Exploring a Statistical Method for Touchscreen Swipe Biometrics

Ada Pozo*, Julian Fierrez*, Marcos Martinez-Diaz*, Javier Galbally†, Aythami Morales*
* BiDA Lab, Universidad Autonoma de Madrid, SPAIN
† European Commission, DG Joint Research Centre, Ispra, ITALY
{julian.fierrez, aythami.morales}@uam.es

***Abstract*—The great popularity of smartphones and the increase in their use in everyday applications has led to sensitive information being carried in them, such as our bank account details, passwords or emails. Motivated by the limited security of traditional systems (e.g. PIN codes, secret patterns), that can be easily broken, this work focuses on the analysis of users normal interaction with touchscreens as a means for active authentication. Given the frequency in which touch operations are performed, characteristic habits, like the strength, rhythm or angle used result in discriminative patterns that can be exploited to authenticate users. In the present work, we explore a statistical approach based on adapted Gaussian Mixture Models. The performance across different kinds of touch operations, reveals that some gestures hold more user-specific information and are more discriminative than others (in particular, horizontal swipes appear to be more discriminative than vertical ones). The experimental results show that touch biometrics have enough discriminability for person recognition and that they are a promising method for active authentication.**

***Index Terms*—Active authentication, biometrics, smartphone, touchscreen, human computer interaction**

I. Introduction

Traditionally, the methods used for authentication on mobile devices have been based on passwords, PIN codes or secret patterns. However, it has been proven that these methods have different problems [1]–[5], of which inconvenience is one of the most remarkable. Because of the need to authenticate each time the device is used, there is a tendency to avoid authentication measures, or to use short and weak passwords and PIN codes, because they are easier to remember and can be entered faster. Additionally, smudge attacks are capable of following the residues left on the device's screen when entering the same pattern repeatedly, thus gaining access to the device and showing that secret patterns are not secure enough. These kind of authentication methods are known as entry-point methods, in which you only authenticate once, when unlocking the device, and this authentication is not performed again until the device is once more locked. Therefore, one cannot detect intruders if the screen is left unlocked or it is not possible to know if the person who is using the phone is the same user who authenticated in the first place.

To overcome these problems, new approaches known as continuous or active authentication methods have arisen. In these systems users are authenticated periodically in the background by passively analysing their biometrics. The usage patterns are studied and compared to the stored templates of the legitimate user, blocking the device if there are not enough coincidences [3], [6].

One method for continuous authentication whose results are only preliminary are touch biometrics on mobile devices. It is based on the idea that every person has characteristic habits and behaves differently from others when swiping the fingers on a touchscreen. As was proven in previous works [2], [3], [5], swiping patterns present high inter-class variability, that is, touch data from different users show great differences and thus can be discriminative among them. However, they also present high intra-class variability, which means that they are not stable biometrics, and hence, may change depending on the user's emotional state or with time, resulting in different patterns of use. Therefore, modelling a subject can be difficult.

In this work, continuous authentication with touch biometrics is investigated. Despite previous works studying and assessing their suitability for authentication, due to the difficulty to model users' touch-behaviour the results are yet preliminary. This work contribution lies on the study of a statistical based approach [7], using Gaussian Mixture Models (GMM) with Universal Background Model (UBM) adaptation [8] for personal authentication. For that purpose, this system models users behaviour and how their data is distributed. Additionally, a comparison between the distinctiveness of the different touch operations is made.

This paper is organized as follows. Section II summarizes related works in touch biometrics. Section III describes the database and Section IV the system architecture, the features used, and the statistical recognition approach evaluated. Experimental results are given in Section V and conclusions are drawn in Section VI.

II. Related works

Touch biometrics studies have mainly considered two approaches so far. The first one includes authentication methods based on touch gestures at an entry point, analysing users' touch behaviour exploiting a set of predefined gestures, for example, the secret pattern on the unlock screen. The second one is continuous authentication while the user freely interacts with the device performing different tasks [1], [3], [5], [6].

