Mobile Active Authentication based on Multiple Biometric and Behavioral Patterns

Alejandro Acien, Aythami Morales, Ruben Vera-Rodriguez, and Julian Fierrez

Abstract In this chapter we evaluate mobile active authentication based on an ensemble of biometrics and behavior-based profiling signals. We consider seven different data channels and their combination. Touch dynamics (touch gestures and keystroking), accelerometer, gyroscope, WiFi, GPS location and app usage are all collected during human-mobile interaction to authenticate the users. We evaluate two approaches: one-time authentication and active authentication. In one-time authentication, we employ the information of all channels available during one session. For active authentication we take advantage of mobile user behavior across multiple sessions by updating a confidence value of the authentication score. Our experiments are conducted on the semi-uncontrolled UMDAA-02 database. This database comprises of smartphone sensor signals acquired during natural human-mobile interaction. Our results show that different traits can be complementary in terms of mobile user authentication and multimodal systems clearly increase the performance when compared to individual biometrics systems with accuracies ranging from 82.2% to 98.0% depending on the authentication scenario.

Keywords Mobile authentication · Multimodal approaches · Behavioral patterns · Behavioral-based profiling · Touch dynamics

The present chapter is adapted from the conference paper A. Acien et al. "MultiLock: Mobile Active Authentication based on Multiple Biometric and Behavioral Patterns", in ACM Intl. Conf. on Multimedia, Workshop on Multimodal Understanding and Learning for Embodied Applications (MULEA), pp. 53–59, Nice, France, October 2019. The new material here includes Table I and Sect. 5.3.

A. Acien (✉) · A. Morales · R. Vera-Rodriguez · J. Fierrez
Biometrics and Data Pattern Analytics (BiDA) Lab, EPS, Universidad Autonoma de Madrid, Madrid, Spain
e-mail: alejandro.acien@uam.es; aythami.morales@uam.es; ruben.vera@uam.es; julian.fierrez@uam.es

T. Bourlai et al. (eds.), *Securing Social Identity in Mobile Platforms*, Advanced Sciences and Technologies for Security Applications,
https://doi.org/10.1007/978-3-030-39489-9_9

1 Introduction

Services are migrating from the physical to the digital domain in the information society. Examples can be found in e-government, banking, education, health, commerce, and leisure. This digital revolution is associated with a massive deployment of mobile devices, including multiple sensors (e.g. camera, gyroscope, GPS, touch screens, etc.), and full connectivity (e.g. bluetooth, WiFi, 4G, etc.). The mobile market has expanded to the point where the number of mobile devices in use is nearly equal to the world's population [1]. Mobile devices are rapidly becoming data hubs, used to store e-mail, personal photos, online history, passwords, and even payment information. Recent studies have shown that about 34% or more users do not use any form of authentication mechanism on their devices [2]. In similar studies, inconvenience is always shown to be one of the main reasons why users do not use any authentication mechanism. In [3], researchers show that mobile device users spent up to 9% of the time they use their smartphone on unlocking their screens, and the 2018 Meeker Report [4] indicated that the average smartphone user checks his/her device 150 times per day. Those factors lead individuals to make less security conscious decisions like leaving their smartphones unprotected or just protecting them using simple to break authentication mechanisms (e.g., simple Google unlock graphical patterns vulnerable to over-the-shoulder attacks [5]).

Biometric technologies improve traditional recognition technologies in several ways based on passwords or swipe patterns. The advantages of biometric systems are many in terms of security and convenience of use, which has led these technologies to take on a leading role in the last years. In fact, there is a growing interest in the biometrics research community towards more transparent and robust authentication methods that make use of the interaction signals originated when using smartphones [6, 7]. Signals generated with the sensors already embedded in mobile devices (e.g., gyroscope, magnetometer, accelerometer, GPS, and touchscreen interactions) along with metadata associated to our use of the technology (e.g. internet point access, browsing history, app usage) could assist in user authentication avoiding the inconveniences of traditional unlocking systems. All this information is originated naturally during the normal usage of the user with a smartphone, and it has been demonstrated that can be used for person identification under certain conditions [7]. By regularly conducting unobtrusive identity checks of the mobile user during a normal session, a continuous authentication system can verify if the device is still being operated by the authorized user. With this active system, if the mobile device is stolen, it should quickly recognize the presence of an unauthorized user.

The aim of this chapter is to analyze multi-modal approaches to improve the performance of mobile authentication. Our experiments include up to four different biometric traits (touch gestures, keystroking, gyroscope, and accelerometer) and three behavioral-based profiling techniques (GPS, WiFi, and app usage). The experiments are conducted on the UMDAA-02 mobile database [8], a challenging dataset acquired under natural conditions.

Previous works have demonstrated the potential of biometric and behavioral-based profiling patterns for user authentication under controlled scenarios. However, the performance of biometric mobile authentication based on human interaction raises doubt under challenging non-supervised scenarios. The contributions of this work are: (i) performance analysis of user authentication based on 4 biometric data channels (touch gestures, keystroking, accelerometer, and gyroscope) and 3 behavior profiling data sources (WiFi, GPS, and App usage), obtained during natural human-smartphone interaction; (ii) study of multimodal approaches for smartphone user authentication based on various combinations of the previous 7 data channels, both for One-Time Authentication and for Active Authentication schemes (i.e., continuously over multiple sessions). The results showed in this chapter suggest that user-profiling techniques can help to improve performances of behavioral-based biometrics authentication systems in all scenarios evaluated.

