A Snapshot of Security and Privacy in Biometrics

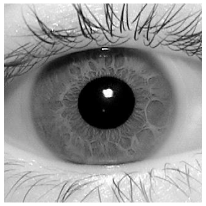
Walter Scheirer

Dir. of R&D at Securics, Inc.
Assistant Prof. Adjoint at the University of Colorado at Colorado Springs

Ethics & Science

- Motivation
  - Biometrics, those methods that can be used to recognize a person based upon physiological features, have become commonplace in recent years.
  - Pros of Biometrics: efficiency, convenience, improved access, improved security
  - Cons of Biometrics: unique identifiers, support unwarranted surveillance, difficulty with storage, questionable security





What must we be aware of?

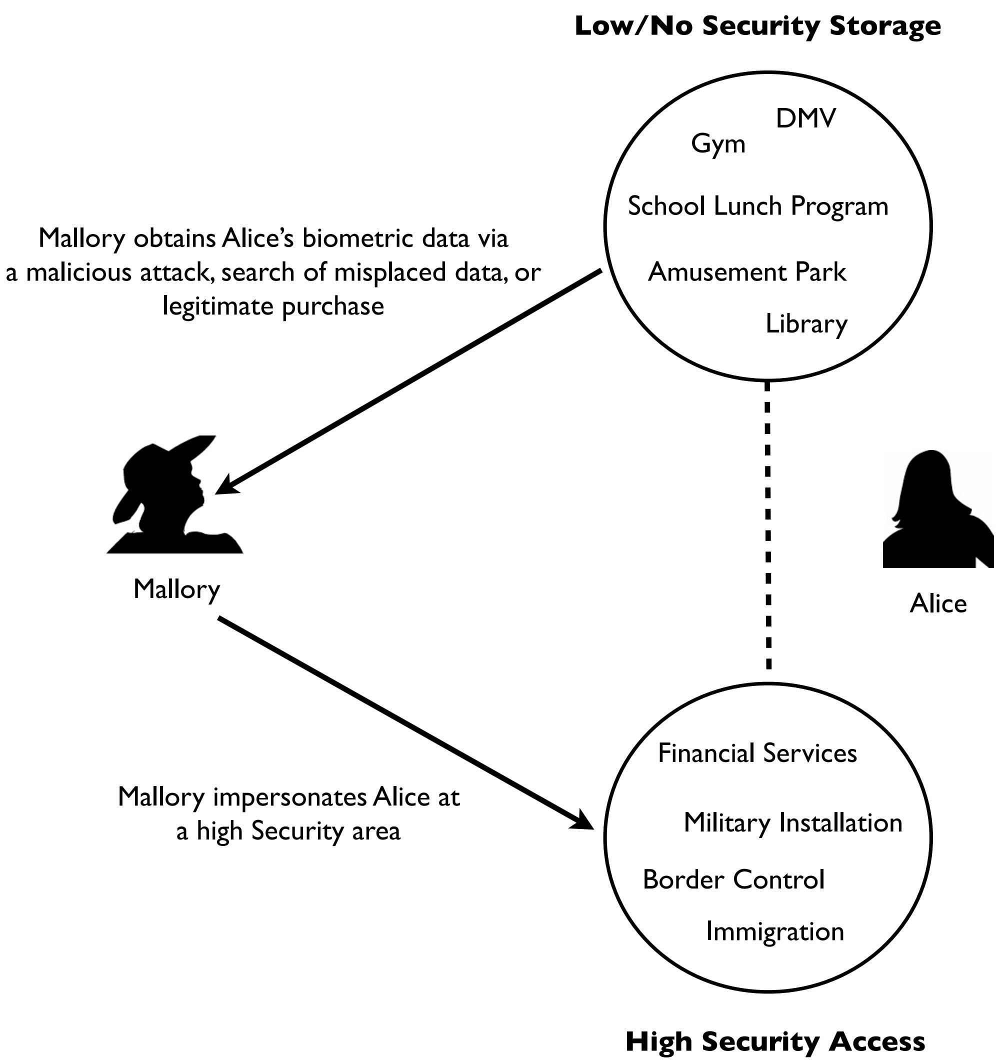
Function Creep

"The expansion of a process or system, where data collected for one specific purpose are subsequently used for another unintended or unauthorized purpose"

- Most familiar example in the US: SSN
- Function Creep and Biometrics: in 2001, Colorado tried to sell face & fingerprint data collected by its DMV[1]

1. http://www.i2i.org/articles/8-2001.PDF

The Biometric Dilemma



Biometrics, Body, and Identity[1]

- The same biometrics can be used in different ways
    - Identification, genetics research, medical monitoring, ethnic categorization
- Serious risk for discrimination based on what is measured from the human body

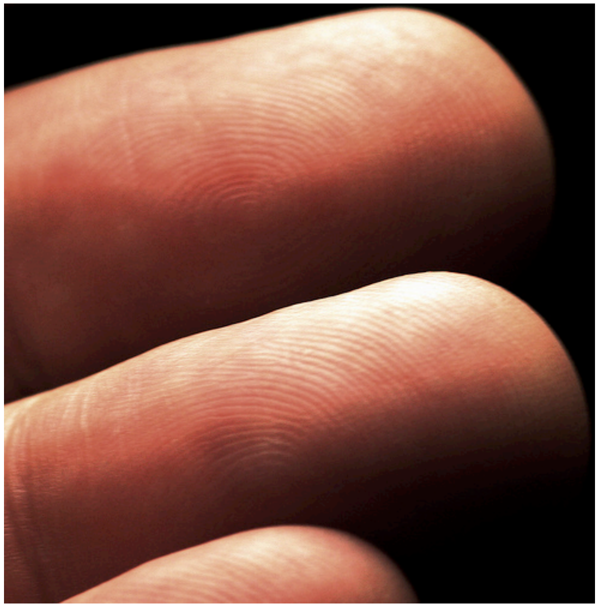
1. E. Mordini, "Ethics and Policy of Biometrics," in M. Tistarelli et al. (eds.), Handbook of Remote Biometrics, 2009.

Security is a Two-way Street

- Biometrics can be incorporated into large security frameworks
  - Identity Assurance
    - Tokens risk a disassociation of the owner from the object
- Biometrics suffer from the same flaws as traditional software security systems (and more!)
  - Limitations of Pattern Recognition

The Doppelganger Threat

- If the FAR is 1 in *X*, then an attacker can try more than *X* different prints
- Lots of public data available!
  – Fingerprint: NIST DB 14, NIST DB 29, FVC 2002, FVC 2004 …
  – Face: MBGC, FRGC, FVT, FERET …
  – Think of this as a biometric dictionary attack

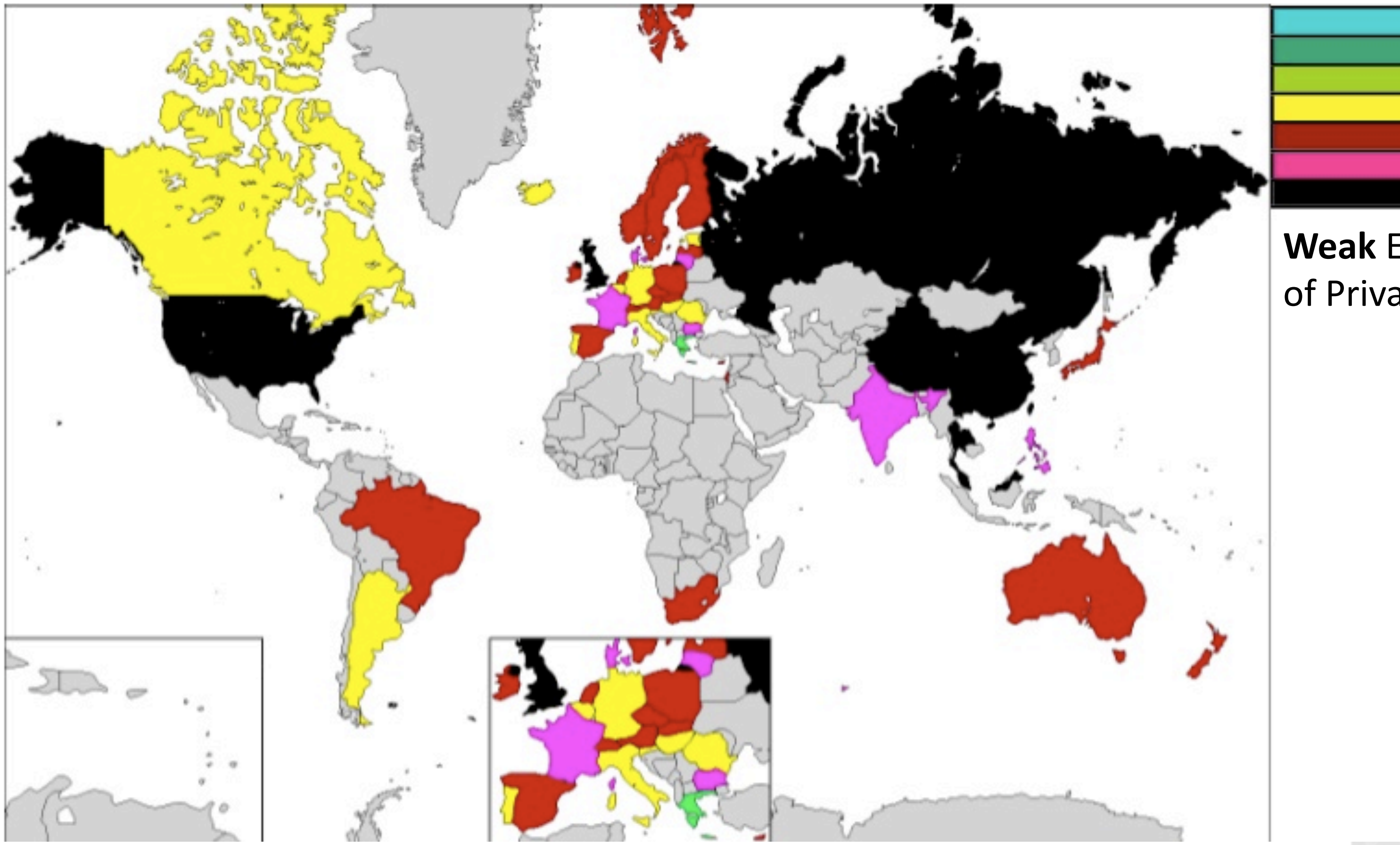


What does this mean for an event requiring strong security?

