Biometric Cryptosystems based Fuzzy Commitment Scheme: A Security Evaluation

Maryam Lafkih¹, Mounia Mikram¹, ², Sanaa Ghouzali¹, ³, Mohamed El Haziti¹, ⁴, Driss Aboutajdine¹

¹LRIT (associated unit with CNRST, URAC 29), Mohammed V-Agdal University, Morocco
²The School of Information Sciences, Morocco
³Department of Information Technology, King Saud University, Saudi Arabia
⁴Higher School of Technology, Morocco

Abstract: Biometric systems are developed in order to replace traditional authentication. However, protecting the stored templates is considered as one of the critical steps in designing a secure biometric system. When biometric data is compromised, unlike passwords, it can’t be revoked. One methodology for biometric template protection is ‘Biometric Cryptosystem’. Biometric cryptosystems benefit from both fields of cryptography and biometrics where the biometrics exclude the need to remember passwords and the cryptography provides high security levels for data. In order to, develop these systems, Fuzzy Commitment Scheme (FCS) is considered as well known approach proposed in the literature to protect the user’s data and has been used in several applications. However, these biometric cryptosystems are hampered by the lack of formal security analysis to prove their security strength and effectiveness. Hence, in this paper we present several metrics to analyze the security and evaluate the weaknesses of biometric cryptosystems based on fuzzy commitment scheme.

Keywords: Biometric systems vulnerabilities, biometric cryptosystems, fuzzy commitment, security analysis.

Received August 18, 2013; accepted May 10, 2014

1. Introduction

Biometric authentication systems are developed in order to address the weaknesses of classical authentication mechanisms such as passwords and tokens. These biometric systems refer to physiological (face, fingerprint, iris, etc.,) and behavioral (signature, gait, etc.,) unique features of individuals. Biometric systems are based on two stages; enrollment where biometric features are extracted and stored in the database, and authentication where biometric features of the query are extracted and compared with stored features. However, storing biometric features as reference without any protection increases the security and the privacy risks. For example, if the database is compromised, it can be used by an adversary to gain unlawful access to user’s information and to the system [16]. Hence, in order to solve these problems, the protection of stored data is a necessity. To this aim, two principal methods are proposed in literature [13]: feature transformation and biometric cryptosystems.

Using feature transformation approach, in the enrollment stage, biometric features are transformed using a specific password to generate a reference template. The reference template is stored in the database instead of the original biometric features. In the authentication stage, biometric features are acquired and transformed using the same password, and then the result is matched with the stored reference template.

Using biometric cryptosystems, a secret key is associated with biometric features to obtain biometric information named ‘helper data’, which is stored in the database in the enrollment stage. In the authentication phase, the query data is used with the stored ‘helper data’ to generate the secret key for successful authentication. The inter-class variability prevents the direct extraction of the key; therefore, the ‘helper data’ allows reconstructing the key in the authentication procedure. Depending on the extraction mechanism of the ‘helper data’, biometric cryptosystems are classified as key binding or key generation systems. Key binding biometric cryptosystems obtain the ‘helper data’ by binding random key to the user’s biometric data; whereas, key generation biometric cryptosystems aim to derive the secret key from the biometric data. Hence, the ‘helper data’ is generated using the biometric data only. In this paper we investigate the security of biometric cryptosystems based on key binding.

There are two principle approaches of key binding systems: Fuzzy vault and fuzzy commitment. The fuzzy vault scheme, introduced by Ari and Madhu [1] aim to lock a key using the biometric features set A. In the enrollment phase, the ‘vault’ is created based on polynomial encoding. In the authentication phase, the secret is reconstructed if the query set A’ overlaps sufficiently with an enrollment set A. The Fuzzy Commitment Scheme (FCS) was developed by Ari and Wattenberg [2] and is considered as one of the first mechanisms of template protection. This method is based on Error Correcting Code (ECC).

