How to Evaluate Transformation Based Cancelable Biometric Systems?

R. Belguechi, E. Cherrier and C. Rosenberger

GREYC Research Lab, ENSICAEN - CNRS – University of Caen, FRANCE

NIST International Biometric Performance Testing Conference 2012
Cancelable biometric systems

- Privacy by design biometric systems,
- Two approaches: crypto-biometrics and transformation based,
- Pionner article: Ratha et al., 2001,
- BioHashing, a popular algorithm: Teoh et al., 2004,
- Difficult to evaluate their security.
Cancelable biometric systems

- Privacy by design biometric systems,
- Two approaches: crypto-biometrics and transformation based,
- Pionner article: Ratha et al., 2001,
- BioHashing, a popular algorithm: Teoh et al., 2004,
- Difficult to evaluate their security.

Contributions

- Proposition of evaluation criteria for privacy and security compliance ⇒ extension of Nagar et al., 2010,
- Illustrations on fingerprints and finger knuckle prints,
- Definition of a Matlab toolbox for the evaluation of BioHashing based cancelable systems
Outline

1. BioHashing algorithm
2. Evaluation framework
3. Experimental results
4. Conclusion & perspectives
Outline

1. BioHashing algorithm
2. Evaluation framework
3. Experimental results
4. Conclusion & perspectives
**Figure 1:** General principle of the BioHashing algorithm
BioHashing algorithm

Random number generation

\[ \begin{bmatrix} r_{1,1} & \cdots & r_{1,m} \\ \vdots & \ddots & \vdots \\ r_{n,1} & \cdots & r_{n,m} \end{bmatrix} \]

M random vectors of dimension n

Gram-Schmidt Orthogonalisation

\[ \begin{bmatrix} o_{1,1} & \cdots & o_{1,m} \\ \vdots & \ddots & \vdots \\ o_{n,1} & \cdots & o_{n,m} \end{bmatrix} \]

Feature extraction of the biometric data

\[ \begin{bmatrix} f_1 & \cdots & f_n \end{bmatrix} \]

Projection on an Orthnormal basis

\[ \begin{bmatrix} o_{1,1} & \cdots & o_{1,m} \\ \vdots & \ddots & \vdots \\ o_{n,1} & \cdots & o_{n,m} \end{bmatrix} \]

\[ \begin{bmatrix} B_1 & \cdots & B_n \end{bmatrix}, B_i \in \{0,1\} \]

Quantification of the result
Properties

- Given the BioCode, the biometric raw data cannot be retrieved,
- Only the BioCode is stored,
- If the BioCode is intercepted, a new one can be generated,
- An individual can have many BioCodes for different applications,
- The BioHashing process improves performances.

Open questions for an attacker

- Is it possible to generate an admissible BioCode without the seed?
- Can we predict a BioCode given previous realizations?
- How different are two BioCodes generated from the same FKPcode?

⇒ Definition of an evaluation framework.
BioHashing algorithm

Properties
- Given the BioCode, the biometric raw data cannot be retrieved,
- Only the BioCode is stored,
- If the BioCode is intercepted, a new one can be generated,
- An individual can have many BioCodes for different applications,
- The BioHashing process improves performances.

Open questions for an attacker
- Is it possible to generate an admissible BioCode without the seed?
- Can we predict a BioCode given previous realizations?
- How different are two BioCodes generated from the same FKPcode?
  ⇒ Definition of an evaluation framework.
Outline

1. BioHashing algorithm

2. Evaluation framework
   - Overview
   - Notations
   - Efficiency
   - Non-invertibility
   - Diversity

3. Experimental results

4. Conclusion & perspectives
Overview

Security properties

- **Performance**: the template protection shall not deteriorate the performance of the original biometric system,
- **Revocability or renewability**: it should be possible to revoke a biometric template.
- **Non-invertibility or irreversibility**: from the transformed data, it should not be possible to obtain enough information on the original biometric data to forge a fake biometric template,
- **Diversity or unlinkability**: it should be possible to generate different biocodes for multiple applications, and no information should be deduced from their different realizations.