Would biometrics be a distraction?

Privacy Around the Globe

Strong Expectation of Privacy



Weak Expectation of Privacy

Image Credit: Privacy International

Opinions on Application Suitability

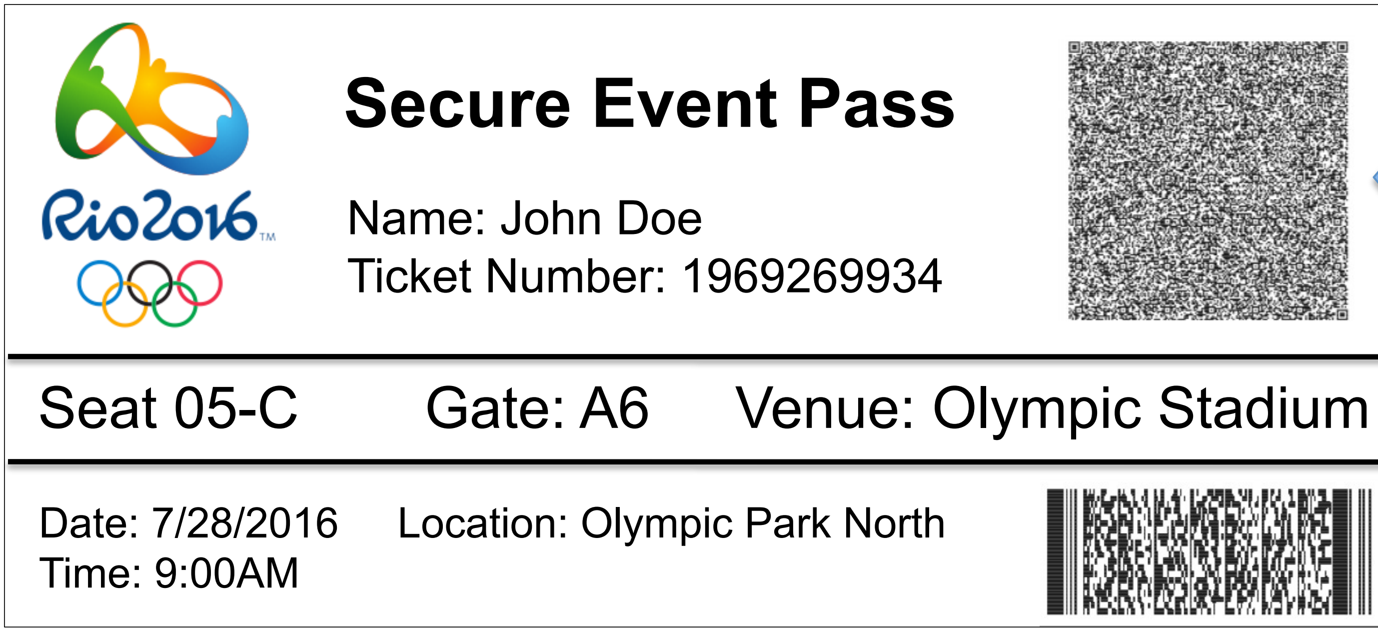
- Elliott, Massie, & Sutton, "The perception of biometric technology: a survey," 2007 IEEE Workshop on Automatic Identification Advanced Technologies.

Biometric Suitability, Aggregated Opinions		
Application	**Yes**	**No**
Identification of Arrested People	92%	7%
Obtaining Passports	91%	8%
Obtaining Drivers License	68%	29%
ID Verification during Credit Card Use	67%	32%
Checking in for A Flight	65%	35%
Scanning Public Places	62%	37%
Entering a School	32%	67%
Time and Attendance at Work	30%	70%

Public Perception of Surveillance

- **Headlines from the London Games**
  - "The 2012 Olympics are set to be the most CCTV-covered sporting event to date. Not everyone is happy about that." E & T Magazine 7.17.12
  - "Olympics: War-like security cordon in London" World News 7.13.12
  - "Will the 2012 Olympics set new surveillance records?" IT Business 7.21.12
  - "London Olympics Security Focuses on Deterrence: Use of Drones, Electric Fences, Missiles and More" Forbes 7.23.12

How can we provide security for an event *and* reassure the user?



What we want for the event:

Identity Assurance for Security Purposes

One-time Use Tickets

London 2012: Tickets were Non-transferrable

Ensure High Throughput

Secure Templates as a Solution

- Protect the Privacy and Security of the Biometric Features
- Revoke and re-issue biometric templates like a password or credit card #
- Match in an encoded space
- Prevent linking across databases (solve the biometric dilemma)
- Prevent the doppelganger attack (multi-factors)

"Getting this right has been much more challenging than we first thought." – Fabian Monrose

Standard Cryptography is a Weak Solution

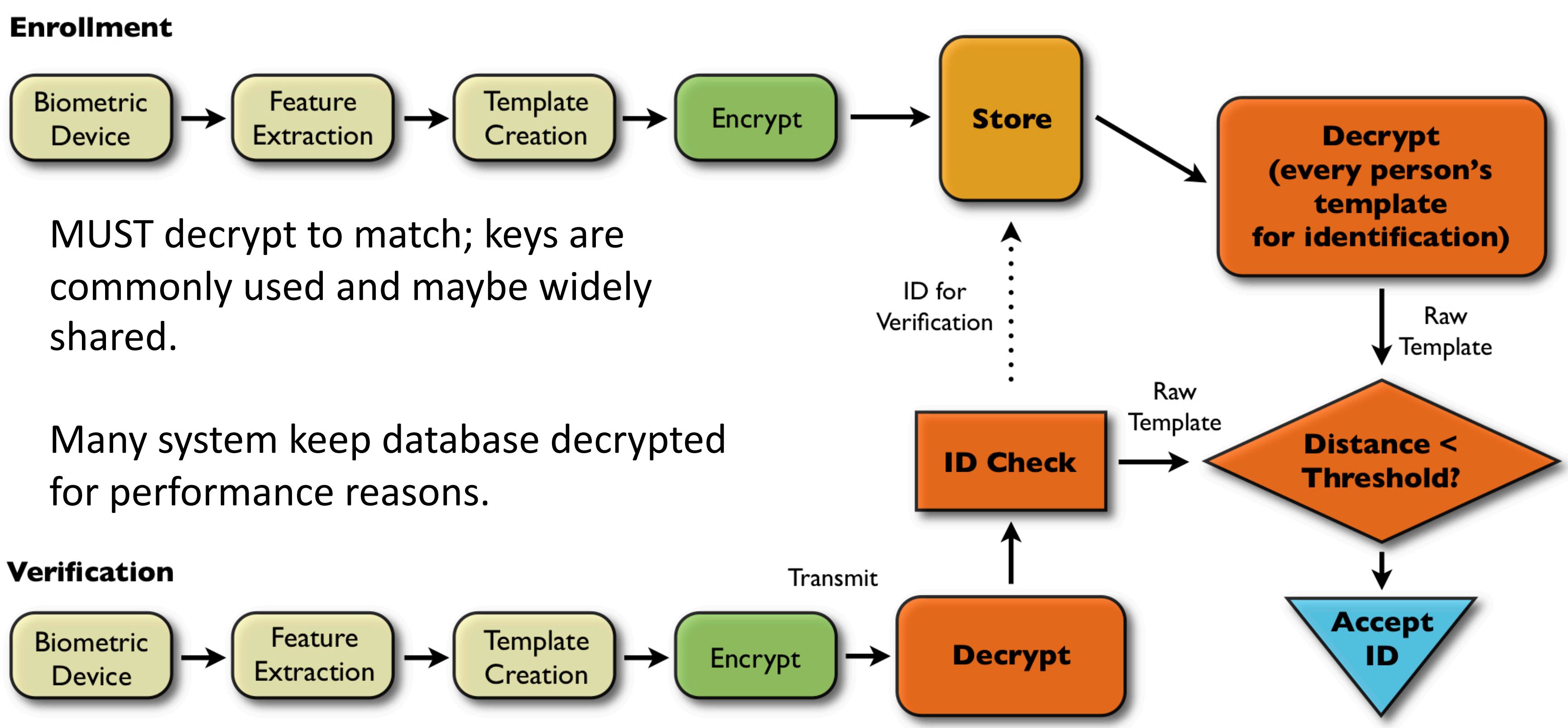
- Hashing/Crypto great for passwords.

Hire Only IEEE Members	1fc486d4b30dd490e044e40a35b6535c
Fire Only IEEE Members	53cc18345f93c390c7469e38c126a13f
Hire Only IEE Members	dfa9d634376d51d311ee55d40722950c

- Minor change results in radically different string (no match)

What does this suggest about potential for Biometrics?

Standard Cryptography is a Weak Solution



MUST decrypt to match; keys are commonly used and maybe widely shared.

Many system keep database decrypted for performance reasons.

Better solutions are out there!

- Biometric Encryption
- Non-invertible Transforms
- BioHashing
- Robust Hashing
- Fuzzy Vaults
- Fuzzy Commitment
- Fuzzy Extractors
- Revocable Biotokens
- Hybrid Combinations

How do they work?

How well do they work?

