Incorporating Touch Biometrics to Mobile One-Time Passwords: Exploration of Digits

Ruben Tolosana, Ruben Vera-Rodriguez, Julian Fierrez and Javier Ortega-Garcia
Biometrics and Data Pattern Analytics (BiDA) Lab
Universidad Autonoma de Madrid, 28049 Madrid, Spain
{ruben.tolosana, ruben.vera, julian.fierrez, javier.ortega}@uam.es

Abstract

This work evaluates the advantages and potential of incorporating touch biometrics to mobile one-time passwords (OTP). The new e-BioDigit database, which comprises online handwritten numerical digits from 0 to 9, has been acquired using the finger touch as input to a mobile device. This database is used in the experiments reported in this work and it is publicly available to the research community. An analysis of the OTP scenario using handwritten digits is carried out regarding which are the most discriminative handwritten digits and how robust the system is when increasing the number of them in the user password. Additionally, the best features for each handwritten numerical digit are studied in order to enhance our proposed biometric system. Our proposed approach achieves remarkable results with EERs ca. 5.0% when using skilled forgeries, outperforming other traditional biometric verification traits such as the handwritten signature or graphical passwords on similar mobile scenarios.

1. Introduction

Mobile devices have become an indispensable tool for most people nowadays [16]. This rapid and continuous deployment of mobile phones around the world has been motivated not only for the high technological evolution and new features incorporated by the mobile phone sector but also to the new internet infrastructures that allow the communications and use of social media in real time, among many other factors. In this way, both public and private sectors are aware of the importance of mobile phones for the society and are trying to deploy their services through user-friendly mobile applications ensuring data protection and a high security access level. However, this idea is difficult to accomplish using only the traditional One-Time Password (OTP) security approaches based on PINs (Personal Identification Numbers). Biometric recognition schemes seem to cope with these problems as they combine both high performance and convenience as they are part of ourselves [12].

Biometric behavioural verification systems are becoming a very attractive way to verify users on mobile devices. One of the most socially accepted traits is the handwritten signature as it has been used in financial and legal agreements scenarios for many years [5, 13, 14]. Biometric verification systems based on on-line handwritten signature are very effective against both skilled (i.e., the case in which impostors have some level of information about the user being attacked and try to forge their signature claiming to be that user in the system) and random impostors (i.e., the case in which no information about the users being attacked is known and impostors present their own signature claiming to be another user of the system). In [18], the authors explored the use of new algorithms based on Recurrent Neural Networks (RNNs) on office-like scenarios for pen-based signature recognition achieving results below 5.0% Equal Error Rate (EER) for skilled impostors. However, a considerable degradation of the system performance with results around 20.0% EER is produced for skilled forgeries when testing on universal mobile scenarios using finger touch for signature generation [2, 15, 17]. The reason for such degradation of the system performance compared to pen-based office-like scenarios is due to the fact that users tend to modify the way they sign, e.g., users who perform their signatures using closed letters with a pen input tend to perform much larger writing executions in comparison with other letters due to the lower precision they are able to achieve using the finger. Besides, users whose signatures are composed of a long name and surname (or two surnames) tend to simplify some parts of their signatures due to the small surface of the screen to sign. In [11], the authors proposed a different approach based on graphical passwords with free doodles for mobiles achieving final results above 20.0% EER for skilled forgery scenarios. The main reason for such degradation of the system performance lays down on the specific task that the user needs to perform to be authenticated, e.g., doodles were difficult to memorize
for most of the users as they didn’t use them on daily basis. Consequently, many researchers are putting their efforts to develop more robust and user-friendly security schemes on mobile devices.

Two-factor authentication approaches have gained a lot of success in the last years in order to improve the level of security. These approaches are based on the combination of two authentication stages. For example, one possible case could be the following: 1) the security system checks that the claimed user introduces its unique password correctly, and 2) its behavioural biometric information is used for an enhanced final verification [9]. This way the robustness of the security system increases as impostors need more than the traditional password to get access to the system. This approach has been studied in previous works. In [1], the authors proposed a two-factor verification system based on dynamic lock patterns, achieving a final average value of 10.39% EER for skilled forgeries. A similar approach based on OTP with dynamical lock patterns was considered in [8] extracting features such as the X and Y position, pressure or finger size with very good results. This approach has also been expanded to periocular biometrics [6].

