An efficient biometric-based continuous authentication scheme with HMM prehensile movements modeling

Feriel Cherifi, Mawloud Omar, Kamal Amroun

An efficient biometric-based continuous authentication scheme with HMM prehensile movements modeling

Feriel Cherifi[a,*], Mawloud Omar[b], Kamal Amroun[a]

[a]*Laboratoire d'Informatique Médicale (LIMED), Faculté des Sciences Exactes Université de Bejaia, 06000 Bejaia, Algérie.*
[b]*LIGM, ESIEE Paris, Université Gustave-Eiffel, Noisy-le-Grand*

Abstract

Biometric is an emerging technique for user authentication thanks to its efficiency compared to the traditional methods, such as passwords and access-cards. However, most existing biometric authentication systems require the cooperation of users and provide only a login time authentication. To address these drawbacks, we propose in this paper a new, efficient continuous authentication scheme based on the newly biometric trait that still under development: prehensile movements. In this work, we model the movements through Hidden Markov Model-Universal Background Model (HMM-UBM) with continuous observations based on Gaussian Mixture Model (GMM). Unlike the literature, the gravity signal is included. The results of the experiments conducted on a public database HMOG and on a proprietary database, collected under uncontrolled conditions, have shown that prehensile movements are very promising. This new biometric feature will allow users to be authenticated continuously, passively and in real time.

Keywords:
Biometric, authentication, prehensile movement, HMM-UBM, GMM.

*Corresponding author
Email addresses: cherififeriel@gmail.com (Feriel Cherifi), mawloud.omar@univ-eiffel.fr (Mawloud Omar), kamalamroun@gmail.com (Kamal Amroun)

1. Introduction

Internet of Things (IoT) is probably one of the most talked about technologies in this decade. Indeed, with the advent of IoT, several connected devices (smartphones, smartwatches, connected glasses, etc.) invaded our daily routine and contribute to the improvement of our daily life due to the manifold applications they offer. However, even if IoT is widely perceived as a revolutionary technology, security and privacy still the biggest obstacles to its development with the general public. It is obvious that the user is concerned about these intelligent objects that constantly transmit data about how they are used and that contain private information. Therefore, user data privacy has become a crucial issue to address. Traditionally, these devices were secured with passwords and PINs that can be easily stolen, forgotten or attacked. More recently, biometric authentication has become a primary focus of academic research and industry adoption/implementation to provide users enhanced security and authentications [1]. Indeed, physiological biometrics, such as fingerprint, iris and face, emerge the industrial devices. However, this type of authentication method is vulnerable to spoofing attacks. By the way, if an attacker succeeds to get the physiological templates (e.g., facial, fingerprint, etc.) of a given user, it will be infeasible to update the security system since we can not modify such biometric data. Furthermore, it provides only a login time authentication and any future change of the user identity will go undetected [2].

Continuous authentication is a way to mitigate this limitation by constantly verifying the user identity and locking the system once a change of the user identity is detected. As such, it is necessary for the system to periodically collect some identifying information about the user. The more frequently, such information is collected the faster a potential intruder can be detected [2]. By the way, physiological biometrics require the user intervention either to scan the fingerprint, iris, face or another, which makes continuous and real-time authentication a hard task. To overcome these impediments, human behavior is studied to achieve continuous authentication. Indeed, research works on the subject are numerous, whether it concerns touch gesture behavior, like in [3], mouse gesture dynamics [4], gait pattern [5] or many other behavioral traits. However, all these biometric features require the collaboration of the user to perform authentication (walking, interacting with the mouse, etc.). Recently, Neverova et al. [6] proposed the

use of human kinematic[1], using the mobile inertial sensors, as a new biometric trait that allows to authenticate a user only by the way he/she moves and holds his/her device without explicit interaction. Although they confirm that natural human kinematics convey sufficient information about mobile user identity and can be used for authentication.

In this paper, investigating the feasibility of utilizing the human kinematic extracted from smartphone sensors for user authentication, we propose an efficient continuous and real-time prehensile movement-based authentication scheme. The latter is passive[2] and allows the security of the user data while putting him/her in a comfortable situation. The user does not intervene at any time and unlike to physiological biometrics, it is harder for someone with malicious intent to successfully capture a natural motion and provide the option of continuous authentication. In this work, unlike [6] which use neural networks for feature extraction, we use a set of statistical features and use Hidden Markov Model-Universal Background Model (HMM-UBM) approach for classification. Using neural networks, we cannot easily determine which variables are the most important contributors to a particular output. Hence, the neural network model may contain a number of unimportant predictor variables, in addition to the high cost of computation, storage and authentication delay compared to the statistical metrics. These criteria are important in the context of connected devices, which must be highly optimized given the limited device resources.

The contribution of this paper is quadruple:

1. The proposition of an unobtrusive approach for continuous and real-time user authentication, which does not require any intervention from the user and no additional hardware, and respects the limited resources in the IoT context.
2. The data-set construction representing prehensile movements.
3. The integration of statistical metrics-based feature extraction from inertial sensor signals.

[1]Kinematics is a branch of physics that describes the motion of bodies or systems of bodies without considering the mass of each or the forces that caused the motion.

[2]A system that runs in the background and does not require any interaction from the user (non-cooperative and non-intrusive)

4. The proposition of a user pattern-based on HMM-UBM with mixture of Gaussian outputs.

We evaluated the proposed system on a proprietary database collected under uncontrolled conditions and on part of HMOG database [13]. The experimental results show that the proposed scheme is promising in terms of accuracy and authentication delay, and confirm that the natural human prehensile movement carries discriminant information and is therefore appropriate as a basis for continuous and real-time human recognition task. Indeed, prehensile movements are a new behavioral biometric trait (according to our knowledge we are the only ones with Neverova et al. [6] that treat this biometric trait) that could improve and facilitate the lives of users by ensuring high security in a continuous and passive way.

The outline of the paper is as follows. Section 2 describes existing work in the area of biometric authentication in mobile devices. In Section 3, we give a detailed description of the proposed scheme. In Section 4, we present the experimental results and discussion. Finally, Section 5 concludes the paper.

2. Related work

Nowadays, several IoT devices are equipped with inertial sensors. The researchers take an interest in these sensors to use them for behavioral biometric authentication that are implemented as a safe, dynamic and simple solution to realize a continuous authentication and addressing the limitations of physiological biometrics. Several works that use these sensors in their authentication scheme are proposed in the literature. Salem et al. [7] use keystroke dynamics as an authentication factor for user verification on touch screen mobile devices. They analyse user's input rhythm when he/she types a password. Roy et al. [8] proposed a continuous authentication scheme for touch interface based mobile devices. Sensors-touch, accelerometer and gyroscope data were used to model the user's gesture patterns based on HMM to model the tap and stroke patterns of a user. The authors of [9] used the information collected from the accelerometer sensor and touchscreen of the smartphone to authenticate its user. In [10], an authentication system based on how a user holds his phone while signing on its touchscreen named Hold & Sign was proposed. It takes into account micro-movements of a phone and movements of the user's finger during writing or signing on the touchscreen. ShakeIn, a smartphone user authentication scheme which authenticates him

by the way he/she shakes his phone, was proposed by Zhu et al. [11]. Data collection was done using accelerometers and gyroscopes. Barra et al. [12], explored the dynamic signals of arm movement produced by the gesture of lifting the phone, to check if the person answering the phone is its legitimate owner. The signals were captured from accelerometer, gyroscope and GPS (Global Positioning System) during the action of answering the phone in one of the following states: 1) standing; 2) sitting; 3) walking; 4) running. Sitova et al. [13] proposed a continuous and unobtrusive user authentication named HMOG (Hand Movement, Orientation, and Grasp) to authenticate smartphone user. They used accelerometer, gyroscope, and magnetometer to capture subtle hand micro-movements and orientation patterns generated when a user taps on the screen.

However, all these proposed methods require the cooperation of users and their interaction with the mobile object to authenticate themselves. In addition, the experiments are conducted under controlled laboratory conditions and the results obtained do not reflect the real performance of these methods.