The rest of this chapter is organized as follows: Section 2 links the present works with related research. Section 3 describes the architecture of our approach. Section 4 explains the experimental protocol, describing the database and the experiments performed. Section 5 presents the final results for single and multimodal architecture and Sect. 6 summarizes the conclusions and future work.

2 Related Works

Mobile authentication based on soft biometrics traits has been extensively studied in the last years [9–11]. In Table 1 we summarize some of the most relevant state-of-the-art works in this field. Swipe dynamics is one of the most popular traits analyzed [9]; however, it has been shown not to have enough discriminative power to replace traditional technologies.

Accelerometer and gyroscope sensors have been studied traditionally for gait recognition, and some works have demonstrated also their utility for user authentication with acceptable performance [12].

Geo-location based verification approaches are scarce in the literature. In [13], Mahbub and Chellappa developed a mobile authentication system using trace histories by generating a confidence score of the new user location taking into account the sparseness of the geo-location data and past locations. For this purpose, they employed modified Hidden Markov Models (HMMs) considering the human mobility as a Markovian motion. In a similar way, in [14] a variation of HMMs was used to develop a user authentication mobile system by exploiting application usage data. They suggest that unforeseen events and unkno applications have more impact in the authentication performance than the most common apps used by the user. The potential of WiFi history data was analyzed in [10] for mobile authentication. They explored: (i) the WiFi networks detected by the smartphone, (ii) when the detection occurs, and (iii) how frequently those networks are detected during a period of time.

Regarding keystroke traits, in [11] a fixed-text keystroking system for mobile user authentication was studied using not only time and space based features (e.g.

Table 1 Summary of the state-of-the-art in biometric mobile authentication. The number of users for each database is in brackets

Study	Modality	Classifier	Database	Acc (%)
Fierrez et al. [9]	Touch gestures	SVM, UBM	Serwadda (190), Antal (71), Frank (41), UMDAA02 (48)	80–90
Li et al. [10]	Accelerometer, WiFi	Templates, random Forest	Prop. DB (321)	90 (3 s)
Busheck et al. [11]	Keystroke	KNN, SVM, NB, LSAD	Prop. DB (28)	64–74
Li et al. [12]	Accelerometer, gyroscope	Random Forest	Prop. DB (304)	77
Mahbub et al. [13]	GPS location	M-HMM	UMDAA02 (48), GeoLife (182)	69–79
Mahbub et al. [14]	App usage	M-HMM	UMDAA02 (48)	70–84
Monaco et al. [15]	Keystroke	POHMM, HMM, SVM	Mutiple Databases (247)	90
Shi et al. [18]	Voice, GPS, touch, gait	NB	Prop. DB	90
Fridman et al. [19]	Stylometry, app usage, web browsing, GPS	Binary classifiers, SVM	Prop. DB (200)	95 (3 s)
Liu et al. [20]	Touch gestures, power consumption, accelerometer, gyroscope, magnetometer	SVM	Prop. DB (10)	95
Li et al. [21]	WiFi, Bluetooth, accelerometer, gyroscope	Random Forest	Prop. DB (321)	90 (3 s)
Deb et al. [22]	Keystroke, GPS, accelerometer, gyroscope, magnetometer, accelerometer, gravity, rotation sensors	Siamese LSTM	Prop. DB (37)	97 (3 s)
Our work	Touch gestures, keystroke, accelerometer, gyroscope, WiFi, GPS, app usage	SVM, templates	UMDAA02 (48)	98 (4 sessions)

hold and flight times, jump angle or drag distance) but also studying the hands postures during typing as discriminative information. In [15], a novel fixed-text authentication system for laptops and mobile devices based on Partially Observable HMMs was studied. This model is an extension of HMMs, in which the hidden state is conditioned on an independent Markov chain. The algorithm is motivated by the idea that typing events depend both on past events and also on a separate process.

Finally, building a multimodal system that integrates all these heterogeneous information sources for mobile user authentication is still a challenge [16]. Noisy data, intra class variation or spoofing attacks [17] are some inevitable problems in unimodal systems that can be overcome by multimodal architectures [7, 16]. In [18], a multimodal user authentication system was based on the fusion at decision level of voice, location, multi-touch, and accelerometer data. Their preliminary results suggest that these four modalities are suitable for continuous authentication. In [19], a fusion was performed also at decision level of behavioral-based profiling signals such as web browsing, application usage, and GPS location with keystroking data achieving 95% of user authentication accuracy using information from one-minute window.