⇒ Definition of 8 evaluation criteria based on Nagar et al., 2010
**Notations**

Verification process

\[
R_z = 1_{\{D_T(f(b_z, K_z), f(b^*_z, K_z)) \leq \epsilon_T\}}
\]

(1)

Where:

- \( R_z \): decision result for the verification of user \( z \) using the cancelable system,
- \( D_T \): distance function in the transformed domain,
- \( f \): the feature transformation function,
- \( b_z, b^*_z \) represent the template and query biometric features of user \( z \),
- \( K_z \): set of transformation parameters,
- \( \epsilon_T \): decision threshold.
Efficiency property

$A_1$ evaluation criterion

\[ A_1 = 1 - \frac{\text{AUC}(\text{FAR}_T, \text{FRR}_T)}{\text{AUC}(\text{FAR}_O, \text{FRR}_O)} \]  \hspace{1cm} (2)

where:

- $\text{AUC}$: area under the ROC curve,
- $\text{FRR}_O$ is the false reject rate and $\text{FAR}_O$ is the false accept rate of the original biometric system (without any template protection),
- $\text{FRR}_T$ is the false reject rate and $\text{FAR}_T$ is the false accept rate of the cancelable biometric system (with template protection).

if $A_1 > 0$, the protection of the template improves the performance.
### A$_2$ to A$_5$ evaluation criteria

\[
FAR_A(\epsilon_T) = P(D_T(f(b_z, K_z), A_z) \leq \epsilon_T)
\]  

Where:

- $FAR_A(\epsilon_T)$: probability of a successful attack by the impostor for the threshold $\epsilon_T$.
- $A_z$: generated biocode by the impostor with different methods,
- We can consider $\epsilon_T = \epsilon_{EER_T}$ ($\epsilon_{EER_T}$: threshold to have the EER functioning point of the cancelable biometric system).
Non-invertibility property

A priori information used by the impostor

- **Zero effort attack (A\textsubscript{2})**: An impostor provides one of its biometric sample to be authenticated as the user \(z\): 
  \[ A_z = f(b_x, K_x), \]
Non-invertibility property

A priori information used by the impostor

- **Zero effort attack** \((A_2)\):
  An impostor provides one of its biometric sample to be authenticated as the user \(z\):
  \[ A_z = f(b_x, K_x), \]

- **Brute force attack** \((A_3)\):
  An impostor tries to be authenticated by trying different random values of \(A\):
  \[ A_z = A, \]
Non-invertibility property

A priori information used by the impostor

- **Zero effort attack** ($A_2$): 
  An impostor provides one of its biometric sample to be authenticated as the user $z$: $A_z = f(b_x, K_x)$,

- **Brute force attack** ($A_3$): 
  An impostor tries to be authenticated by trying different random values of $A$: $A_z = A$,

- **Stolen token attack** ($A_4$): 
  An impostor has obtained the token $K_z$ of the genuine user $z$ and tries different random values of $b$ to generate: $A_z = f(b, K_z)$,
### Non-invertibility property

#### A priori information used by the impostor

- **Zero effort attack (A2)**:  
  An impostor provides one of its biometric sample to be authenticated as the user $z$: $A_z = f(b_x, K_x)$,

- **Brute force attack (A3)**:  
  An impostor tries to be authenticated by trying different random values of $A$: $A_z = A$,

- **Stolen token attack (A4)**:  
  An impostor has obtained the token $K_z$ of the genuine user $z$ and tries different random values of $b$ to generate: $A_z = f(b, K_z)$,

- **Stolen biometric data attack (A5)**:  
  An impostor knows $b_z$ and tries different random numbers $K$ to generate: $A_z = f(b_z, K)$. 
A\textsubscript{6} evaluation criterion

\[ A_6 = \frac{1}{N} \sum_{z} \sum_{j=1}^{M} \max(I(f(b_z, K_z), f(b^j_z, K_z))) \]

\[ I(X, Y) = \sum_{x} \sum_{y} P(x, y) \log\left(\frac{P(x, y)}{P(x)P(y)}\right) \]

