PIN-based cancelable biometrics

Patrick Lacharme, Aude Plateaux

PIN-based cancelable biometrics

Patrick Lacharme, Aude Plateaux
GREYC Research lab,
Ensicaen - UCBN - CNRS
6 Boulevard Maréchal Juin, 14000 Caen, France.
Aude Plateaux is also with BULL SAS, France.
patrick.lacharme@ensicaen.fr, aude.plateaux@ensicaen.fr

Abstract

Biometric authentication systems are more and more deployed in replacement of traditional authentication systems. Security of such systems is required in real-world applications and constitutes a major challenge in biometric field. An efficient approach of this issue is realized by cancelable biometrics. However, the security of such biometric systems is often overestimated or based on restrictive assumptions. This paper presents and investigates a PIN-based variant for cancelable biometrics applied to fingerprint, achieving 0-EER and templates diversity.

1 Introduction

Biometric authentication mechanisms are widely used in many security systems, with a large diversity of applications such as e-government or e-commerce. Biometric data are considered as unique for each person and are directly related to its owner. Biometric technologies are consequently supposed to provide stronger authentication than traditional mechanisms or can be deployed in complement of them. Nevertheless, these biometric characteristics are personal data and constitute very sensitive data. Moreover, original biometric data can not be revoked if they are stolen or compromised and this weakness is a critical threat for privacy. Recent results propose a new approach of this problem, using a new biometric template for each applications, without possibilities to recover the original template. This technique, called cancelable biometrics, increases the security and conveniance for users of such biometric systems.

Biometric authentication is traditionally subject to two types of errors, false acceptation and false reject [JRP04]. Performances of such systems are generally described by the two following parameters. The False Acceptance Rate (FAR) relates the possibility for an impostor to be authenticated by the authentication system, whereas the False Rejection Rate (FRR) reports the possibility for a genuine person to be rejected. The level of these parameters is a trade-off, because the reduction of the FAR would naturally increase the FRR, and conversely. Additionally, when these rates are equal, the corresponding rate is called Equal Error Rate (ERR), with: ERR = (FAR+FRR)/2.

Thus, perfect biometric systems should have ERR equal to zero, whereas a system with a high ERR could not be used in an operational authentication.

Principle of cancelable biometrics was introduced by Ratha et al. in [RCB01] and later developped with details in many papers, as [TN05], [NNJ10a], [RCCB07] or [KTT10]. It attemps to propose an alternative to classical biometric system, for privacy-preserving issues. The approach of this method is to use only cancelable templates for authentication and storage. These cancelable templates are computed from the original biometric template and an additional random value called seed. Consequently, if the cancelable template is leaked or stolen, a new template is computed from the original template and from a new random seed. Moreover, cancelable biometrics allows the generation of different templates for a person, for different applications, with the generation of a new seed. This randomized transformation should verify several criteria as cancelability, diversity and non-invertibility [TNG06], [JNN08]. The randomized transform is one-way, consequently, this technique achieves the security of the original biometrics template. It should be impossible to recover the original biometric template with the knowledge of the cancelable template and the seed (non invertibility criteria). Moreover, recognition performance must not be reduced by this mechanism and must achieve low error rates. Cancelable biometrics can also be combined in hybrid systems with fuzzy vault [NNJ10b] or fuzzy commitments [BCK08].

During the verification procedure, cancelable biometrics technique is able to realize very low error rates, comparing to traditional biometric approach. Thus, Teoh et al. [TNG04] perform a biohashing algorithm on fingercode and achieve a 0-EER in a model where the random seed is not compromise. However, the knowledge of the random seed by an impostor is a realizable assumption, because the seed is used during the verification procedure and is generally stored with the cancelable template. The theft of the random seed and its use by an impostor is a classical assumption for the security analysis of cancelable biometric systems. In this situation, the EER is much higher, as presented by Kong et al. [KCZ+06], Teoh et al. [TKL08] and Lumini and Nanni [LN06].

