Privacy Compliant Multi-biometric Authentication on Smartphones

Alexandre Ninassi, Sylvain Vernois and Christophe Rosenberger

Normandie Univ., UNICAEN, ENSICAEN, CNRS, GREYC, 14000 Caen, France

Keywords: Authentication, Multi-biometrics, Biometric Template Protection.

Abstract: Smartphones are more and more used by Internet users for different services such as social networks, e-commerce or email. User authentication with passwords on such devices is not user-friendly and does not offer a high security level for this task. Biometrics is becoming one popular solution to achieve this goal with the embedding of fingerprint scanners in smartphones. In this paper, we propose a new protocol combining fingerprint and behavioral biometrics to enhance the security of user authentication while preserving usability and privacy. The behavior when entering a pattern based authentication on the smartphone touch screen is considered as a fast and usable solution for users. We think the proposed multi-biometric solution offers great advantages for many applications such as e-payment in terms of security, usability and privacy. We show through experimental results the efficiency of the proposed method.

1 INTRODUCTION

User authentication with smartphones is more and more envisaged for applications on the Internet and electronic transactions. A recent survey in 2016 (Sukhraj, 2016) showed that over 50% of smartphone users grab it immediately after waking up. As a smartphone embeds more and more personal information (contacts, mail content, media...) and is used as preferred device to access distant services, a strong authentication is necessary for logical access control. PIN code authentication is a common solution on smartphones. Even if this solution is simple, it does not constitute a strong identity proof as anybody looking at the user typing it could use it.

In order to solve this problem, biometrics is more and more used to increase the level of confidence of user authentication. Nevertheless, biometric data is sensitive and requires a particular attention in terms of security and privacy. Biometric data protection should be realized during all the life cycle of the data including the storage and the handling. Standard cryptography as symmetric encryption (or hash functions) does not ensure the data protection during the comparison step because two biometric data from the same individual are not exactly identical. Consequently, the comparison cannot be realized on the encrypted/hashed domain. Several ways are proposed to achieve the protection of biometric data, including adapted cryptographic schemes (fuzzy commitment, homomorphic encryption) (Juels and Wattenberg, 1999; Barni et al., 2010) feature transformations (including the Biohashing algorithm) (Teoh et al., 2004; Nagar et al., 2010). All these schemes should ensure security, diversity and revocability of the biometric data. We do not intend in this paper to propose a new scheme for biometric template protection. For more details on these schemes, we refer the reader to the detailed survey (Rathgeb and Uhl, 2011).

A biometric authentication is realized in two steps: the enrollment and the verification phases. The first one consists in generating the biometric reference template of one user and to store it for further comparison. During verification, a query biometric template is compared to the reference one for decision. In order to enhance security of user authentication, it is required in general to combine different authentication factors. This can be realized by using different biometric data to define a multi-biometric system. Note that Multi-Factor Authentication (MFA) Market is expected to reach USD 9.60 Billion by 2020 (Mar, 2016).

The first contribution of this paper is to propose an efficient and usable multi-biometric system to enhance the security of user authentication on smartphones. We combine two biometric modalities namely fingerprint and behavioral biometrics. We assume in this work the used smartphone has a finger-
print scanner. This hypothesis is realistic as a recent survey estimates that 67% of smartphones in 2018 will have a fingerprint scanner (Statista, 2016). The reference fingerprint template in such smartphones is stored in a secure element to ensure its protection. Second, we use a behavioral biometric modality with a template protection scheme: biometric pattern drawn on a touch screen. This solution has the advantage to be very simple to use and very quick. The biometric reference is stored in the smartphone as a BioCode that can be canceled in case of attack. Experiments are carried out on a home made chimeric (own made) benchmark dataset with 34 users (fingerprint and biometric pattern) and show the benefit of the proposed solution face to the literature.

This paper is organized as follows. Section 2 provides a literature review on existing solutions for user authentication on smartphone. The proposed method is described in Section 3. We present its general principle of user, the concept of feature transformation template protection schemes and the Biohashing algorithm. The computation of biometric features from pattern drawing is also described. Section 4 illustrates through experimental results the benefit of the proposed solution. Finally, we conclude and give some perspectives in Section 5.