Neverova et al. [6] proposed a non-cooperative and non-intrusive method for on-device authentication. According to the authors, they present the first method for active biometric authentication with mobile inertial sensors. They explore the capability of temporal deep neural networks to interpret natural human kinematics and confirmed that natural human kinematics convey necessary information about mobile user identity and can be used for authentication. The system acheives 20.52% equal error rate (EER) and the user is either authenticated or rejected after 30 seconds, we believe that this period is large enough for an imposter to access personal data of the legitimate user. The authors use neural networks for feature extraction which is a black box, we cannot easily determine which variables are the most important contributors to a particular output, so the neural network model may contain a number of unimportant predictor variables. In addition to the high cost of computation and memory compared to statistical metrics. These two criteria are very important in the context of IoT, and they must be minimized given the limited resource.

3. The proposed scheme

A malicious attacker may have several opportunities to violate a mobile phone owner's privacy. The device could be unprotected (no PIN, no password, no biometric authentication) or the attacker could be someone who knows the authentication secret (e.g., shoulder surfing, spoofing attack). The aim of our work is to handle such situations by analyzing prehensile behavior biometric of the user throughout the use of his/her smartphone. To this end, we propose a new efficient and discrete authentication scheme based on continuous HMM to continuously authenticate the user through its prehensile movements which describe the combined movement of reaching, grasping and manipulating objects. Since is quite hard to distinguish the intention of making authentication to the general action, the proposed approach does not support the operation of smartphone opening/closing. The proposed approach operates after unlocking the smartphone to provide an additional line of defense, designed as a nonintrusive and passive security countermeasure. The proposed system operates without user intervention and without additional hardware. The prehensile movements are measured only by the device inertial sensors as behavioral trait.

The proposed scheme consists of the following modules:

1. Data acquisition module, where the prehensile movements are passively collected from the inertial sensors by the user device.
2. Feature extraction module, where the inertial signals are pre-processed and used to generate the biometric feature vectors.
3. Training module, where a model based on continuous HMM-UBM is trained to characterize the device owner in a unique way.
4. User verification module, where the user identity is checked and the user is either authenticated or not.

In Table 1, we present the important notations used in this paper, and in the following subsections, we give the description of each module.

3.1. Data acquisition

This module allows to create the data-set, which describes how the user moves and holds his/her device. The data-set is collected from signals extracted from the multiple 3-dimensional inertial sensors: the gyroscope that measures the rotation rate along the device three axes in radians per second (rad/s), the accelerometer that measures the proper acceleration in meter

Table 1: Notations

Notation	Description
$\overrightarrow{M}, \overrightarrow{G}, \overrightarrow{P}$	Acceleration, angular velocity and gravity vectors
$[M], \lvert M\rvert, M^{-1}$	Diagonal, determinant and inverse of the matrix M
$\overline{X}, \sigma_X$	Mean and standard deviation of the sequence X
$\lVert \overrightarrow{U} \rVert$	Magnitude of vector $\overrightarrow{U}$
Π	Initial state distribution of HMM
A	Transition probability matrix of HMM
B	Multimodal mixture of Gaussian function
c	Weight component of GMM
Σ	Covariance matrix of GMM
μ	Mean vector of GMM
d_X	Dimension of the sequence X
Θ	Multidimensional Gaussian density function
λ	GMM-HMM model
$LLR_X(Y)$	Log-likelihood ratio with the sequence X of the user Y
$P(X\lvert\lambda)$	Probability that the sequence X is generated by the model λ

per second square (m/s^2), and the gravity that measures the gravitational component. The gravity is taken into account in the proposed work, unlike authentication systems using inertial sensors, because we believe that it contains important information about the user's identity, such as his/her force. In physics gravity is a linear acceleration that pulls us towards the ground. In order to retrieve gravity data, the Android sensor API offers a software sensor (an algorithm derived from a mathematical model which is able to reconstruct state variables (whose measurement is technically difficult or relatively expensive) from variables that can be captured constantly from common instruments [14]) that filters the raw data from the accelerometer using the Kalman filter to show only gravity. The data-set should be collected without any explicit intervention of the user. The collected data should represent at the best the user natural behavior.

3.2. Signal pre-processing

Before the feature extraction phase, the raw data collected from the accelerometer and gyroscope, which are sensitive to interference, are pre-processed as follows:

1. First, to have a reliable authentication system and high recognition

rates, we filter the noise caused by calibration imperfections from the signal that represents the user's movements. To mitigate such noise, an obfuscation method has been proposed in [15], having the advantage of independence from the calibration operation, which requires user intervention. The obfuscation introduces additional noise to the sensor readings to hide the natural errors. In this work, we consider a small obfuscation value. After obfuscation process the accelerometer and gyroscope outputs ($\overrightarrow{M}^\circ$ and $\overrightarrow{G}^\circ$) can be expressed as follows:

$$\overrightarrow{M}^\circ = \overrightarrow{O}^\circ + [S^\circ]\overrightarrow{M} \tag{1}$$

and

$$\overrightarrow{G}^\circ = \overrightarrow{O}^\circ + [S^\circ]\overrightarrow{G}, \tag{2}$$

where S° and O° are the obfuscation gain and offset[3], respectively.

2. For each vector $\overrightarrow{U} \in \{\overrightarrow{M}, \overrightarrow{G}, \overrightarrow{P}\}$, which are orientation sensitive, the magnitude $\|\overrightarrow{U}\|$ is computed and added to the existing three dimensions of each sensor as a fourth dimension. The magnitude has the advantage of making the sensors data independent of device orientation.

$$\|\overrightarrow{U}\| = \sqrt{U_x^2 + U_y^2 + U_z^2}, \tag{3}$$

3. Each projected component is then normalized as follows

$$U'_x = \frac{U_x}{\|\overrightarrow{U}\|}, U'_y = \frac{U_y}{\|\overrightarrow{U}\|}, U'_z = \frac{U_z}{\|\overrightarrow{U}\|}. \tag{4}$$

4. Before feature extraction, for each component, the time series signal should be segmented into a fixed-size sliding window. As the length of the sliding window affects recognition accuracy, we have to do several tests to choose the size that gives the optimal recognition rates. The results are presented and discussed in section 4.

[3]Gain and offset are two measurement errors. The deviation from the ideal voltage is the offset error. The deviation from the ideal over the full output range is known as the gain error.

3.3. Feature extraction

Feature extraction phase allows us to characterize each user and distinguish him from an impostor. For this purpose, we selected two different settings from the time domain, namely the mean and standard deviation. This approach does not require many computational resources, which is extremely important when using mobile devices [16]. For each window and for each $\{x, y, z, \|\overrightarrow{U}\|\}$ dimensions, the mean $\overline{X}$ and the standard deviation σ_X are computed in order to form the feature vectors. The parameter $\overline{X}$ is used in each dimension representing the speed/angle of the motion of that section on average [1]. The parameter σ_X represents both the variability of the data-set and the probability distribution within each section. The resulting vector is finally applied to the training module. At last, the feature vector is of the form

$$\{\overline{X}_{m_x}, \sigma_{m_x}, \overline{X}_{m_y}, \sigma_{m_y}, \overline{X}_{m_z}, \sigma_{m_z}, \overline{X}_{\|\overrightarrow{M}\|}, \sigma_{\|\overrightarrow{M}\|}, \overline{X}_{g_x}, \sigma_{g_x}, \overline{X}_{g_y}, \sigma_{g_y}, \overline{X}_{g_z}, \sigma_{g_z},$$
$$\overline{X}_{\|\overrightarrow{G}\|}, \sigma_{\|\overrightarrow{G}\|}, \overline{X}_{p_x}, \sigma_{p_x}, \overline{X}_{p_y}, \sigma_{p_y}, \overline{X}_{p_z}, \sigma_{p_z}, \overline{X}_{\|\overrightarrow{P}\|}, \sigma_{\|\overrightarrow{P}\|}, \}$$

3.4. HMM-UBM system

In this section, we first present some basics of HMM, and then we present how the HMM-UBM is used in our system for user authentication.

3.4.1. Hidden Markov Model (HMM)

An HMM is a collection of finite states connected by transitions. A typical HMM is defined by n states $S = \{S_1, S_2, \cdots, S_n\}$. A state at a time t is denoted by q_t and m, which is the number of distinct observation symbols per state, expressed by $V = \{V_1, V_2, \cdots, V_m\}$. The HMM is denoted by $\lambda = (\Pi, A, B)$. The parameter $\Pi = \{\pi_i\}$ represents the initial state distribution, where

$$\pi_i = P(q_1 = S_i), \tag{5}$$

with

$$\sum_{i=1}^{n} \pi_i = 1. \tag{6}$$

The parameter $A = \{a_{ij}\}$ represents a $n \times n$ transition probability matrix, where

$$a_{ij} = P(q_t = S_j | q_{t-1} = S_i), i = 2..n, j = 2..n, \tag{7}$$

and

$$\sum_{j=1}^{n} a_{ij} = 1, \forall i. \tag{8}$$

The parameter $B = \{b_j(k)\}$ represents a $m \times n$ observation symbol probability matrix in the state j, where

$$b_j(k) = P(V_k = t | q_t = S_j), j = 1..n, k = 1..m \tag{9}$$

and

$$\sum_{k=1}^{m} b_{jk} = 1, \forall j. \tag{10}$$

$b_j(k)$ is the probability of emitting V_k at the time t in the state S_j.