This work proposes a novel OTP system, where the users perform handwritten numerical digits on the screen of a mobile device. This way, the traditional OTP is enhanced by incorporating biometric dynamic handwritten information. Two different aspects of the security system are analysed. First, the analysis of the OTP regarding which are the most discriminative handwritten digits and how robust the system is when increasing the number of them in the user password. Second, the analysis of the biometric system in terms of which are the best features extracted for each handwritten numerical digit. One example of use that motivates our proposed approach is focused on internet payments by means of credit cards. Banks usually send a numerical code (typically between 6 and 8 digits) to the user mobile phone. This numerical code must be inserted by the user in the security platform in order to complete the payment. Our proposed approach enhances such scenario by including a second authentication factor based on the user biometric information while performing the handwritten digits.

The main contributions of this work are the following:

- We incorporate touch biometrics to mobile OTP. An exhaustive analysis of the OTP regarding which are the most discriminative handwritten digits and how robust the system is when increasing the number of them in the user password is carried out.

- An analysis of our proposed system regarding the best features extracted for each handwritten digit.

- The new e-BioDigit database, comprising on-line handwritten numerical digits from 0 to 9 for a total of 93 users, captured on a mobile device using finger touch interactions. Handwritten digits were acquired in two different sessions in order to capture the intra-user variability. This database is publicly available to the research community.

The remainder of the paper is organized as follows. Sec. 2 describes our proposed OTP system including touch biometrics for mobile scenarios. In Sec. 3, we describe the new e-BioDigit database which comprises on-line handwritten numerical digits from 0 to 9. Sec. 4 and 5 describes the

---

1https://atvs.ii.uam.es/atvs/e-BioDigit.html
experimental protocol and results achieved using our proposed approach, respectively. Finally, Sec. 6 draws the final conclusions and points out some lines for future work.

2. Proposed System

In this work we propose an OTP system which includes touch biometrics for mobile scenarios, as shown in Fig. 1. In our proposed approach, users have to perform the handwritten numerical digits (one at a time) of the traditional OTP on the screen to be authenticated. This group of handwritten digits is then compared to the enrolment data of the claimed user comparing one by one each digit. This way the final score is produced after averaging the different one by one digit score comparisons. First, we analyse the case of just using one digit for user verification and then we analyse the discriminative power of the combination of several digits. Only the behavioural information of the user while performing the handwritten digits is analysed in this work making the assumption that impostors pass the first stage of the security system (i.e., they know the password of the attacked users).

2.1. Feature Extraction and Selection

In this work we propose a biometric verification system based on time functions (a.k.a. local system) [19]. Signals captured by the digitizer (i.e., X and Y spatial coordinates) are used to extract a set of 21 time functions for each numerical digit sample (see Table 1). Information related to pressure, pen angular orientations or pen ups broadly used in other biometric traits such as the handwritten signature is not considered here as this information is not available in universal mobile scenarios using finger touch as input.

Sequential Forward Feature Selection (SFFS) is considered in some of the experiments so as to select subsets of time functions that improve the system performance in terms of EER (%). In addition, SFFS is also used in the experimental work to analyse the discriminative power of digit combinations.

2.2. User Verification

Dynamic Time Warping (DTW) is used to compare the similarity between time functions from handwritten digit samples. Scores are obtained as:

\[ \text{score} = e^{-D/K} \]  

(1)

where \( D \) and \( K \) represent respectively the minimal accumulated distance and the number of points aligned between two digit samples using the DTW algorithm [10].

