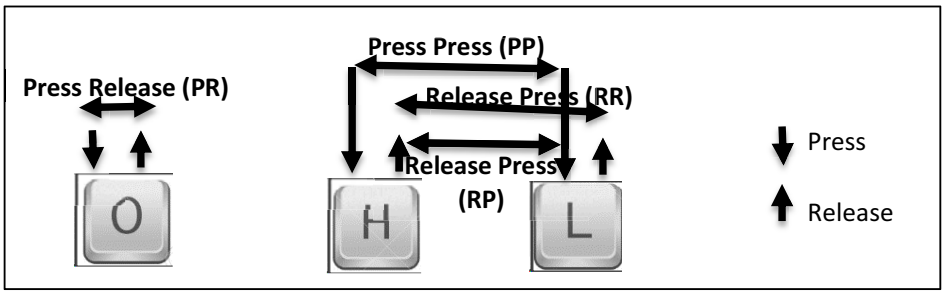

Figure 6: A detailed diagram of the time recording method for KD

Performance measurement for these benchmarks were calculated using mathematical formulas including mean, standard deviation, distance and distance comparison. The results of the performance calculations for these benchmarks were translated using the ROC (Receiver Operating Characteristic) graph and EER value. The ROC obtained is the result of the Imposter Score and Genuine Score calculations using mathematical formulas to obtain FAR and FRR from the 5 words tested in this study. The ROC curve is obtained by assigning the probability of correct acceptance (FRR) and the probability of incorrect acceptance FAR to the vertical and horizontal axes, respectively, and varying the decision threshold, as shown in Figure 7. A detailed calculation for obtaining FAR and FRR in this study is described in this section.

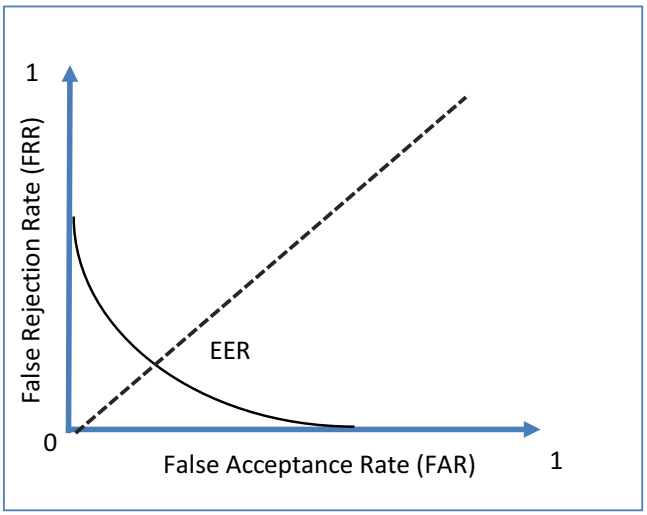

Figure 7: Diagram showing ROC graph pattern and EER value position

These are the steps taken to measure benchmarks. The value in this first approach measurement was obtained from the analysis of user's five sentences. The selected raw data is a vector data as the result of the fusion of PR, RR, RP and PP times for each sentence.

The data collection process requires, each user to type 10 times correctly for each given sentence. The raw data for each sentence from the user was divided into two sections, five records are used as test data and five more as template data. The sampling of 5 data for each section (test and template) was made randomly using the bootstrapping method.

A template data was obtained by computing the average time and standard deviation of the five data. This record will be labeled as the reference data. The calculation was done by comparing each score for reference data with the test data for each user. Distance score was calculated based on the differences between reference data and test data. In general, the distance score obtained between reference data and test data from the same individual is low compared to the distance score with two different individuals. Figure 8 below illustrates an illustration of how measurements will be made