Several studies have investigated whether touch data is discriminative and stable enough for authenticating users. In Frank *et al.* [3], we can see one of the earliest yet more comprehensive works on continuous authentication using

978-1-5386-1585-0/17/$31.00 ©2017 IEEE

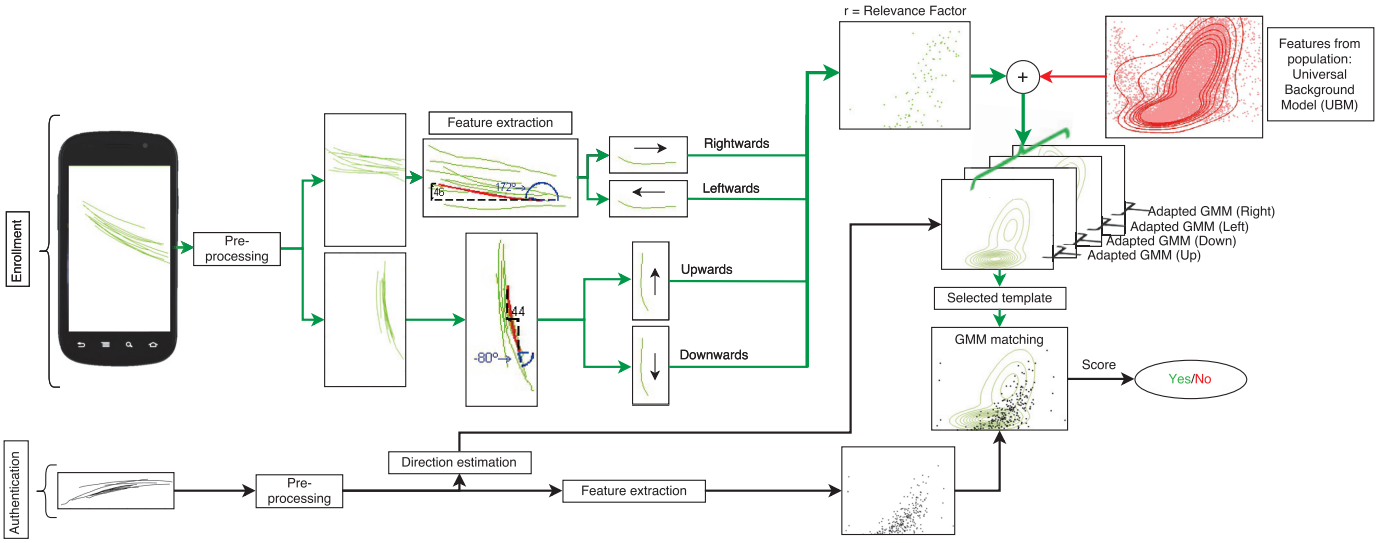


Fig. 1: System architecture

touch data. Using a database consisting of 41 users and a 27-dimensional feature vector, Support Vector Machines with a RBF-kernel (SVM) and k-Nearest Neighbors (kNNs) classifiers were used for classification, resulting in a performance with Equal Error Rate (EER) between 2% and 4% for eleven consecutive strokes and 13% for one single stroke, proving that using blocks of strokes resulted in a better performance than a single stroke, which was also observed in other articles such as [1], [5].

Single touch operations were also studied in Serwadda *et al.* [1] and Shen *et al.* [5]. In Serwadda *et al.* [1] a benchmark of ten different algorithms is analysed, concluding that among them the better suited for continuous authentication of touch operations are SVM with RBF-kernel, random forest and logistic regression, all reporting 10% to 20% EER. In Shen *et al.* [5], less than 5% EER is obtained when over ten strokes are combined with SVM, kNNs, neural networks and random forest classifiers. Additionally, the performance with specific tasks is compared to the one obtained with free tasks, concluding that specific tasks are more discriminative.

In addition to focusing on single touch operations, a fusion of biometrics is studied in Xu *et al.* [2] (keystroke, slide, and pinch) and Kumar *et al.* [9] (typing, swiping, and phone movement). In the first one, an EER below 10% is obtained over an SVM classifier with an RBF kernel, whereas in the second one kNNs with Euclidean distance and random forest are used to classify, obtaining accuracies up to 90%.