2.3. Database e-BioDigit

The new e-BioDigit database was captured in order to perform the experimental work included in this article. This database is comprised of on-line handwritten numerical digits from 0 to 9 acquired using a Samsung Galaxy Note 10.1 general purpose tablet. This device has a 10.1-inch LCD display with a resolution of 1280 × 800 pixels. Information related to pressure (1024 levels) and pen-ups trajectories are also available when using the pen stylus. However, as this work is focused on universal mobile scenarios, samples were acquired using only the finger as input so only X and Y spatial coordinates are used.

Regarding the acquisition protocol, data subjects had to perform handwritten numerical digits from 0 to 9 one at a time. Some examples of the handwritten numerical digits acquired for the e-BioDigit database are depicted in Fig. 2. Additionally, samples were collected in two sessions with a time gap of at least three weeks between them in order to consider inter-session variability, very important for behavioural biometric traits. For each session, users had to perform a total of 4 numerical sequences from 0 to 9. Therefore, there are a total of 8 samples per numerical digit and user.

The software for capturing handwritten numerical digits was developed in order to minimize the variability of the user during the acquisition process. A rectangular area with a writing surface size similar to a 5-inch screen smartphone was considered. A horizontal line was represented in the bottom part of the rectangular area, including two buttons OK and Cancel to press after writing if the sample was good or bad respectively. If the sample was not good, then it was repeated.

---

<table>
<thead>
<tr>
<th>#</th>
<th>Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>X-coordinate: ( x_n )</td>
</tr>
<tr>
<td>2</td>
<td>Y-coordinate: ( y_n )</td>
</tr>
<tr>
<td>3</td>
<td>Path-tangent angle: ( \theta_n )</td>
</tr>
<tr>
<td>4</td>
<td>Path velocity magnitude: ( v_n )</td>
</tr>
<tr>
<td>5</td>
<td>Log curvature radius: ( \rho_n )</td>
</tr>
<tr>
<td>6</td>
<td>Total acceleration magnitude: ( a_n )</td>
</tr>
<tr>
<td>7-12</td>
<td>First-order derivative of features 1-7: ( \dot{x}_n, \dot{y}_n, \theta_n, v_n, \rho_n, a_n )</td>
</tr>
<tr>
<td>13-14</td>
<td>Second-order derivative of features 1-2: ( \ddot{x}_n, \ddot{y}_n )</td>
</tr>
<tr>
<td>15</td>
<td>Ratio of the minimum over the maximum speed over a 5-samples window: ( \nu'_n )</td>
</tr>
<tr>
<td>16-17</td>
<td>Angle of consecutive samples and first-order derivative: ( \alpha_n, \alpha'_n )</td>
</tr>
<tr>
<td>18</td>
<td>Sine: ( s_n )</td>
</tr>
<tr>
<td>19</td>
<td>Cosine: ( c_n )</td>
</tr>
<tr>
<td>20</td>
<td>Stroke length to width ratio over a 5-samples window: ( r^5_n )</td>
</tr>
<tr>
<td>21</td>
<td>Stroke length to width ratio over a 7-samples window: ( r^7_n )</td>
</tr>
</tbody>
</table>
The database is comprised of a total of 93 users. Fig. 3 shows the statistics for the distribution of user population in e-BioDigit database. Regarding the age distribution, the majority of the subjects (85.0%) are between 17 and 27 years old, as the database was collected in a university environment. Fig. 3 also shows the handedness and the gender distributions. The gender was designed to be as balanced as possible, having 66.7% of males and 33.3% of females whereas for the handedness distribution, 89.2% of the population was righthanded.

4. Experimental Protocol

The experimental protocol is designed in order to assess the potential of our proposed digit-based biometric verification system in real mobile scenarios. Thus, the e-BioDigit database is divided into development (the first 50 users) and evaluation (the remaining 43 users) datasets.

For the development of our proposed approach, the SFFS algorithm is applied to each handwritten numerical digit in order to select the most discriminative time functions for each digit. For that, the 4 available genuine samples from the first session are used as enrolment samples, whereas the remaining 4 genuine samples from the second session are used for testing. Impostor scores are obtained by comparing the enrolment samples with one genuine sample of each of the remaining users.