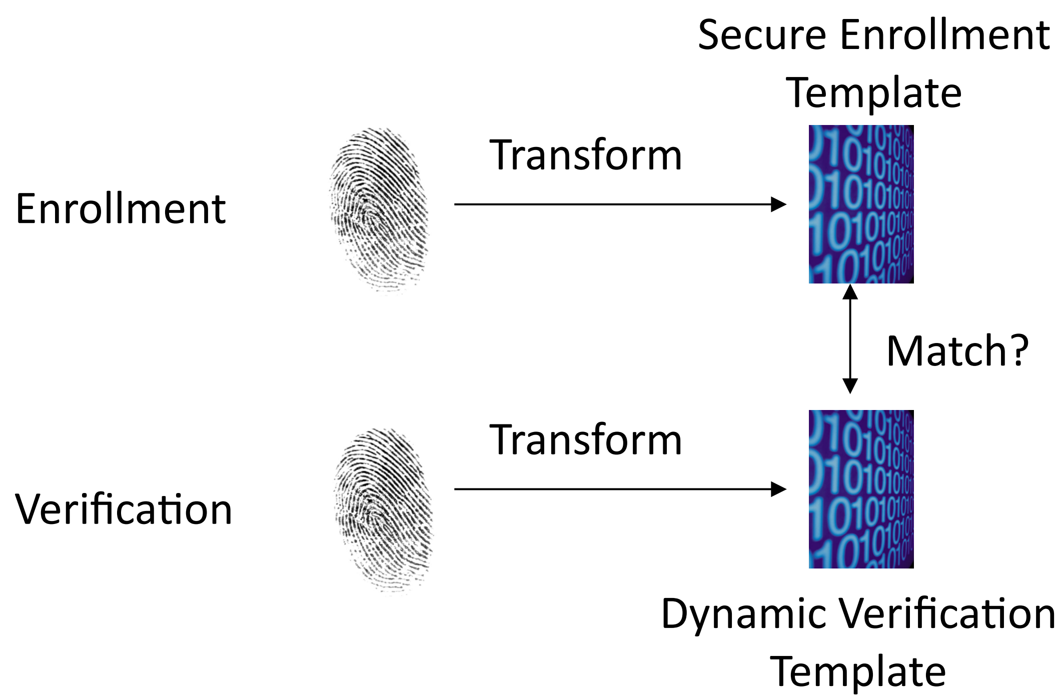
How secure are they?

Secure Template Technology

- Transformation of features that can be revoked and re-issued like a password or PIN
- Additional factors (PINs, passwords) used in transformation improve security
- Two interesting classes for crypto protocols
  - Key-generating biometric cryptosystems
    - Derive key data from biometric data; Ex. Fuzzy Extractors
  - Key-binding biometric cryptosystems
    - Bind any key data with biometric data; Ex. Fuzzy Commitment, Fuzzy Vault, Revocable Biotokens

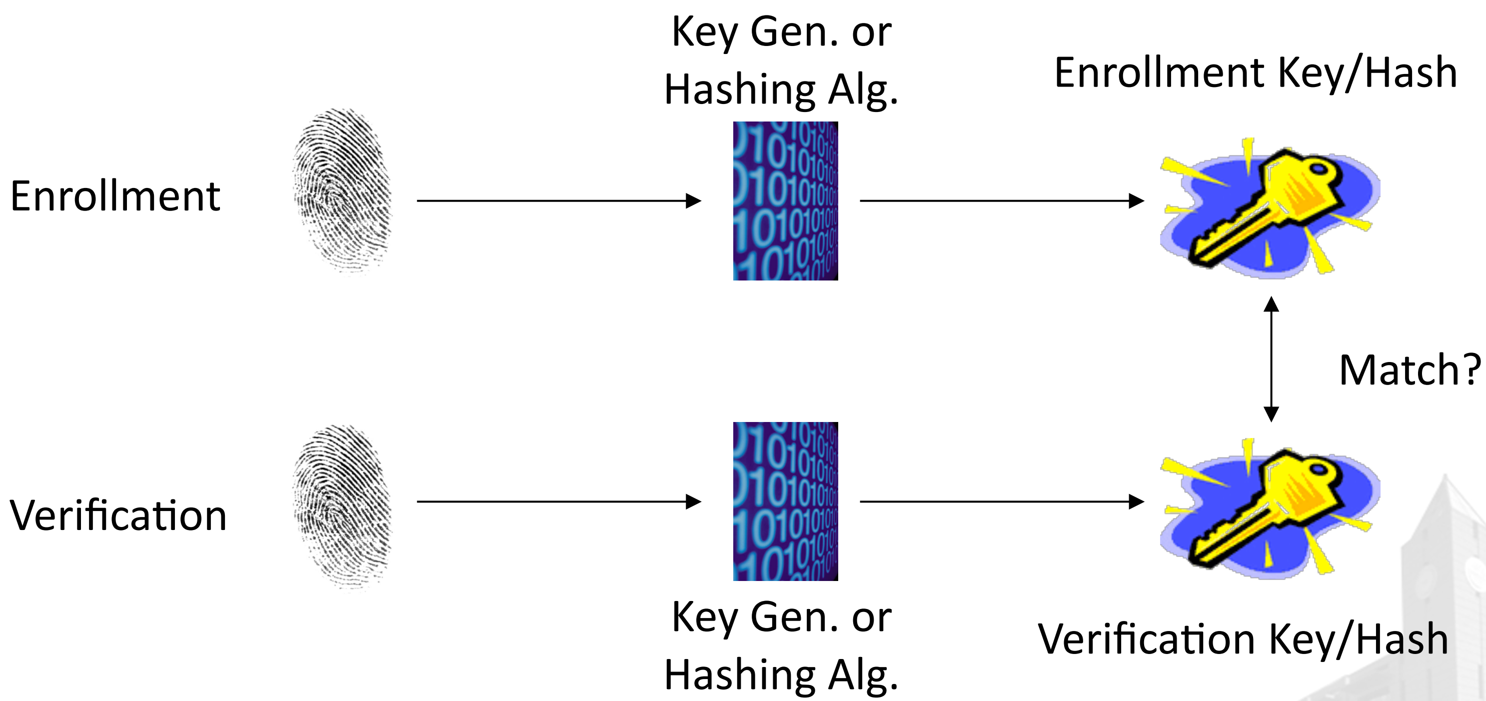
Secure Template Architectures

- Simply protect the original biometric features using some transformation that allows matching in encoded space



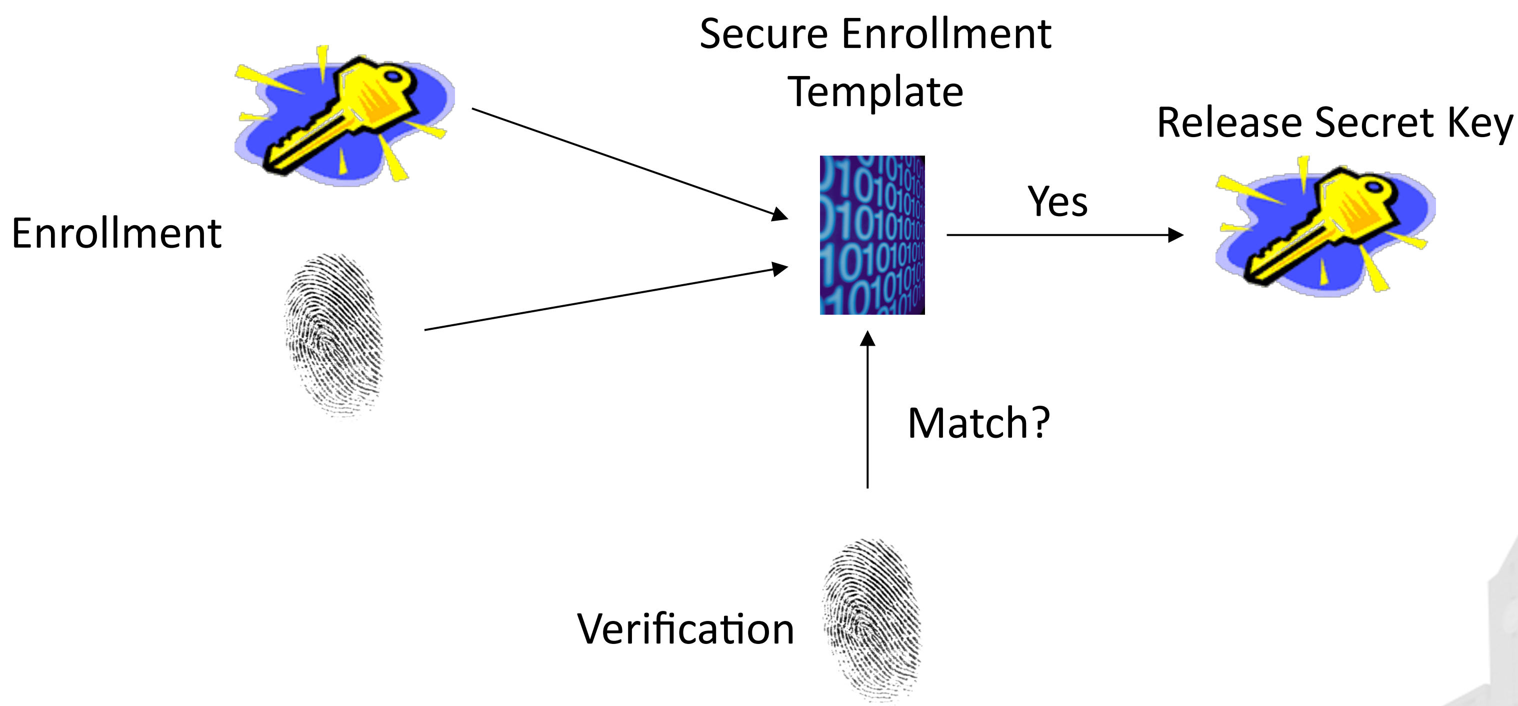
Secure Template Architectures

- Key-generating: Biometric cryptosystem that derives a key from the biometric data