In the enrollment stage, biometric features are extracted in real values format. Hereafter, binary
biometric features $X$ of length $n$ are created such that $X=(X_1, X_2, ..., X_n)\in\{0, 1\}^n$. Furthermore, a binary secret key $K$ is generated randomly and encoded using ECC into a codeword $C=(C_1, C_2, ..., C_n)$ of length $n$. XOR function is then applied to commit the features set with the codeword in order to create the ‘helper data’ $H=HXOR C$ that will be stored in the database. The hash of codeword $h(C)$ can also be stored in the database; in this case the commitment is the couple $(H, h(C))$ instead of the ‘helper data’ and the hashed codeword. In the authentication phase, query binary biometric features $X'$ are extracted such that $X'=(X'_1, X'_2, ..., X'_n)\in\{0, 1\}^n$. Next, the stored ‘helper data’ is Xored with the query features to generate the codeword $Z=H\oplus X'$. The authentication is successful if the query features are close enough to enrolled features. Figure 1 shows the biometric cryptosystem based FCS.

Figure 1. Biometric cryptosystem based FCS.

Biometric cryptosystems have several limitations such as: Poor performance [15] and the constraint of falsification [8]. A major challenge of biometric cryptosystems is the security analysis that allows comparing different systems. Lafkikh et al. [10] have discussed the key elements of the security in the key binding biometric cryptosystems. Lafkikh et al. [11] have proposed a security analysis framework for biometric cryptosystems based on the fuzzy vault scheme. In this paper, we propose a framework to evaluate the security of biometric cryptosystems based on the FCS.

The outline of this paper is as follows: In section 2, an overview of security analysis of FCS is briefly presented. Section 3 will reveal proposed scenarios of attacks and different metrics to evaluate the performance and security of biometric cryptosystems based on the FCS. Section 4 shows the results of the proposed framework. Finally, the conclusion and perspectives are drawn in section 5.

2. Security Analysis of the FCS: An Overview

Security analysis is an important element to compare different cryptosystems in several aspects such as usability and resistance against attacks. In literature, many papers discussed the security analysis of the FCS [17, 21, 22]. Rathgeb and Uhl [17] discussed the key elements of the security in biometric cryptosystems. Zhou et al. [22] studied the security of the FCS based on 3D face systems using information-theoretical measures. Their work focused on measuring the security and the privacy using the entropy to evaluate the independence and the distribution of biometric features. Zhou et al. [21] investigated the distribution within iris codes and disclosed their Markov property in order to show the importance of the independence of iris features. The results of [21, 22] showed that features dependency affects the security of the FCS.

Nagar et al. [14] measured the security of the FCS based on multi-biometric cryptosystems and then studied the security taking into account the distribution of biometric features and the estimation to break a ‘helper data’ using the entropy. Ignatenko and Willems [7] discussed the secrecy of the FCS based on the information leakage and the maximum secret-key rate. Wang et al. [20] proposed a security framework to provide a comparative information-theoretical analysis of two methods, secure ‘helper data’ and the FCS, using False Acceptance Rate (FAR), percentage of adversaries accepted by the system, False Reject Rate (FRR), percentage of users rejected by the system, Successful Attack Rate (SAR), percentage of false detection when an adversary is enhanced with knowledge of some combination of stored data, biometric features and the key and the leaked information about the user [7]. Rathgeb and Uhl [18] applied statistical attack on iris-FCS. This attack is based on running error correction codes in a decoding mode in order to produce the nearest codeword [19]. The decoder corrects more errors, resulting in decreased FRR and increased FAR. Kelkboom et al. [9] tested cross matching attack using fingerprint features. They discussed cross matching attack based on exhaustive search approach which consists on determining whether two protected templates (of different systems generated from the same biometric trait) belong or not to the same user.