Where:
- \( b_z \): denotes the reference of the individual \( z \) in the database,
- \( b^j_z \): denotes the \( j^{th} \) test data of the individual \( z \) in the database,
- \( N \): the number of individuals in the database,
- \( M \): the number of generated biocodes for each individual,
- \( P \): the estimation of the probability.
A7 to A8 evaluation criteria

For each template of the genuine user:

- Generation of Q biocodes \( B_z = \{f(b_z, K_z^1), \ldots, f(b_z, K_z^Q)\} \) for user \( z \),
- Prediction of a possible biocode value by setting the most probable value of each bit given \( B_z \),
- Computation of equation (2).

\[ \implies A_7 \text{ value for } Q = 3 \text{ and } A_8 \text{ for } Q = 11 \]
A\textsubscript{7} to A\textsubscript{8} evaluation criteria

For each template of the genuine user:
- Generation of $Q$ biocodes $B_z = \{f(b_z, K_z^1), \ldots, f(b_z, K_z^Q)\}$ for user $z$,
- Prediction of a possible biocode value by setting the most probable value of each bit given $B_z$,
- Computation of equation (2).

$\Rightarrow$ $A_7$ value for $Q = 3$ and $A_8$ for $Q = 11$

Summary

The security and robustness of a cancelable biometric system are characterized by an eight-dimensional vector $(A_i, i = 1, \ldots, 8)$
Outline

1. BioHashing algorithm
2. Evaluation framework
3. Experimental results
   - Protocol
   - Robustness to attacks
   - Summary
4. Conclusion & perspectives
### Benchmark databases

- **PolyU FKP Database** Lin Zhang, 2009: 4 fingers of 165 volunteers, each individual has provided 12 images,

- **FVC2002 benchmark** Maio et al., 2002 (dB3): composed of 8 fingerprints (resolution 355 x 390 pixels) for 100 individuals.
Experimental protocol

Benchmark databases

- PolyU FKP Database Lin Zhang, 2009: 4 fingers of 165 volunteers, each individual has provided 12 images,
- FVC2002 benchmark Maio et al., 2002 (dB3): composed of 8 fingerprints (resolution 355 x 390 pixels) for 100 individuals.

Feature computation

Gabor descriptors
Size: 128 parameters (16 scales, 8 orientations)
Computation: single enrolment, Hamming distance verification
Robustness to attacks: fingerprint case

**Figure 2:** Analysis on fingerprints (FVC 2002)
Robustness to attacks: FKP case

Figure 3: Analysis on finger knuckle prints (POLY FKP)
Synthesis

- Evaluation is done on a functioning point,
- The more *a priori* information the attacker knows, the more the attack is efficient,
- It is possible to compare attacks (same algorithm and biometric data).

<table>
<thead>
<tr>
<th>Modalities</th>
<th>$A_1$</th>
<th>$A_2$</th>
<th>$A_3$</th>
<th>$A_4$</th>
<th>$A_5$</th>
<th>$A_6$</th>
<th>$A_7$</th>
<th>$A_8$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fingerprint</td>
<td>1.0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.44</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>FKP</td>
<td>0.10</td>
<td>0.25</td>
<td>0.15</td>
<td>0.54</td>
<td>0.25</td>
<td>0.58</td>
<td>0.51</td>
<td>0.59</td>
</tr>
</tbody>
</table>

**Table 1:** Evaluation results of the cancelable biometric systems.
Outline

1. BioHashing algorithm
2. Evaluation framework
3. Experimental results
4. Conclusion & perspectives
Conclusion & perspectives

Contributions

- Evaluation framework for cancelable biometric systems,
- Simulation of different attacks,
- Illustration on a FKP and fingerprint generic biometric system.
Conclusion & perspectives

Contributions
- Evaluation framework for cancelable biometric systems,
- Simulation of different attacks,
- Illustration on a FKP and fingerprint generic biometric system.

Perspectives
- More complex attacks
  ⇒ generation of the biocode based on the listening attack
  ⇒ impact of the random generator
http://www.epaymentbiometrics.ensicaen.fr/