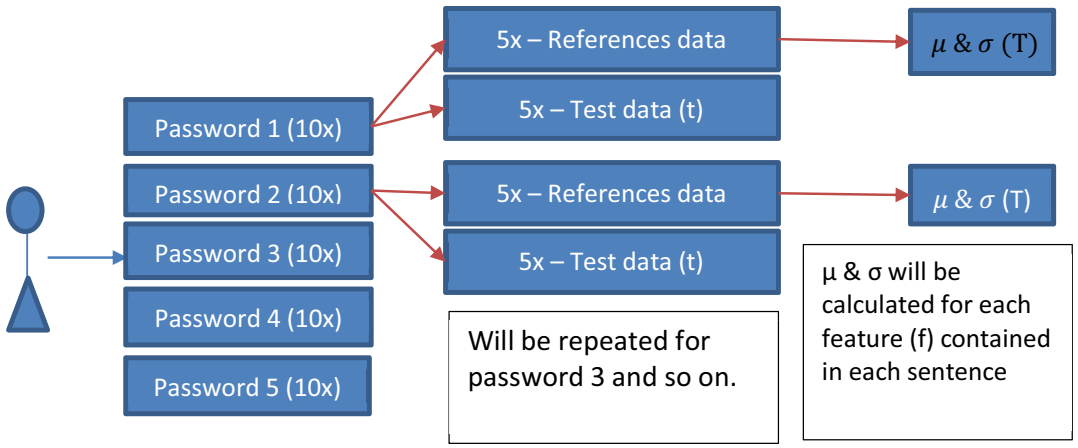

Figure 8: Shows how data fragments are created to enable the user authentication process

The distance score calculation used in this study is called Scaled Manhattan Distance (SMD)[35]. This distance calculation method was chosen because it has the potential to be used for high-dimensional data. Features derived from this dynamic keystroke consist of various time differences from Press Press (PP) time, Press Release (PR) time, Release Press (RP) and Release Release (RR). The basic formula for calculating distances using SMD method is:

$$| \mathbf{x1 - x2} | + | \mathbf{y1 - y2} |.$$

In order to calculate the distance score in this study, the sum of the differences for each of the existing features was added. Suppose f represents the number of features contained in each sentence used, while T represents the value of each feature estimated based on the mean and standard deviation of the five sentences selected as reference data. T= (μ_1, σ), (μ_2, σ_2), …., (μ_f, σ_f).

Subsequently, t which represents test data is the input data. The distance score can then be calculated based on the sum of the different features of each test data compared to the reference data. A summary of the Equation (3.0) to use for this distance calculation is

$$Distance(T,t) = \sum_{i=1}^{f} \frac{|ti - \mu i|}{\sigma i} \qquad (3.0)$$

There are two types of distance scores that were calculated: distance score for legitimate users (FAR) and distance score for impostors (FRR). As mentioned previously, the entire keystroke data for the 250 users was used. Each user needed to record 10 typing patterns for the same sentence. Therefore, the total keystroke data obtained was (250 users X 10 = 2500 keystroke data) for each

sentence. Five out of ten typing patterns of a user were used as reference data or template data. Then the remaining five typing patterns for the same individual were counted as five genuine scores or as test data. The rest of the 249 user keystroke data was treated as data impostors (249 users X 10 = 2490 impostors). This process was repeated for each user, and the final result obtained was 1250 genuine user records and 311250 records as an imposter. A summary of the formulas used is as follows:

i. Genuine record number = 5 (reference data) x 250 (user) = 1250
ii. Total Impostor record = 5 (test data) x 250 (user) x 249 (user) = 311250

The distance score difference for the respective user against the others was calculated. The range of allowable score differences was calculated based on the threshold value. The threshold is a point or distance that can be used to identify an individual based on the referenced data of that user. Practically, the probability of overlap or data or record similarity is high as the number of users recorded increases. For example, the score values of references or templates obtained for individual A are 200 points. This score is stored in the system as a benchmark for the user. The threshold value allowed by user A is 30, which means that if user A types the same sentence then the allowed difference is between 170 and 230. If the distance score obtained falls within this allowable difference, then the record is considered valid for the respective user. If the value of the score obtained is outside this allowed range, it is considered an impostor. The combination of five passwords performed in this study.

3.3. Score Calculation for Fusion Technique with Soft Biometric Elements

The second approach is the combination of soft biometric culture (Malay, Chinese, Indian and Others), gender, educational level (CGPA - Cumulative Grade Point Average) and region of birth for improving authentication scores obtained. Distance scores from benchmark and soft biometric scores were combined into one value and used as analysis process later. There are three methods of incorporation that were used in this study. The data used in this second approach is also divided into two as in step 3.2 above, only reference data is used.

