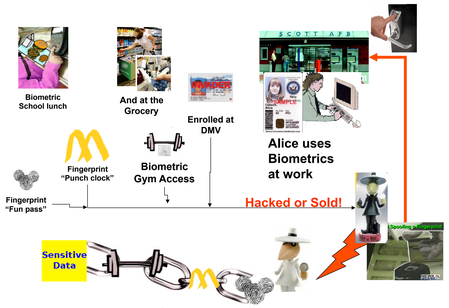
Revocable Fingerprint Biotokens: Accuracy and Security Analysis

Terrance E. Boult[1,2] , Walter J. Scheirer[1] and Robert Woodworth[2]
[1] VAST Lab University of Colorado at Colorado Springs and [2] Securics, Inc

The Biometric Dilemma



Who's the weak link in your security chain?

The key properties of biometrics, those unique traits that do not change significantly over a lifetime, are also their Achilles heel. *The biometric dilemma is that while biometrics can initially improve security, as biometric databases become widespread, compromises will ultimately undermine biometrics' usefulness for security.*

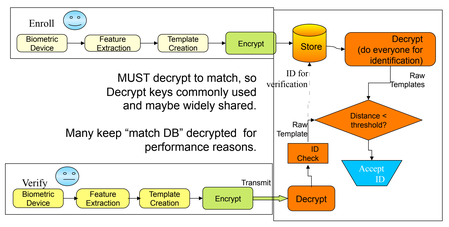
"Spoofing" with gummy fingers or reply injection is not the only issue. While many people like to think of biometrics as "unique", operationally they are not. Even FBI examiners have made high-profile misidentifications with fingerprints, e.g. [Cole- 05] documents 22 examples.

The best fingerprint systems tested by the US government have only 98% true acceptance rates, when set to reject 99.99% of false matches. At 99.99%, finding a false match in a database of millions is likely, leading to what we call the ***doppelganger threat***, where compromised databases with millions of users will allow an intruder to find a few "close enough" matches they can directly impersonate.

At least 40 million "financial records" were compromised or illegally sold in 2005, and over 50 million more financial/identity records lost or stolen in 2006. A database with millions of permanent "non-revocable" biometric records will become more significant cyber-target.

No one serious about security would use accounts, or tokens that could not be revoked.
Why except less from biometrics?

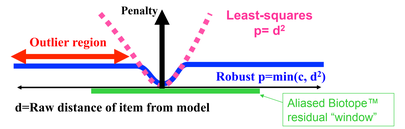
Traditional Encrypted Biometric Privacy Vulnerabilities "Levels"



The Biotope™ Transform Key Idea

Separate each field into 2 parts, as in robust distance operator.
Stable part can be encrypted, the inherent variation left unencrypted.

- Robust distance measures with a "window" effect.
- Use "modulus-like" computations to hide location of window and still allow local distance measure!
- "Window" location can be encrypted to protect.

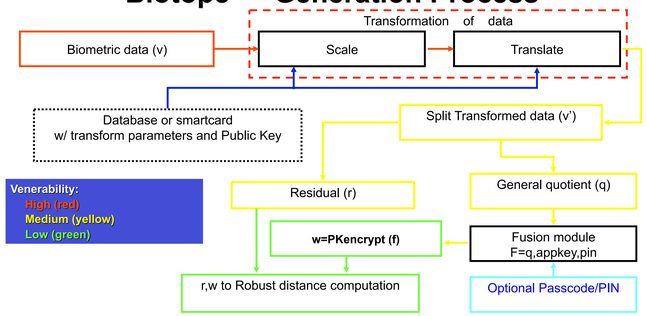


The penalty in the "similarity function". For weighted least squares errors, the penalty is a constant times distance, and grows quadratically. As in line fitting, a single outlier (feature) significantly impacts the similarity and hence the overall matching.

For a robust similarity metric (M-estimator), the penalty is limited to a maximum value e.g. outliers have a constant, and thus limited, impact on the overall measure.

Given measurements p,q, we can define a robust measure
$m_s(p,q) = c$ if $abs(r(p) - r(q)) > b$, and
$m_b(p,q) = (r(p) - r(q))^2$ otherwise.

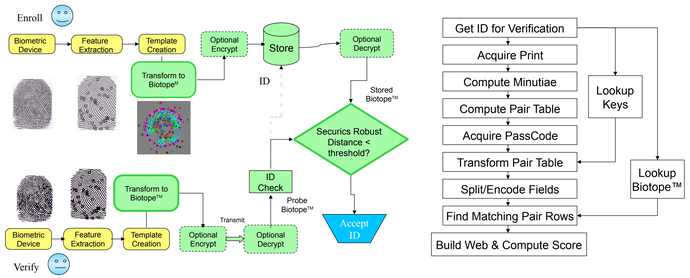
Biotope™ Generation Process



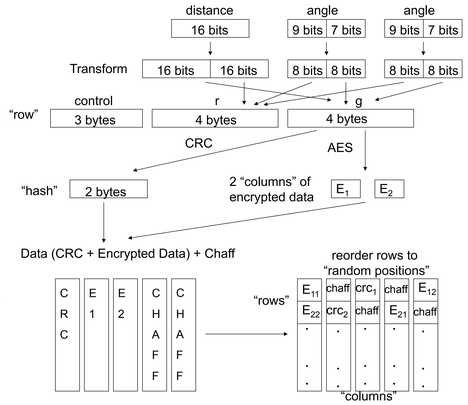
$d(p,q) = c$ if $w(p) != w(q) \;||\; abs(r(p) - r(q)) >= b$
$d(p,q) = (r(p) - r(q))^2$ otherwise

Let $e_{k,j}$ be the j^{th} biometric signature for user k. If s_k and t_k are such that $b\ s_k < r_k(e_{k,j}) < (1-b\ s_k)\ \forall\ j$ then $d(p, e_k) = m_{sb}(p, e_k)$, and we prove Biotopes can only improve accuracy.