For the evaluation of our proposed approach, the following two scenarios are considered: 1) having just one genuine sample per digit as enrolment (i.e., 1vs1), and 2) performing the average score of four one-to-one comparisons (i.e., 4vs1) when the number of enrolment samples is four genuine digit samples per user. In addition, for both scenarios, in case of using passwords comprised of several digits, the final score is produced after averaging the different one by one digit score comparisons.

It is important to highlight that the inter-session variability problem is also considered in the experimental protocol carried out in this work as genuine digit samples from different sessions are used as enrolment and testing samples respectively. This effect has proven to be very important for many behavioural biometric traits such as the case of the handwritten signature.

5. Experimental Results

5.1. Experiment 1. Baseline System: One-Digit Results

This section analyses the potential of each numerical digit (i.e., from 0 to 9) in terms of EER(%) for the task of user verification. In order to provide an easily reproducible
framework, we first consider in this section a Baseline System with the same fixed time functions for all numerical digits. Therefore, no development through the use of the SFFS algorithm is considered in this first experiment. Table 2 shows the time functions selected for the Baseline System. We select this set of time functions as baseline as they are commonly used as baseline in other biometric traits such as the handwritten signature [3, 17].

The system performance results in terms of EER(%) obtained for each numerical digit on the evaluation dataset using the Baseline System are depicted in Table 3. Overall, very good verification results are obtained in this first experiment taking into account that only one numerical digit and a Baseline System are considered for verification.

Analysing the extreme scenario of having just one available digit sample during the enrolment (1vs1), the numerical digit 7 achieves the best result with 22.5% EER. In addition, other numerical digits such as 4 or 5 achieve similar results below 25.0% EER. This first experiment puts in evidence the different user verification capacity achieved by each numerical digit. Fig. 4 shows examples of the numerical digit 7 performed for two different users in order to see the low intra- and high inter-user variability of this number. This effect is produced because each person tends to perform numerical digits in a different way, i.e., starting from a different stroke of the numerical digit or even removing some of them such as the crossed horizontal stroke of the number 7.

Analysing the scenario of having four digit samples during the enrolment (4vs1), an average absolute improvement of 3.2% EER is achieved compared to the 1vs1 scenario showing the importance of acquiring as much information as possible during the enrolment stage. For this scenario, the digit 4 achieves the best result with 18.0% EER.

5.2. Experiment 2. Proposed System: One-Digit Results

We apply SFFS over the development dataset in order to enhance the biometric verification system through the selection of specific time functions for each numerical digit. Fig. 5 shows the number of times each time function is selected in our Proposed System from the 21 total time functions described in Table 1. In general, we can highlight the importance of $x_n$, $y_n$ time functions as they are selected for 70% of the numerical digits. In addition, time functions $\dot{x}_n$, $\dot{y}_n$ related to $X$ and $Y$ time derivatives seem to be very important as they are selected for half of the digits. Other time functions such as $\rho_n$, $\dot{\rho}_n$, $\alpha_n$, and $s_n$ related to geometrical aspects of the numerical digits are proven not to be very useful to discriminate between genuine and imposter samples.

Table 4 shows the results achieved for each digit using our Proposed System over the evaluation dataset. In general, better results are achieved compared to the Baseline System (Table 3). Analysing the 1vs1 scenario, our Proposed System achieves an average absolute improvement of 2.0% EER, being the numerical digit 5 the one that provides the best result with a 21.7% EER. Analysing the 4vs1 scenario, our Proposed System achieves an average absolute improvement of 1.6% EER, being again the numerical digit 5 the one that achieves the best result with a 16.9% EER. These results put in evidence the importance of considering different time functions for each numerical digit in order to develop more robust biometric verification systems based on handwritten digits.