The first method, the distance score value from benchmark was combined with the soft biometric score value obtained. The method of merging is as follows. Assuming that Ds represents the distance score obtained, SBs represents the soft biometric score obtained.

Soft biometric scores have two important reading values; the first is the value of class classification, whether they are categorized as male or female for gender categories. This value is called the soft biometric class label or known as SBc. The second value read in soft biometric is the value that determines which record represents the category of male or female. This value is between 0 and 1. For example, if the reading value is below 0.5, it will fall into the female category whereas if the value is 0.5 and above, it will fall into the male category. This second soft biometric value is labeled as SBv.

Finally, the sum of distance score and the absolute value obtained from the difference SBc with SBv were tagged as a new distance known as the verification combine score (VCS) to avoid confusion with the existing distance score terms. The mathematical Equation (3.1) are as below:

$$VCS1 = Ds + |SBs(SBc) - SBs(SBv)| \tag{3.1}$$

The second method is to combine the scores obtained for each culture category and region of birth. Assume that e represents the elements of each soft biometric category, and SBz is the sum of the total soft biometric scores within the same group, culture and region of birth without involving the data to be tested. This method is similar to the method of obtaining biometric scores in Equation (3.1). The formula for combining scores in these two categories is as in Equation (3.2)

$$SBz = \sum_{i=1}^{e} |SBs(SBc) - SBs(SBv)| \tag{3.2}$$

The third method of combining is to combine the absolute masses (SBz) obtained from each soft biometric category individually. The absolute value (SBz) of each of these categories was added to Ds. The Equation (3.3) obtained for the merger with the third second are as follows:

$$VCS2 = Ds + |SBz1| + |SBz2| + |SBz3| + |SBz4| \quad (3.3)$$

Table 6 below shows the fusion and Equation methods used for each combination of benchmark scores with soft biometric scores. The results of the merger were recorded and described in this article.

Table 6: Fusion level and Equation methods to be used for each combination

No	Fusion	Fusion level	Equation
1	Benchmark + Culture (Chinese VS Indian)	First Fusion	3.1
2	Benchmark + Culture (Others VS Chinese)		
3	Benchmark + Culture (Others VS Indian)		
4	Benchmark + Culture (Malay VS Chinese)		
5	Benchmark + Culture (Malay VS Indian)		
6	Benchmark + Culture (Malay VS Others)		
7	Benchmark + ROB (Central VS South)		
8	Benchmark + ROB (Central VS East)		
9	Benchmark + ROB (East VS South)		
10	Benchmark + ROB (North VS South)		
11	Benchmark + ROB (North VS Central)		
12	Benchmark + ROB (North VS East)		
13	Benchmark + Gender (Female VS Male)		
14	Benchmark + Educational Level (3.0 Above VS 3.0 Below)		
15	(Chinese VS Indian) + (Others VS Chinese) + (Others VS Indian) + (Malay VS Chinese) + (Malay VS Indian) + (Malay VS Others)	Second Fusion	3.2
16	(Central VS South) + (Central VS East) + (East VS South) + (North VS South) + (North VS Central) + (North VS East)		
17	Benchmark + (Culture Score)	Combine Fusion Score	3.3
18	Benchmark + (Region of Birth Score)		
19	Benchmark + (ROB Score) + (Culture Score)		
20	Benchmark + (ROB Score) + (Culture Score) + (Female VS Male)		
21	Benchmark + (ROB Score) + (Culture Score) + (Gender Score) + Education Level (CGPA 3.0 Above VS CGPA 3.0 Below)		

4. Results

The results obtained for the user's authentication process using two different approaches are described in this section. As a reminder, these two methods are user authentication without combining with soft biometric and authentication user by combining soft biometric elements. This section describes the five findings from this study on user authentication. Determining whether new soft biometric elements can help reduce EER rates is mentioned in this chapter.

4.1. Benchmarking for User Authentication

Figure 9 shows the ROC graph for each password or word analyzed using the first method. The graph values shown for each paragraph are benchmarks for this study that were compared later with the combination of soft biometric methods. The curved graphs were calculated based on the intra and inter class values obtained using the calculations in section 4.5. From Figure 9 also, a straight line has been drawn to indicate the position of the EER. In other words, ERR is a value where FAR is equal to FRR. This ROC graph is plotted using MATLAB software and ROC functions. Intra and inter class values were entered into this function for analysis to obtain EER.