There are two types of HMM characterized by two sets of probabilities: discrete and continuous. Those denominations are related to the outputs given by a single HMM state. In the discrete HMM, a discrete output probability distribution defines the condition probability of emitting each output symbol from a finite alphabet. In the case of continuous HMM, continuous output probability density function defines the condition probability of emitting each output symbol from a continuous random vector.

3.4.2. Our HMM-UBM based behavior model

When registering users for the first time, the user's model training should be done in a reasonable time as the registration is done in real time. However, the HMM is a complex model and cannot be initialized with limited training data. Since we are dealing with short enrollment data, we adopted the hidden Markov model-universal background model (HMM-UBM) approach. The UBM is a single common model, which represents the whole population. It is an off-line trained model using prehensile sequences from multiple persons. In our work, the UBM, that we note λ_{UBM}, is a large HMM trained to

represent the distribution of users' prehensile movements features.

In order to generate the user model, during the enrollment phase, we use the UBM as a baseline and adapted it using Maximum-a-Posteriori (MAP) approach to derive the user model. The obtained optimal model represents the enrolled user. This adaptation allows a faster scoring method. Figure 1 summarizes the registration steps of a user. The user is then described by a set of n distinct hidden states. The emission probability distribution of each state of the HMM is modeled as a multimodal mixture of Gaussian (GMM). In this case, B represents a multimodal mixture of Gaussian function of a set of observation data such as

$$B = \{b_j(X), j = 1..n\}, \tag{11}$$

where

$$b_j(X) = \sum_{k=1}^{m} c_{jk} b_{jk}(X) = \sum_{k=1}^{m} c_{jk} \Theta(X, \mu_{jk}, \Sigma_{jk}). \tag{12}$$

Θ is the multidimensional Gaussian density function such as

$$\Theta(X, \mu_{jk}, \Sigma_{jk}) = \frac{1}{\sqrt{(2\pi)^{d_X} |\Sigma_{jk}|}} e^{-\frac{1}{2}(X-\mu_{jk})^T \Sigma_{jk}^{-1} (X-\mu_{jk})}. \tag{13}$$

The idea of using continuous HMM has been greatly influenced by the successful application of HMM in gesture recognition. Moreover, for their efficiency to model gestures as a continuous motion phenomenon on a sequential time series [17]. Furthermore, it is computationally inexpensive and is sensitive to the temporal aspects of the human movement.

The continuous HMM with GMM outputs model of the proposed scheme, is defined by $\lambda = (\Pi, A, c, \mu, \Sigma)$. In the training phase, the feature sequence is considered as the observation sequence of GMM. Therefore, the parameters of GMM are estimated by using the Expectation-Maximization (EM) algorithm [18], which iteratively refines the GMM parameters. Afterwards, the Baum-Welch algorithm [19] is executed to estimate the matrix A in such a way that they can maximize the likelihood probability for the given training data.

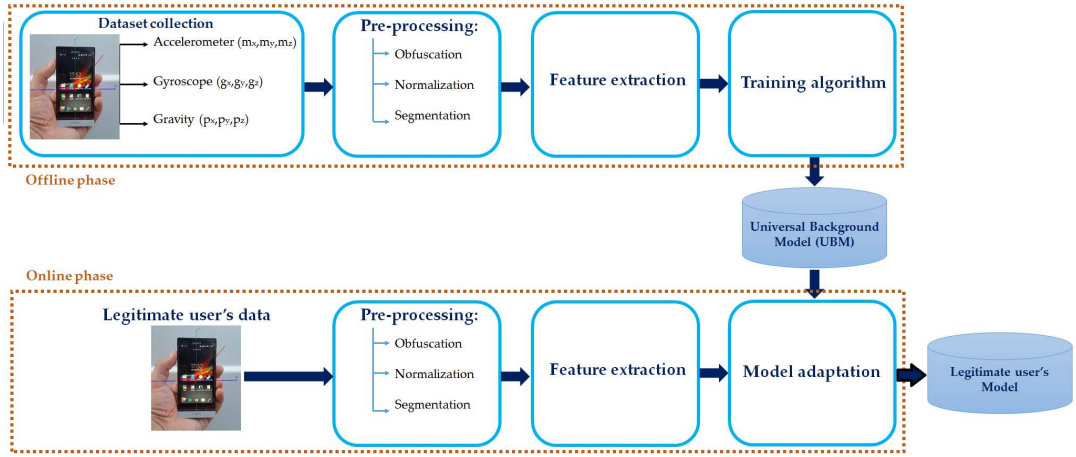


Figure 1: Enrollment phase of the proposed system. First, in offline phase, prehensile movement data are collected from inertial sensors including information from all users to traine an Universal Background Model (UBM). In online phase, user model is generated by adapting the UBM.

3.5. User verification module

The log-likelihood is calculated to determine if a segment of moves X is produced by a hypothesized user Y. For verification, a constructed model from a particular user is scored against the UBM model, which is trained with a large amount of data, used to represent general user-independent characteristics. The log-likelihood ratio (LLR) of the segment X is computed as follows

$$LLR_Y(X) = \log\left(P(X|\lambda_Y)\right) - \log\left(P(X|\lambda_{UBM})\right) \tag{14}$$

The result is compared with a predetermined threshold θ, which is defined during the training phase according to the inter- and intra-user distance distributions. The user Y is authenticated if and only if $LLR_Y(X) \geq \theta$. Finally, we applied the zt-score normalization [20] to compensate for inter-session and inter-person variations.

$$LLR(Y)_{X,ZTnorm} = \frac{LLR_X(Y) - \alpha(X|\lambda_{UBM})}{\beta(X|\lambda_{UBM})}. \tag{15}$$

where α and β are normalization statistics. A schematic diagram is given in

Figure 2 which illustrates the user verification process.

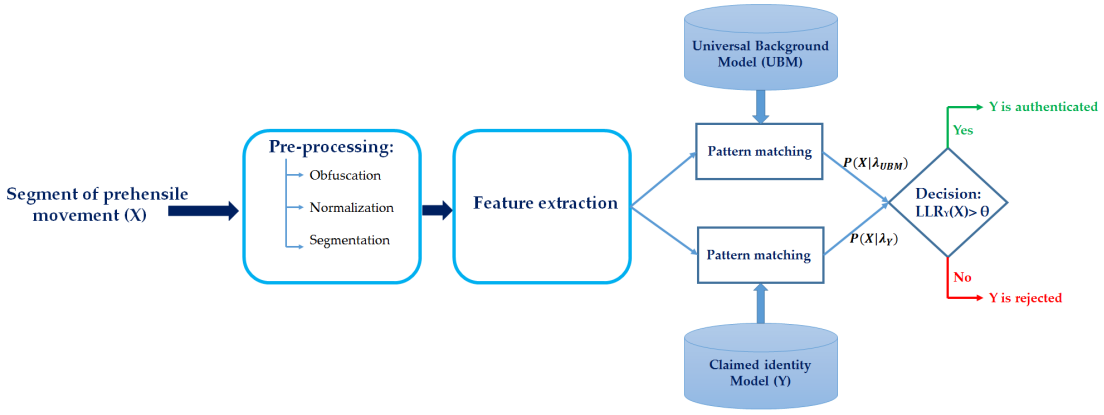


Figure 2: Verification phase of the proposed system. First, prehensile movement sequence (X) is collected from inertial sensors. Specific features, relevant for recognition, are extracted. The features vector is matched with both the UBM model, which represent the negative class, and the claimed user model. Finally a matching decision is made.

The pseudo code 1, resumes all steps of our proposition.

4. Performance evaluation

In this section, we first give details about the collected data-set, then we present and discuss the obtained results.