3. Proposed Security Analysis Framework of Biometric Cryptosystems based FCS

Previous studies on security analysis are mostly based on information-theoretical measurements (such as: Entropy and leakage rate) which are difficult to estimate in the case of unknown biometric features distribution. Hence, our contribution is to offer simple, yet theoretically and practically detailed and rigorous, security evaluation framework. To this end, in this paper we define different scenarios of attacks that can be launched against biometric cryptosystems based FCS including intrusion, correlation, combination and injection attacks and we propose several criteria to evaluate the performance and the security strength of these biometric cryptosystems.
3.1. Evaluation of Performance

In order to evaluate the performance of biometric cryptosystems we use the FAR, which indicates the percentage of attackers who have gained access to the system and FRR, which indicates the percentage of users who have been rejected by the system.

3.1.1. Evaluation of the Original Biometric System

In the original biometric system, FRR\(\_O\) is the probability that the Euclidean distance \(D\) between the user template stored as reference \(X_{R(U)}\) and the query features \(X_{R\_Query(U)}\) is superior or equal to a threshold \(\epsilon\).

\[
FRR\_O(\epsilon) = P(\text{D}(X_{R(U)}, \; X_{R\_Query(U)}) \geq \epsilon)
\]  

(1)

FAR\(\_O\) is the probability that the distance \(D\) between the user template stored as reference and the adversary features \(X_{R\_Query(U)}\) is lower than a threshold \(\epsilon\).

\[
FAR\_O(\epsilon) = P(D(X_{R(U)}, \; X_{R\_Query(U)} < \epsilon)
\]  

(2)

3.1.2. Evaluation of the Biometric Cryptosystem based on the FCS

In the biometric cryptosystem based FCS, the ‘helper data’ is stored in the database instead of the original biometric features of the user. FRR\(\_FC\) is the probability that the Hamming distance \(D_H\) between the stored ‘helper data’ \(H\) XORed with the user’s query binary features \(X_{R\_Query(U)}\) and the enrolled codeword \(C\) is superior or equal to a threshold \(\epsilon\).

\[
FRR\_FC(\epsilon) = P(D_H(X_{R\_Query(U)}) \oplus H, \; C) \geq \epsilon)
\]  

(3)

On the other hand, the FAR\(\_FC\) is the probability that the distance \(D_H\) between the stored ‘helper data’ \(H\) XORed with the adversary's binary features \(X_{R\_Query(U)}\) and the enrolled codeword \(C\) is lower than a threshold \(\epsilon\).

\[
FAR\_FC(\epsilon) = P(D_H(X_{R\_Query(U)}) \oplus H, \; C < \epsilon)
\]  

(4)

3.2. Analysis of the FCS Security

In order to evaluate the security strength of the FCS against several threats including intrusion, correlation, combination and injection, we have defined different metrics.

3.2.1. Intrusion Threat

The adversary tries to access a system \(S_1\) based on the information of another system \(S_1\) (the ‘helper data’ and the key), assuming that both systems use the same biometric modality. The adversary can generate biometric features of \(S_1\) (i.e., \(EX_{R\_Query(U)}(C) = XOR(encode(K_{S1}), H^{S1})\)) and use them to access to the second system \(S_2\) (assuming the user is enrolled in both \(S_1\) and \(S_2\)). We name this criterion Cryptosystem Intrusion Rate in Different system (CIRD) and measure it by the probability that the distance \(D_H\) between the ‘helper data’ of \(S_2\) \((H^{S2})\) XORed with the estimated query features of \(S_1\) \((EX_{R\_Query(U)}(C))\) and the enrolled codeword \(C\) is lower than a threshold \(\epsilon\). The formula is as follows:

\[
CIRD_{FC}(\epsilon)=P(D_H(X_{R\_Query(U)}(C)), \; C < \epsilon)
\]  

(5)

3.2.2. Correlation Threat

Nagar et al. [13] proposed cross matching attack in order to determine whether two ‘helper data’ are generated from the same user. In our study, correlation attack has as objective to link different ‘helper data’ of different systems to estimate the biometric features or the secret key used in the protection process (both systems use the same biometric trait). We can evaluate the vulnerability of the FCS to this attack by the probability that the distance \(D_H\) between different ‘helper data’ is lower than a threshold \(\epsilon\):

\[
CR_{\_FC}(\epsilon) = P(D_H(H^{S2}, H^{S1}) < \epsilon)
\]  