5.3. Experiment 3: Digit Combinations

This section evaluates the robustness of our proposed digit-based biometric verification approach when increasing the number of digits that comprise the user password. Fig. 6 shows the time functions selected for the Proposed System when increasing the number of digits during the enrolment stage. In general, better results are achieved compared to the Baseline System (Table 3). Analysing the 1vs1 scenario, our Proposed System achieves an average absolute improvement of 2.0% EER, being the numerical digit 5 the one that provides the best result with a 21.7% EER. Analysing the 4vs1 scenario, our Proposed System achieves an average absolute improvement of 1.6% EER, being again the numerical digit 5 the one that achieves the best result with a 16.9% EER.
Table 3. **Experiment 1**: System performance as EER(%) of each numerical digit using the **Baseline System** on the evaluation dataset.

<table>
<thead>
<tr>
<th>Numerical Digit</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1vs1 scenario</td>
<td>34.9</td>
<td>32.3</td>
<td>32.8</td>
<td>35.0</td>
<td>23.5</td>
<td>24.4</td>
<td>36.9</td>
<td><strong>22.5</strong></td>
<td>26.0</td>
<td>29.6</td>
</tr>
<tr>
<td>4vs1 scenario</td>
<td>33.1</td>
<td>28.5</td>
<td>30.2</td>
<td>32.6</td>
<td><strong>18.0</strong></td>
<td>20.3</td>
<td>36.6</td>
<td>19.2</td>
<td>22.7</td>
<td>25.0</td>
</tr>
</tbody>
</table>

Table 4. **Experiment 2**: System performance as EER(%) of each numerical digit using the **Proposed System** on the evaluation dataset.

<table>
<thead>
<tr>
<th>Numerical Digit</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1vs1 scenario</td>
<td>33.0</td>
<td>34.0</td>
<td>30.9</td>
<td>32.3</td>
<td>22.0</td>
<td><strong>21.7</strong></td>
<td>33.6</td>
<td>21.8</td>
<td>21.8</td>
<td>27.0</td>
</tr>
<tr>
<td>4vs1 scenario</td>
<td>31.4</td>
<td>33.1</td>
<td>27.9</td>
<td>29.7</td>
<td>19.2</td>
<td><strong>16.9</strong></td>
<td>29.7</td>
<td>20.3</td>
<td>18.6</td>
<td>23.3</td>
</tr>
</tbody>
</table>

Figure 5. **Experiment 2**: Histogram of functions selected by SFFS for our Proposed System. Functions described in Table 1.

6 shows the evolution of the system performance in terms of EER (%) on the evaluation dataset when increasing the number of handwritten numerical digits of the password.

Analysing the 1vs1 scenario, a considerable improvement of the system performance is achieved when adding more handwritten digits to the password. A password that is comprised of just two handwritten numerical digits achieves a 16.0% EER, an absolute improvement of 5.8% EER compared to the case of using a password with just one handwritten digit. This result is further improved when increasing the number of handwritten digits of the password with a final 9.0% EER for the case of considering a handwritten numerical password of 8 digits.

For the scenario of acquiring 4 genuine samples per digit during the enrolment (i.e., 4vs1), much better results are achieved because the intra-user variability is better modelled. In this case, a password comprised of just two handwritten numerical digits achieves a value of 11.0% EER, an absolute improvement of 7.6% EER compared to the case of using a password with just one handwritten digit.

Our proposed approach is now compared to other existing state-of-the-art biometric verification approaches for mobile scenarios. Information related to the verification method, and verification performance in terms of EER is included in Table 5 for each study.