4.1. data-set collection

In our experiments, we have used a smartphone as a connected device. In order to collect motion data, we have designed and implemented a realtime Android service. The service runs without any user interaction in the background during the period between the unlocking to locking of the phone screen. The data were acquired simultaneously from the accelerometer, gyroscope and gravity sensors based on the coordinate system used by the Android Sensor API (Application Programming Interface), illustrated in Figure 3. We have set the sampling frequency to 50Hz. During the data acquisition, 7 volunteers within an age ranging from 19 to 56 years used our research smartphone, Sony Xperia P model. We collected the data using a single smartphone to ensure that the use of different phone models would not impact the authentication results. The session time ranged between 15 to 25 minutes per volunteer. The data collection was done in two days, the first day we collected data from 4 volunteers and the second day we collected data from 3 volunteers. In our work, we did not impose any tasks or interaction with the user's smartphone. The volunteers were free to manipulate the smartphone as they wished to be sure that the collected data were representative of the natural and regular movements. The collected data is available online `https://figshare.com/articles/Prehensile_movements_when_handling_a_smartphone/11855709`.

To preserve the volunteers' privacy, we did not collect any information that can be used to identify them. For each sensor, we collect 3-dimensional values denoting the user motion. Figure 14 illustrates a sample of the data collected by the 3-dimensional sensors: gyroscope, accelerometer and gravity in x, y and z direction, composed of 500 points with a window of 1s for a given user. For a given timestamp t, we get a vector of the form $X^{(t)} = (m_x, m_y, m_z, g_x, g_y, g_z, p_x, p_y, p_z) \in \mathbb{R}^9$, that we record in our database, as illustrated in Figure 4. The parameters x, y and z denote the projections of $\overrightarrow{M}$, $\overrightarrow{G}$ and $\overrightarrow{P}$ respectively, on the corresponding axes aligned with the smartphone.

The pseudo code of our proposition

1: Data collection from accelerometer $\overrightarrow{M}$, gyroscope $\overrightarrow{G}$ and gravity $\overrightarrow{P}$ in X, Y and Z directions; the collected data are of the form $\{m_x, m_y, m_z, g_x, g_y, g_z, p_x, p_y, p_z\} \in \mathbb{R}^9$.

// Signal pre-processing

2: Add magnitude $\|\overrightarrow{U}\|$ as a fourth dimension to all of these sensors.

3: Apply the process of obfuscation.

$$\overrightarrow{M}^{\circ} = \overrightarrow{O}^{\circ} + [S]\overrightarrow{M}$$

and

$$\overrightarrow{G}^{\circ} = \overrightarrow{O}^{\circ} + [S^{\circ}]\overrightarrow{G}$$

4: Normalize each projected component, $U'_x = \frac{U_x}{\|\overrightarrow{U}\|}, U'_y = \frac{U_y}{\|\overrightarrow{U}\|}, U'_z = \frac{U_z}{\|\overrightarrow{U}\|}$.

5: For each dimension X, Y, Z and magnitude, segment the collected signal into fixed-size sliding window of size h.

// Feature extraction

6: For each window, compute separately the mean $\overline{X}$ and the standard deviation σ_X; the feature vector is of the form

$$\{\overline{X}_{m_x}, \sigma_{m_x}, \overline{X}_{m_y}, \sigma_{m_y}, \overline{X}_{m_z}, \sigma_{m_z}, \overline{X}_{\|\overrightarrow{M}\|}, \sigma_{\|\overrightarrow{M}\|}, \overline{X}_{g_x}, \sigma_{g_x}, \overline{X}_{g_y}, \sigma_{g_y}, \overline{X}_{g_z}, \sigma_{g_z},$$
$$\overline{X}_{\|\overrightarrow{G}\|}, \sigma_{\|\overrightarrow{G}\|}, \overline{X}_{p_x}, \sigma_{p_x}, \overline{X}_{p_y}, \sigma_{p_y}, \overline{X}_{p_z}, \sigma_{p_z}, \overline{X}_{\|\overrightarrow{P}\|}, \sigma_{\|\overrightarrow{P}\|}, \}$$

// Model learning

7: UBM learning offline.

8: UBM adaptation for user model construction, online.

// User verification

9: compute $P(X|\lambda_Y)$.

10: compute $P(X|\lambda_{UBM})$.

11: compute $LLR_Y(X) = \log\left(P(X|\lambda_Y)\right) - \log\left(P(X|\lambda_{UBM})\right)$.

12: zt-score normalization

if $LLR_Y(X) \geq \theta$ **then**

The user is the legitimate one.

else

The user is an imposter.

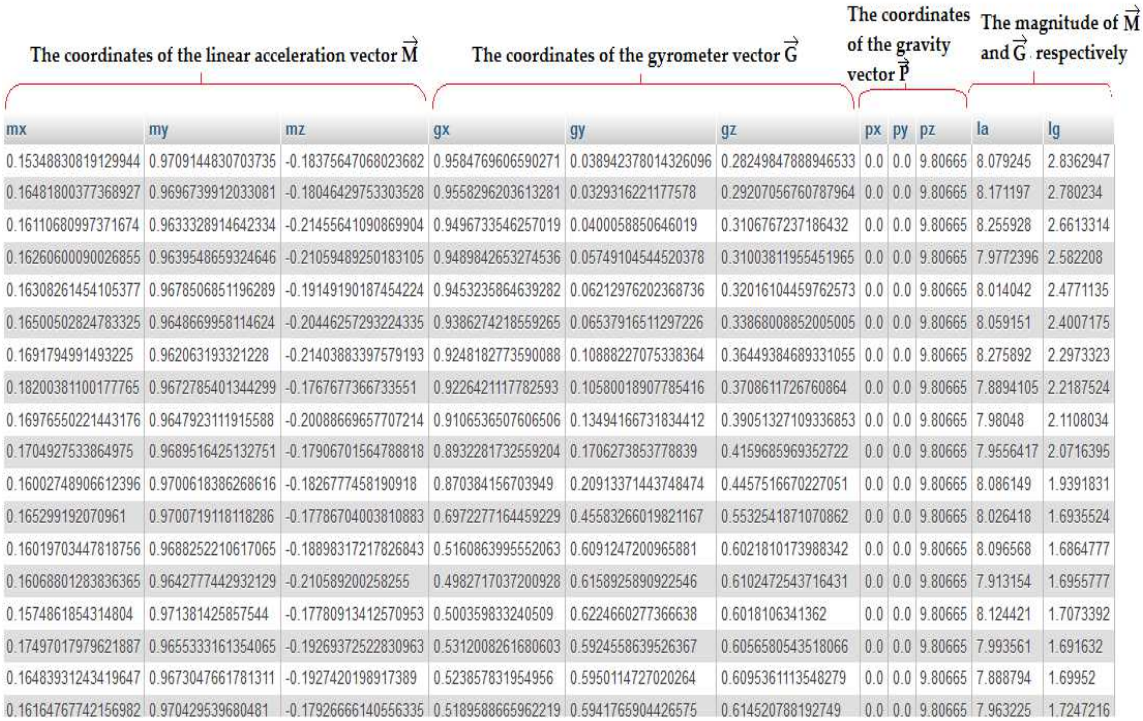
end if

Figure 3: Coordinate system used by the Android Sensor API.

4.2. Experiments

For the evaluation, we have divided the collected data-set into two sub-classes. 60% of the data were used for HMM-UBM training and the remaining 40% were used for the test phase. At the beginning of the development of our system, we were confronted with the choice of the HMM state number and the number of Gaussian mixtures per state. In order to set the optimal values, we have tested different combinations of mixtures per state and states per model. The obtained experimental results are presented in Figure5, in which the optimal values are 3 states and 2 Gaussian mixtures.