Biotope™ Generation and Matching



Security Analysis for Finger Biotopes™



To support full PK inversion, we use PK to encrypt an AES key, a random index, plus padding, which produces two "columns" of data. For the real data, after transform, we have 3 control bytes that are not protected (or transformed), 4 bytes of residuals, i.e., *r* values, and 4 bytes of *q* values. The process by which a "row" is transformed uses 64 different potential sets of transforms. The CRC folds the data producing a *p*-fold ambiguity per field, with $p=2^{24}$ or $p=2^{16}$.

With a total of *c* possible match positions for the data in the columns of data+chaff, this produces a $(64*pc)$-fold ambiguity a would-be attacker must resolve to recover the data on that row. To recover a print (if its even possible) needs at least recovering n rows. Thus, a brute force search would require $\boldsymbol{n}^{64pc}$ attempts.

In our current implementation $64\boldsymbol{pc} = 2^6 \cdot 2^{16} \cdot 2^7 = 2^{29}$, so for a brute force attack to recover 16 minutiae would require a minimum $n=2^4$, $\boldsymbol{n}^{pc} = 2^{(4 \cdot 25)} = 2^{100}$ and more realistically it would be $\boldsymbol{n}=2^7$, $\boldsymbol{n}^{pc} = 2^{(7 \cdot 25)} = 2^{175}$ brute force attempts to recover 16 original minutiae. This presume that after generating hypotheses for each of the unknown items in a row there is a testable hypothesis to confirm the collection of rows is correct, then invert to a print. No such algorithm is known.

Finger Biotope™ Visualization

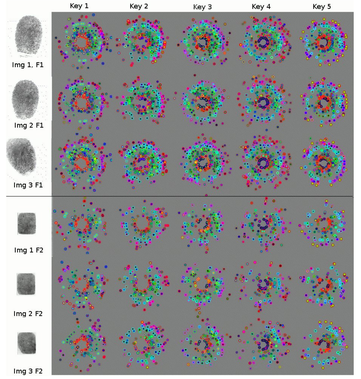
3 images of 2 different fingers

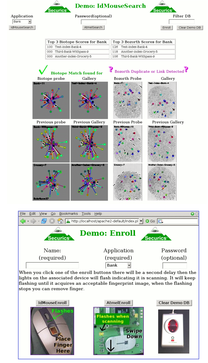
Same image per row

Same key per column

Note the differences across different keys for the same image!

Different keys cause greater variations than different people!

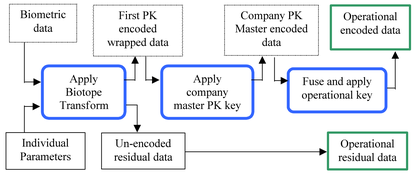




Ask us about a demo

Why Non-invertible Is Neither Sufficient nor Necessary

- $Y=X^2$ is non-invertible, but has only 2 point ambiguity.
  Ever do a cryptogram or other Puzzle? Significant levels of "ambiguity" can be overcome with knowledge and the use of constraints.
- Let Z=RSA(X;N); The RSA transform is fully "invertible" (given the private key), but without the key is computationally intractable to recover X from Z.
- Privacy/security requires "cryptographically" secure transformations, not simply non-invertible ones.
- Important to consider "how to reissue." If it's difficult then will only get canceled if lost for sure. How can non-invertible tokens be reissued?

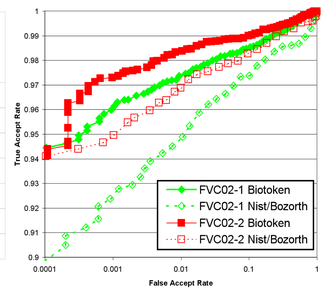


Biotope generation followed by multi step PK encoding to provide easily reissued biotopes. Data in dotted boxes is not stored.

Accuracy Analysis: The Biotope™ Process Actually Improved Performance!

Dataset	Biotoken Verification EER	Improvement Over EER of NIST VBT
FVC 2000 db1	.029	30%
FVC 2000 db2	.025	37%
FVC 2002 db1	.021	34%
FVC 2002 db2	.012	30%
FVC 2004 db1	.086	39%
FVC 2004 db2	.075	33%

Table 1: Finger Biotope™ accuracy



Implementation Based on NIST/FBI Bozorth matcher (NFIS2). For 380x380 image yielding a max of 150 A Pentium 4 1.6Ghz processor takes
0.394 sec to extract minutiae
0.029 sec for Biotope™ transform/match
0.021 sec for standard Nist/FBI Bozorth

Equal Error Rates and ROC curves comparing Biotope™ and the NIST/Bozorth matcher

This work supported in part under NSF STTR "Improving Privacy and Security in Biometrics," Award Number OII-0611283