2 RELATED WORKS

Biometric based mobile authentication is an emerging issue, with increasing references in the literature. The NIST report (Orandi and McCabe, 2009) details some recommendations concerning portable biometric acquisition station and considers the following modalities: fingerprint, face and iris. Most of papers in the literature are devoted to a particular modality. Researchers considered morphological modalities to solve this issue. Face recognition is dealt with in the paper (Hadid et al., 2007), along with eye detection, or in (Choi et al., 2011), where a real time training algorithm is developed for mobile devices. The authors propose to extract local face features using some local random bases and then to incrementally train a neural network. Image processing also concerns hand biometrics on mobile as in the reference (de Santos-Sierra et al., 2011), where hand images are acquired by a mobile device without any constraint in orientation, distance to camera or illumination. Other papers (Clarke and Furnell, 2007 ; Changa et al., 2012) consider keystroke dynamics based recognition. The first paper makes a study about user identification using keystroke dynamics-based authentication (KDA) on mobile devices, relying on 11-digit telephone numbers and text messages as well as 4-digit PINs to classify users. Many papers propose to use touch screen to capture biometric data (Sae-Bae et al., 2012). Most of these studies use methods used for keystroke or signature dynamics. As for example, the notion of TapPrint has been proposed by Miluzzo et al. (Miluzzo et al., 2012) where the concept of keystroke dynamics is generalized to touch screen. The proposed method is based on the location of the tap on the key associated to a letter or by analyzing gyroscope information. The system has been tested on 10 volunteers with a total number of 40000 taps. The recognition efficiency is between 80% and 90%. The work done by Luca et al. (Luca et al., 2012) is very interesting because it combines pattern based password and biometrics. They proposed a system and test it with 34 users. Authors obtained a performance of 19% for the FRR value (False Rejection Rate) and 21% for the FAR (False Acceptance Rate). In 2013, a method has been proposed (Beton et al., 2013) combining multiple information compared with the Pearson Correlation and the Dynamic Time warping (DTW) methods. The equal error rate (EER) is near 17% which is the best result for this biometric modality. Apart from the literature dedicated to biometric solutions for mobile authentication related to a specific modality, some recent papers propose to use more than one biometric modality (Alzubaidi and Kalita, 2016; Buriro, 2017). The work proposed in (Vildjiounaite et al., 2006) combines the voice and accelerometer information from the mobile phone. They obtain a low performance with an EER value equals to 9%. Another work combined different behaviors of the user to authenticate him/her with an EER equals to 9.2% (Saevanee et al., 2014). An interesting work done by galdi et al. (Galdi et al., 2016) combined iris recognition and information on the camera (sensor fingerprinting) with an excellent EER value (i.e. 0.05%), unfortunately, none protection of biometric templates is provided. A recent paper (Stokkenes et al., 2016), combined face and eyes images for user authentication. The obtained performance in the best case (without any attack of the protection scheme) equals 1.8%.

We can see that many works have been done to propose biometric systems for user authentication on mobile devices. Based on this literature review, we propose in the next section a new authentication solution combining fingerprint and biometric pattern. The first modality is present in most smartphones and the biometric reference template is safely stored in a secure element (SIM card or the one associated to the fingerprint sensor). Using the biometric pattern allows to enhance the security level while providing a practical authentication solution (fast interaction) and
a good protection of the biometric data by applying a template protection scheme.

3 PROPOSED METHOD

The general principle of the proposed method is given by Figures 1 and 3. During enrollment phase, Alice has to provide her fingerprint to the smartphone scanner to generate her reference fingerprint. The extracted minutiae template is stored on a dedicated secure element to protect it. She has also to enter a pattern on the touch screen. The application computes many features based on her behavior. We present in the next section the details on the computation of behavioral parameters. We apply then the BioHashing algorithm in order to protect this template, this algorithm is presented in the following sections. For this algorithm, we need a secret key in order to be able to revoke it in case of attack. This secret key could be a password, a random value, a binary representation of the pattern. This secret can also be concatenated with other information such as: IMEI number (International Mobile Equipment Identity), Alice’s name, random value... During the verification phase, Alice has to provide her fingerprint to authenticate herself. The captured fingerprint is compared with Alice’s reference in the secure element. If her identity is verified, she has to enter the pattern on the touch screen. A capture BioCode is computed and compared to Alice’s Reference BioCode in the smartphone. If both biometric systems accept Alice’s proof, she is authenticated.

3.1 Fingerprint Authentication

We propose to use in this paper the fingerprint as first biometric modality. Most of smartphones embed a fingerprint sensor. Alice has to enroll herself by providing one or more fingerprint captures. The reference fingerprint template (a set of minutiae) is stored in a secure element (SE) associated to the smartphone or the fingerprint sensor hardware. Figure 2 illustrates the architecture on Android smartphones embedding a fingerprint sensor. The matching between a fingerprint template and Alice’s reference template is also realized in the SE and a decision value is provided (the score is not available for security reasons).