Another important value to define, which could have an impact on the performances of our system, is the size of the window used for feature extraction. To decide the window-size to use and to show the contribution of gravity data in the performance of our proposition, we tested it according to these two parameters. Figure 6 shows the obtained variation of our scheme in terms of Half Total Error Rate (HTER) in function of the window size (the one used during the feature extraction module) in two cases, with and without inclusion of the gravity sensor data. We note that the performance results with gravity data inclusion are clearly better.
Indeed, HTER = 16.10 % is the lowest rate obtained when we consider a window-size of 20 points, and taking into account gravity. We explain this by the fact that the gravity reveals important information about the smartphone user. The force with which the smartphone is worn is naturally



mx	my	mz	gx	gy	gz	px	py	pz	la	lg
0.15348830819129944	0.9709144830703735	-0.18375647068023682	0.9584769606590271	0.038942378014326096	0.28249847888946533	0.0	0.0	9.80665	8.079245	2.8362947
0.16481800377368927	0.9696739912033081	-0.18046429753303528	0.9558296203613281	0.032931622117578	0.29207056760787964	0.0	0.0	9.80665	8.171197	2.780234
0.16110680997371674	0.9633328914642334	-0.21455641090869904	0.9496733546257019	0.0400058850646019	0.3106767237186432	0.0	0.0	9.80665	8.255928	2.6613314
0.16260600090026855	0.9639548659324646	-0.21059489250183105	0.9489842653274536	0.05749104544520378	0.31003811955451965	0.0	0.0	9.80665	7.9772396	2.582208
0.16308261454105377	0.9678506851196289	-0.19149190187454224	0.9453235864639282	0.06212976202368736	0.32016104459762573	0.0	0.0	9.80665	8.014042	2.4771135
0.16500502824783325	0.9648669958114624	-0.20446257293224335	0.9386274218559265	0.06537916511297226	0.33868008852005005	0.0	0.0	9.80665	8.059151	2.4007175
0.1691794991493225	0.962063193321228	-0.21403883397579193	0.9248182773590088	0.10888227075338364	0.36449384689331055	0.0	0.0	9.80665	8.275892	2.2973323
0.18200381100177765	0.9672785401344299	-0.1767677366733551	0.9226421117782593	0.10580018907785416	0.3708611726760864	0.0	0.0	9.80665	7.8894105	2.2187524
0.16976550221443176	0.9647923111915588	-0.20088669657707214	0.9106536507606506	0.13494166731834412	0.39051327109336853	0.0	0.0	9.80665	7.98048	2.1108034
0.1704927533864975	0.9689516425132751	-0.17906701564788818	0.8932281732559204	0.1706273853778839	0.4159685969352722	0.0	0.0	9.80665	7.9556417	2.0716395
0.16002748906612396	0.9700618386268616	-0.1826777458190918	0.870384156703949	0.20913371443748474	0.4457516670227051	0.0	0.0	9.80665	8.086149	1.9391831
0.165299192070961	0.9700719118118286	-0.17786704003810883	0.6972277164459229	0.45583266019821167	0.5532541871070862	0.0	0.0	9.80665	8.026418	1.6935524
0.16019703447818756	0.9688252210617065	-0.18898317217826843	0.5160863995552063	0.6091247200965881	0.6021810173988342	0.0	0.0	9.80665	8.096568	1.6864777
0.16068801283836365	0.9642777442932129	-0.210589200258255	0.4982717037200928	0.6158925890922546	0.6102472543716431	0.0	0.0	9.80665	7.913154	1.6955777
0.1574861854314804	0.971381425857544	-0.17780913412570953	0.500359833240509	0.6224660277366638	0.6018106341362	0.0	0.0	9.80665	8.124421	1.7073392
0.17497017979621887	0.9655333161354065	-0.19269372522830963	0.5312008261680603	0.5924558639526367	0.6056580543518066	0.0	0.0	9.80665	7.993561	1.691632
0.16483931243419647	0.9673047661781311	-0.1927420198917389	0.523857831954956	0.5950114727020264	0.6095361113548279	0.0	0.0	9.80665	7.888794	1.69952
0.16164767742156982	0.970429539680481	-0.17926666140556335	0.5189588665962219	0.5941765904426575	0.614520788192749	0.0	0.0	9.80665	7.963225	1.7247216

Figure 4: An overview of the data-set structure

different from a user to another.

Based on these results, we set our system parameters to a 20 points window and an HMM with three states and two Gaussian mixtures per state, Figure 7 gives an overview of the chain. Our scheme is suitable for real-time applications on smartphones since it requires only 42 kilobytes of memory to save the UBM and the user models.

In our experiments, an objective evaluation of the prehensile movement as biometric is performed by measuring the False Acceptance Rate (FAR), False Rejection Rate (FRR) and Equal Error Rate (EER) performed by the system. EER indicates that the proportion of FAR is equal to the proportion of FRR. The lower the EER value, the higher the accuracy of the biometric system. On a graph depicting the FAR versus the FRR the EER is the intersection of both lines of the graph.

To investigate the feasibility of continuous authentication using only the

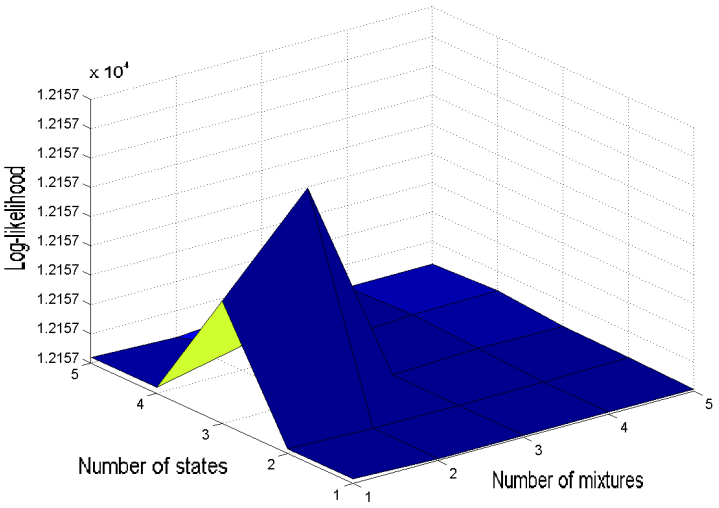


Figure 5: Log-likelihood according to number of states and mixtures number per state

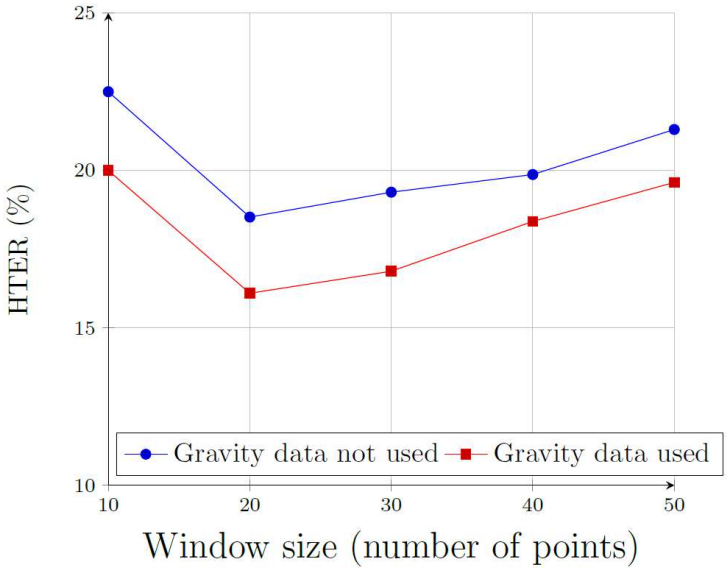


Figure 6: HTER according to the window size

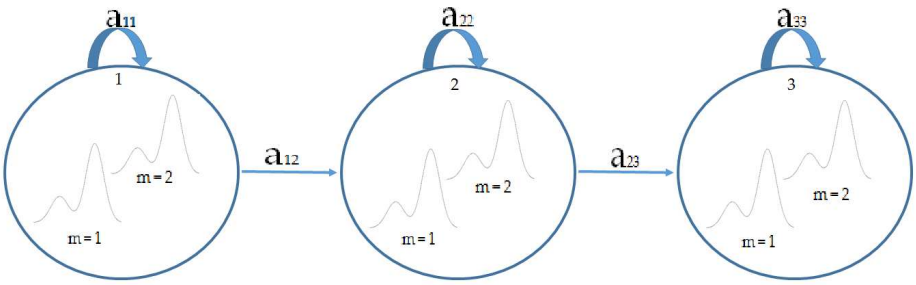


Figure 7: An overview of the HMM model used in our system, it consists of three states and two Gaussian mixtures per state.

prehensile movements, we provided experiments on two different databases: HMOG [13] and our database.

