Better know your limits and adversaries

Julien Bringer
julien bringer (at) morpho.com
Better know your limits and adversaries

A practical view on various template protection and key binding schemes

This talk is based on several joint works with various co-authors, in particular Hervé Chabanne and Constance Morel from Morpho, and that have been partially funded by European FP7 projects FIDELITY and BEAT.
This talk is NOT about

- Classical *on-the-shelf* crypto
- Homomorphic encryption
- Cryptographic protocols (e.g. SMC, private retrieval)
- PET (e.g. *k*-anonymity, *l*-diversity, privacy protection of the link between ID & bio)
- HW-based solution
- Formal Models for PbD
- …

It is about

- Template Protection Schemes (TPS) or TPS-like
TPS principles come from both crypto and biometrics community

- Helper data, cancelable biometrics, biometric key, …
- FCS, FV, Code offset, SSK, FE …

Image courtesy of M. Favre
Secure Sketches (Dodis, Reyzin & Smith – 2004)

- SSK: secure sketch function
- Rec: correction function
- \( \text{Rec}(b', \text{SSK}(b)) = b \) if \( d(b, b') \leq t \)
Secure Sketches after binarization of biometrics

Concept introduced in late 90’s
PROBLEM SOLVED?

→ ...

- Need to find a representation compatible with TPS algorithm
  - Usually binary & fixed-length vector
- Correcting large amount of errors
- Finding nice trade-off between accuracy and security
- Impact of storage & computational cost on operational constraints

→ To date, still very important challenges: security vs performances vs use cases (functionality & cost)
one of the most accurate published solution but…

*Related to papers @ BTAS 2010, SPIE 2011 with V. Despiegel & M. Favre
FINGERPRINT EXAMPLE

<table>
<thead>
<tr>
<th></th>
<th>FRR@10^{-3} FA FVC 2002 DB2</th>
<th>FRR@10^{-3} FA FVC 2000 DB2</th>
</tr>
</thead>
<tbody>
<tr>
<td>one COTS</td>
<td>1.25 %</td>
<td>0.81 %</td>
</tr>
<tr>
<td>FV(Feature-Vector)-based</td>
<td>14.1 %</td>
<td>15 %</td>
</tr>
</tbody>
</table>

→ Accuracy drop of 1 order of magnitude

→ Usual size
  - of a template w/o TPS: 100-200B
  - w/ the FV representation: ~29kB
Issued ISO/IEC 24745:2011, Information technology — Security techniques — Biometric information protection

On-going ISO CD 30136, Information Technology — Performance Testing of Template Protection Schemes
TPS PROPERTIES: 101

Match?  
Yes/no

TPS

IBvEIV93RPlgtGkZsH3
uvZf63k8gKm

IBvEIV93RPlgtGkZsH3
uvZf63k8gKm

irreversibility

Image courtesy of Jens Hermans
TPS PROPERTIES: 101

TPS

JBvflV93RPI
j1GkZsH3uv
Zf63k8gKm

X

TPS

MNB8e35frjP
QPehukjs4SX
UAa2j7nn

Image courtesy of Jens Hermans
TPS PROPERTIES: 101

TPS w/ key

IBvElV93RP1gtGkZsH3
uvZf63k8gKm

Match?

Yes / No

Image courtesy of Jens Hermans
TPS PROPERTIES: 101

Also

- False Match Rate (FMR) / False Accept Rate (FAR)
- False Non-Match Rate (FNMR) / False Reject Rate (FRR)
- Failure-To-Enroll (FTE) Rate
- Failure-To-Acquire (FTA) Rate
- Successful Attack Rate (SAR)
- Accuracy Variation
- Template Diversity
- Storage Requirement per Registered User, speed…
THREAT MODELS (ISO 30136)

**Naive Model**
- No information, black box, no access to any biometric data.

**Collision Model**
- Adversary possesses a large amount of biometric data.

**General Models**
- Full knowledge of the underlying TPS
- **Standard Model**
  - none of the secrets.
  - related to known-ciphertext attack.
- **Advanced Model**
  - augmented with the capability of the adversary to execute part of or all submodules that make use of the secrets.
  - related to chosen-plaintext attack and chosen-ciphertext attack
- **Full Disclosure Model**
  - augmented by disclosing the secrets to the adversary (e.g. malicious insider)