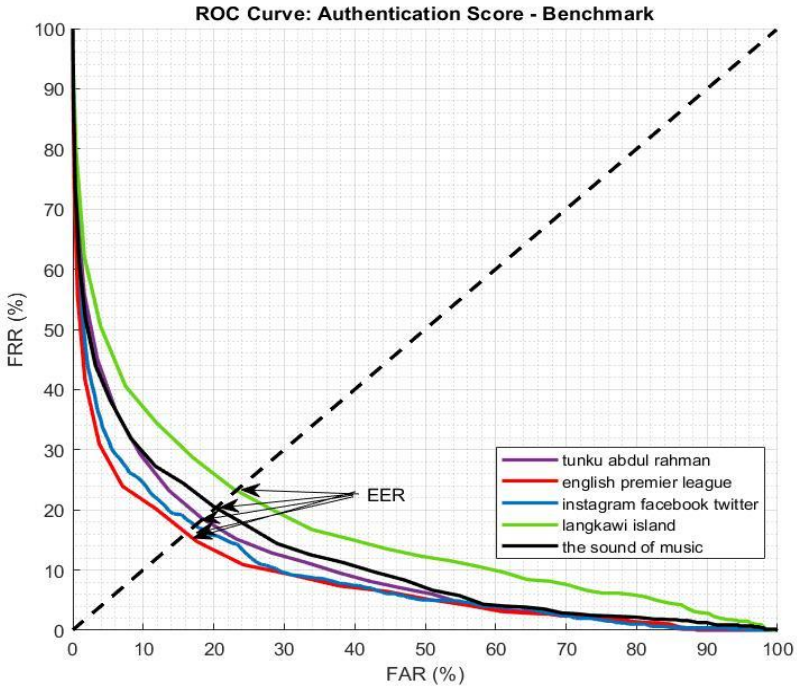

Figure 9: ROC Curve for benchmark authentication

The obtained EER is showed in Table 7. Based on the table, the EER values obtained range from 16.17% to 23.13%. The lowest EER value for password 1 and the highest EER is for Password 3.

Table 7: EER obtained for benchmark computation for each sentence

	Sentences	**EER Value**
P1	english premier league	16.17%
P2	instagram facebook twitter	17.40%
P3	langkawi island	23.13%
P4	the sound of music	20.27%
P5	tunku abdul rahman	18.45%

4.2. Fusion with Single Soft Biometric Score

Table 8 shows the results for user authentication obtained by combining the keystroke dynamic with soft biometric elements using Equation (3.1). The table shows that the EER results were obtained by combining each of the soft biometric elements separately. There are four categories of soft biometric tests to be combined with user authentication for KD. The categories are gender, culture, educational level and region of birth. The categories of culture and state of birth have six minor divisions according to culture and region of birth. For the first sentence, P1, the baseline obtained was 16.17%, but after integration of the soft biometric elements one by one, the EER readings ranged from 13.2% to 15.92%. For the second sentence, P2, the baseline EER reading recorded before the fusion technique was 17.40%. However, it showed a reduction when combined with the soft biometric element with a reading between 12.97% and 16.67%. In general, it was found that the baseline EER reading for the third sentence, P3, had the highest EER of 23.13%, but after the fusion technique was used, the reading decreased from 17.67% to 21.71%. The fourth and fifth sentences, P4 and P5, showed a decrease in EER reading compared with the basic reading obtained. Overall, the EER reading decreased for all of the combined soft biometric elements compared to previously calculated benchmarks

The observation on the approach of combining soft biometric element into the user authentication process using KD is able to reduce EER reading lower than EER benchmark. This proved that fusion method can provide better accuracy readings.