III. Database

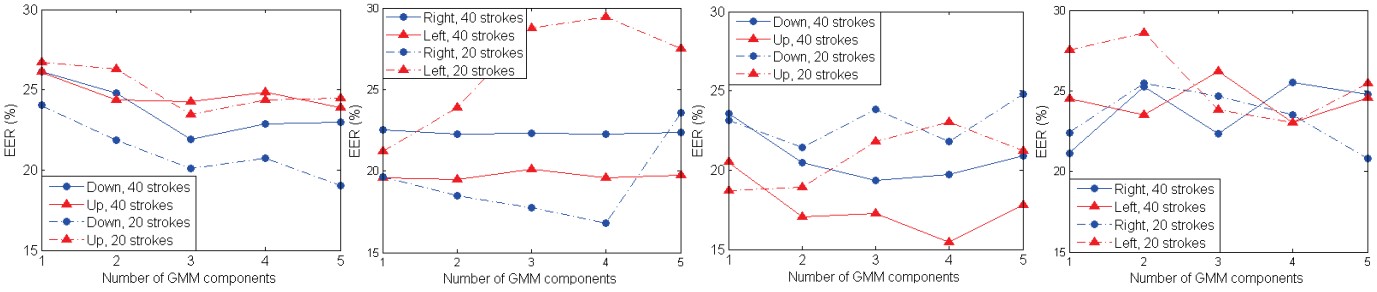
In the present work, the dataset acquired in [1] will be used. According to [1], this database consists of data from 190 users, all of whom are students, faculty or staff at Louisiana Tech University. The data was collected over two sessions, at least one day apart, using for all users only one phone model (*Google Nexus S*), running on Android 4.0. The training set contains the first session data, whereas the test set comprises the second session. Two Android applications were used for data collection, in which users answered a series of multiple choice questions. These questions were different in each session and allowed the user to move freely, scrolling through the short paragraphs and/or images, on which questions were based. Both portrait and landscape orientation of the phone were allowed.

For each point of a stroke, the application recorded the x and y coordinates, the pressure on the screen, the area occupied by the finger, the timestamp of the data point and the phone orientation. Only two types of interactions were recorded: horizontal strokes and vertical strokes. All other gestures, such as zooming or rotations, were ignored.

IV. System description

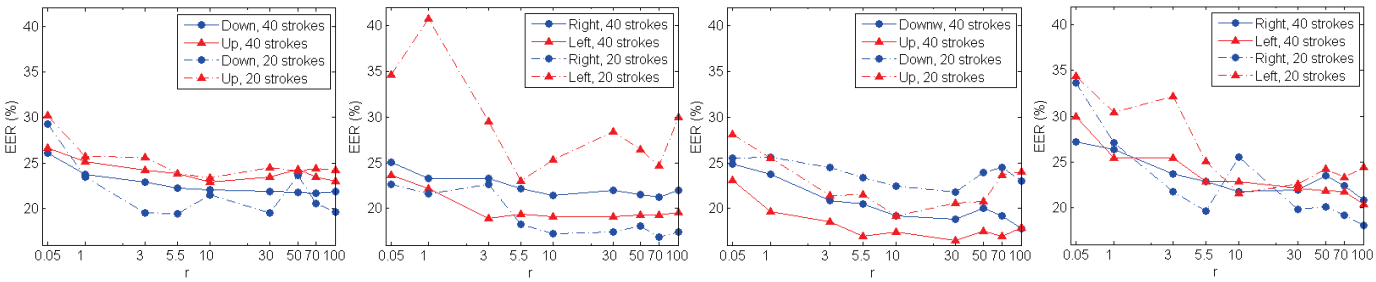
A. System architecture

The system architecture is depicted in Fig. 1. Landscape and portrait data are processed separately, because some features, e.g. the coordinates from start and end-points or the velocity, may change for the same user depending on the orientation. The strokes from each orientation are first classified as vertical or horizontal. Moreover, they are further divided based on their direction as: upwards, downwards, leftwards or rightwards. The motivation is that despite being, for example, both vertical gestures, upwards and downwards strokes are performed differently (e.g. with different fingers) and present their own characteristic features, in the same way that one does not walk equally forward and backwards. This division guarantees that the particularities of how each user makes each gesture is taken into account. For each stroke, a feature vector is extracted and, afterwards, similarity scores are obtained using averaged blocks of ten scores, as it has been proven in the literature that blocks of strokes work better [1], [3], [5].