Considering privacy, this solution is suitable as the fingerprint reference is stored in the SE. In terms of security, the solution is interesting even if the enrollment process is realized by the user without any control. We can expect a smartphone is a personal object, the enrollment process should be done by the proprietor of the smartphone (Alice in this case). Concerning performance, it is described by the False Acceptance Rate (FAR) value corresponding of the percentage of successful attacks by an impostor. For smartphones, the targeted level of the FAR is lower than 0.005% (Burr et al., 2013), corresponding to the security level 3. It is difficult to verify this value without the score provided by the matching algorithm. The associated False Rejection Rate (FRR) corresponding to errors to recognize legitimate users is supposed to be under 2%. No study exists in the state of the art on the evaluation of commercial fingerprint sensors on smartphones, indeed, the number of users should be very important to provide significant results.

3.2 Biometric Pattern

The biometric system we propose to use in this work intends to increase security for a quick authentication on the mobile device. It corresponds to a two factor approach. We intend to first recognize the user by the knowledge of a password represented by a pattern. We use the classical pattern based unlock screen approach. This way of entering a password is quicker and is more user-friendly on a mobile device. Second, the user behavior while drawing the pattern is analyzed. Many information could be collected during the capture process:

- X position: the horizontal position of the finger on the touch screen is recorded during the capture,
- Y position: the vertical position of the finger on the touch screen is also recorded,
- Pressure: the pressure of the finger on the touch screen is captured (provided by the Android OS),
- Touch size: ratio of pixels where the finger is in contact with the touch screen,
- Tilt: orientation information from the accelerometer sensor,
- Accelerometers: three angles corresponding to the orientation of the smartphone.

As the time needed to draw the same pattern can be different for each capture, signals are undersampled to a fixed length. A constant size description is necessary to use this template as input of the BioHashing algorithm that we detail in the next section.

3.3 BioHashing

A feature transformation is a function $F$ using a key $K$ (that is typically a random seed or a password), applied to a biometric template $T$. The transformed template $F_K(T)$ is stored in a database or in a personal
device. The Biohashing algorithm, described below, belongs to this class of transformation. During the authentication step, the same transformation is applied to the query template \( T' \) with the same key \( K \) and a comparison is realized between \( F_K(T) \) and \( F_K(T') \). It is generally considered that, given the transformed template \( F_K(T) \) and the key \( K \), it is not possible to recover the original template \( T \) (or a close approximation) as presented in (Nagar et al., 2010) (mainly because the transformation is not invertible). The key constitutes an important secret. The performance of the authentication system is generally estimated with FMR (False Match Rate computing the ratio of false positive verification) FNMR (False Non Match Rate calculating the ratio of false negative verification) rates and the feature transformation should not decline the performance of the system. In fact, this approach tends to improve the performance of the biometric system without any protection (but the key \( K \) is necessary). Indeed, the projection of similar biometric templates from two distinct individuals with two different keys is in general very different. The Biohashing algorithm is applied to biometric templates, represented by real-valued vector of fixed length (the metric used to evaluate the similarity between two biometric features is the Euclidean distance) and generates binary templates of length lower or equal to the original length (the metric used to evaluate the similarity between two transformed templates is the Hamming distance). This algorithm has been originally proposed for face and fingerprints by Teoh et al. in (Teoh et al., 2004). The Biohashing algorithm generates a binary template called BioCode. At the end of the enrollement phase, the biometric raw data is discarded and the BioCode (with the associated seed) is stored. The biohashing algorithm can be applied on any biometric modalities, that can be represented by a real values vector of fixed length.

The Biohashing algorithm transforms the biometric template \( T = (T_1, \ldots, T_n) \) in a binary template \( B = (B_1, \ldots, B_m) \), with \( m \leq n \), as follows:

1. \( m \) pseudorandom orthonormal vectors \( V_1, \ldots, V_m \) of length \( n \) are generated from the random seed \( K \) typically with the Gram Schmidt algorithm.
2. For \( i = 1, \ldots, m \), compute the scalar product \( x_i = \langle T, V_i \rangle \).
3. Compute the binary template \( B = (B_1, \ldots, B_m) \) with the quantization process:

\[
B_i = \begin{cases} 
0 & \text{if } x_i < \tau \\
1 & \text{if } x_i \geq \tau,
\end{cases}
\]

where \( \tau \) is a given threshold, generally equal to 0.