More recently, in [20] a fusion also at decision level of touch dynamics, power consumption, and physical movements modalities achieved 94.5% of accuracy with a dataset that was captured under supervised conditions. In [21], an unobtrusive mobile authentication application is designed for single and multimodal approaches. They collected data from WiFi, Bluetooth, accelerometer, and gyroscope sources in unsupervised conditions and fused them at score level achieving up to 90% of accuracy in the best scenario. In [22], they propose a Siamese Long ShortTerm Memory network architecture to merge up to 8 modalities (keystroke dynamics, GPS location, accelerometer, gyroscope, magnetometer, linear accelerometer, gravity, and rotation sensors) for mobile authentication, achieving 97.15% of accuracy using data from a 3 s window for each of the modalities considered individually.

Previous works fusing different modalities ([19, 21, 22]) have focused their approach on obtaining time windows from the different modalities and then carry out the fusion. However, this does not represent a realistic scenario due to not all modalities fused can always be captured in a specific time windows. In this work we go a step forward by merging the modalities at session level (time during an unlock and the next lock of the device), and therefore fusing only the modalities available at each session.

3 System Description

In this chapter we analyze 4 biometric data channels (touch gestures, keystroking, gyroscope, and accelerometer) and 3 behavior data sources (GPS, WiFi, and app usage). We study 2 approaches for user authentication (see Fig. 1):

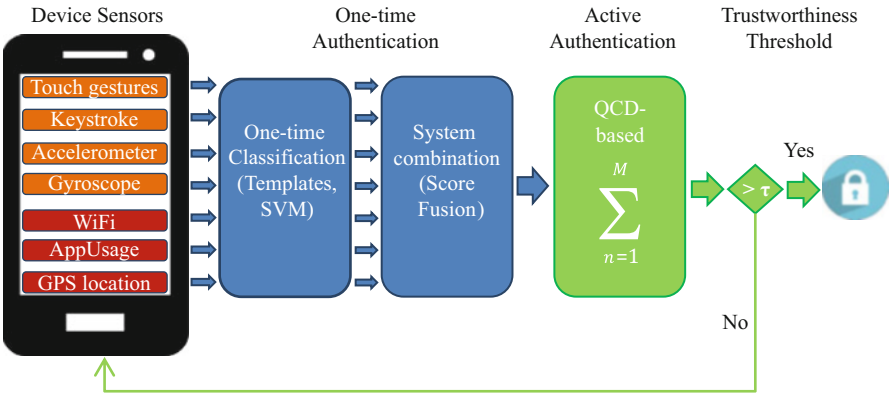

Fig. 1 System architecture. Blue boxes correspond to one-time authentication, and green boxes are add-on modules for active authentication

- The first approach (blue boxes in Fig. 1), referred to as One-Time Authentication (OTA) is based on unimodal systems trained with the information extracted from the mobile sensors during a user session. Remember that a session is defined as the elapsed period between the device unlock and the next lock. Therefore, sessions have a variable duration and information obtained from sensors varies depending on the usage of the device during the session. The information provided by the sensors is employed to model the user according to the seven systems mentioned before: keystroking, touch gestures, accelerometer, gyroscope, WiFi, app usage, and GPS location. Each system provides a single authentication score and these scores are combined to generate a unique score for each session.
- The second approach, called Active Authentication (green boxes in Fig. 1), is based on updating a confidence value generated from the One-Time Authentication during consecutive sessions.

The seven systems are categorized into two main groups according to the nature of the information employed to model the user: biometric and behavior-based profiling systems. In this work, biometric systems refer to the top 4 channels in the Sensors Data module of Fig. 1 (red boxes). The way we carry out touch gestures, typing, or handle the device is determined by behavioral aspects (e.g. emotional state, attention) and neuromotor characteristics of users (e.g. ergonomic, muscles activation/deactivation timing, motor abilities). Behavioral-based profiling refers to those systems that identify the owners of the device according to the services they use during their daily habits (orange boxes in Fig. 1, bottom 3 channels in the Sensors Data module).

Table 2 Example of an app-usage user template generated according the data captured during 6 days

Event	Time slot	Frequency
WhatsApp	4	5
Navigator	4	3
YouTube	5	1
WhatsApp	5	1
Facebook	7	2

3.1 Behavioral-Based Profiling Systems

WiFi, app usage, and GPS location system are based on a similar template-based matching algorithm. A user template is defined as a table containing the time stamps and the frequency of the events [10]. For this, we divided the time (24 h of the day) into N equal time slots (e.g. if we choose N = 48 we will have 48 time slots of 30 min), giving to each time slot a number ID. Then, we store in the template the event's name, the number ID of the time slot and the occurrence frequency of that event (number of times this event occurs during this particular time slot on a window of consecutive days). Table 2 shows an example of the app-usage template for a given user generated according the data obtained during 6 days; in this case the WhatsApp application, for instance, is detected in the fourth time slot during five out of six total days considered, meanwhile the same app is detected only one out of 6 days in the fifth time slot. Note that multiple detections of the same event in the same time slot and day are ignored, but they are stored if they belong to different time slots or days. Depending on the system, the event could be the name of the WiFi network, latitude and longitude of a location (with two decimals of accuracy), or the name of a mobile app for WiFi, GPS location, and app usage systems, respectively.