(a) Portrait Vertical (b) Portrait Horizontal (c) Landscape Vertical (d) Landscape Horizontal

Fig. 2: Effect on the performance of the number of Gaussian components, with $r = 5.5$



(a) Portrait Vertical (b) Portrait Horizontal (c) Landscape Vertical (d) Landscape Horizontal

Fig. 3: System performance as a function of r with 4 GMM components

TABLE I: Optimal feature set selected by the SFFS algorithm

Scenario	Best performing features
Vertical	θ(finger-down to finger-up), $(x_{max} - x_{min})/x_{max\_range}$, std of a_x, $(\bar{x} - x_{min})/\bar{x}$, $(y_{max} - y_{min})/y_{max\_range}$
Horizontal	θ(finger-down to finger-up), $(y_{max} - y_{min})/y_{max\_range}$, std of a_y, $(\bar{y} - y_{min})/\bar{y}$, $(x_{max} - x_{min})/x_{max\_range}$

B. Feature extraction

Prior to computing the feature vectors, short strokes of less than five touch points, which probably come from taps on the screen, are considered outliers and discarded.

The feature set used in this work is adapted from the 100-dimensional feature vector studied in [10], which contains a high proportion of the best performing global features from the signature biometrics literature. The use of this feature vector is motivated by the fact that similarly to touch biometrics, temporal features of gestures made on a surface are extracted in online signatures and graphical passwords with doodles [11]. A complete description of these features can be found in [10]. However, signatures are much more sophisticated descriptors than swipes on a touchscreen and hence, need more complex features. This way, to adapt the feature vector, several features that cannot be applied to the current problem were not considered, such as those relating to pen-ups and pen-downs. Additionally, feature selection is performed using the Sequential Forward Floating Search (SFFS) algorithm [12]. The best 5 features in terms of EER chosen by this algorithm are depicted in Table I. This algorithm is used with vertical strokes in portrait orientation, the most commonly performed. Despite the small dimension of the resulting feature vector, it should be noted that the gestures used for classification can be easily described due to their simplicity. Lastly, feature vectors normalization into the interval (0,1) is computed using tanh-estimators [13].

C. Similarity computation

The UBM is a general model that describes the behaviour of a population in a feature space. It is computed once for all users using full covariance matrices and all data from the training set. Once calculated, it is adapted to the legitimate subject's model using his training samples via a relevance factor *r*, which trade-offs between the general information of the UBM and the specificities of the user training data. The higher *r* the more importance is given to the UBM general data. A complete description of the steps followed to obtain the user's model can be found in [8]. The adapted GMM was implemented using Matlab's statistics toolbox (Version 8.3).

V. Experiments

A. Results

In this section, the best parameters for authentication are studied. The effect on the performance of the number N of GMM components with a fixed relevance factor $r = 5.5$ is first studied. In Fig. 2 it is shown that the number of components that result in the best performance varies depending on the type of stroke and the number of training samples considered to adapt the UBM. In most cases, the performance improves as the number of components grows. With 40 training samples, the performance is better and more

stable in most cases. Downwards strokes perform better than upwards in portrait orientation, while the opposite happens in landscape orientation. This means that these strokes are more distinctive for each of the orientations. Horizontal strokes performance presents a larger variability depending on the number of training samples and it barely improves as the number of GMM components grows. This may be because data points of horizontal strokes are very condensed in one or two clusters, and, when using more components, the models represent outliers that do not show the general behaviour of the current user.

The effect of the relevant factor r in the adaptation process is studied for $N = 4$ GMM components, which resulted in a good performance in most scenarios. In Fig. 3 the mean EER across all users in the system is depicted as a function of r. It can be observed that, as r increases, the EER has a slight initial descent, with some rises and falls, until it stabilizes. This can be caused by the model becoming more general as r increases, thus being less adapted to the user's specific characteristics. Nevertheless, it should also be noted that even with a large r, the adapted model still holds enough user specific information to be discriminative. As was also observed when changing the number of GMM components, downwards strokes perform better than upwards in portrait orientation, while the reverse occurs in landscape orientation. For horizontal strokes, the performance is slightly better in leftwards strokes for portrait orientation, whereas in landscape orientation it is similar.

The performance obtained for all the described configurations is around 20% EER, with the best values ranging from 15% to 22%. These values are obtained with different parameters in each of the operations, but using a relevance factor $r = 30$ and 40 training samples to adapt the UBM results in a better performance than other configurations.

B. Discussion of the performance across touch operations

In portrait orientation, horizontal strokes perform better than the vertical ones. Additionally, it should be noted, the best performing gestures are, in most experiments, the downwards and rightwards strokes over the upwards and leftwards strokes. This means that they hold more discriminative information. A possible reason may be that they are performed more often and are more stable and hence users swipe the screen following more distinctive patterns in these gestures.