The performance of this algorithm is ensured by the scalar products with the orthonormal vectors, as detailed in (Teoh et al., 2008). The quantization process of the last step ensures the non-invertibility of the data (even if \( n = m \), because each coordinate of the input \( T \) is a real value, whereas the coordinates of the output \( B \) is a single bit). Finally, the random seed \( K \) guaranties the diversity and revocability properties.
3.4 Implementation

The traditional approach to Android development is to use the native SDK published by Google. If this is an efficient approach at first sight, it does not allow further reuse for a different mobile platform from Android because the code is tightly connected to this SDK. As we may want to be able to later evaluate the authentication performance on any platforms the proposed application is developed using Xamarin framework.

Android touch screen raises touch events every 10 to 30ms based on the gestures. Each event defines raw values such as horizontal X and vertical Y coordinates, pressure and size of the touch on the screen, an orientation indicator dedicated to stylus use and a time stamp. These values highly depend on mobile model used during the acquisition process:

- \((X,Y)\) position depend on screen resolution when the velocity vector \((X',Y')\) also depend on screen size.
- Pressure and size depend on screen resolution but also on screen capabilities.
- No guaranties are given by Android concerning the frequency of events.

Each Android raised event also makes available some historical events that occurred since the last raised event, allowing a higher sampling. Concrete tests show an increase of available events of up to 4 times when drawing a pattern. These values also depend on the operating system: iOS mobiles (such as iPhone and iPad), raise events defining similar but not identical raw values and furthermore the pressure data availability depends on iOS version. The frequency of events is similar to the ones observed with Android mobiles but no additional history is available. The implementation is based only on actually raised events, Android historical events being unused. Raw data sequence is then re-sampled to get the desired input data length to feed the BioHashing algorithm. The random seed \(K\) used to generate the pseudo random orthonormal vectors is in fact the result of the combination of a random seed chosen at the time of enrollment that is stored on the device for further use and a value generated by the drawing of the pattern. For this last value, each dot used as crossing point is given as a string value. A pattern string is then created by aggregating the value of each dot encountered during the drawing and finally hashed when finished. The seed is then the result of the \(XOR\) of the random device value and the hash. This choice is to allow the concrete seed not to be fully stored on the device.

4 EXPERIMENTAL RESULTS

In this section, we present the experimental results we realized for the validation of the proposed system.

4.1 Protocol

In this work, we first used a biometric dataset of data captured when users draw a single pattern:
• Data have been collected on a Nexus 4 mobile phone with a touch screen having a resolution of 800 x 1280 pixels,
• The pattern was the same for all users and is defined by the following pattern code “1235987”. This experimental setup can be considered as the worst case where an attacker knows the pattern to draw. 34 users participated to this experiment,
• Each user provided 15 samples described by 4 signals undersampled to 200 values (time normalization). Four more signals have been computed by considering the first and second derivative of X and Y signals. We also added the total time to draw the secret path as additional parameter. Figure 4 presents the data of the first sample for user 1. The x-axis corresponds to the time and the y-axis to feature value. So, the template size is 1601 (by concatenating all undersampled signals) and the single time value,

In total, we have a subset of $34 \times 15 = 510$ biometric templates of size 1601 real values for biometric pattern. Considering the BioHashing setup, we set the parameter values as following:
• Template size: $n=1601$,
• BioCode size: $m=750$ for the Reference BioCode and Capture BioCode,
• As the pattern is the same for all users, in the computation of the Reference BioCode, the secret $K$ for all users is randomly drawn,
• Matching algorithm: Hamming distance.