Finally, we test the systems by calculating a behavior-based confidence score [10] for each test session as:

$$score = \sum_{i=1}^{S} f_i^2 \tag{1}$$

where *fi* is the frequency of the event stored in the template that match with the test event i in the same time slot and S is the total number of events detected in that test session. For example, if the test session includes the usage of *WhatsApp* and *Navigator* apps during the fourth slot, the score confidence will be $52 + 32 = 34$ (according to the template showed in Table 1). Based on this explanation, a higher score in the test session implies higher confidence for authentication.

3.2 Biometric Systems

For touch gestures, keystroking, accelerometer and gyroscope systems, the feature extraction and classification algorithms are adapted to model the user information.

In the touch gestures system, the feature set employed is a reduced set of the global features presented in [23] (commonly used for online handwriting sequence modeling) and adapted for swipe biometrics in [7]. Mean velocity, max acceleration, distance between adjacent points, or total duration are some examples of this subset of 28 features extracted (see [23] for details).

For accelerometer and gyroscope, the data captured are comprised of the x, y, and z coordinates of the inclination vector of the device (gyroscope) and the acceleration vector (accelerometer) in each time stamp. For these 2 sensors we use the feature set proposed in [12]: mean, median, maximum, minimum, distance between maximum and minimum, and the standard deviation for each array of coordinates. Moreover, we propose the 1 and 99 percentiles[1] and the distance between them as additional features.

Regarding keystrok dynamics, the keys pressed were encrypted in order to ensure users' privacy. Thus, systems based on graphs were discarded and we adopted traditional timing features: hold time, press-press latency, and press-release latency as in [24, 25]. Finally, we propose a feature set based on six statics (mean, median, standard deviation, 1 percentile, 99 percentile, and 99-1 percentile). Note that UMDAA-02 keystroke data can be considered as a free text scenario. However, the limited samplesper session and the encrypted keys difficult the application of popular free-text keystroke authentication methods.

For classification we train different Support Vector Machines (SVM) with a radial basis function (RBF) kernel, one for each feature set and user with an optimization of both hyperparameters (C, σ).

4 Experiments

4.1 Database

The experiments were conducted with UMDAA-02 database [8]. This database comprises 141.14 GB of smartphone sensor signals collected from 48 Maryland University students over a period of 2 months. The participants used a smartphone provided by the researchers as their primary device during their daily life (unsupervised scenario). The sensors captured are touchscreen (i.e. touch gestures and keystroking), gyroscope, accelerometer, magnetometer, light sensor, GPS, and WiFi, among others. Information related to mobile user's behavior such as lock and unlock time events, start and end time stamps of calls and app usage are also stored. Table 3 summarizes the characteristics of the database. During a session, the data collection application stored the information provided by the sensors in use.

[1] Indicate the value below which a given percentage of observation (samples in this case) in a group of observation falls.

Table 3 General UMDAA-02 dataset information

Description	Statistics
Gender	36 M/12F
Age	22–31 years
Avg. days/user	10 days
Avg. sessions/user	248 sessions
Avg. time/session	224 s
Avg. systems/session	5.2 systems[a]

[a]Systems: refers to the number of systems available out of the 7 studied in this work

4.2 Experimental Protocol

The experiments are divided into two different scenarios: One-Time Authentication (OTA) and Active Authentication (AA). In OTA the performance is calculated using only one session to authenticate the user meanwhile in AA we employ multiple consecutive sessions in order to improve the confidence in the authentication. For all experiments the dataset is divided into 60% days for training (first sessions) and the remaining 40% days for testing in order to have train and test sets as more balanced as posible. This means that we employ 6 days in average to model the user and 4 days in average to test such a model. The performance for both scenarios is presented in terms of average correct classification rate computed as $100 - EER$ (Equal Error Rate).[2]

4.2.1 One-Time Authentication

In OTA experiments, all 7 systems are trained separately for each user and the scores are calculated at session level, generating 7 scores for each test session as maximum (note that the number of systems available during a session varies). The 4 biometric systems considered can produce more than one score per session (e.g. multiple gestures or multiple keystroking sequences during a text chat). In those cases, the scores available during the session are averaged to obtain one score for each biometric system and session. Finally, we normalize with *tan* normalization and fuse the scores (mean rule) to calculate a single score [14] according to the different fusion set-ups proposed. The scores from the best fusion set-up will be used in the AA scenario (details are provided in Sect. 4.2.2 below).

[2]EER refers to the value where False Acceptance Rate (percentage of impostors classified as genuine) and False Rejection Rate (percentage of genuine users classified as impostors) are equal.