4.3. Experiments provided on HMOG dataset

HMOG (downloadable on: http://www.cs.wm.edu/ qyang/hmog.html) is a public database. It contains the record of touch, sensor and key press data invoked by 100 users during document reading, text production and navigation on a map to locate a destination. For each of the 100 subjects, 24 sessions have been considered. For our experiments, we have selected only accelerometer and gyroscope data as we are interested in authentication using only initial sensors data. As the data was retrieved in 24 sessions we have concatenated all the sessions to have only one session which includes all activities (reading, writing and map navigation). The process of pre-processing (data filtering, magnitude calculation, data normalization and data segmentation into fixed-size sliding window) and feature extraction (as described in Section 3.3) is the same as for our protocol. HMOG was collected in constrained settings as a part of a lab study and its size is larger than our dataset (the HMM-UBM training would be better) so we presume that the ERRs would be better than the ones we get from our database. The result of this experiment is shown in Figure 8. The EER = 14.8 %.

4.4. Experiments provided on our dataset

In this part we will present the results of the experiments applied on our dataset. Figure 9 resumes the quantitative results achieved by the HMM-UBM system trained with the raw data where the gravity data were included.

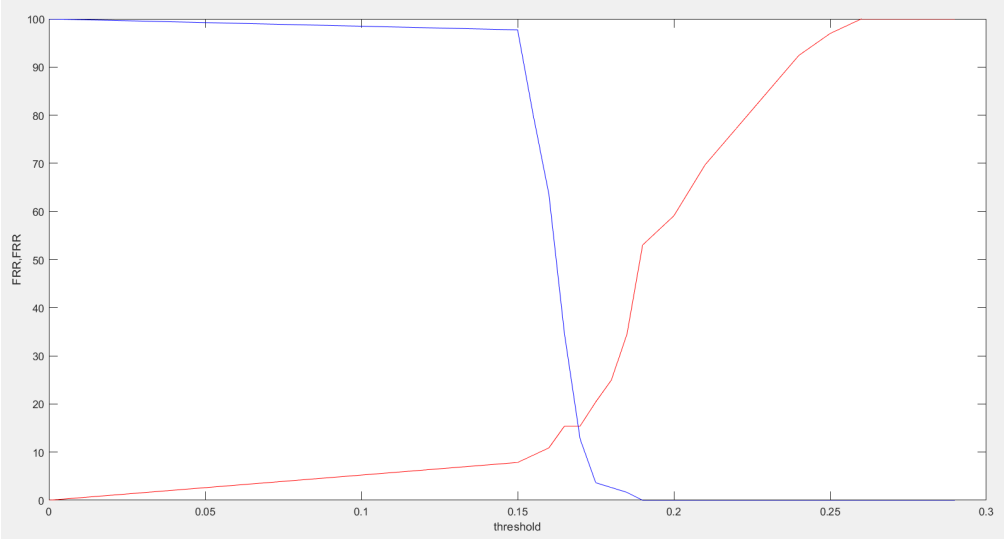


Figure 8: FAR vs FRR performed by the trained systems using HMOG dataset.

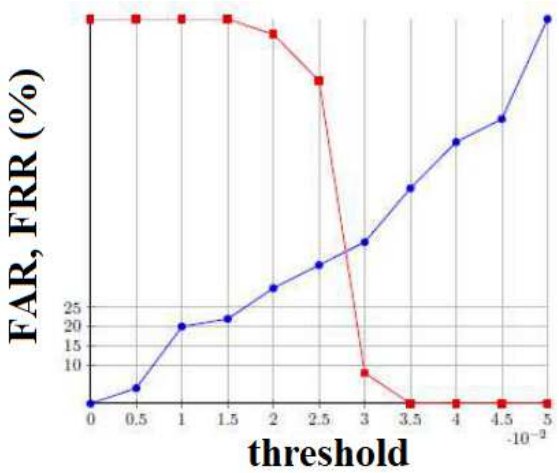


Figure 9: FAR vs FRR performed by the trained systems using raw data including gravity data.

As illustrated in Figure 9, the EER = 32.8%. Figure 10 shows the case where HMMs were trained without using gravity data, we used only accelerometer and gyroscope data, the results show an EER = 39.6%. As

expected, using gravity data improves system performance by providing additional information on the user's strength.

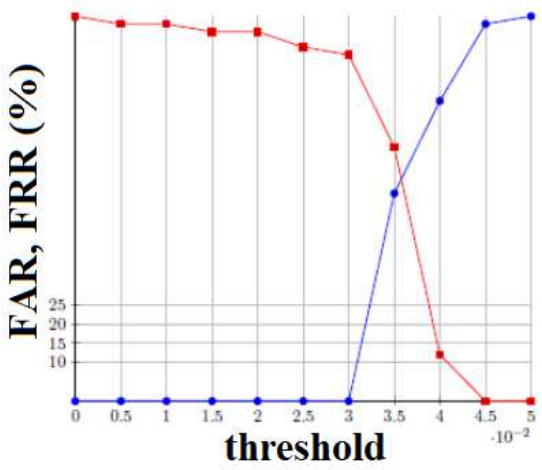


Figure 10: FAR vs FRR performed by the trained systems using raw data without using gravity data.

4.4.1. Influence of sensors data fusion

To improve the authentication rates, we also investigate the contribution of combining the data of accelerometer, gyroscope and gravity by fusing the corresponding captured data using a weighted sum. The results are reported in figure 11 and 12. Looking at the obtained results, we notice that fusing the sensors data together improves the system performance, whether it is when gravity data is included or not. EER = 19.2% is the lowest rate obtained when gravity data are used and fused with accelerometer and gyroscope data. In the case where gravity data are not included, EER is significantly higher is reaching 35.2%.

4.5. Discussion

In this section we depict the outcome of all the experiments. We tested the performance of the proposed system on two databases: a public database called HMOG [13] and a personal database [21], by measuring EERs. The EER obtained on HMOG (shown in Figure 8) is equal to 14.8 %. We then conducted a set of experiments on our database evaluating the impact of

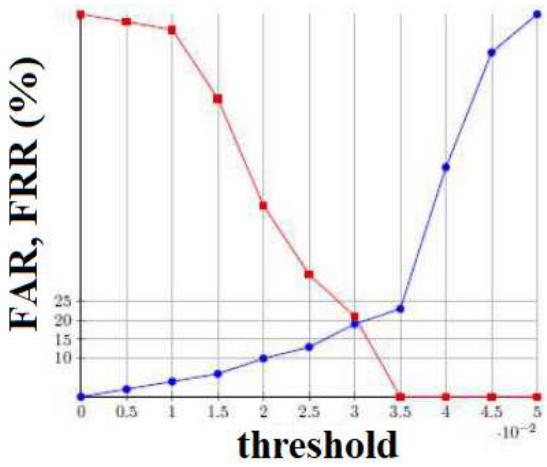


Figure 11: FAR vs FRR performed by the trained systems using the fusing data including gravity data.

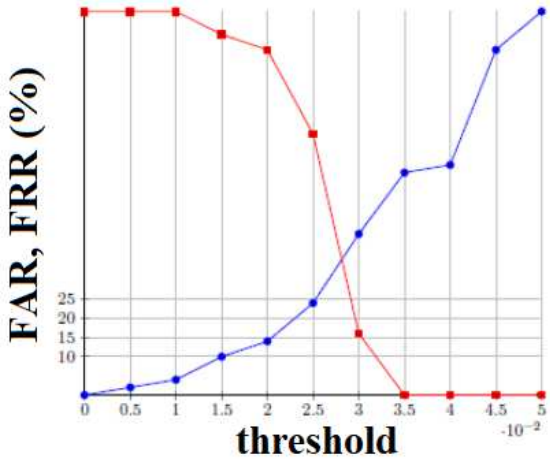


Figure 12: FAR vs FRR performed by the trained systems using fusing data without using gravity data.

gravity data and sensor data fusion on the system performances to answer two questions: *Does gravity improve the system performances? Does the fusion of accelerometer, gyrometer and gravity data improve the system performances?* To answer the first question, we can analyze the results obtained

in Figures 9 and 10, which illustrate the EERs obtained with and without the inclusion of gravity data, respectively. As shown in Figure 9, EER = 32.8% as opposed to EER = 39.6% obtained when excluding gravity in the tests. Based on these results we can conclude that gravity enriches the authentication model and improves recognition rates. With respect to sensor data fusion (to answer the second question) we analyze the results obtained in Figures 11 and 12 which represent the EERs obtained in the case of sensor data fusion by including and excluding gravity data, respectively. In this case the results show that fusing the accelerometer, gyroscope and gravity data improves the performance of the system by achieving EER = 19.2 %.