**FA attack issue**
SOME PRACTICAL CONCERNS

=> With ECC based construction
  - Use of non-perfect codes => if one decodes, it is most probably that \( d(b,b') < t \)
  - \( \Rightarrow \) unlinkability attacks (Simoens et al. 2009)

=> FAR attack
  - Linkability issue
  - Pseudo-reversibility issue
    - With SSK construction, enables to retrieve \( b \)

=> Biometric data and errors between data may NOT be uniformly distributed
  - Can we do more?
  - Statistical attacks possible
Shuffling is not sufficient

*Related to IJCB 2014 Security Analysis of Cancelable IrisCodes based on a Secret Permutation with H. Chabanne & C. Morel*
USE OF APPLICATION-SPECIFIC TRANSFORM

- Cancelable biometrics / Ratha et al., 2001
- Application-specific bio / Cambier et al. 2002
- Also as user-specific secret, e.g. biohashing / Goh et al. 2004
- Also combined with other techniques, e.g. with fuzzy commitment scheme (Bringer et al. 2007, Kelkboom et al. 2011)
SHUFFLING ON IRIS

Images from Rathgeb & Uhl, A survey on biometric cryptosystems and cancelable biometrics. EURASIP J. of. Inf. Sec. 2011

→ **Iriscode**: 256-byte iris + 256-byte mask
  - Mask indicates (in)exploitable data: eyelids, eyelashes, blurred pixels...

$$score((I1, M1), (I2, M2)) = \frac{|(I1 \oplus I2) \cap M1 \cap M2|}{|M1 \cap M2|}$$

SHUFFLING ON IRIS

Acquisition

Permutation of bytes

Permutation of bits

Shuffling

Application-specific

Secure DB

Store in DB

$(\pi(I), \pi(M))$
SHUFFLING

- Naive Model, Collision Model
  - ok …

- Full Disclosure Model
  - NOK

- Advanced Model (execution)
  - FAR attacks
  - Statistics with know (plaintext or matching-plaintext, ciphertext) couples
    => good approximate of permutation

- Standard Model
  - ?
SHUFFLING IS NOT SUFFICIENT

→ **Same** transformation applied to the whole reference DB

→ **Biometric data are not uniformly random**
  
  ▪ Correlated bits
    

    ▫ For instance, on iris information part
      
      – Transition 0 → 0  \( \text{proba} > 0.40 \)
      – Transition 1 → 1  \( \text{proba} > 0.20 \)

  ▪ Non-random masks
ATTACK WITH STOLEN DB) 
(ON BYTE PERMUTATION)

-> **Method**
- Assign a probability of being neighbors for each couple of bytes

-> **Results :**
- Blue - Percentage of the permutation retrieved : 39% and Matching : 0.33
- Blue+Black - Percentage of the permutation retrieved : 58% and Matching : 0.20
Compression is neither sufficient

*Related to ICB 2015  Security analysis of Bloom Filter-based Iris Biometric Template Protection w/ C. Morel & C. Rathgeb
From Rathgeb et al.’s ICB 2013

→ Claimed properties (even with T public): unlinkability & irreversibility

→ Full Disclosure Model = Advanced Model = Standard Model
  - FAR attacks ⇒ linkability & pseudo-reversibility
  - Can we do more?
Unlinkability analysis

- Methods: $|BF(X, T^1)| = |BF(X, T^2)|$
- Results: 96% of success

Irreversibility analysis

- Methods: analysis based on uniformly random data
- Results: reconfirm Rathgeb et al.’s irreversibility security analysis

UNLINKABILITY ANALYSIS

\[ BF(X^1, T^1) \]

\[ BF(X^2, T^2 \oplus 0) \]

\[ BF(X^2, T^2 \oplus 1) \]

\[ BF(X^2, T^2 \oplus 1023) \]

DS

DS

DS

Min

Score
IRREVERSIBILITY ANALYSIS

- General irreversibility attack

\[ BF(X, T) \rightarrow \text{Irreversibility Analysis} \rightarrow \text{Information about X} \]