4.2.2 Active Authentication

For AA experiments we consider the QCD algorithm (Quickest Change Detection) as explained in [26]. The QCD-based algorithm updates a confidence score based on previous events (sessions in this work) by performing a cumulative sum of scores. This cumulative sum will be almost zero if the scores belong to the genuine user, and will grow if an impostor takes the control, until it reaches a certain threshold that would detect the intruder. The cumulative sum is calculated as follow:

$$score_j^{AA} = \max\left(score_{j-1}^{AA} + L_j, 0\right) \tag{2}$$

where j means the actual session and $score_{j-1}^{AA}$ is the previous cumulative score. L_j is the contribution of the actual session calculated as the log-likelihood ratio between score distributions:

$$L_j = \log\left(\frac{f_I\left(score_j\right)}{f_G\left(score_j\right)}\right) \tag{3}$$

where f_G and f_I are the probability distributions of the genuine and impostor scores respectively calculated previously in the OTA fusion scenario, and *scorej* is the OTA fused score of the actual session. According to (3), the log-likelihood ratio L_j will be negative if $score_j$ belongs to a genuine user and positive in the opposite case and, therefore, multiple consecutive sessions of an impostor in control will increase the cumulative sum ($score_j^{AA}$). Figure 2 depicts an example of $score_j^{AA}$ evolution. At the time the mobile starts to be operated by an intruder (session number sixteen in Fig. 2) the $score_j^{AA}(j > 16)$ will tend to increase until reaching the threshold. The time elapsed between the intrusion start and the intrusion detection is known as Detection Delay (DD) measured in number of sessions.

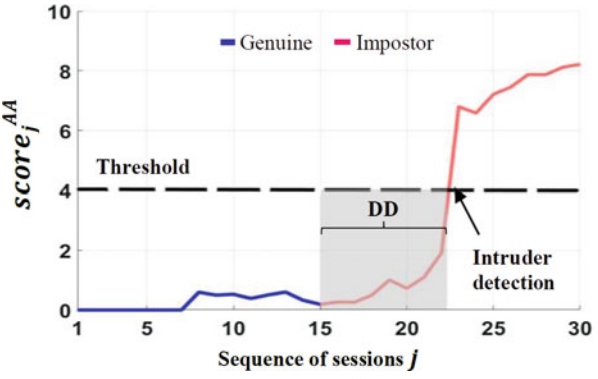

Fig. 2 An example of QCD-based curve with a sequence of 30 sessions (15 genuine and 15 impostors). The dashed line is the intruder detection threshold and the grey box shows the Detection Delay (DD)

5 Results and Discussion

5.1 One-Time Authentication

In this section we analyze the OTA scenario: the accuracy for the 4 biometric systems and the fusion with each behavior-based profiling system. Table 4 summarizes the final results by ranking from the best individual biometric system performance to the worst one. The first column shows the performance obtained for each single biometric system. From the second to the fourth column, we show the performance for the fusion of each biometric system with each behavior-based profiling system, and the fifth column shows the fusion with all of them. Firstly, the poor performance achieved by some biometric systems can be caused by the uncontrolled acquisition conditions of the database and the limited number of samples per session (e.g. free text keystroke usually requires large sequences) but the combination of all of them (last row in Table 4) shows acceptable performance for unsupervised scenarios. Secondly, we can observe that behavior-based profiling systems always improve biometric systems performances in all fusion schemes. In fact, the combination of all behavior-based profiling approaches with each biometric system achieves the most competitive performance, improving them in more than 18% of accuracy in the best case. If we analyze each single behavior-based profiling fusion, we can observe that the GPS system achieves the best improvements, boosting biometric systems performances in more than 13% of accuracy. Finally, in Fig. 3 we plot the ROC curves for each single biometric system and the best fusion set-up, i.e. the fusion of all behavior-based profiling systems with each biometric system (column 5 in Table 4). The results in OTA scenario suggest that behavior-based profiling systems always improve the biometric ones and the best performance is achieved by fusing with all of them, and therefore, the scores obtained from this fusion scheme will be used in the AA scenario.

Table 4 Results achieved for both One-Time and Active Authentication (AA) scenarios in terms of correct classification rate (%) according to different number of biometric systems and their fusion with behavior-based profiling systems. In brackets, average number of sessions employed (ADD)

System	Acc.	+WiFi	+ GPS	+ App usage	All	AA
Touch gestures	72.0	78.2	78.3	75.4	**83.1**	95.0 (6)
Keystroking	62.5	72.6	70.9	67.8	**79.1**	92.9 (7)
Accelerometer	61.3	70.8	77.3	64.7	**78.7**	93.7 (7)
Gyroscope	59.5	69.7	72.6	63.4	**78.4**	92.3 (6)
Combined	73.2	77.3	78.9	75.3	**82.2**	**97.1 (5)**

Bold indicates the best accuracy achieved for each system

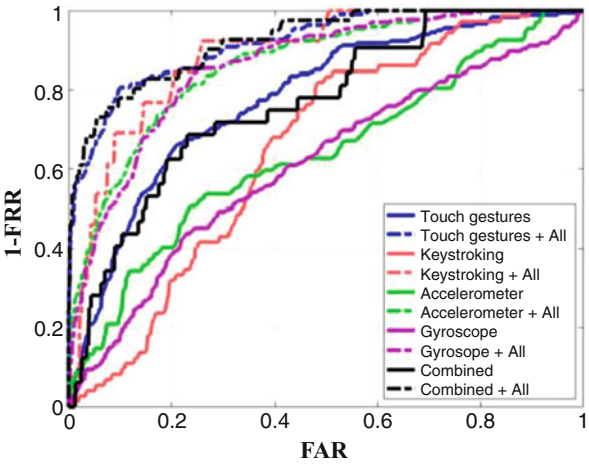

Fig. 3 ROC curves (One-Time Authentication) for individual biometrics and the best fusion set-up incorporating the 3 considered behaviour profiling sources (All = WiFi + GPS + App usage)