(6)

We assume that the adversary knows both ‘helper data’ \(H^{S1}\) and \(H^{S2}\) of both systems \(S_1\) and \(S_2\). Using correlation attack, the adversary can estimate the distance between both biometric features of the user in both systems. Hence, he/she can retrieve the original biometric features using any codeword and choose the codeword with minimal distance. If the adversary knows the secret key \(K_{S1}\) and a codeword \(C\) such as \(C = XOR(K_{S1}, C_{S2})\), he/she can recover the secret key \(K_{S1}\) using \(K_{S1}\) and \(C\) (i.e., \(C_{S2} = XOR(K_{S1}, C\)) where \(C_{S1} = encod(K_{S1})\) and \(C_{S2} = encod(K_{S2})\)). Hence, the adversary can recover the original biometric features of the second system (i.e., \(X^{S2} = XOR(H^{S2}, C_{S2})\)).

If the codeword \(C\) is not known, we have the same scenario as CIRD attack, but in this case the goal of the adversary is to recover the biometric features or the secret key of the second system \(S_2\). Then, he/she can make the correlation between both ‘helper data’ and estimate the distance between both biometric features \(X^{S1}\) and \(X^{S2}\) where \(X^{S2}\) can be estimated based on searching the nearest codeword \(C_{S2}\) (i.e., \(D_H(X^{S1}, \; XOR(H^{S2}, C_{S2}))\) is minimal). If \(K_{S1}\) and \(K_{S2}\) are unknown and cannot be estimated, the adversary can estimate only the distance between the original biometric features of both systems without retrieving their real values.

3.2.3. Combination Threat

In this attack, we assume the adversary knows part of the user biometric features, and then he/she extracts part of his/her own biometric features in order to complete the biometric vector used in the system. We consider also the case when the adversary completes the biometric vector using the ‘helper data’ instead of his/her own biometric features in order to have a high probability of acceptance. We calculate this probability using the distance \(D_H\) between the ‘helper data’ of the system XORed with the combined biometric features \(X_{R\_Query(U)}(C)\) and the enrolled codeword \(C\) is inferior to a threshold \(\epsilon\) as follows:
\[ C_{A,F}(\varepsilon) = P(D_f(X_{B,Query}(A)), C < \varepsilon) \] (7)

3.2.4. Injection Threat

The adversary can also inject his/her biometric features in the database in order to be accepted by the system (the adversary replaces the stored ‘helper data’ by a falsified ‘helper data’ that contains the injected elements). We suppose that the adversary replaces some elements of the ‘helper data’ by his/her own features \( H(X, A) = \text{replace}(H_{uc}, X_{uc}) \). We measure this criterion based on the probability that the distance \( D_{ij} \) between the ‘helper data’ created by the adversary containing injected data \( H(X, A) \) Xored with the adversary query binary features \( X_{B,Query}(A) \) and the enrolled codeword \( C \) is lower to a threshold \( \varepsilon \).

\[ IA_{FC}(\varepsilon) = P(D_f(X_{B,Query}(A)), C < \varepsilon) \] (8)

4. Experimental Results

To evaluate the proposed security analysis framework we used face biometric authentication systems based on different features extraction approaches. Laplacian Smoothing Transform (LST) [5] combined with Linear Discriminant Analysis (LDA) are used in the first system and LST combined with Support Vector Discriminant Analysis (SVDA) [4] are used in the second system [12]. Experiments are conducted using yale face database considering 44 adversaries images (4 identities and 11 images per identity), 66 users test images (11 identity and 6 images per identity) and 55 users reference images (11 identity and 5 images per identity). In the enrollment phase, firstly, for both systems biometric features vectors are extracted of a size 30 (i.e., \( n = 30 \)). Secondly, FCS based on Reed Solomon Error Correcting Code (RS-ECC) [6] is used to secure both systems. We generated random secret keys \( K \) with different sizes in order to create several ECC capabilities\(^1\) (i.e., we varied parameters of RS-ECC \( (n, K) \) to obtain several tolerance values between the enrolled biometric features and the query biometric features).

