Biometrics - Trust But Test

Tomáš Rosa

crypto.hyperlink.cz
Biometric Identification/Verification

...automated establishment of a human identity based on their physical or behavioral characteristics.
Biometric System Topology

It would be nice if we had a simple true/false result.

As in conventional crypto.

But we cannot...

All we have is a value of random variable $X$ that follows two conditional distributions.

- $f(x \mid \text{impostor})$
- $f(x \mid \text{genuine})$
Base “Camel” Graph

![Graph showing two distributions: one for impostors and one for genuine attempts.](image-url)
Signal Detection Approach

Density distribution with parameters:
- Impostor distribution: $\mu = -20, \sigma = 15$
- Genuine distribution: $\mu = 40, \sigma = 30$
- Detection threshold: $\eta = 12$
False Match Rate

\[ \mu = -20, \sigma = 15 \]

\[ \eta = 12 \]

\[ \mu = 40, \sigma = 30 \]

FMR

\[ \int_{x > \eta} f(x \mid \text{impostor}) \, dx \]

\[ = 0.0164486958227453 \]
False Non-Match Rate

\[ \mu = -20, \sigma = 15 \]

\[ \eta = 12 \]

\[ \mu = 40, \sigma = 30 \]

\[ \text{FNMR} = \int_{x < \eta} f(x | \text{genuine}) \, dx \]

\[ = 0.175323944852229 \]
Error Distribution Functions

\[ \mu = -20, \sigma = 15 \quad \mu = 40, \sigma = 30 \]

Equal Error Rate
Receiver Operating Characteristics

imp: \( \mu = -20, \sigma = 15 \)
gen: \( \mu = 40, \sigma = 30 \)

EER \approx 0.09121122, \eta = 0
Detection Error Trade-Off

EER ≈ 0.09121122 for $\eta = 0$
ISO/IEC 19795

- Performance test methodologies for different life-cycle phases:
  - technology evaluation
  - scenario evaluation
  - operational evaluation

- We get comparable results with plausible confidence intervals.
Bunch of Parameters

- False Match Rate / False Non-Match Rate
  - attempt oriented
- False Acceptance Rate / False Rejection Rate
  - transactional version of FMR/FNMR
- Failure To Acquire
- Failure To Enroll
  - both attempt and txn-oriented versions
Biometric Data Mining

- In any life-cycle phase, we shall gather as much data as we can to estimate the performance or check we are still operating in expected margins.
- Anomalies may indicate a component malfunction or even a fraud.
- Again, be careful about confidence.
- Misleading statistics can be worse than none!
DET Estimation Simulation

ind. samples : 100
experiments : 1
Confidence Intervals?!
Any Confidence, Yet?

ind. samples: 1000
experiments: 150
Fair Confidence

ind. samples: 10000
experiments: 150
We Can be Proud

ind. samples: 1e+05
experiments: 150
Just a Dream…

ind. samples: 1e+06
experiments: 150
Biometric Menagerie

- To further complicate biometrics testing, those score distributions are usually *not* person-independent.
  - That means the performance is *not* the same for all people.
- There are plenty of anomalies we shall be aware of to interpret the system behaviour correctly.
Sheep: An Ordinary User
Goat: Problematic FNMR
Lamb/Wolf: Easy Target and-or Effective Predator
Worms: Both FNMR and FMR Increased
Dove: Excellent User
Chameleon: Excellent Scores, Anyway(!)
Phantom: Problematic Matching, Anyway
Secret Files on Biometrics
Reactive Forensics

I am solving criminals recognition and this just works...
Reactive Forensics

I am solving criminals recognition and this just works...

Alphonse Bertillon, 1883
Turned Into Proactive Security

...such a massive invocation of the hidden algorithm design!
Turned Into Proactive Security

...such a massive invocation of the hidden algorithm design!

Auguste Kerckhoffs, 1883
Contrasting Design Approach

- Classic cryptography
  - infeasible mathematical problems
- Quantum cryptography
  - intractable physical problems
- Biometric identification
  - statistical signal detection
  - intractability is usually *not* the prime concern
  - we hope the Mother Nature complexity *somehow* guarantees the security
Open Problems

- Convincing Algorithms
- Risk Assessment
- Template Revocation
- Liveness Detection
Convincing Algorithms?
Safe Template Revocation?
Internal Experts Are Ready
Consultants Always Eager to Help!