5.2 Active Authentication

Even performance metrics used for Active Authentication and One-time Authentication can be similar, we want to highlight some important differences:

- *Probability of False Detections (PFD):* is the percentage of genuine users detected as intruder during a sequence of genuine sessions. It means that $score_j^{AA}$ reaches the intruder detection threshold during a genuine session sequence (genuine curve in Fig. 2). PFD is similar to FMR (False Match Rate) in One-Time Authentication.
- *Probability of Non-Detection (PND):* is the percentage of intruders not detected during a sequence of intruder sessions. It means that $score_j^{AA}$ does not reach the intruder detection threshold during the intruder sessions sequence (impostor curve in Fig. 2). PND is similar to FNMR (False Non-Match rate) in One-Time Authentication.
- *Average Detection Delay (ADD):* is the average number of impostor sessions needed to detect an intruder (the grey box in Fig. 2).

To calculate the correct classification rate in AA we plot in Fig. 4 the PND vs. PFD and ADD vs. PFD curves. The PND-PFD curves are similar to FMR-FNMR curve in one-time authentication with the main difference that those results are obtained from a sequence of stacked scores instead of only one. The equal error rate (EER) will be the value where PND and PFD are equal and the correct classification rate will be computed as $100 - EER$. The ADD-PFD curve shows the number of sessions needed to detect an intruder according to the PFD. This curve allows us to know how many sessions are needed to achieve the EER reported. For instance, the PND-PFD curves in Fig. 4 (right) show that the EER in Active Authentication is 2.9% and the ADD to achieve that EER is 5 sessions. This means that we can

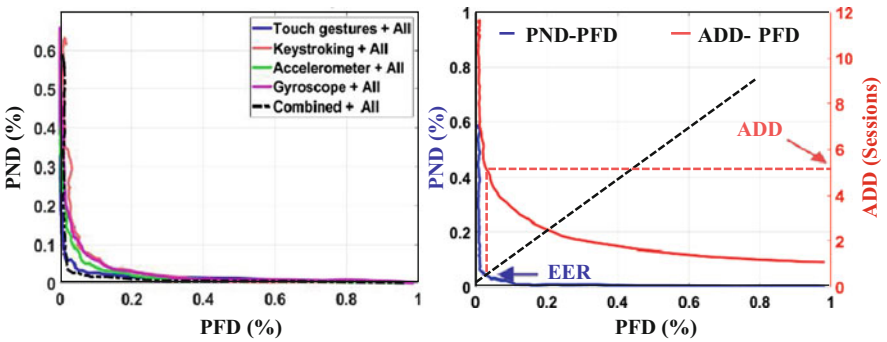

Fig. 4 PND vs PFD curves of active authentication for the best fusion schemes (left), PND vs PFD and ADD vs PFD curves for the best fusion set-up (right). The dark dashed line shows the EER and the red one shows the Average Detection Delay for that EER in the right plot

improve OTA results at the cost of having more sessions to detect an intruder. All curves were calculated for each user and averaged.

Finally, all AA results are summarized in the last column of Table 4. Remember that scores employed in the QCD-based algorithm come from the fusion scores of the best OTA scenario (fusing with all behavior-based profiling systems) so both performances are correlated. Each performance in Table 4 for AA is followed by the average detection delay in brackets needed to achieve it. As we expected, in all different fusion set-ups the AA algorithm improves the accuracy at the cost of needing more sessions to detect the intruder. In fact, for the best fusion set-up the performance improves from 82.2% to 97.1% by using 5 consecutive intruder sessions to detect the impostor. Comparing all scenarios, the greatest improvement occurs with all biometric systems combined (14.9% of improvement in the last row of Table 4) with an average 5 sessions.

The cost of need up to 5 sessions to detect an intruder could be unacceptable in some real life scenarios (e.g. prevent unauthorized use of devices during distractions). However, as some recently surveys suggest [27], the market of secondhand mobile phone is constantly growing and some of these devices have a provenance of dubious legality. In these scenarios Active Authentication approaches can serve to persuade burglars and unathorized usages.

5.3 Temporal Dependency in Behavioral-Based Profiling Systems

As mentioned in [10], the performance of behavioral-based profiling systems could be affected by differences in our routines in our daily life. For example, the places we visit during the week could vary at weekends or the WiFi signals detected are different if we are at work (working hours) or at home (leisure hours). In order to

Table 5 Results achieved for behavioral-based profiling systems and the Combined + All fusion scenario according to the temporally division in week/weekend and working/leisure time

Profiling System	Acc.	Week	Weekend	Working time	Leisure time
WiFi	77.5	77.1	77.9	74.4	**85.0**
GPS	75.4	74.0	80.1	70.1	**83.7**
App usage	67.4	67.6	69.2	66.2	**69.7**
Combined + all	82.2	81.6	82.0	82.1	**86.7**

study these assumptions, we divide the score sessions of the OTA scenario in two groups depending on when the session was performed: week time (from Monday to Friday) or weekend (Saturday and Sunday), and working time (from 9 a.m to 6 p.m from Monday to Friday) or leisure time (the remaining hours for all days of the week). The results are shown in Table 5 for all behavioral-based profiling systems separately (i.e. using only the scores of profiling systems in each sessions) and the best fusion scenario of OTA (Combined + All).