This paper proposes and investigates an alternative to the use of cancelable biometrics for fingerprint template applications, with an additional random value (e.g. a PIN code or a password). Introduction of a secret key or a password, instead of the random seed, was already proposed by several authors as [ASNM05] and [NNJ10a], but the security of the system is only based on an additional device (e.g. token, smart card), as for the random seed. Our solution uses a traditional random seed and a PIN code, without assumptions on a secure storage. The PIN code is known by the user, and is only used as an additional security in a scenario where the random seed is stolen and used by an impostor, without the knowledge of the PIN code. This secret is combined with a pseudorandom number generator or a cryptographic hash function and the system achieves 0-EER, cancelling the statistical advantage of an impostor with the knowledge of the random seed on previous cancelable biometric systems.

This article is organized as follows: Section 2 describes the procedure for the construction of a biocode from a fingercode. Two classical attack scenarios: with and without the knowledge of the random seed are presented in Section 3. Finally, Section 4 proposes and evaluates a PIN-based (or password-based) countermeasure for an impostor knowing the random seed.

2 Cancelable biometrics for fingerprint templates

General process of biometric authentication mechanism requires at high level two operations: Enrollment phasis, when a user is added to a biometric system with acquirement and pre-processing of a biometric template. Verification phasis, when a user presents a new biometric data, attempting to be compared with the template stored during the enrollment phasis.

The biohashing enrollment mechanism requires two steps. In a first time, the biometric feature is computed in an extraction and discretization process for the enrollment procedure. In a second time, this biometric feature is transformed in a template, called biocode, in a Biohashing process, using a one-way function. This process uses a random seed, which should be stored, in order to be re-used during the verification phasis. Furthermore, this transformation is randomized and every enrolled biometric feature uses a different function, in order to create a specific biocode. If this template is compromised, a new biocode is created with a new random seed.

More precisely, the initial biometric feature F is represented by a vector of length n with real numbers values. As for the biocode, it is computed from the random seed and from the biometric feature in a four-step process, as described in [TNG04] and presented in the following simplified way :

1. For $i = 1, \ldots, n$, generate n pseudorandom vectors v_i of length n, called pseudorandom matrix, from the random seed.

2. Use the Gram-Schmidt algorithm on the n vectors v_i into n orthonormal vectors $r_1, \ldots, r_n$.

3. For $i = 1, \ldots, n$, compute the n scalar products $p_i =< F, r_i >$ between the biometric feature F and the n orthonormal vectors r_i .

4. Compute a n-bit biocode $B = (b_0, \ldots, b_n)$ using the following threshold :

$$b_i = \begin{cases} 0 & \text{if} \quad p_i < \tau \\ 1 & \text{if} \quad p_i \geq \tau, \end{cases}$$

   where τ is a given threshold.

The verification procedure uses the same random seed, as for the enrollment phasis. Consequently, this latter need to be stored with the biocode template computed in the enrollment phasis. A new biometric feature is extracted for the calculation of a new biocode, using the same algorithm as previously.

Thus, the resulting biocode is compared with the Hamming distance of the original biocode, computed during the enrollment procedure. If the Hamming distance between the two biocodes is lower than a given acceptance rate R (with $0 \leq R \leq n$), then the new biocode is accepted. Otherwise, the biocode is rejected. Therefore, the value of the acceptance rate R is directly related to FAR and FRR rates. Indeed, if the acceptance rate R is low, then the probability that a genuine biocode is rejected is high. Converselly, if R is high, then the probability that an impostor is accepted is high.

R	12	16	20	24	28	32	36	40
FRR	38.9	18	6.8	2.1	0.6	0.2	0	0
FAR	0	0	0	0	0	0	0	0

Table 1: FRR and FAR without knowledge of the random seed

R	12	16	20	24	28	32	36	40
FAR	7	22	43	65	83	93	98	99

Table 2: FAR with knowledge of the random seed

3 Basic experiments on fingerprint

This experiment uses the fingerprint database with 800 images of 100 fingers (eight images per finger) issues from the public-domain FCV2002 database [FCV02]. For each of the 100 fingers (or persons) the first template is used for the enrollment phasis and the seven others for the verification phasis. All impressions of the FCV2002 database are used in our experiments, unlike to other experiments realized on the same database, as [LN06] or [BRA10]. A general reference on fingerprint recognition is given by the book of Maltoni et al. [MMJP09].