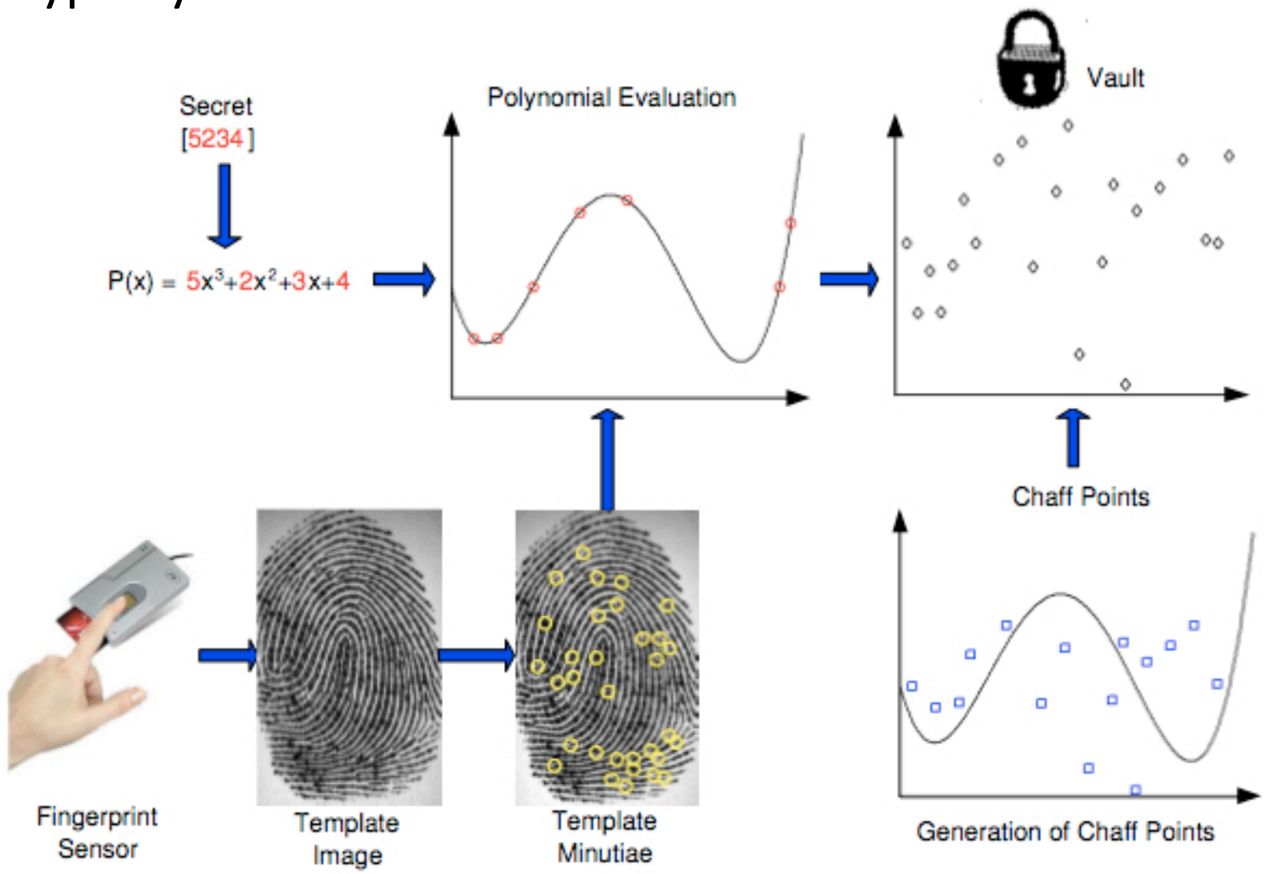
Secure Template Architectures

- Key-binding: Biometric cryptosystem that binds key data with the biometric data



Fuzzy Vaults[1]

- Not specific to biometric data, but typically applied to minutiae based fingerprint matchers as a key binding biometric cryptosystem

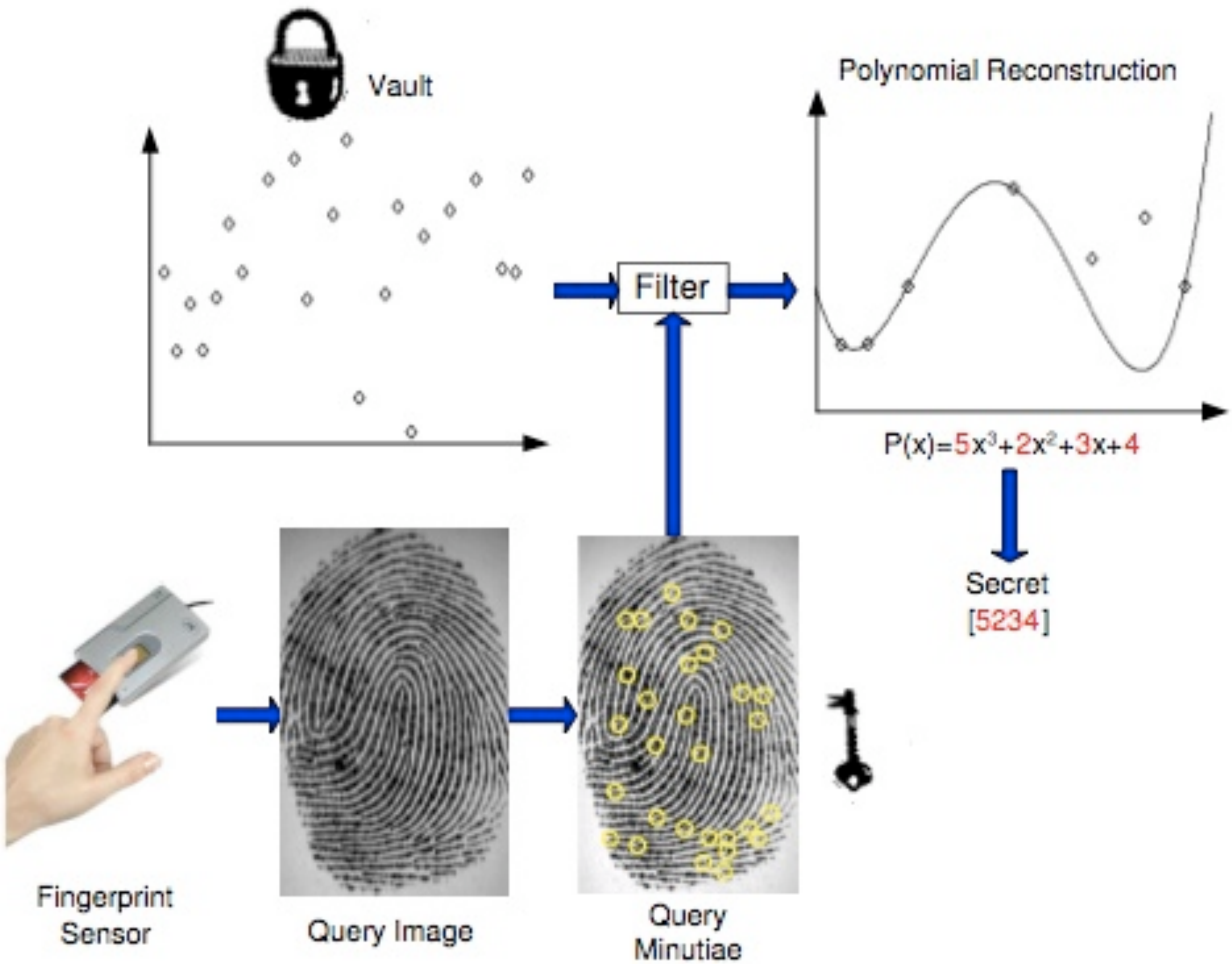


Encoding

Image credit: K. Nandakumar

1. A. Juels and M. Sudan, "A Fuzzy Vault Scheme," IEEE International Symposium on Information Theory, 2002.

Fuzzy Vaults



Decoding

Image credit: K. Nandakumar

Performance Numbers

	112 Bits		128 Bits		160 Bits	
	GAR	FAR	GAR	FAR	GAR	FAR
F.P. Fuzzy Vaults[1]	89	0.13	89	0.01	84	0
F.P. FV, Mosaic with 2 Queries[1]	96	0.24	95	0.04	89	0
Password Vault[2]	88	?	86	?	79	?

1. K. Nandakumar, A. K. Jain and S. Pankanti, "Fingerprint-based Fuzzy Vault: Implementation and Performance", In IEEE TIFS, vol. 2, no. 4, 2007

2. K. Nandakumar, A. Nagar and A. K. Jain, "Hardening Fingerprint Fuzzy Vault Using Password", in Proc. of ICB 2007

Fuzzy Vaults: Security Problems

- Chaff Point Identification[1]
- Improved Brute Force Attack[2]
- Correlation Attack, Known Key Attack, Correlation Attacks[3]

1. W. Chang, R. Shen, and F. W. Teo, "Finding the Original Point Set Hidden Among Chaff," in Proc. of the ACM Symposium on Information, Computer And Communications Security, 2006.
2. P. Mihailescu, "The Fuzzy Vault for Fingerprints is Vulnerable to Brute Force Attack," 2007.
3. W. Scheirer and T. Bout, "Cracking Fuzzy Vaults and Biometric Encryption," Biometrics Symposium, 2007.