From all these experiments we can conclude that it is possible to distinguish users based only on their prehensile movements. This allows real-time biometric authentication to be carried out, continuously and passively (without requiring the user's collaboration). The lowest EER = 19.2% is obtained when fusing the accelerometer, gyroscope and gravity data. The study is proposed as a first step in the development of a secure non-obtrusive authentication system. The obtained result can also mean that 80.8% of the time the user is authenticated only by the way she/he holds and moves the smartphone, which is in our opinion quite encouraging. The obtained EER, that may seem high, is due to the fact that the data were collected in an uncontrolled environment. In real-word applications, the biometric data is affected by several sources of noise that affects recognition rates.

4.6. Comparison results with Neverova et al.'s approach

We also assessed the performances of the proposed scheme in comparison to [6], which is the closest one to our system as it uses the prehensile movements as biometric trait. The comparison was done with respect of two important metrics, namely the EER and the authentication delay. The authentication delay represents the time spent by the system to accept or reject the authentication of the user. First, we present the results obtained in relation to the authentication delay. As can be seen in Table 2, the system takes 8 seconds (s) to authenticate a user when using a 20 point window (where the best HTER is performed), this is much better than the delay achieved by [6] where the authentication delay is 30 seconds.

Table 3 shows the obtained results through the proposed scheme compared to [6]. EER = 19.2% are the lowest rates and are obtained when we consider a window of 20 points. In this case, as shown in Table 2, only 8s

Table 2: Authentication delay in function of feature window size.

Feature window size	10	20	30	40	50
Delay	2s	8s	18s	32s	50s

of the inertial data are used to authenticate the user. This means that the user verification test can be done after every 8s, which allows to improve the security level. Indeed, the more frequently the user identity verification is carried, the faster a potential intruder can be detected. Regarding the results according [6], 20.52% of EER and 30s of authentication delay are obtained. This period could be sufficient for an impostor to access the personal data of a legitimate user without be detected.

To make a comparison with [6] on the same data set, we calculated the EERs obtained by the two protocols using the HMOG database. To develop the protocol of Neverova et al. [6], we used the parameters indicated in Neverova's thesis [22].

Table 3: Comparison.

	EER	Authentication delay	EER on HMOG dataset
Neverova et al. [6]	20.52%	30s	17.52
The proposed scheme	19.2%	8s	14.8

4.7. Testing the protocol on touchscreen data

After testing the performance of the proposed protocol using prehensile movements as a biometric trait, we will study its performance on another biometric trait. In this section we investigate whether our protocol can continuously authenticate users based on the way they interact with the touchscreen of a smartphone. To do so, we used the public database proposed by Frank et al. [23] where they propose a set of raw touch data collected from 41 users interacting with a smartphone using basic navigation maneuvers.

For our tests we have extracted the same features as those proposed by Frank et al. [23] and the same classifier as discussed in section 3.4. The

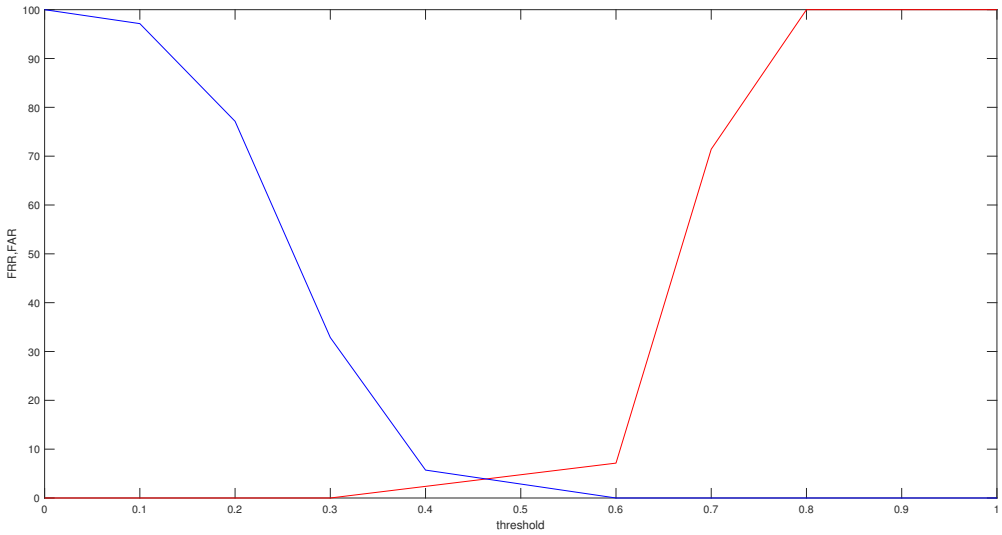


Figure 13: FAR vs FRR performed by the proposed protocol using the touchscreen data

results of the experiments are illustrated in the Figure 13. EER = 3.74 % is the rate obtained by applying the proposed protocol on touchscreen data. The table 4, summarizes the EERs obtained by our protocol and the one proposed by Neverova et al. [6] applied to three different datasets. Our protocol achieves better recognition rates on all three datasets.

Table 4: Comparison results with Neverova et al. [6].

	EER	EER on HMOG dataset	EER on Frank at al. dataset
Neverova et al. [6]	20.52%	17.52%	8.25%
The proposed scheme	19.2%	14.8%	3.74%

5. Conclusion

In this paper, we have proposed an efficient continuous and realtime authentication scheme for smartphones based on user prehensile movements

through the inertial signal sensors. We have elaborated a data-set of prehensile movements and integrated a Hidden Markov Model-Universal Background Model (HMM-UBM) to evaluate the identity of users. The proposed scheme collects in realtime the prehensile movement data with gravity signals, extracts the feature vectors, trains the model HMM-UBM and verifies the user's identities. For the performance evaluation, we have implemented a prototype of the proposed scheme and we have tested its efficiency in terms of Equal Error Rate (EER) and authentication delay. The proposed scheme demonstrates a considerable gain in terms of accuracy and delay. We plan to improve the authentication rate by proposing a multimodal biometric system combining prehensile movement with touch movements and location.

acknowledgements

This work was carried out in the framework of research activities of the laboratory LIMED, which is affiliated to the Faculty of Exact Sciences of the University of Bejaia. The authors are grateful to all the anonymous volunteers who participated in the process of the data-set elaboration.

funding sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declarations of interest

None.

References

[1] J. Maghsoudi and C. C. Tappert, *A Behavioral Biometrics User Authentication Study Using Motion Data from Android Smartphones*, European Intelligence and Security Informatics Conference (EISIC), 2016, pp. 184–187, Uppsala, Sweden. https://doi.org/10.1109/EISIC.2016.047.

[2] S. Eberz, K. B. Rasmussen, V. Lenders and I. Martinovic, *Evaluating Behavioral Biometrics for Continuous Authentication: Challenges and Metrics*, ASIA CCS '17 Proceedings of the 2017

ACM on Asia Conference on Computer and Communications Security, 2017, pp. 386–399, Abu Dhabi, United Arab Emirates. https://doi.org/10.1145/3052973.3053032.

[3] A. A. Alariki, A. A. Manaf and S. Khan, *A study of touching behavior for authentication in touch screen smart devices*, International Conference on Intelligent Systems Engineering (ICISE), 2016, pp. 216–221, Islamabad, Pakistan. https://doi.org/10.1109/INTELSE.2016.7475123.

[4] B. Sayed, I. Traoré, I. Woungang and M. S. Obaidat, *Biometric Authentication Using Mouse Gesture Dynamics*, IEEE Systems Journal, vol. 7(2)(2013) 262–274. https://doi.org/10.1109/JSYST.2012.2221932.

[5] N. Al-Naffakh, N. Clarke, F. Li and P. H. Dowland, *Unobtrusive Gait Recognition Using Smartwatches*, International Conference of the Biometrics Special Interest Group (BIOSIG), 2017, pp. 1–5, Darmstadt, Germany. https://doi.org/10.23919/BIOSIG.2017.8053523.

[6] N. Neverova, C. Wolf, G. Lacey, L. Fridman, D. Chandra, B. Barbello and G. Taylor, *Learning Human Identity From Motion Patterns*, IEEE Access, vol. 04(2016) 1810–1820. https://doi.org/10.1109/ACCESS.2016.2557846.

[7] A. Salem, D. Zaidan, A. Swidan and R. Saifan, *Analysis of Strong Password Using Keystroke Dynamics Authentication in Touch Screen Devices*, Cybersecurity and Cyberforensics Conference, 2016, pp. 15–21, Amman, Jordan. https://doi.org/10.1109/CCC.2016.11.