In this paper, a classical procedure for the minutiae-based extraction procedure, using Gabor filters [JPHP00], is adopted. Characteristics of these minutiae are stored in a feature vector with n components, composed of real numbers. This feature vector is called fingercode. In this experiment, a fingercode of size $n = 128$ is extracted, where each components are real numbers in the range $[-128, 128]$. Consequently, a pseudorandom matrix composed of 128 pseudorandom vectors of length 128 is generated from the random seed, during the biocode generation (step 1). The threshold τ, used at step 4, is 0.

The first simulation gives the FAR and FRR in a scenario where the random seed is unknown. It is realized with several acceptance rate $R : 12, 16, 20, 24, 28, 32, 36$ and 40, which means an acceptance rate R between 9.3% and 31% for a feature vector of length 128. Without knowledge of the random seed and with these acceptance rates, the FAR is zero. For an acceptance rate $R = 36$ or 40, the FRR and consequently the EER is 0, as claimed in [TNG04]. In the Table 1, the detailed FRR are presented for several rates R.

The second simulation investigates the scenario where the impostor has the entire knowledge of the random seed. In this case, the FAR is not null and is sometimes high. The random seed is used during the verification phasis. Consequently, without additional protections, it is not realistic to suppose that this seed is secret. Clearly, the FRR does not depend to the knowledge of the impostor, but only to the acceptance rate R and is the same as presented in Table 1. In this scenario, FAR are given in Table 2 for the same acceptance rates as previously and EER is around 20.

Type of password	(1)	(2)	(3)	(4)	(5)
Entropy	13	19	38	26	52

Table 3: Entropy (in bits) of PIN code and passwords

4 PIN-based experiments on fingerprint

In this section, we realize a biometric authentication with an additional secret: a PIN code or different types of passwords. This secret value is considered as known only by the user. Consequently, it can not be stolen or compromised by an impostor because it is not stored with the biometric template or anywhere else. This additional secret is completly independant to the random seed and is not used for the generation of the seed or the random matrix. It is introduced and used inside the biohashing algorithm, in addition to the random seed in order to improve the security of the biohashing mechanism, when the seed is compromise by an impostor.

The first experiment directly performs the bit wise addition modulo 2 of the password and the biocode. Let $p = (p_1, \dots p_l)$ be the password described in a binary vector of length l and $b = (b_1, \dots b_n)$ the biocode obtained at the end of the step 4 of the biohash process. Then, we realize a fifth step in this process in order to have the new biocode $b' = (b'_1, \dots, b'_n)$ by

$$b'_i = \begin{cases} b_i \oplus p_i & \text{if} \quad 1 \leq i \leq l \\ b_i & \text{if} \quad l+1 \leq i \leq n. \end{cases}$$

Five types of password are considered in this first experiment: a PIN code with 4 digits (Type 1), a password with 4 or 8 letters (Type 2 or 3) and a password with 4 or 8 letters, ciphers or other characters (Type 4 or 5). Table 3 gives the entropy of these five passwords.

This simulation uses the same fingerprint FCV2002 database as presented in the previous section. It is realized for all acceptance rates R between 12 and 40 (corresponding to acceptance rates R between 9.3% and 31% for a feature vector of length 128). Clearly, the FRR does not change in this scenario from a classical authentication, because the PIN code or the password is known by the user. FAR are presented in Figure 1 for acceptance rates between 12 and 40.

