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 $\Rightarrow$ 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

Outline

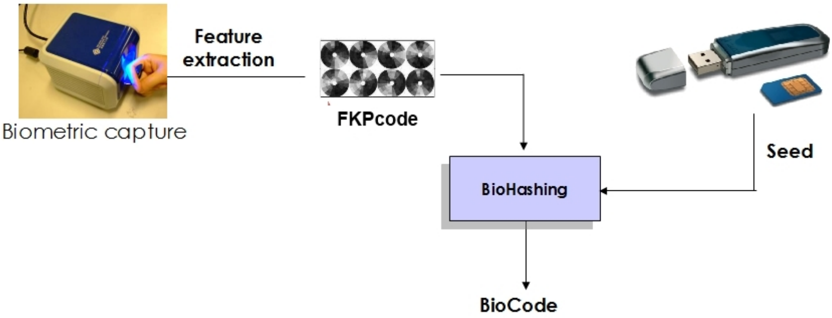
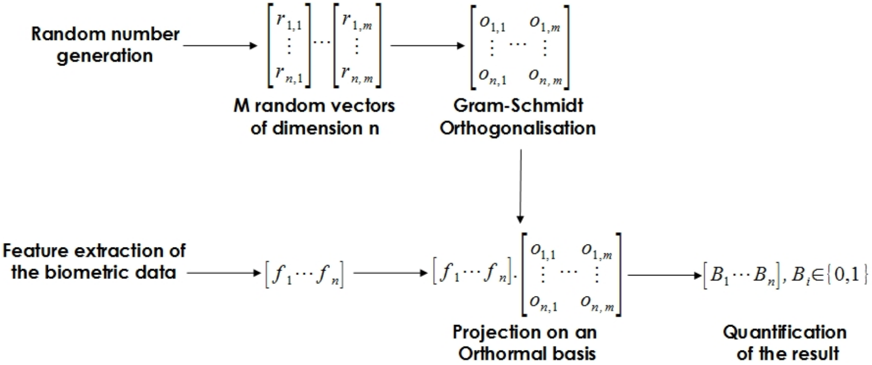


FIGURE 1: General principle of the BioHashing algorithm


Random number generation
M random vectors of dimension n
Gram-Schmidt Orthogonalisation
Feature extraction of the biometric data
Projection on an Orthormal basis
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.

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 ?

  $\Rightarrow$ Definition of an evaluation framework.

Outline

Security properties

- **Performance** : the template protection shall not deteriorate the performance of the original biometric system,
- **Revocability or renewability** : it shoud 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

Verification process

$$R_z = 1_{\{D_T(f(b_z,K_z),f(\acute{b_z},K_z))\leq\epsilon_T\}} \quad (1)$$

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, $\acute{b_z}$ represent the template and query biometric features of user z,
- K_z : set of transformation parameters,
- ϵ_T : decision threshold.

A_1 evaluation criterion

$$A_1 = 1 - \frac{\mathrm{AUC}(\mathrm{FAR_T}, \mathrm{FRR_T})}{\mathrm{AUC}(\mathrm{FAR_O}, \mathrm{FRR_O})} \quad (2)$$

where :

- *AUC* : area under the ROC curve,
- FRR_O is the false reject rate and FAR_O is the false accept rate of the **original biometric system** (without any template protection),
- FRR_T is the false reject rate and 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

$$FAR_A(\epsilon_T) = P(D_T(f(b_z, K_z), A_z) \leq \epsilon_T) \quad (3)$$

Where :

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

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 : $A_z = f(b'_x, K_x)$,

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 : $A_z = A$,

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 : $A_z = f(b, K_z)$,

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 : $A_z = f(b, K_z)$,
- *Stolen biometric data attack* (A_5) :
  An impostor knows b'_z and tries different random numbers K to generate : $A_z = f(b'_z, K)$.