Fuzzy Vaults: Correlation Attack

- Without a matching sample, the polynomial reconstruction problem is infeasible to solve
- What if we have *two or more* BFV instances?
  - Take the intersection of the abscissa values $(x, P(x))$ for the BFV instances
  - The result is the original template data
  - Some chaff points are likely to match - but the error correcting code is designed for this possibility

Implication: stolen biometric data

Fuzzy Vaults: Known Key Attack

- From the key, the polynomial P is directly reconstructed
- Sets of points may be directly enumerated to separate the template data, in the form $(x, \mathrm{P}(x))$, from the chaff
- Again, the error correcting code will help us if some chaff matches

Implication: stolen biometric data

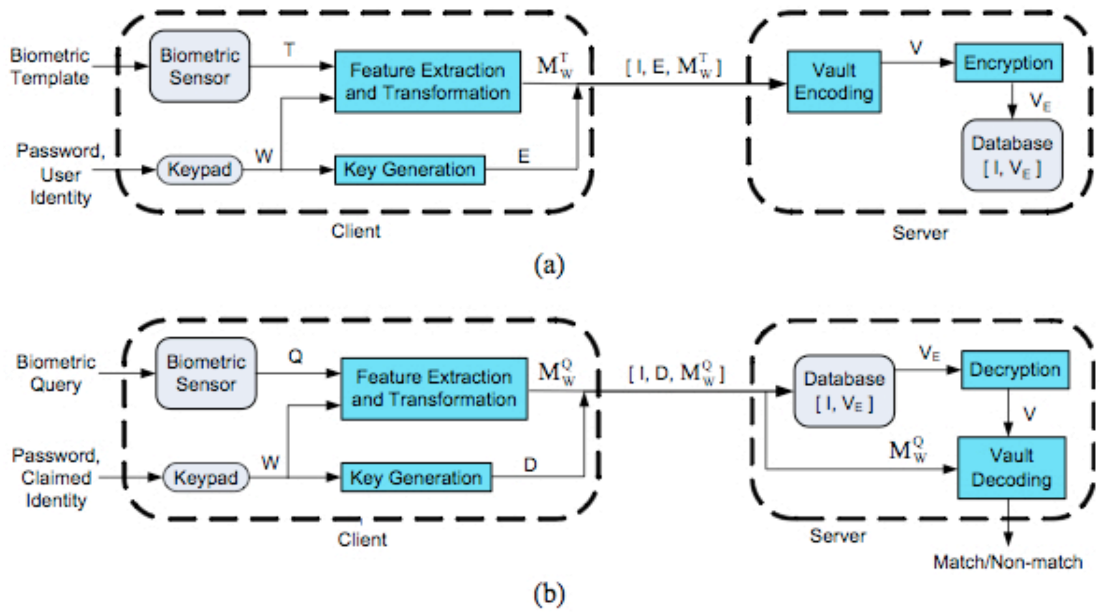
Fuzzy Vaults: Substitution Attacks

- Most of the vault is chaff. Matching uses only a small fraction of real data hidden in it.
- Overwrite chaff lines with attacker's template data
- Resulting template has both the user's and attacker's data.
- Insidious attack - attacker encodes their data with the user's key

Implication: backdoor for attacker

Response To Vulnerabilities in Fuzzy Vaults

- Password Hardened Fuzzy Vault[1]



Enrollment

Authentication

1. Karthik Nandakumar, Abhishek Nagar and Anil K. Jain, "Hardening Fuzzy Vault Using Password", in Proc. of ICB 2007 (and image credit).

Response to Vulnerabilities in Fuzzy Vaults

- Fuzzy Commitment to "encrypt" polynomial evalutions[1]
- Carefully chosen chaff[2]
- Incorporate local ridge information of minutiae (also incorporates a password)[3]
- Distance preserving hash functions[4]

1. A. Nagar et al. "Securing Fingerprint Template: Fuzzy Vault with Minutiae Descriptors," ICPR 2008.
2. S. Lee et al. "Secure Fuzzy Fingerprint Vault Against Correlation Attack," IEICE Electronics Express, Vol. 6, No. 18, 2009.
3. P. Li et al. "Security-Enhanced Fuzzy Fingerprint Vault Based on Minutiae's Local Ridge Information," ICB, 2009.
4. C. Orencik et al. "Securing Fuzzy Vault Schemes Through Biometric Hashing," Turk. J. Elec. Eng. & Comp. Sci., Vol. 18, No. 4, 2010.

Fuzzy Commitment

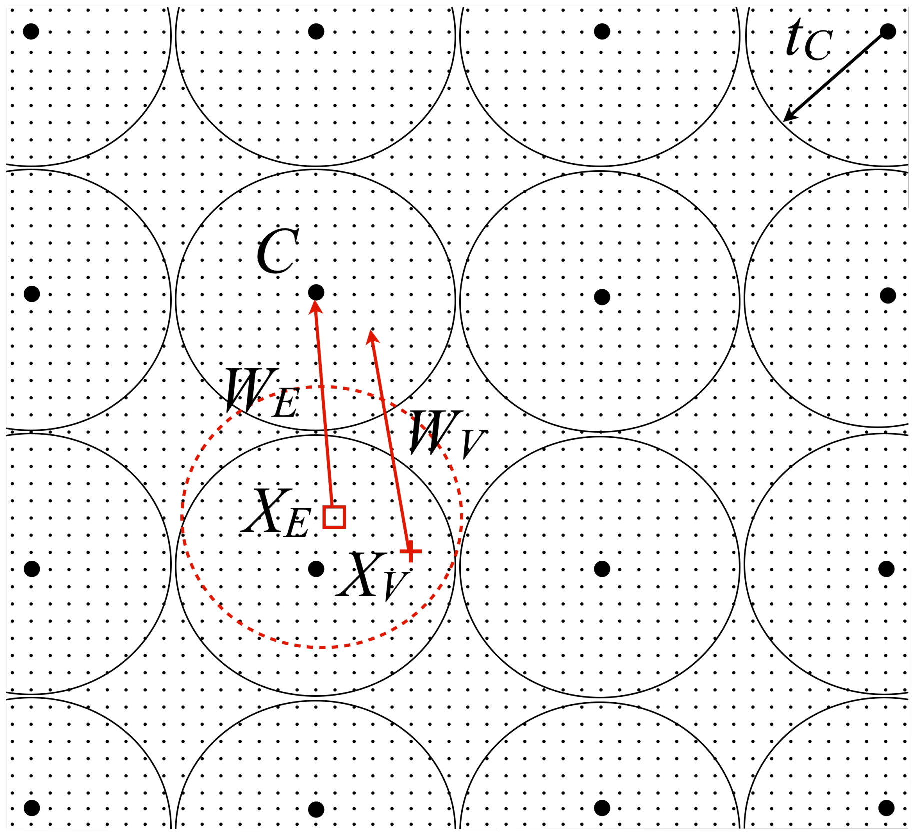
- Another well known key binding approach[1]
- Enrollment
  - Commit a codeword C (acts as the key) of an error correcting code using a fixed length biometric feature vector X as a witness
  - Store a hash h of C as "helper data"
  - Fuzzy Commitment: $X \oplus C, h(C)$

1. A. Juels and M. Wattenberg, "A Fuzzy Commitment Scheme," 6th ACM Conf. on Computer and Communication Security, 1999.

Fuzzy Commitment

- Verification
  - User presents a biometric, producing feature vector X'
  - X' is then used to unlock the codeword
    - $(X \oplus C) \oplus X' = C' = C \oplus e$
    - Hamming distance d_H indicates the number of errors corrupting C
      - $\epsilon = d_H(X, X') = ||e||$
    - An ECC Decoder can correct errors, yielding an extracted candidate key K
    - A successful match occurs when $h(K) = h(C)$

Illustration of Fuzzy Commitment

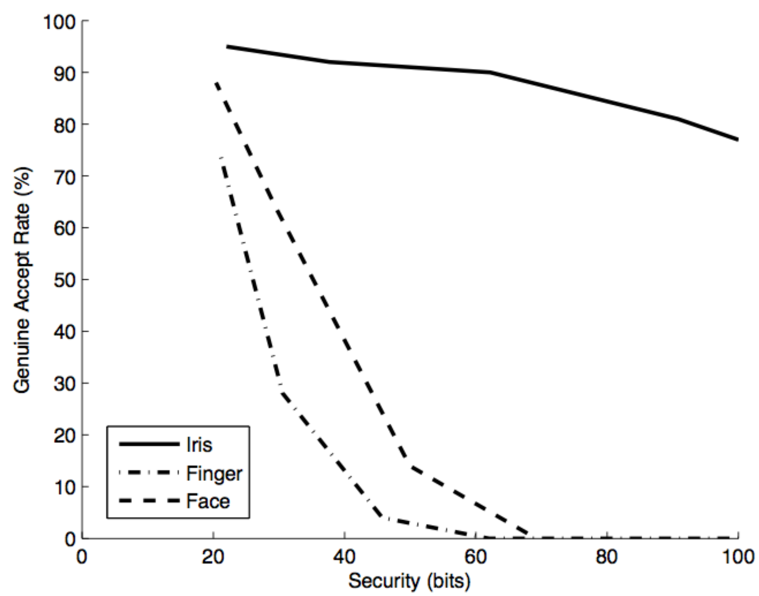


Grid of small dots: word space $\{0,1\}^{n_c}$

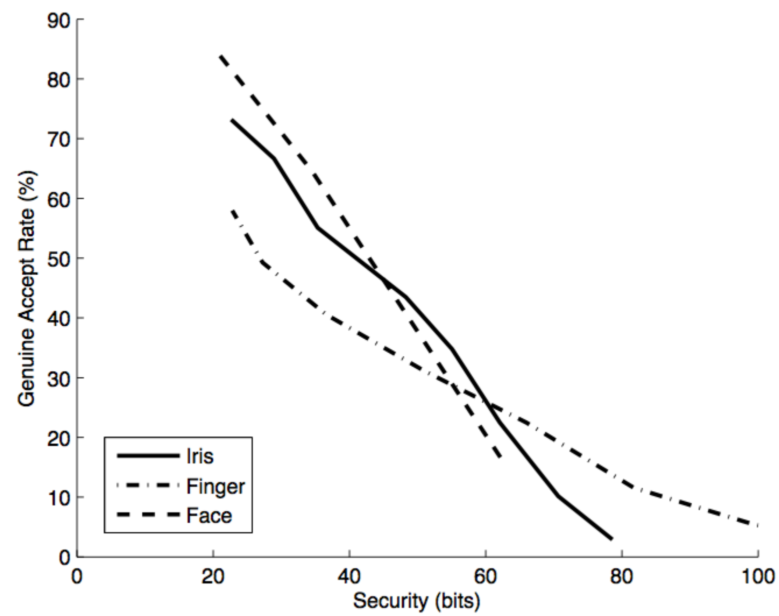
Bigger dots: codewords from C with the error correcting capability of the circles with radius t_c

Image adapted from: Kelkboom et al. "Preventing the Decodability Attack Based Cross-Matching in a Fuzzy Commitment Scheme," T-IFS, March 2011.