The second experiment uses a PIN code of 4 digits with the pseudorandom number generator of L'Ecuyer [L'E99], producing a pseudorandom binary sequence with large period. This generator combines several linear feedback shift registers (LFSR) and is not cryptographically secure. The following generator, called LFSR113, has a period approximately of 2^{113}. It generates random 32-bits numbers and is described by the following source code in C language, given by L'Ecuyer, where z_1, z_2, z_3 and z_4 represent the current state of the generator:

```
unsigned long lfsr113(void){
  unsigned long b;
```

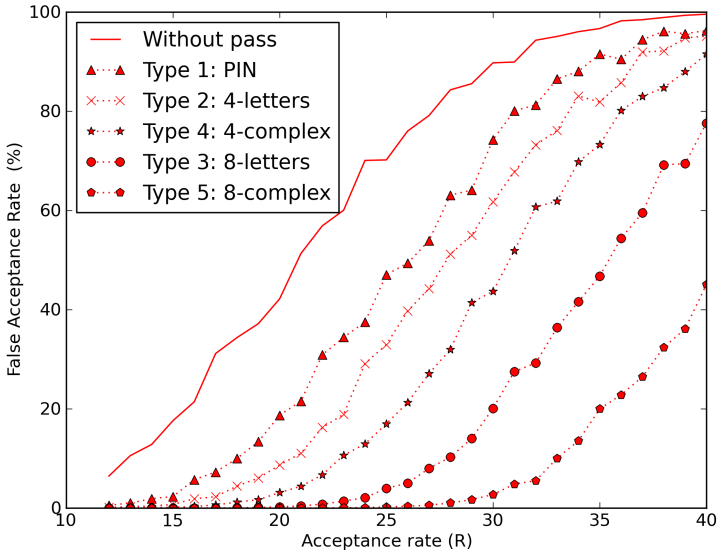

Figure 1: PIN-based and password-based cancelable biometrics

```
  b = (((z1 << 6) ^ z1) >> 13);
  z1 = (((z1 & 4294967294) << 18) ^ b);
  b = (((z2 << 2) ^ z2) >> 27);
  z2 = (((z2 & 4294967288) << 2) ^ b);
  b = (((z3 << 13) ^ z3) >> 21);
  z3 = (((z3 & 4294967280) << 7) ^ b);
  b = (((z4 << 3) ^ z4) >> 12);
  z4 = (((z4 & 4294967168) << 13) ^ b);
  return (z1 ^ z2 ^ z3 ^ z4);
}
```

The PIN-code is completed by zero's and is used as the initial state of the generator, where first outputs are deleted. Then, the n outputs of this generator are words of 32 bits, and p_i's are defined as the less significant bits of each output. Then, $p_1, \ldots, p_n$ are introduced in the biohashing process in the same way as previously: the new biocode $b' = (b'_1, \ldots, b'_n)$ id defined for $1 \leq i \leq n$, by

$$b'_i = b_i \oplus p_i.$$

Simulation is realized with acceptance rates R between 12 and 40 and the same fingerprint FCV2002 database, in the scenario where the random seed is known. In all cases, a FAR between 0 and 0.02 is reached, which corresponds to the same rate than with the scenario where the random seed is unknown, as presented in Table 1.

The combination of a PIN code of 4 digits with a pseudorandom number generator cancels the knowledge of the random seed by an impostor, because the advantage of this scenario is only statistical for him. Even if the size of the passwords space is not high (e.g. 10000 for a PIN code of 4 digits), an active impostor has to test each possibilities in order to come back in the situation of the classical attack scenario with knwon seed, which could be not possible even for a PIN code of limited size.

Other pseudorandom number generators, as proposed by the Estream European project for stream ciphers [eST], or a cryptographic hash function as Sha-1 or Sha-256 could be also used on the original PIN code, instead of the LFSR 113 generator.

5 Conclusion

Biometric authentication systems require additional security schemes for templates revocation and privacy preserving issues. This paper analyzes the security of cancelable biometrics using fingerprinting, in the scenario where the random seed is known. We propose to use an additional PIN code or a password, combined with a pseudorandom number generator or a cryptographic hash function as a countermeasure on this scenario. This mechanism achieves a 0-FAR and a 0-EER. A PIN code is not required to be stored somewhere, and consequently compromise. The choice and the type of passwords depends obviously of the application.