First of all, the results of week/weekend division suggest that only GPS system improves their performance by using the scores of the weekend sessions due to at weekends the users move to more locations than during the week, making the GPS system more discriminant at weekends. Secondly, if we divide the sessions according to whether they belong to working time or leisure time we find out that all behavioral-based profiling systems improve their performance in leisure hours, specially GPS and WiFi systems. The invariance across daily hours of App usage system suggest that users usually employ almost the same mobile applications at work as they do at home.

It is worth noting that these conclusions relate heavily on the nature of the database considered, where most of the users are students of the same university, who therefore share similar location patterns during weekdays and working hours. On a much broader database we would probably expect to have different trends, maybe having better results for working hours where probably users follow a more constant location pattern.

Finally, the last row in Table 5 shows the variation of the best fusion set-up in OTA scenario according to the proposed time division. As we expected, the best improvement is achieved during leisure time (see Fig. 5 for details) due to the improvement of the behavioral-based profiling systems, raising the accuracy up to 86.7%.

Regarding AA scenario, if we employ only the sessions from leisure time in the best fusion OTA set-up (combined + All), we achieve an accuracy of 98.0% with 4 sessions, improving the best result of Table 4, where we achieved a 97.1% of accuracy for AA with 5 sessions, but considering data from all hours and not only leisure time.

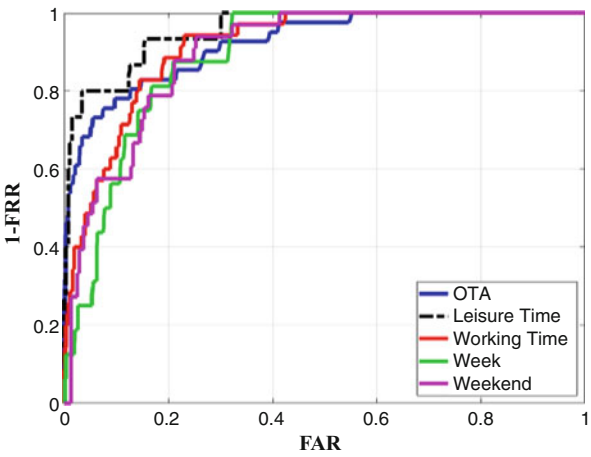

Fig. 5 ROC curves for the best OTA scenario (combined + ALL) according to the proposed time division: week/weekend, working/leisure time

6 Conclusions and Future Work

In this chapter, we have studied user mobile active authentication based on multiple biometric and behavior-based profiling systems. For this, we studied two scenarios according to the number of sessions used: one session (One-Time Authentication) and multiple sessions (Active Authentication). The results suggest that some swipe and keystroking modalities work better than accelerometer and gyroscope in the scenarios evaluated in this work. The fusion with behavior-based profiling systems improves the results of single biometric modalities, achieving accuracies up to 82.2% in the best case for an OTA scenario. Our experiments also suggest that Active Authentication improves the accuracy of One-time Authentication scenario with up to 14% of enhancement using information from 5 sessions. As we mentioned in the section before, Active Authentication algorithms are useful in those scenarios where the intruder attempt to use the mobile phone during mid-term periods (e.g. to use it as their personal device, reselling in the second-hand mobilephone market, etc). According to this, a continuous usage of the stolen mobilephone in which One-Time verification system has already been hacked and intruders has no limitations regarding device's usage. Active Authentication continuously monitorizes and check the identity of users. These approaches can serve to persuade robberies and unauthorized usages.

For future works we will work to improve the performance of individual systems, especially biometrics systems. Better individual performances will produce better fused schemes. The combination of heterogeneous data at data and feature level will be evaluated in order to merge correlations between systems (e.g. touch gestures and apps used are highly correlated).

Regarding the temporary dependency in behavioral-based profiling systems, note that all participants of UMDAA02 database are students from the same university

so probably some of them share work places and leisure activities, and therefore, these variations reported could be greater in other mobile databases with users from different places and habits.

Acknowledgments This work was funded by the projects BIBECA (RTI2018-101248-B-I00 MINECO/FEDER) and Bio-Guard (Ayudas Fundación BBVA a Equipos de Investigación Científica 2017), and by CECABANK.