On the other hand, in landscape orientation horizontal strokes present only a slightly better performance than vertical ones. Considering this also occurred in portrait orientation, horizontal gestures probably hold more discriminative information. In addition to the frequency in which these strokes are performed, a possible cause may be related to being much shorter gestures. Thus, horizontal gestures have less degrees of freedom and users tend to show a more stable behaviour.

The overall performance is better in landscape orientation than in portrait orientation. One of the reasons for this could be that the users who swiped the screen in landscape orientation almost never use portrait orientation, and thus, present more consistent and stable habits.

VI. Conclusions and future work

In this work, continuous authentication using the most common touch operations has been studied using a statistical approach. Users' touch behaviour can be modelled using a GMM adapted from a UBM, that represents an "average user". The results have shown that touch biometrics are discriminative and stable enough to be used for user recognition and that they are a promising method for active authentication. In addition, it has been found that horizontal strokes hold more user-specific information and are more discriminative than vertical strokes. Gestures made in landscape orientation are also more discriminative than in portrait orientation.

Nevertheless, these results are only preliminary. Further work includes studying this system with other databases and comparing it to other approaches [1], as well as extending the considered gestures to multitouch operations [14].

Acknowledgment

Project CogniMetrics (TEC2015-70627-R) and contract IJCI-2015-24742 from MINECO/FEDER provide funding.

References

[1] A. Serwadda, V. V. Phoha, and Z. Wang, "Which verifiers work?: A benchmark evaluation of touch-based authentication algorithms," in *Proc. IEEE BTAS*, 2013, pp. 1–8.

[2] H. Xu, Y. Zhou, and M. R. Lyu, "Towards continuous and passive authentication via touch biometrics: An experimental study on smartphones," in *Proc. SOUPS*, 2014, pp. 187–198.

[3] 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 Transactions on Information Forensics and Security*, vol. 8, no. 1, pp. 136–148, 2013.

[4] V. M. Patel, R. Chellappa, D. Chandra, and B. Barbello, "Continuous user authentication on mobile devices: Recent progress and remaining challenges," *IEEE Signal Processing Magazine*, vol. 33, no. 4, pp. 49–61, July 2016.

[5] C. Shen, Y. Zhang, X. Guan, and R. A. Maxion, "Performance analysis of touch-interaction behavior for active smartphone authentication," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 3, pp. 498 – 513, March 2016.

[6] U. Mahbub, S. Sarkar, V. M. Patel, and R. Chellappa, "Active user authentication for smartphones: A challenge data set and benchmark results," in *Proc. IEEE BTAS*, 2016.

[7] J. Fierrez-Aguilar, D. Garcia-Romero, J. Ortega-Garcia, and J. Gonzalez-Rodriguez, "Bayesian adaptation for user-dependent multimodal biometric authentication," *Pattern Recognition,*, vol. 38, no. 8, pp. 1317–1319, August 2005.

[8] M. Martinez-Diaz, J. Fierrez, and J. Ortega-Garcia, "Universal background models for dynamic signature verification," in *Proc. IEEE BTAS*, September 2007, pp. 1–6.

[9] R. Kumar, V. V. Phoha, and A. Serwadda, "Continuous authentication of smartphone users by fusing typing swiping and phone movement patterns," in *Proc. IEEE BTAS*, 2016, pp. 1–8.

[10] M. Martinez-Diaz, J. Fierrez, R. P. Krish, and J. Galbally, "Mobile signature verification: Feature robustness and performance comparison," *IET Biometrics*, vol. 3, no. 4, pp. 267–277, December 2014.

[11] M. Martinez-Diaz, J. Fierrez, and J. Galbally, "Graphical password-based user authentication with free-form doodles," *IEEE Trans. on Human-Machine Systems*, vol. 46, no. 4, pp. 607–614, August 2016.

[12] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*, 4th ed. Academic Press, 2008.

[13] A. Jain, K. Nandakumar, and A. Ross, "Score normalization in multimodal biometric systems," *Pattern Recognition*, vol. 38, no. 12, pp. 2270 – 2285, 2005.

[14] N. Sae-Bae, N. Memon, K. Isbister, and K. Ahmed, "Multitouch gesture-based authentication," *IEEE Trans. on Information Forensics and Security*, vol. 9, no. 4, pp. 568–582, April 2014.