- Our irreversibility attack

\[ BF(X^1, T) \rightarrow \text{Irreversibility Analysis} \rightarrow X_{rec} \rightarrow HD_{rot}(X, X_{rec}) \]

\[ X^1, X^2 \text{ from the same iris} \]
MEAN COLUMN OF EACH BLOCK

Protected template $b_1, b_2, ..., b_K$

$\begin{bmatrix}
1 & 0 & \cdot & \cdot & \cdot & \cdot & 0 \\
0 & 1 & \cdot & \cdot & \cdot & \cdot & 1 \\
0 & 0 & \cdot & \cdot & \cdot & \cdot & 1 \\
\vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\
1 & 0 & \cdot & \cdot & \cdot & \cdot & 1 \\
1 & 0 & \cdot & \cdot & \cdot & \cdot & 0 \\
\vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\
0 & 1 & \cdot & \cdot & \cdot & \cdot & 0 \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot
\end{bmatrix}$

$|b_2| \text{ columns}$

$T=0$

Reconstructed iriscode $c_1^m, ..., c_1^m, c_2^m, ..., c_2^m, c_K^m, ..., c_K^m$
IRREVERSIBILITY ATTACK - EXPERIMENTATIONS

\[ BF(X^1, T) \rightarrow X^1_{rec} \]

\[ X^2 \]

\[ HD_{rot} \rightarrow \text{Score} \]

\[ T=0 \]

**Graph:**
- **Probability density** axis.
- **Hamming distances** axis.
- Lines for different values of \( L \): L=16, L=32, L=64.
- Lines for different values of FA: FA 10-3, FA 10-4.
Optimal security?
Goal: ensuring FAR attack = worst case situation

Seems realistic for error-correcting code (ECC) based TPS

One of our solutions

- Product codes
- + randomly permuted biometric binary vector (interleaving)
- + iterative soft decoding algorithm

Underlying idea: to tend toward the worst-possible FAR

- Use near-optimal decoding algorithm (vs Shannon)
- And use i.i.d. bits for messages or break correlations

*Related to IEEE TIFS 2008 Theoretical and Practical Boundaries of Binary Secure Sketches w/ H. Chabanne, G. Cohen, B. Kindarji & G. Zémor*
APPLICATION TO DIFFERENT MODALITIES

➢ Preliminary step:
  ▪ embedding into a Hamming space
  ▪ constraints: amount of errors, low FAR with usable FRR

➢ Almost direct for iris (cf. IEEE TIFS 08, BTAS 2007 w/ Chabanne, Cohen, Kindarji & Zémor)

➢ Works well for vein recognition
  ▪ (with specific dedicated alignment-based techniques) (cf. ICISP 2015 w/ Chabanne & Favre & Picard)

➢ Face: quite okay as fixed length feature vectors in Euclidean space

➢ Fingerprint still a challenge
APPLICATION TO KEY BINDING

 Goals: low FAR and « valuable » key length
FUSION FOR DECREASING FAR

- Fusion of binary vectors and application to fuzzy vault

Results on FVC2000 DB2

<table>
<thead>
<tr>
<th>Number of fused templates</th>
<th>FRR (%) at 0.01% FAR</th>
<th>Computational Security (bits)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>32.87</td>
<td>31</td>
</tr>
<tr>
<td>2</td>
<td>17.30</td>
<td>36</td>
</tr>
<tr>
<td>4</td>
<td>2.10</td>
<td>53</td>
</tr>
<tr>
<td>8</td>
<td>0.30</td>
<td>143</td>
</tr>
</tbody>
</table>

Security

- More costly FA attacks due to fusion
- High computational security in fuzzy vault setting
CONCLUSION

Design of TPS

- Need to take into account practical constraints
- Security analysis is a critical task in the design
  - FAR and Intrinsic properties of data MUST be taken in account
- Progresses in the last years on trade-offs between security vs accuracy/efficiency
  - Decreasing FAR for some modalities still desirable
- Applications to key generation

1st layer of protection for “stand-alone” use case

To be combined or replaced with more robust cryptographic techniques in a system-oriented approach