Concerning fingerprint, we extract a subset of well known datasets namely FVC2002 DB2, FVC2004 DB1 and FVC2004DB3 (FVC, 2002). We can see that fingerprints are quite different and representative of the different types of fingerprint (acquired with sensors using different technologies). These datasets have fingerprints from 100 individuals with 8 samples per individual. In order to constitute a chimeric multi-biometric dataset (synthetic database), we took into account the fingerprints from the 34 first individuals. For each FVC dataset, we use so $34 \times 8 = 272$ fingerprint samples. In order to evaluate the performance of the proposed method, we use the following methodology:
• We use the first sample of each user as reference template, for the biometric pattern, we use this data to compute the Reference BioCode,
• As we do not have access to the fingerprint sensor hardware (i.e. the value of the matching score) on a mobile phone, we simulate the result of the matching score by considering the Bozorth3 algorithm provided by the NIST. This algorithm does not provide as good results as commercial on card comparison (OCC) algorithms, it can be considered as an estimate of the worst performance,
• We compute genuine scores as follows. We consider all reference fingerprints and we compare them with each available samples belonging to the same individual. We consider two times these scores because biometric pattern has 14 samples. For biometric pattern, we compare the Reference Biocode with all other BioCodes from the same individuals. We obtain $14 \times 34 = 476$ fingerprint and biometric pattern genuine scores for each FVC dataset,
• We have a similar process to simulate impostor attack by considering all biometric samples belonging to another user. We obtain $14 \times 34 \times 33 = 15708$ fingerprint and biometric pattern impostor scores for each FVC dataset,
• Given these two sets of scores, we can compute their distribution in order to estimate in which measure impostor scores are different than legitimate ones. Second, we compute the Equal Error Rate (EER) value that is a well known metric in biometrics that measures the behavior of the biometric system when the decision threshold is set to have the same number of false rejected users and false accepted ones.

4.2 Results

First, we try to estimate the efficiency of each biometric system we combine. The EER value is between 5.2% and 8%. We could expect by using a commercial OCC a much better performance, this value estimates an upper bound of error. We compute the performance of the system based on biometric pattern. We consider 200 threshold values for estimating False Acceptance Rate (FAR) and False Rejection Rate (FRR) in all scenarios. First, use the Euclidean distance as matching score. In this case, we see many errors as the associated EER value is 27.4%. This value is very high but we use the BioHashing algorithm to enhance this performance. When applying the BioHashing algorithm in the best case (secret only known by the legitimate user), we obtain a perfect recognition with an EER value of 0%. The worst case scenario (the secret is known by the impostor) with a performance similar to raw features performance. This error could be considered as too high but without using the biometric pattern, impostors would be able to enter the system knowing the pattern (used as password). We limit this attack by considering biometric features. Figure 5 provides the score distribu-
Table 1: Comparison with existing multi-biometric systems for mobile phones.

<table>
<thead>
<tr>
<th>authors</th>
<th>biometric modalities</th>
<th>protection</th>
<th>EER</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Vildjounaite et al., 2006)</td>
<td>voice and accelerometer</td>
<td>no</td>
<td>9%</td>
</tr>
<tr>
<td>(Saevanee et al., 2014)</td>
<td>behaviors</td>
<td>no</td>
<td>9.2%</td>
</tr>
<tr>
<td>(Stokkenes et al., 2016)</td>
<td>eyes and face</td>
<td>yes - best case</td>
<td>1.8%</td>
</tr>
<tr>
<td>(Galdi et al., 2016)</td>
<td>iris and sensor fingerprinting</td>
<td>no</td>
<td>0.05%</td>
</tr>
<tr>
<td>Contribution</td>
<td>fingerprint and biometric pattern</td>
<td>yes - best case</td>
<td>0%</td>
</tr>
<tr>
<td>Contribution</td>
<td>fingerprint and biometric pattern</td>
<td>yes - worst case</td>
<td>2.3%</td>
</tr>
</tbody>
</table>

4.3 Comparisons

Table 1 presents the performance of some multi-biometric systems for mobile phones in the literature. Most methods do not provide any protection of biometric templates. This is an important issue as mobile phones are vulnerable to attacks with malwares. The proposed contribution provides an excellent result in the best case and a good result in the worst case (when the BioHashing key is known by the imposter).

5 CONCLUSION AND PERSPECTIVES

In this paper, we propose a multi-biometric system for mobile devices by combining the fingerprint recognition using its embedded sensor and a behavioral biometric system. The proposed system is very fast (a few seconds for the capture and about 200 milliseconds for comparison) and practical for users as all these verification systems are commonly used. The combination of fingerprint recognition and biometric pattern allows to limit the possible attack and increase the security of user authentication. The privacy protection of biometric templates is ensured by using at the same time a secure element and template protection schemes. In the best case, we obtain a perfect recognition on our own made database and a FAR lower than 2.3% in the worst case (the impostor needs the mobile devices, knows the secret path and the secret key $K$ associated to the BioHashing algorithm). We intend in the future to embed other biometric systems such as face and voice recognition systems. We also work on embedding the template protection scheme in the secure element.

ACKNOWLEDGMENTS

Authors would like to thank the United Biometrics company for financial support of this work.
REFERENCES

(2002). Fingerprint verification competition databases.