[8] A. Roy, T. Halevi and N. Memon, *An HMM-based Multi-sensor Approach for Continuous Mobile Authentication*, IEEE Military Communications Conference, 2015, pp. 1311–1316. https://doi.org/10.1109/MILCOM.2015.7357626.

[9] D. Shih, C. Lu and M. Shih, *A Flick Biometric Authentication Mechanism on Mobile Devices*, International Conference on Informative and Cybernetics for Computational Social Systems (ICCSS), 2015, pp. 31–33, Chengdu, China. https://doi.org/10.1109/ICCSS.2015.7281144.

[10] A. Buriro, B. Crispo, F. Delfrari and K. Wrona, *Hold and Sign: A Novel Behavioral Biometrics for Smartphone User Authentication*, IEEE

Security and Privacy Workshops (SPW), 2016, pp. 276–285, San Jose, CA, USA. https://doi.org/10.1109/SPW.2016.20.

[11] H. Zhu, J. Hu, S. Chang and L. Lu, *ShakeIn: Secure User Authentication of Smartphones with Single-Handed Shakes*, IEEE Transactions on Mobile Computing, vol. 16(10)(2017) 2901–2912. https://doi.org/10.1109/TMC.2017.2651820.

[12] S. Barra, G. Fenu, M. De Marsico, A. Castiglione and M. Nappi, *Have you permission to answer this phone?*, International Workshop on Biometrics and Forensics (IWBF), 2018, pp. 1–7, Sassari, Italy. https://doi.org/10.1109/IWBF.2018.8401563.

[13] Z. Sitova, J. Sedenka, Q. Yang, G. Peng, G. Zhou, P. Gasti and K. S. Balagani, *HMOG: New Behavioral Biometric Features for Continuous Authentication of Smartphone Users*, IEEE Transactions on Information Forensics and Security, vol. 11(5)(2016) 877–892. https://doi.org/10.1109/TIFS.2015.2506542.

[14] E. Aguilar-Garnica, J. P. García-Sandoval, *Software sensors design and selection for the production of biodiesel from grease trap wastes*, Computer Aided Chemical Engineering, vol. 37(2015), pp. 1589–1594. https://doi.org/10.1016/B978-0-444-63577-8.50110-8.

[15] A. Das, N. Borisov and M. Caesar, *Exploring Ways To Mitigate Sensor-Based Smartphone Fingerprinting*, arXiv:1503.01874, Cornell university library (2015).

[16] D. Figo, P. C. Diniz, D. R. Ferreira and J. M. P. Cardoso, *Preprocessing techniques for context recognition from accelerometer data*, Personal and Ubiquitous Computing, vol. 14 (2010) 645–662. https://doi.org/10.1007/s00779-010-0293-9.

[17] S. Saha, R. Lahiri, A. Konar, B. Banerjee and A. K. Nagar, *HMM-based gesture recognition system using kinect sensor for improvised humancomputer interaction*, International Joint Conference on Neural Networks (IJCNN), 2017, pp. 2776–2783, Anchorage, AK, USA. https://doi.org/10.1109/IJCNN.2017.7966198.

[18] A. Dempster, N. Laird, and D. Rubin, *Maximum likelihood from incomplete data via the EM algorithm*, Journal of the royal statistical society. Series B (methodological) (1977) 1–38.

[19] L. Baum, T. Petrie, G. Soules, and N. Weiss, *A maximization technique occurring in the statistical analysis of probabilistic functions of markov chains*, The Annals of Mathematical Statistics (1970) 164–171. https://doi.org/10.1214/aoms/1177697196.

[20] R. Auckenthaler, M. Carey, and H. Lloyd-Thomas, *Score normalization for text-independent speaker verification systems*, Digital Signal Processing, vol. 10(1)(2000) 42–54. https://doi.org/10.1006/dspr.1999.0360.

[21] Cherifi, Prehensile movements when handling a smartphone. figshare. Dataset. (2020). https://doi.org/10.6084/m9.figshare.11855709.v1

[22] Natalia Neverova, *Deep Learning for Human Motion Analysis*, PhD thesis, UNIVERSITE DE LYON, 2016, France.

[23] M. Frank, R. Biedert, E. Ma, I. Martinovic, and D. Song, *Touchalytics: On the applicability of touchscreen input as a behavioral biometric for continuous authentication*, IEEE Trans. Inf. Forensics Security, vol. 8(1) (2013) 136–148.

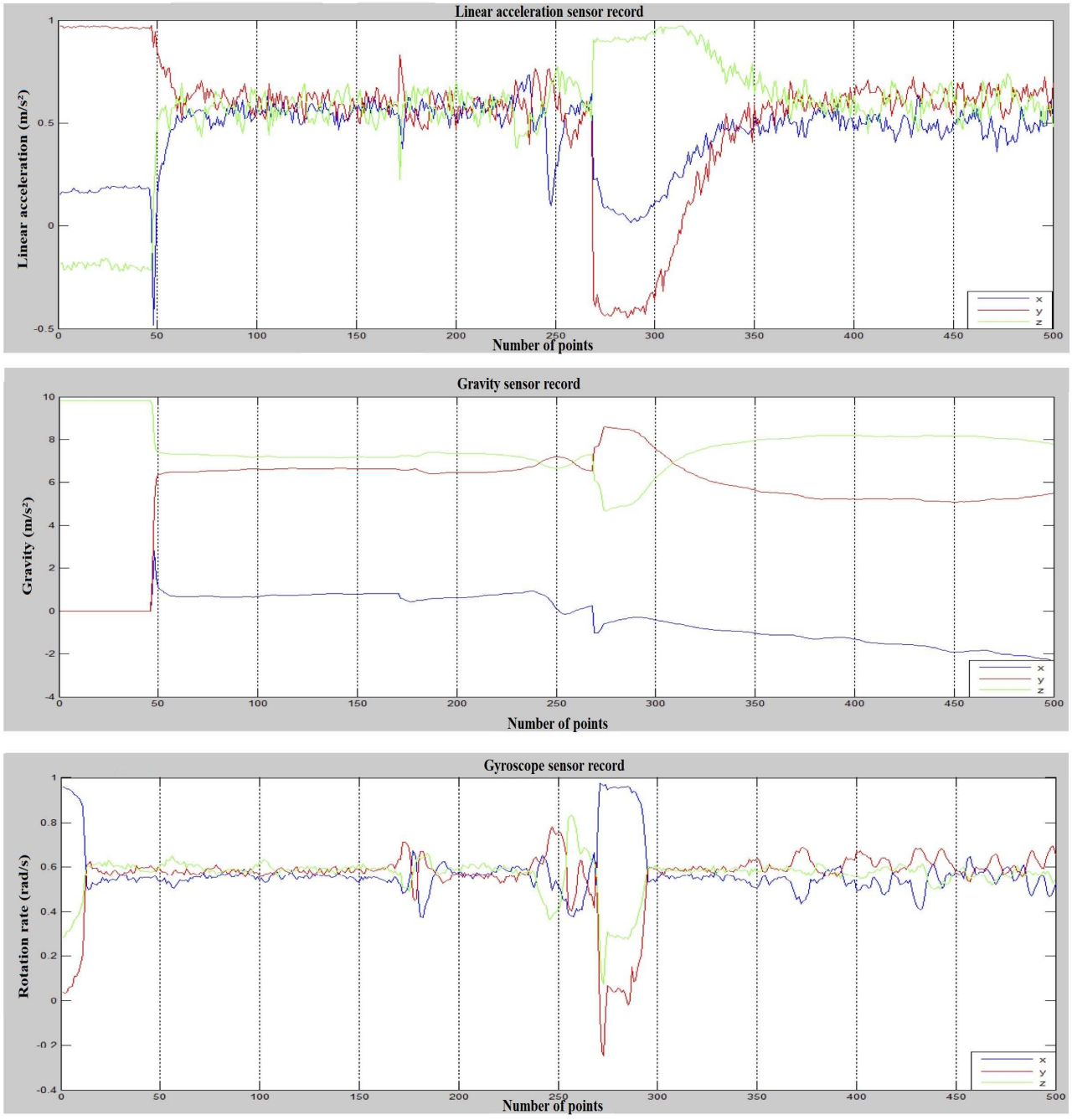


Figure 14: Sample data of 500 points with a window of 1s