A_6 evaluation criterion

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

$$I(X,Y) = \sum_{x}\sum_{y} P(x,y)\log(\frac{P(x,y)}{P(x)P(y)})$$

Where :

- b_z : denotes the reference of the individual z in the database,
- b_z^j : 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.

A_7 to A_8 evaluation criteria

For each template of the genuine user :

- Generation of Q biocodes $B_z = \{f(b_z, {K_z}^1), .., f(b_z, {K_z}^Q)\}$ 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 $Q = 11$

Diversity property

A_7 to A_8 evaluation criteria

For each template of the genuine user :

- Generation of Q biocodes $B_z = \{f(b_z, {K_z}^1), .., f(b_z, {K_z}^Q)\}$ 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 $Q = 11$

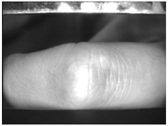
Summary

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

Outline

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.



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

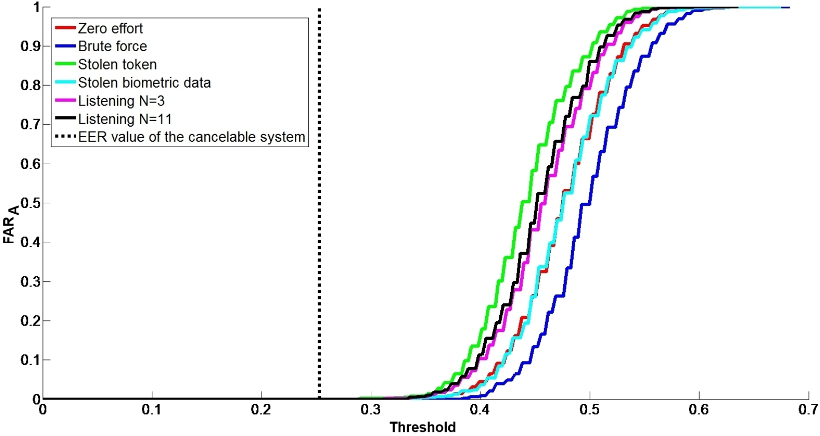


FIGURE 2: Analysis on fingerprints (FVC 2002)

Robustness to attacks : FKP case

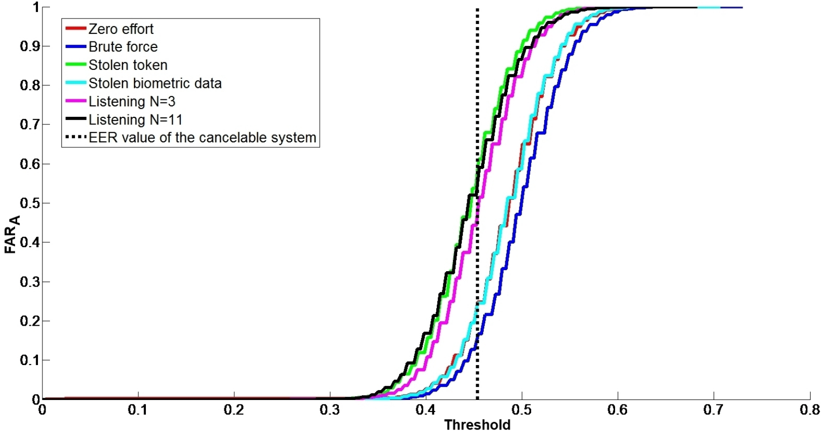


FIGURE 3: Analysis on finger knuckle prints (POLY FKP)

Synthesis

- Evaluation is done on a functionning 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).

Modalities	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8
Fingerprint	1.0	0	0	0	0	0.44	0	0
FKP	0.10	0.25	0.15	0.54	0.25	0.58	0.51	0.59

TABLE 1: Evaluation results of the cancelable biometric systems.

Outline

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
  $\Rightarrow$ generation of the biocode based on the listening attack
  $\Rightarrow$ impact of the random generator

Questions

http ://www.epaymentbiometrics.ensicaen.fr/