4.3. Fusion with Multiple Soft Biometric Scores

Table 9 shows the results for user authentication obtained by combining the keystroke dynamic with soft biometric elements using Equation (3.3). The table shows that the EER results were obtained by combining each of the soft biometric elements. Generally, the method of combining the Equation (3.1) and the Equation (3.3) is the same like using the summation. The difference between the Equation (3.1) and the Equation (3.3) is the technique of combining soft biometric elements. Each soft biometric element is incorporated as shown in Table 9 and the EER readings are recorded. prior to equation (3.3) carried out. Fusion on soft biometric score level for culture and region of birth category was performed using equation (3.2). From Table 9 also, Section A shows the result of EER obtained by combining two biometric soft elements namely gender and CGPA. The biometric scores obtained from each of these elements will go through the mathematical process through equations 3.2 and 3.3. The same process will be repeated for each section until section G. The results in Section D indicate the overall combination of the elements studied in this study (CGPA), GENDER, ROB and MC.

The EER results obtained are good compared to the EER benchmark obtained. The results obtained from the combination of two soft biometric elements, Educational Level and Gender, were able to reduce the EER reading of the whole sentence except the first sentence (P1). Next, the integration of soft biometric elements is done by adding culture elements (main residents of Malaysia, Malay, Chinese, Indian and Others). The results obtained from the combination of these three elements have a better reading than the combination of the two elements. The merger is then performed with all possible combinations of soft biometric elements shown in Table 9. With regard to Table 9, the EER readings improved for each of the combinations performed except for only P1 sentence that had readings beyond the basic readings for the combination of the two elements, Educational Level and gender.

Table 8: EER results for user authentication obtained by combining the keystroke dynamic with soft biometric elements using Equation (3.1).

		MALAYSIA CULTURE (MC)						ROB						GENDER	EDUCATIONAL LEVEL (CGPA)
	BENCHMARK	Chinese VS Indian	Others VS Chinese	Others VS Indian	Malay VS Chinese	Malay VS Indian	Malay VS Others	Central VS South	Central VS East	East VS South	North VS South	North VS Central	North VS East	Male VS Female	3.0 ABOVE VS 3.0 BELOW
P1	16.17	15.92	15.33	14.67	14.47	14.66	13.20	14.85	14.99	14.29	14.71	14.99	14.57	14.29	14.42
P2	17.40	15.79	16.00	16.67	14.85	12.97	14.72	15.45	14.29	14.85	14.29	14.85	13.87	14.85	16.83
P3	23.13	18.95	19.33	20.00	20.49	17.67	19.48	20.76	20.87	19.47	19.75	20.87	21.71	18.63	21.67
P4	20.27	15.66	15.67	16.33	14.85	15.60	14.94	15.41	15.41	15.41	15.41	15.41	15.41	15.41	15.41
P5	18.45	16.97	17.00	18.00	15.41	16.73	14.50	16.36	15.97	15.97	15.97	15.27	15.13	15.83	15.92

Table 9: EER results for user authentication obtained by combining the keystroke dynamic with soft biometric elements using Equation (3.3).

		MULTIPLE FUSION RESULTS						
		SECTION A	**SECTION B**	**SECTION C**	**SECTION D**	**SECTION E**	**SECTION F**	**SECTION G**
	BENCHMARK	EDUCATIONAL LEVEL (CGPA) + GENDER	EDUCATIONAL LEVEL (CGPA) + GENDER + MC	EDUCATIONAL LEVEL (CGPA) + GENDER + ROB	**EDUCATIONAL LEVEL (CGPA) + GENDER + ROB + MC**	EDUCATIONAL LEVEL (CGPA) + MC	EDUCATIONAL LEVEL (CGPA) + ROB	EDUCATIONAL LEVEL (CGPA) + ROB + MC
P1	16.17	20.87	16.00	16.67	**15.00**	17.00	19.55	16.67
P2	17.40	17.02	18.00	16.67	**11.33**	16.67	15.76	15.67
P3	23.13	17.23	17.00	16.52	**14.67**	17.33	19.85	16.00
P4	20.27	16.25	14.00	16.67	**13.00**	15.00	16.36	14.67
P5	18.45	17.37	13.33	16.52	**14.67**	14.00	15.30	14.33

5. Conclusion

Table 10 shows the comparisons on EER benchmarks, single fusion and multiple fusion. In conclusion, the results presented in this research can be used to improve user verification based on keystroke dynamics by combining soft biometric information using multiple fusion technique. Multiple fusion is the combination of scores derived from the soft biometric classification from the gender, culture, region of birth and educational level category. In addition, there are three integration techniques introduced in this study that help improve the accuracy of KD validation. The results obtained in the previous sections found that the EER reading of all five words was reduced when this fusion technique was performed. Nonetheless, there are also some results in combination acquired higher EER readings than benchmark scores, but the lowest EER readings obtained after the overall combination were 11.33%.