Performance Numbers



CASIA Ver-1, FVC 2002 DB2, XM2VTS



WVU Multimodal

	FVC/CASIA/XM2VTS	WVU
Iris	37%	91%
Face	30%	2%
Finger	33%	12%

Comparison of GAR at 53 bits of security

Images and Results: A. Nagar, "Biometric Template Security," Thesis Proposal, Michigan State University, 2011.

Performance Numbers

- 3-layer coding scheme[1]: ERR of 6.5% for 1032 bit key on FVC2000 DB2
- Multibiometric Fusion[2]:

	FVC/CASIA/XM2VTS	WVU
AND Rule	27%	89%
"Multibiometric Cryptosystem"	75%	99%

Comparison of GAR at 53 bits of security

- Bringer et al. 2008[3] for 2028 bit keys:
  - ICE: FRR 5.62%, FAR $< 10^{-5}$
  - CASIA: FRR 6.65%, FAR 0%
  - FVC 2000: FRR 2.73%, FAR 5.53%

1. X. Shao et al., "A 3-layer Coding Scheme for Biometry Template Protection Based on Spectral Minutiae", ICASSP, 2011.
2. A Nagar et al., "Technical Report: Multibiometric Cryptosystem", MSU Tech. Report, 2011.
3. J. Bringer et al., "Theoretical and Practical Boundaries of Binary Secure Sketches", IEEE T-IFS, 2011.

Fuzzy Commitment: Security Problem

- Decodability Attack[1]
  - Codewords: C_1, C_2
  - Biometric Data: X_1, X_2
  - $W_1 = C_1 \oplus X_1; W_2 = C_2 \oplus X_2$
  - $W_1 \oplus W_2 = (C_1 \oplus C_2) \oplus (X_1 \oplus X_2) = C_3 \oplus (X_1 \oplus X_2)$
  - If $(X_1 \oplus X_2)$ is small, the result of the XOR will be close to another codeword (decodes)

Implication: match users across databases

1. F. Carter and A. Stoianov, "Implications of Biometric Encryption on Wide Spread Use of Biometrics," EBF Biometric Encryption Seminar, June 2008.

Response to Vulnerabilities in Fuzzy Commitment[1]

- Incorporate random bit permutation process
- Prior to the XOR operation of the biometric data X with the code word C, randomize X with a bit permutation matrix M_r
- The new template: $W = C \oplus M_r X$
- M_r is not considered a secret

1. Kelkboom et al. "Preventing the Decodability Attack Based Cross-Matching in a Fuzzy Commitment Scheme," T-IFS, March 2011.

Fuzzy Extractors

- Key generating biometric cryptosystem[1]
- Attractive proposition, but difficult due to intra-user variability
- Goal: Extract a uniformly random string R from its input w in a noise-tolerant way
  - If the input changes to some w', but remains close, the string R can still be reproduced exactly

1. Dodis et al., "Fuzzy Extractors: How to Generate Strong Keys from Biometrics and Other Noisy Data," EUROPCRYPT, 2004.

Secure Sketch[1]

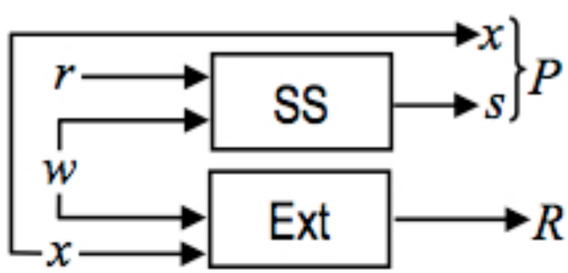
- "Helper Data" for Fuzzy Extractors
- A *secure sketch* produces public information about its input w that does not reveal w, and yet allows exact recovery of w given another value that is close to w.

1. Y. Dodis, L. Reyzin and A. Smith, Fuzzy Extractors," In Security with Noisy Data: Private Biometrics, Secure Key Storage and Anti-Counterfeiting, P. Tuyls, B. Skoric and T. Kevenaar, Eds., Springer-Verlag, 2007.

Fuzzy Extractors

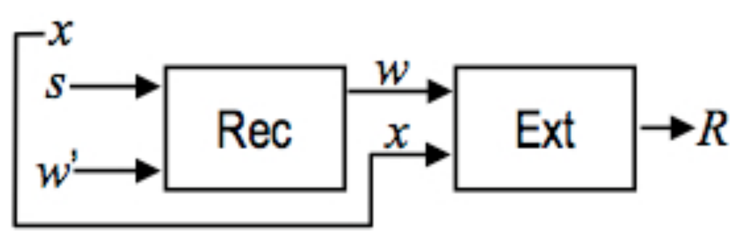
- A secure sketch SS producing a string s bound with a random number x forms the basis of the helper string *P*
- Recovery procedure allows matching with a “close” string w'
- Extractor returns a string R*, the key*, when approximate input matching is successful
- P assists in the reproduction of R

Sketching Procedure



r is some randomness

Recovery Procedure



Security Analysis: Fuzzy Extractors

- Security analysis of the fuzzy extractor scheme made in terms of the *min-entropy*
- An adversary's best strategy is to guess the most likely value
  - Predictability of a random variable
  - Min-entropy is the "worst case" entropy
- Information theoretical balance between stability and suitable randomness

*Analysis is not made with consideration to FAR/GAR!

Practical Concerns

- At the present, fuzzy extractors exist in the realm of theory
- Fuzzy extractors may suffer from practical constraints during error-prone data collection; difficulty for key generation[1]
  - Unclear whether known constructions can correct the errors typically generated by humans
  - Require biometric inputs with high min-entropy, but haven't discussed feature selection

1. Ballard, S. Kamara and M. Reiter, "The Practical Subtleties of Biometric Key Generation", in Proc. of the USENIX Security Symposium, 2008.

Revocable Biotokens

- We want two different things:
  - Robust distance/matching
  - Security/Revocability

→Break data into two parts:

Stable and *Unstable*

5ft (stable)
2in unstable

6ft (stable)
1in unstable

- Stable part is encrypted/hashed to provide security/ privacy and revocability - straight feature protection
- Two parts together provide robust distance measure, which we can prove will not decrease accuracy

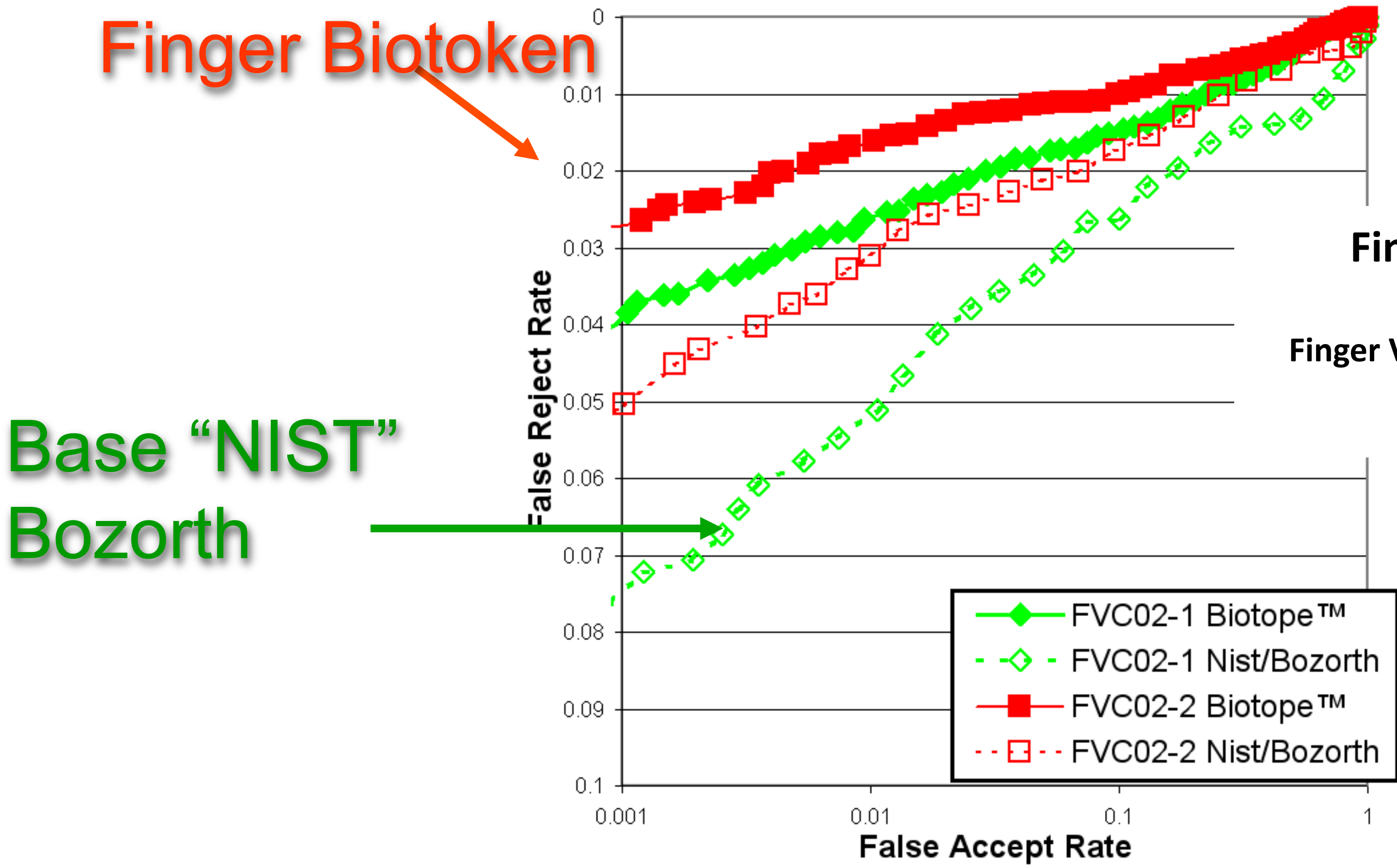
Revocable Biotokens[1]