They fought like seven hundred
Anyway, do the Pentest!
BIO Brute Force Attack

- Randomly generate plausible circa 1/FMR samples and put them to the test.
  - Also termed “Zero-Effort”, denoting that the attacker makes no special effort to imitate the original person characteristic.
- Synthetic samples generation is quite feasible today.
Cryptanalysis-Like Attacks

- Masquerade attacks, can be a variant of “Hill-Climbing” denoting the attacker iteratively improves the BIO sample data based on:
  - scoring feedback *(side channels)*
  - stolen template *(pre-image attacks)*
  - independent template trained from intercepted BIO samples *(correlation attacks)*
  - known scoring anomaly *(differential analysis. etc.)*
  - implementation faults *(general hacking)*
Spoofing

- The process of defeating a biometric system through the introduction of fake biometric samples.  
  - (Schuckers, Adler et al., 2010)
- Particular modus operandi on how to deploy the attacking data vectors.
  - Can be seen as being orthogonal to the aforementioned ways of gaining fake samples.
Sensor-Bypass Attacks

- Do not expose API service for unrestricted automated sample verification!
- Recall the zero-effort attack complexity is often trivial.
- Furthermore, masquerade attacks can shift FMR significantly.
Sample Generator + Spoofing

- Spoofing techniques are, however, not “just helpers”.
- They are tightly interconnected with the fake samples generator to create complex attacks, as e.g.:
  - Text-To-Speech Synthesis
  - Voice Conversion
  - Artificial Signals
Conversion Attack Example

Kinnunen et al., ICASSP 2012
Reporting Attack Impact

Kinnunen et al., ICASSP 2012
Artificial Signals Impact

(b) FA system

Alegre et al., EUSIPCO 2012-13
Biometric Cryptography?
Cryptography Exactness

Let \( y = AES_K(x) \) for a random \( K \).

Then \( AES_K^{-1}(y) = x \), while
\[
AES_{K \oplus 1}^{-1}(y) \neq x \text{ (probability } \approx 1)\.
\]

- The better the algorithm is the more randomized response we get for even one-bit error.
Biometrics Fuzziness

- We seldom get the same data in the subsequent scans of the very same person.
- Actually, this is usually a clear sign of a spoofed sample.
- To overcome this (intra-user) variability, we can employ the biometric cryptography.
1. analyse the entropy gain from inter-class variation
2. use an error-correction code to cope with intra-class noise
1. Analyse the entropy gain from inter-class variation.
2. Use an error-correction code to cope with intra-class noise.

Claude Elwood Shannon, 1948-49
ISO/IEC 24745 Requirements

- **Renewability**
  - allows multiple independent *biometric references* created ad hoc
  - a particular leaked template does not compromise the other ones (provably!)

- **Revocability**
  - user can revoke the ability of being successfully verified by a particular template from now on

- Biocryptography is an effective way on how to achieve these goals.
Is It Enough?

- **Template protection in contemporary systems is often quite questionable (to be polite).**
- **On the other hand, is it the only one problem?**
  - **No.** We shall not push the concept of bio-keys too hard anyway.
One Key to Rule Them All…

- Conventional cryptographic keys can be freely discarded and re-generated from a scratch.
- There is no nature-wide master-key that would compromise all these keys at once (hopefully).
- On the other hand, for all your bio-keys, You are the “master” key!
Once the problem of template protection is solved, this will become a new attack vector.

- Attackers use a fake sensor (or hack into an original one) to skim the “bio-master-key”.
- At the end of the day, how many eyes, fingers, faces, vocal tracts (etc.) do we have?
- It is like having few master-keys for a whole life.
- Furthermore, we prove the master-key possession by simply handing it over to almost any device that asks so (again, again, and again…).
Spoofing Still Matters!

- That said, liveness detection will be always important!
  - Remember, biometrics is nothing but a signal detection.
  - It all works as long as we can assume the signal is coming from a live human being!
Tamper-Resistant Sensors

• It signs the output samples with its private key to indicate it already has sampled the signal from a living individual.
  • Furthermore, the sample shall be then processed as soon as possible.
  • Otherwise, we have to mitigate the risk of a sensor compromise in the intermediate time by a further time-tamping ("LongTerm Verifiable bio-samples").
  • This concept is all too often neglected in the emerging handwritten signature biometrics!
Conclusion

- We shall require ISO 19795 methodology during biometric application selection, comparison, and operation testing.
- Use independent penetration test to verify:
  - zero-effort attack complexity (*look for automated APIs!*)
  - masquerade attacks
  - spoofing possibilities
  - template security
  - system security in general (*threshold settings and template tampering, etc.*)
Thank You For Attention

Tomáš Rosa, Ph.D.
http://crypto.hyperlink.cz
Movie Snapshots Taken From

- Tajemství hradu v Karpatech, ČR, 1978
- Císařův pekař, ČR, 1951
- The Magnificent Seven, United Artists, USA, 1960
- Slunce, seno, jahody, ČR, 1983
All the quotations of Alphonse Bertillon, Auguste Kerckhoffs, and Claude E. Shannon were purely fictional.