The ECC capability becomes higher where the secret key has lower length. The secret key \( K \) is encoded using RS-ECC encoder to a codeword \( C \) of size \( n \) (\( n \) is also, the size of biometric features). Biometric features are binarized using the median of biometric features followed by thresholding [8]; thereafter, \( XOR \) function is applied to construct the ‘helper data’ \( H \) from the codeword and the binary biometric features. The ‘helper data’ is then stored in the database. In the authentication phase, query features are extracted and then binarized using the same process used in the enrollment phase. Stored ‘helper data’ \( H \) is then Xored with the query binary features in order to generate a codeword \( Z \). The authentication is successful if \( Z \) can be corrected to \( C \) using the RS-ECC. The proposed criteria are applied to evaluate the performance and the security of FCS biometric cryptosystems. In the security analysis framework, ECC capability is used as a threshold in order to plot different curves resulted by varying RS-ECC parameters \( n \) and \( K \).

4.1. Performance Measurement

In order to compare the overall performance, Receiver Operating Characteristic (ROC) [3] curves are obtained by computing the performance of systems in multiple operating points based on variation of FAR and FRR with tolerance values; hence, the decisions for FAR and FRR depend on the choice of a threshold \( \varepsilon \). In Figure 2-a we notice that the performance of the system based SVDA is better compared to the system based LDA which is confirmed in [4]. The performance of the biometric cryptosystems based on the FCS is considerably degraded compared to the original domain as shown in Figure 2-b due to the use of ECC. In particular, the performance of the FCS biometric cryptosystem based on LDA is decreased compared to the SVDA protected system.

4.2. FCS Security Analysis Framework

Figure 3 presents the CIRD curve. The adversary uses the user’s features stolen from a first system and tries to gain access to a second system (both systems use the same modality). We remark that the CIRD rate is increased in accordance with the threshold values; hence the adversary is rejected by the system if the error correction capability is minimal. The breach can be explained by the number of errors in the codeword (generated by the adversary using the ‘helper data’ of the first system and the ‘helper data’ of the second system) that is greater than ECC capability. The intra-

\(^1\)ECC capability is the number of errors that the ECC tolerates.
class variability also plays an important role to prevent this attack from being successful.

In correlation attack the adversary links two ‘helper data’ of two different systems using the same modality of the same user (i.e., face). As shown in Figure 4, the adversary can easily link both systems as the system can correct the distance between both ‘helper data’ if the capability of ECC is equal or superior to 11.

Figure 5 shows the combination attack where the adversary can randomly combine both biometric features without prior knowledge of the system. We considered also the case when the adversary knows the conception of the biometric system and then tries to use the correlation of biometric features in order to identify the order of insertion of the falsified data. Instead of combining the real values of biometric features with the adversary’s features, we propose another way to combine both data when the adversary uses the ‘helper data’ $H$ of the user along with the known part of biometric features in order to find the closest codeword in the authentication and then gain access to the system with a high probability.

We remark that even if the threshold is minimal (equal to 7), the adversary can have access to the system using combined features with a probability of 20%. Hence, the codeword generated using the ‘helper data’ $H$ and the features created by the adversary $X_{B_{\text{query}}(i+1)}$ contain a number of errors inferior to the ECC capability. The adversary can then successfully have access to the system when the capability of ECC is superior or equal to 11 if the combination is done randomly. In the case of the correlation of biometric features, we remark that the adversary can have access with higher probability compared to the random combination. The vulnerability is increased when the adversary uses the ‘helper data’ and biometric features in the combination, making it possible for the adversary to access the system with approximately 60% even if the threshold is minimal.