Table 10: Comparison of EER readings based on benchmark, single fusion and multiple fusion.

	Benchmark	Single Fusion	Multiple Fusion - 4 Elements soft biometrics **(CGPA + GENDER + ROB + MC)**
P1	16.17	13.2	15.00
P2	17.40	12.97	11.33
P3	23.13	17.67	14.67
P4	20.27	14.94	13.00
P5	18.45	15.13	14.67

The finding of this writing can be applied in a daily environment where this method can be used as a second layer of security filtering after username and password. For example, when user A uses a university system, personal data relating to the individual is recorded during the registration process. Confirmation of individual soft biometric details is made by the university by instructing the user to include a copy of their identity card. For the process of enrolling usernames and passwords for each user, users must enter a username and password at least 10 times to allow the system to identify individual typing patterns. Typing intervals between these letters are recorded and stored in the database.

To implement the authentication method using this KD, when a user uses the system, there are two check filters that the system will perform. The first check will compare the username and password the respective user typed and logged. After passing the first filter, the system will use this KD authentication method to determine if the correct username and password are keyed-in by the exact individual.

Undeniably, there will be some differences when users use different keyboards. To address this, the user's enrollment process needs to use the same type of keyboard that the user is using. For example, if a user uses two different types of keyboards, the enrollment process should record both typing patterns for these users on both keyboards.

References

1. Tsiakis, T. and G. Sthephanides, *The concept of security and trust in electronic payments.* Elsevier International Journal of Computers & Security, 2005. **24**(1): p. 10-15.
2. Flowerday, S. and R. von Solms, *Real-time information integrity=system integrity+data integrity+continuous assurances.* Elsevier International Journal of Computers & Security, 2005. **24**(8): p. 604-613.
3. Roberts, C., *Biometric attack vectors and defences.* Elsevier International Journal of Computers & Security, 2007. **26**(1): p. 14-25.
4. Weir, C.S., et al., *User perceptions of security, convenience and usability for ebanking authentication tokens.* Computers & Security, 2009. **28**(1): p. 47-62.
5. Ng, E., J.J.S. Oh, and K.C.F. Tan, *Reduced keyboard system that emulates QWERTY-type mapping and typing*. 2007, Google Patents.
6. Idrus, S.Z.S., et al. *Soft biometrics database: A benchmark for keystroke dynamics biometric systems*. in *2013 International Conference of the BIOSIG Special Interest Group (BIOSIG)*. 2013. IEEE.
7. Nonaka, H. and M.J.I.J.o.C.I. Kurihara, *Sensing pressure for authentication system using keystroke dynamics.* 2004. **1**(1): p. 19-22.
8. Trojahn, M., F.J.I.J.o.C.S. Ortmeier, and I. Technology, *Biometric authentication through a virtual keyboard for smartphones.* 2012. **4**(5): p. 1.
9. Lin, D.-T. *Computer-access authentication with neural network based keystroke identity verification*. in *Neural Networks, 1997., International Conference on*. 1997. IEEE.
10. Epp, C., M. Lippold, and R.L. Mandryk. *Identifying emotional states using keystroke dynamics*. in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2011. ACM.
11. Bergadano, F., D. Gunetti, and C. Picardi, *User authentication through keystroke dynamics.* ACM Transactions on Information and System Security (TISSEC), 2002. **5**(4): p. 367-397.
12. Yu, E. and S. Cho. *Novelty detection approach for keystroke dynamics identity verification*. in *International Conference on Intelligent Data Engineering and Automated Learning*. 2003. Springer.
13. Saini, B.S., N. Kaur, and K.S. Bhatia, *Keystroke dynamics for mobile phones: A survey.* Indian Journal of Science and Technology, 2016. **9**(6).
14. Monaco, J.V., *Robust Keystroke Biometric Anomaly Detection.* arXiv preprint arXiv:1606.09075, 2016.
15. Idrus, S.Z.S., et al. *Soft biometrics for keystroke dynamics*. in *International Conference Image Analysis and Recognition*. 2013. Springer.
16. Rumelhart, D.E. and D.A.J.C.s. Norman, *Simulating a skilled typist: A study of skilled cognitive-motor performance.* 1982. **6**(1): p. 1-36.
17. Saevanee, H. and P. Bhattarakosol. *Authenticating user using keystroke dynamics and finger pressure*. in *2009 6th IEEE Consumer Communications and Networking Conference*. 2009. IEEE.
18. Zhao, G., et al., *Keystroke Dynamics Identification Based on Triboelectric Nanogenerator for Intelligent Keyboard Using Deep Learning Method.* 2019. **4**(1): p. 1800167.
19. Lu, L., et al. *KeyLiSterber: Inferring Keystrokes on QWERTY Keyboard of Touch Screen through Acoustic Signals*. in *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*. 2019. IEEE.
20. Idrus, S.Z.S., et al., *Soft biometrics for keystroke dynamics: Profiling individuals while typing passwords.* Elsevier International Journal of Computers & Security, 2014. **45**: p. 147-155.
21. Idrus, S.Z.S., et al. *Keystroke dynamics performance enhancement with soft biometrics*. in *IEEE International Conference on Identity, Security and Behavior Analysis (ISBA 2015)*. 2015. IEEE.