- Assume a biometric produces a value v that is transformed via scaling and translation
  - $v' = (v - t) * s$
- Split v' into stable component q and residual component r
- For user j, leave the residual un-encoded (base scheme)
  - $r_j(v')$
- Encrypt q with public key P
  - $w_{j,1}(v', P)$

Brute Force Attack to revert biotoken back to original features: 2^{108} for insider, 2^{120} without access to all keys/data

1. T. Boult, W. Scheirer and R. Woodworth, "Revocable Fingerprint Biotokens: Accuracy and Security Analysis," CVPR 2007.

Revocable Biotoken Performance

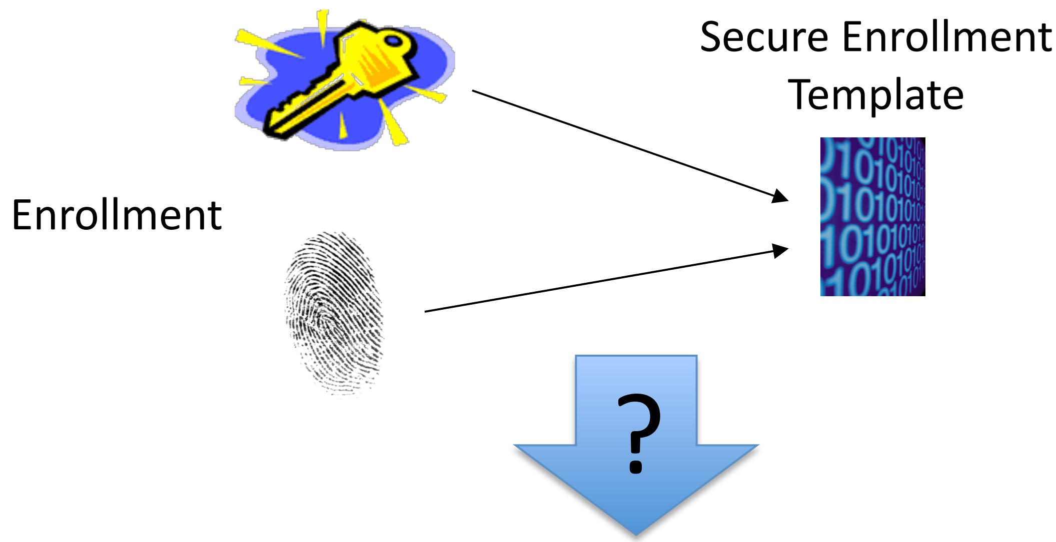


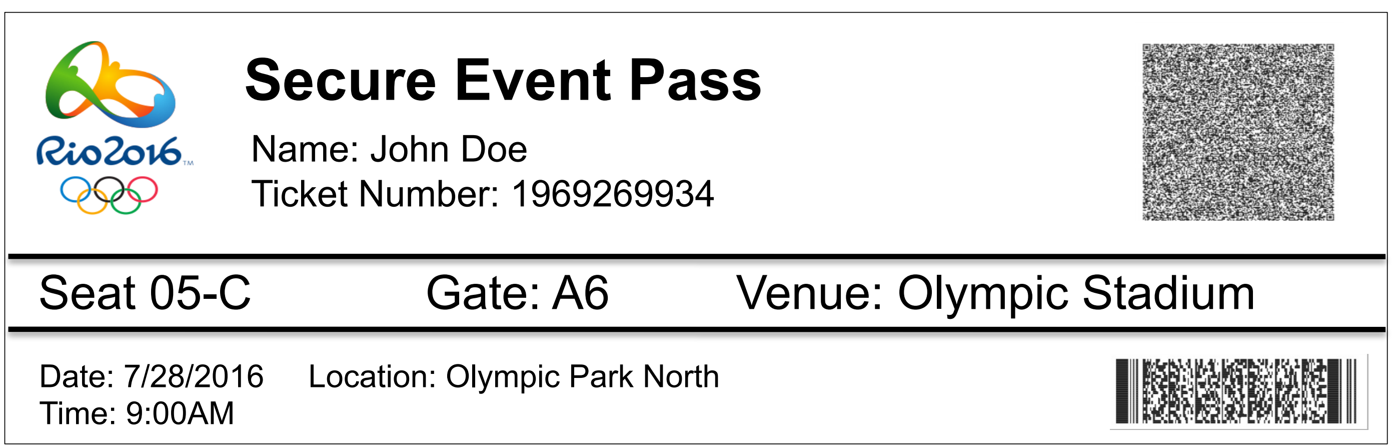
Finger Biotoken

on

Finger Verification Challenge datasets

How do we go from a secure template algorithm to something we can use?





ISO/IEC 24745

- Security requirements for securely binding between a biometric reference and an identity reference
- Biometric system application models
- Scenarios for storage and comparison of biometric references
- Guidance on privacy protection for users

Potentially compliant solutions: Fuzzy Commitment and Revocable Biotokens

First Step: Barcodes

http://www.securics.com

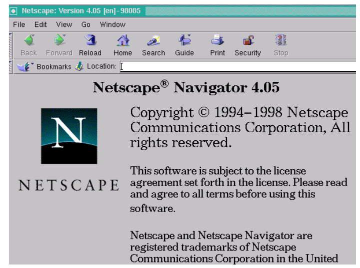
QR Code

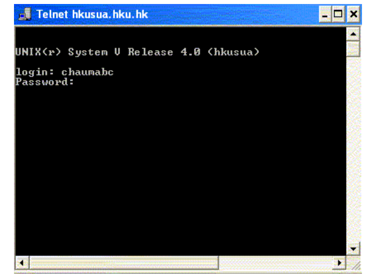
Revocable Biotoken Template can fit in 3KB

http://www.priv-id.com

Fuzzy Commitment Template can fit in 180 Bytes

Second Step: Protocols



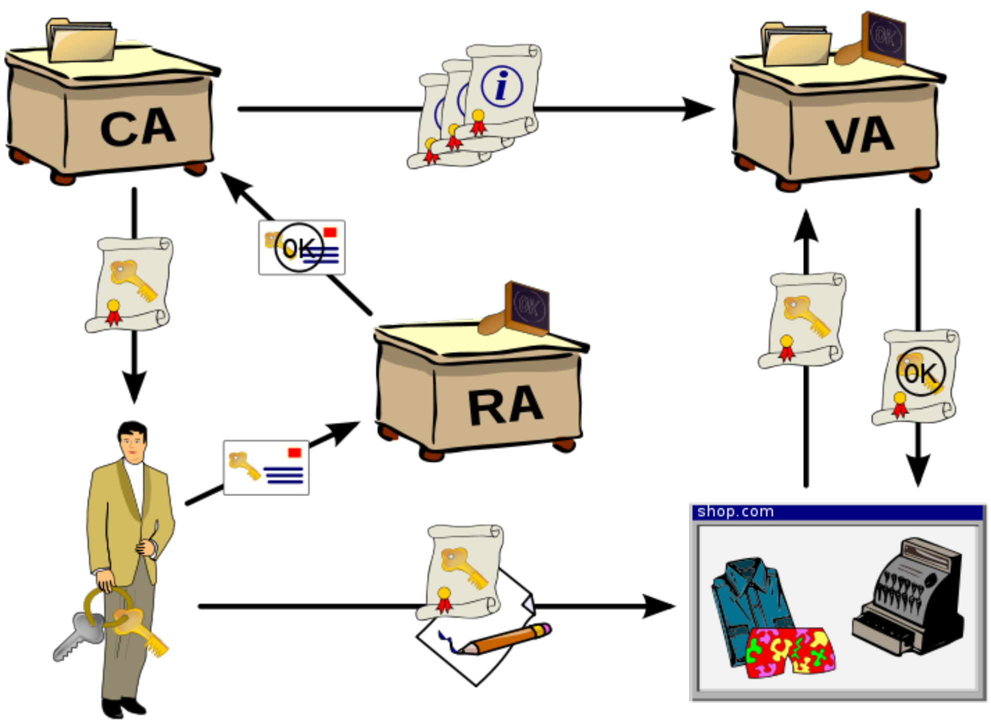


- Recall the 1990s: Huge explosion in new network protocols for e-commerce, electronic record keeping, access control, etc.
- Security of these protocols was an afterthought!
  - We need cryptography to protect insecure channels
  - How can Alice verify a public key?

Solution: Public Key Infrastructure

Public Key Infrastructure

- PKI is the infrastructure for handling the complete management of digital certificates (x.509 compliant)
  - Certificates contain trusted information: a public key



Problems with PKI

- Ellison and Schneier (2000)[1]
  - "Risk #1: Who do we trust, and for what?"
  - "Risk #2: Who is using my key?"
  - "Risk #4: Which John Robinson is he?"
  - "Risk #6: Is the user part of the security design?"
  - "Risk #8: How did the CA identify the certificate holder"?