References

1. Radicati S (2018) Mobile statistics report, 2014–2018. The Radicati Group, INC. A Techonology Market Research Firm, Palo Alto
2. Cho G, Huh JH, Cho J, Oh S, Song Y, Kim H (2017) SysPal: system-guided pattern locks for android. In: Proceedings of IEEE Symposium on Security and Privacy, California, UE
3. Harbach M, von Zezschwitz E, Fichtner A, Luca AD, Smith M (2014) It's a hard lock life: a field study of smartphone (un)locking behavior and risk perception. In: Proceedings of symposium on usable privacy and security, California, USA
4. Molla R (2018) Mary Meeker's 2018 internet trends report: all the slides, plus analysis. In Recode
5. Martinez-Diaz M, Fierrez J, Galbally J (2016) Graphical password-based user authentication with free-form doodles. IEEE Trans Human-Machine Syst 46(4):607–661
6. Crouse D, Han H, Chandra D, Barbello B, Jain AK (2015) Continuous authentication of Mobile user: fusion of face image and inertial measurement unit data. In: Proceedings of IAPR international conference on biometrics, Phuket, Thailand
7. Patel VM, Chellappa R, Chandra D, Barbello B (2016) Continuous user authentication on mobile devices: recent progress and remaining challenges. IEEE Signal Process Mag 33(4):49–61
8. Mahbub U, Sarkar S, Patel VM, Chellappa R (2016) Active user authentication for smartphones: a challenge data set and benchmark results. In: Proceedings of IEEE 8th international conference on biometrics theory, applications and systems, New York, USA
9. Fierrez J, Pozo A, Martinez-Diaz M, Galbally J, Morales A (2018) Benchmarking touchscreen biometrics for Mobile authentication. IEEE Trans Inf Forensics Sec 13(11):2720–2733
10. G. Li and P. Bours (2018). Studying WiFi and accelerometer data based authentication method on mobile phones. In: Proceedings of 2nd international conference on biometric engineering and applications, Amsterdam, Netherlands
11. Buschek D, De Luca A, Alt F (2015) Improving accuracy, applicability and usability of keystroke biometrics on mobile touchscreen devices. In: Proceedings of 33rd annual ACM conference on human factors in computing systems, Seoul, Republic of Korea
12. Li G, Bours P (2018) A novel mobilephone application authentication approach based on accelerometer and gyroscope data. In: Proceedings of 17th international conference of the biometrics specials interest group, Fraunhofer, Germany
13. Mahbub U, Chellappa R (2016) PATH: person authentication using trace histories. In: Proceedings of ubiquitous computing, electronics & mobile communication conference, IEEE, New York, USA
14. Mahbub U, Komulainen J, Ferreira D, Chellappa R (2018) Continuous authentication of smartphones based on application usage. IEEE Transactions on Biometrics, Behavior, and Identity Science 1(3):165–180
15. Monaco JV, Tappert CC (2018) The partially observable hidden Markov model and its application to keystroke dynamics. Pattern Recogn 76:449–462

16. Fierrez J, Morales A, Vera-Rodriguez R, Camacho D (2018) Multiple classifiers in biometrics. Part 2: trends and challenges. Inf Fusion 44:103–112
17. Marcel S, Nixon MS, Fierrez J, Evans N (2019) Handbook of biometric anti-spoofing. presentation attack detection, Advances in computer vision and pattern recognition. Springer, Cham
18. Shi W, Yang J, Jiang Y, Yang F, Xiong Y (2011) Senguard: passive user identification on smartphones using multiple sensors. In: Proceedings of 7th IEEE international conference on wireless and mobile computing, networking and communications, Shangai, China, pp 141–148
19. Fridman L, Weber S, Greenstadt R, Kam M (2015) Active authentication on mobile devices via stylometry, GPS location, web browsing behavior, and application usage patterns. IEEE Syst J 11(2):513–521
20. Liu X, Shen C, Chen Y (2018) Multi-source interactive behavior analysis for continuous user authentication on smartphones. In: Proceedings of Chinese conference on biometric recognition, Urumchi, China
21. Li G, Bours P (2018) A mobile app authentication approach by fusing the scores from multi-modal data. In: Proceedings of 21st international conference on information fusion, Cambridge, UK
22. Deb D, Ross A, Jain AK, Prakah-Asante K,Prasad KV (2019) Actions Speak Louder Than (Pass) words: Passive Authentication of Smartphone Users via Deep Temporal Features. In: Proceedings of the 12th IAPR International Conference on Biometrics, Crete, Greece
23. Martinez-Diaz M, Fierrez J, Krish RP, Galbally J (2014) Mobile signature verification: feature robustness and performance comparison. IET Biometrics 3(4):267–277
24. O'Neal M, Balagani K, Phoha V, Rosenberg A, Serwadda A, Karim ME (2016) Context-aware active authentication using touch gestures, typing patterns and body movement. Louisiana Tech University, Ruston
25. Morales JF, Tolosana R, Ortega-Garcia J, Galbally J, Gomez-Barrero M, Anjos A (2016) Keystroke biometrics ongoing competition. IEEE Access 4:7736–7746
26. Perera P, Patel VM (2018) Efficient and low latency detection of intruders in mobile active authentication. IEEE Trans Inf Forensics Secur 13(6):1392–1405
27. Ernst R (2019) Mobile phone afterlife – why the second-hand market will be all the rage in 2019. In RCR Wireless News