22. Tsai, C.-J. and K.-J.J.A.S.C. Shih, *Mining a new biometrics to improve the accuracy of keystroke dynamics-based authentication system on free-text.* 2019. **80**: p. 125-137.
23. Delac, K. and M. Grgic. *A survey of biometric recognition methods*. in *Proceedings. Elmar-2004. 46th International Symposium on Electronics in Marine*. 2004.
24. Hadid, A., et al., *Biometrics Systems Under Spoofing Attack: An evaluation methodology and lessons learned.* IEEE Signal Processing Magazine, 2015. **32**(5): p. 20-30.
25. Porwik, P., R. Doroz, and K.J.E.S.w.A. Wrobel, *An ensemble learning approach to lip-based biometric verification, with a dynamic selection of classifiers.* 2019. **115**: p. 673-683.
26. Friedman, L., et al., *Relationship between Number of Subjects and Biometric Authentication Equal Error Rates.* 2019.
27. Horn, S.S.J.C.D.P., *Sexual Orientation and Gender Identity-Based Prejudice.* 2019. **13**(1): p. 21-27.
28. He, M., et al., *Performance evaluation of score level fusion in multimodal biometric systems.* 2010. **43**(5): p. 1789-1800.
29. Jain, A., K. Nandakumar, and A.J.P.r. Ross, *Score normalization in multimodal biometric systems.* 2005. **38**(12): p. 2270-2285.
30. Jain, Y.K., S.K.J.I.J.o.C. Bhandare, and C. Technology, *Min max normalization based data perturbation method for privacy protection.* 2011. **2**(8): p. 45-50.
31. Aksu, G., C.O. Güzeller, and M.T.J.I.J.o.A.T.i.E. Eser, *The Effect of the Normalization Method Used in Different Sample Sizes on the Success of Artificial Neural Network Model.* 2019. **6**(2): p. 170-192.
32. Hampel, F.R., *Robust Statistics: The Approach Based on Influence Functions*. 1986: Wiley.
33. Sauri, J., *Cours complet de mathématiques*. Vol. 1. 1774: Ruault.
34. Larose, D.T., *Data Mining and Predictive Analytics*. 2015: Wiley.
35. Dziech, A., M. Leszczuk, and R. Baran, *Multimedia Communications, Services and Security: 8th International Conference, MCSS 2015, Kraków, Poland, November 24, 2015. Proceedings*. 2015: Springer International Publishing.
36. Idrus, S. Z. S., *Soft Biometrics for Keystroke Dynamics (Biométrie douce pour la dynamique de frappe au clavier)*(Doctoral dissertation, University of Caen Normandy, France), 2014.
37. S Syed Idrus, E Cherrier, C Rosenberger, Fusion et biométrie douce pour la dynamique de frappe au clavier. *Colloque COmpression et REprésentation des Signaux Audiovisuels (CORESA)*, 2016.