References

[ASNM05] R. Ang, R. Safavi-Naini, and L. McAven. Cancelable key-based fingerprint templates. In *ACISP in LNCS 3574 Springer*, 2005.

[BCK08] J. Bringer, H. Chabanne, and B. Kindarji. The best of both worlds: Applying secure sketches to cancelable biometrics. *Sciences of Computer Programming*, 74(1-2):43–51, 2008.

[BRA10] R. Belguechi, C. Rosenberger, and S.A. Aoudia. Biohashing for securing fingerprint minutiae templates. In *IAPR International conference for pattern recognition*, 2010.

[eST] eSTREAM, the ECRYPT stream cipher project.

[FCV02] FCV2002. Second international fingerprint verification competition, 2002. http://bias.csr.unibo.it/fvc2002/.

[JNN08] A.K. Jain, K. Nandakumar, and A. Nagar. Biometric template security. *EURASIP J. Advances in signal processing*, 8(2):1–17, 2008.

[JPHP00] A.K. Jain, S. Prabhakar, L. Hong, and S. Pankanti. Filterbank-based fingerprint matching. *IEEE Trans. Image Process*, 5(9):846–859, 2000.

[JRP04] A.K. Jain, A Ross, and S. Prabhakar. An introduction to biometric recognition. *IEEE Trans. Circuits Suts. Video Technol.*, 14(1):4–20, 2004.

[KCZ+06] A. Kong, K.H. Cheung, D. Zhang, M. Kamel, and J. You. An analysis of biohashing and its variants. *Pattern recognition*, 39(7):1359–1368, 2006.

[KTT10] Y.S. Kim, A.B.J. Teoh, and K.A. Toh. A performance driven methodology for cancelable face templates generation. *Pattern recognition*, 43(7):2544–2559, 2010.

[L'E99] P. L'Ecuyer. Tables of maximally-equidistribued combined lfsr generators. *Mathematics of computation*, 68(225):261–269, 1999.

[LN06] A. Lumini and L. Nanni. An improved biohashing for human authentication. *Pattern recognition*, 40(3):1057–1065, 2006.

[MMJP09] D. Maltoni, D. Maio, A.K. Jain, and S. Prabhakar. *Handbook of Fingerprint Recognition*. Springer (second edition), London, 2009.

[NNJ10a] A. Nagar, K. Nandakumar, and A.K. Jain. Biometric template transformation: A security analysis. In *Media Forensics and Security*, 2010.

[NNJ10b] A. Nagar, K. Nandakumar, and A.K. Jain. A hybrid biometric cryptosystem for securing fingerprint minutiae templates. *Pattern recognition letters*, 31(8):733–741, 2010.

[RCB01] N. Ratha, J. Connell, and R.M. Bolle. Enhancing security and privacy in biometrics-biased authentication systems. *IBM syst.*, 40(3):614–634, 2001.

[RCCB07] N. Ratha, S. Chikkerur, J. Connell, and R.M. Bolle. Generating cancelable fingerprint templates. *IEEE Transactions on pattern analysis and machine intelligence*, 29(4):561–572, 2007.

[TKL08] A. Teoh, Y. Kuan, and S. Lee. Cancellable biometrics and annotations on biohash. *Pattern recognition*, 41(6):2034–204, 2008.

[TN05] B.J.A. Teoh and C.L.D Ngo. Cancellable biometrics featuring with tokenised random number. *Pattern recognition Letters*, 26(10):1454–1460, 2005.

[TNG04] A. Teoh, D. Ngo, and A. Goh. Biohashing: two factor authentication featuring fingerprint data and tokenised random number. *Pattern recognition*, 37(11):2245–2255, 2004.

[TNG06] A. Teoh, D. Ngo, and A. Goh. Random multispace quantisation as an analytic mechanism for biohashing of biometric and random identity inputs. *IEEE Trans. Pattern Anal Mach. Intell.*, 28(12):1892–1901, 2006.