In the injection attack, we consider the case when the adversary submits fake biometric features in the database as shown in Figure 6. The first case taken into consideration is when the adversary injects randomly the fake biometric features. In the second case, we suppose that the adversary knows the conception of the biometric system and uses the correlation of the elements of the stored ‘helper data’ in order to determine the place of injection of the falsified biometric features. We remark that despite minimal threshold and random injection, the adversary can gain access to the system with a probability of 40%. The rate of the vulnerability can amount to 83% if the system has high correcting capability. The rate of vulnerability is increased especially if the adversary uses the correlation of ‘helper data’. Hence, the adversary can be accepted with the probability of 58% and a minimal threshold. In this attack, the biometric system can similarly refuse trusted users by the fact that the stored ‘helper data’ is modified compared to the enrollment process (stored ‘helper data’ also, contains the falsified biometric features injected by the adversary).

5. Conclusions

In this paper, we proposed a security analysis framework based on several scenarios of threats that can affect biometric cryptosystems and applied this analysis on FCS. Our study confirms theoretically and practically that cryptosystems based on FCS do not ensure high level of security or protection of privacy. As future work, different settings will be studied and
other metrics will be proposed to analyze the security level of different biometric cryptosystems.

Acknowledgment

This research project was supported by grant RGPV-VPP-157 from the Deanship of Scientific Research of King Saud University.

References


Maryam Lafkih is a PhD student in Engineering Sciences at Mohammed V-Agdal University, Morocco. She received the Master’s degree in 2011 from the same University. Her research interests include Biometrics and security. Her current work focuses on security in Biometric Cryptosystems.
**Mounia Mikram** is an Assistant Professor of computer Sciences and Mathematics at school of information Sciences, Rabat since 2010. She received a Master degree from Mohammed V-Agdal University (Rabat, Morocco) in 2003 and joint PhD degree from Mohammed V-Agdal University (Rabat, Morocco) and Bordeaux I University (Bordeaux, France) in 2008. Her research interests include pattern recognition, computer vision and biometrics security systems.

**Sanaa Ghouzali** received both the Master's and the PhD degrees in computer science and telecommunications from Mohamed V-Agdal University (Rabat, Morocco) in 2004 and 2009, respectively. She was a Fulbright visiting student at Cornell University (Ithaca, NY, USA) between 2005 and 2007. She was an Assistant Professor at ENSA (the National school of Applied Sciences), within the University Abdelmalek Essaadi (Tetuan, Morocco), between 2009 and 2011. In 2012, she joined the College of Computer and Information Sciences at King Saud University (Riyadh, Saudi Arabia) where she is an Assistant Professor in the department of Information Technology. Her research interests include statistical pattern detection and recognition, Biometrics, Biometric Security and Protection.

**Mohamed El Haziti** received the Doctorat de 3ᵉ Cycle and the Doctorat d’Etat-es-Sciences degrees in image processing from Mohammed V-Agdal University (Rabat, Morocco) in 1997 and 2003, respectively. He is an Assistant Professor at Higher School of Technology (Sale, Morocco). His research interests include image, compression, watermarking and complex networks.

**Driss Aboutajdine** received the Doctorat de 3ᵉ Cycle and the Doctorat d’Etat-es-Sciences degrees in signal processing from Mohammed V-Agdal University (Rabat, Morocco) in 1980 and 1985, respectively. He joined the same university in 1978, first as an assistant professor, then as an associate professor in 1985, and full Professor since 1990, where he is teaching, Signal/image Processing and Communications. Over 30 years, he co-supervised more than 50 PhD theses and published over 300 journal papers and conference communications. He is associate editor and member of the editorial board of numerous international journals. He was elected member of the Moroccan Hassan II Academy of Science and technology on 2006 and fellow of the TWAS academy of sciences on 2007. He received several awards such: ”Chevalier de l’Ordre des Palmes Académiques” by the French Prime Minister. Currently he is the head of CNRST (Centre National pour la Recherche Scientifique et Technique), Rabat, Morocco.