In general, our proposed approach achieves good results in comparison to other approaches. For the case of lock pattern dynamic systems [1], the best system performance reported was a 10.39% EER whereas for other methods such as the handwritten signature or graphical password [11, 15, 17], results in the range of 20.0% EER were achieved for skilled forgeries. Additionally, our proposed approach has been compared to other state-of-the-art approaches based on handwritten passwords. In [7], the authors proposed the use of handwritten passwords with a fixed length of 8 characters. Regarding the acquisition process, users had to perform all 8 characters at the same time on the screen of a tablet device. For each handwritten password, a total of 25 static and dynamic features were extracted achieving a final False Acceptance Rate (FAR) of 10.42% with a KStar classifier and using a total of 12 training samples per user. However, that approach seems to be only applicable to devices with big screens such as tablet devices but not to smartphones as it would be very difficult for the users to perform such a long password on a much smaller size screen. Our proposed approach achieves a final value of 5.5% EER for skilled forgeries and is able to mitigate the effect commented before about the size of the screen devices as users have to perform numerical digits one at a time. Additionally, only 4 training samples and not 12 are considered in our proposed approach so it would be easier to deploy on real applications.
Figure 6. **Experiment 3**: Evolution of the system performance in terms of EER (%) on the evaluation dataset when increasing the number of handwritten numerical digits of the password.

Table 5. Comparison of different finger touch biometrics approaches for mobile scenarios.

<table>
<thead>
<tr>
<th>Work</th>
<th>Method</th>
<th>Verification Performance (EER)</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Random Forgeries</td>
<td>Skilled Forgeries</td>
<td></td>
</tr>
<tr>
<td>Angulo and Wastlund (2011)</td>
<td>Lock Pattern Dynamics</td>
<td>-</td>
<td>32</td>
</tr>
<tr>
<td>Martinez-Diaz <em>et al.</em> (2016)</td>
<td>Graphical Passwords</td>
<td>3.4%</td>
<td>100</td>
</tr>
<tr>
<td>Sae and Memon (2014)</td>
<td>Handwritten Signatures</td>
<td>5.6%</td>
<td>180</td>
</tr>
<tr>
<td>Tolosana <em>et al.</em> (2017)</td>
<td>Handwritten Signatures</td>
<td>0.5%</td>
<td>65</td>
</tr>
<tr>
<td>Kutzner <em>et al.</em> (2015)</td>
<td>Handwritten Characters</td>
<td>FAR = 10.42%</td>
<td>32</td>
</tr>
<tr>
<td>Proposed Approach</td>
<td>Handwritten Digits</td>
<td>-</td>
<td>93</td>
</tr>
</tbody>
</table>

These results show the benefits of our proposed handwritten digit-based scheme not only in terms of accuracy but also usability for real applications on mobile scenarios.

### 6. Conclusions

This work evaluates the advantages and potential of incorporating touch biometrics to mobile one-time passwords (OTP). The new e-BioDigit database which has been acquired comprising handwritten numerical digits from 0 to 9 is used in the experiments reported in this work and it will be made publicly available to the research community. Data was collected in two sessions with a time gap of at least three weeks between them for a total of 93 subjects. Handwritten numerical digits were acquired using the finger touch as the writing input on a Samsung Galaxy Note 10.1 general purpose tablet device.

For the new e-BioDigit database, we report a benchmark evaluation using our proposed digit-based system. The following three different experiments are considered: 1) a Baseline System comprised of a set of simple and fixed time functions for all numerical digits in order to make our work easily reproducible; 2) an study of the best features for each handwritten numerical digit through the SFFS algorithm on the development dataset; and 3) an analysis of the OTP system regarding which are the most discriminative handwritten digits and how robust the system is when increasing the number of digits included in the OTP.

Our proposed approach achieves remarkable results with EERs ca. 5.0% when using skilled forgeries, outperforming other traditional biometric verification traits such as the handwritten signature or graphical password on similar mobile scenarios.

Future work will be oriented to investigating how the different discriminative performance shown by individual digits can be exploited to design robust passwords, i.e., the OTP Digit Selection module in Fig. 1. Additionally, the integration of individual digits into a combined biometric decision [4] is subject of further investigation, i.e., the Biometric Comparison module in Fig. 1. The core matcher in that module can be also improved following recent advances from the machine learning community exploiting deep learning for on-line handwriting biometrics [18, 20].

### Acknowledgments

This work has been supported by project TEC2015-70627-R MINECO/FEDER and by UAM-CecaBank Project. Ruben Tolosana is supported by a FPU Fellowship from Spanish MECD.
References