1. C. Ellison and B. Schneier, "Ten Risks of PKI: What You're Not Being Told About Public Key Infrastructure," *Computer Security Journal*, 16(1):1-7, 2000.

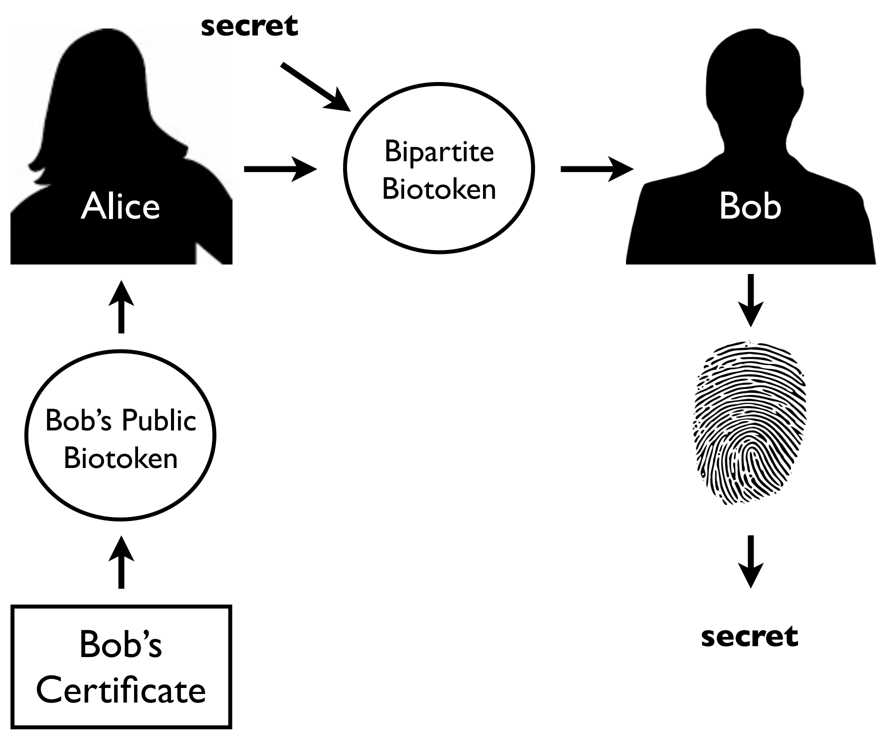
Biometric Solution?

- By adding a second factor, we can mitigate the inherent trust problems with PKI
- What about Biometrics?
    - Improved non-repudiation
    - Strong verification for actors in a transaction, certificate authority establishment, and general certificate issue

Address the trouble with Biometrics using Secure Templates. Case Study: Revocable Biotokens

Benefit of a BKI

- Ability to store public biotokens in digital certificates
  - Any entity in the infrastructure can send secret data that only the owner of the biotoken can unlock



Requirements for a Biocryptographic Key Infrastructure

1. Cryptographically strong protection of the underlying biometric features
2. Ability to revoke and re-issue templates
3. Nested re-encoding, allowing a hierarchy of templates to be generated from a single base template

Potential for Rapid Ticketing {

4. Support for public templates
5. Key-binding capability without the need of intervention by the person associated with the template

Nesting Property

- Protected template w_j is re-encoded using a transformation function T

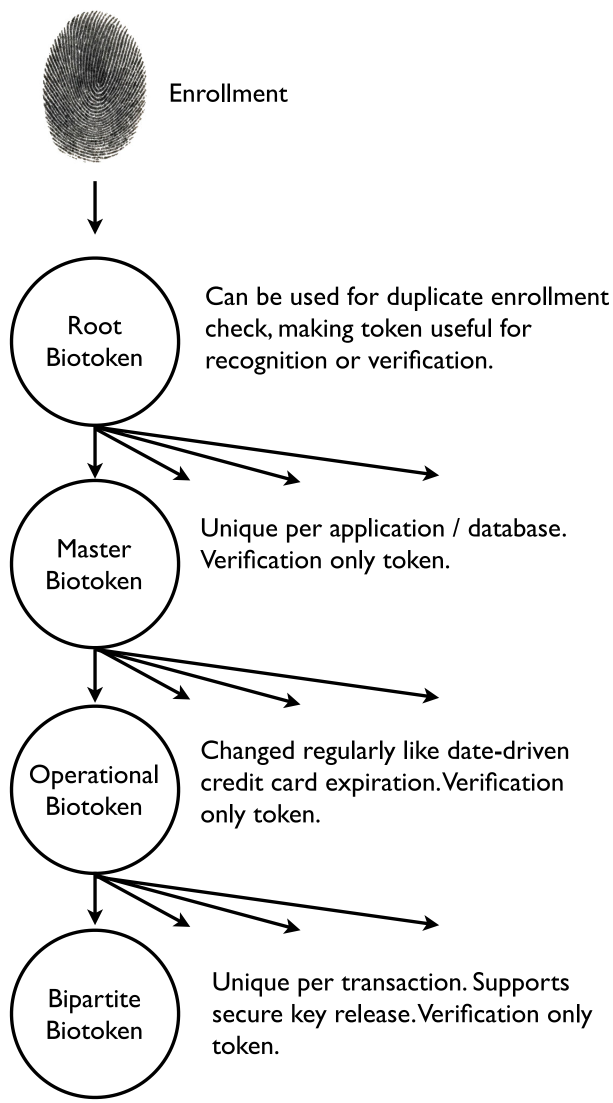
  1st encoding: $w_{j,1}(v', P)$

  2nd encoding: $w_{j,2}(w_{j,1}, T_2)$

  nth encoding: $w_{j,n}(w_{j,n\text{-}1}, T_n)$

- The nesting process is formally invertible via the keys, but cryptographically secure

Biotoken Issue/Re-Issue Tree



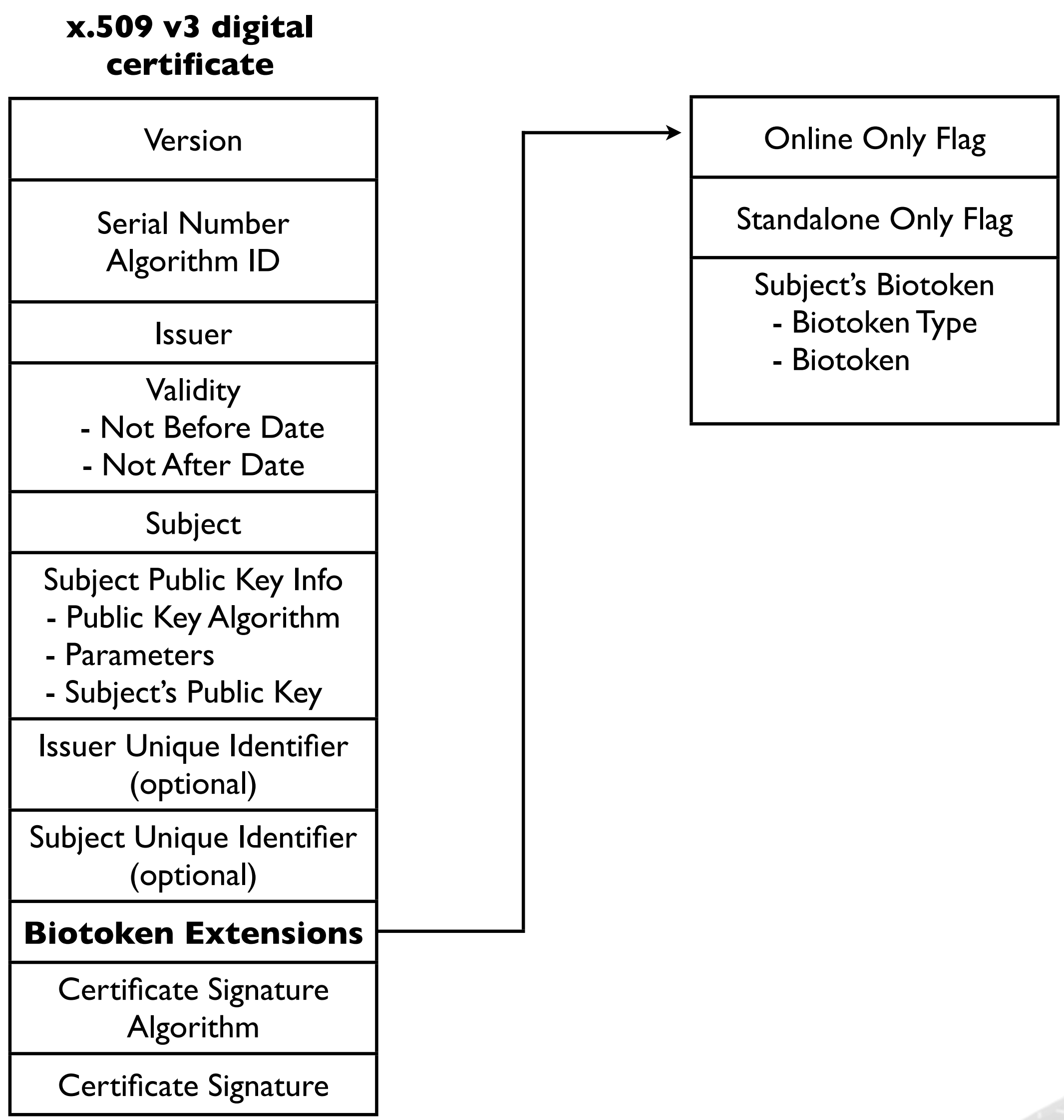
This biotoken is encoded in the barcode

Bipartite Biotokens

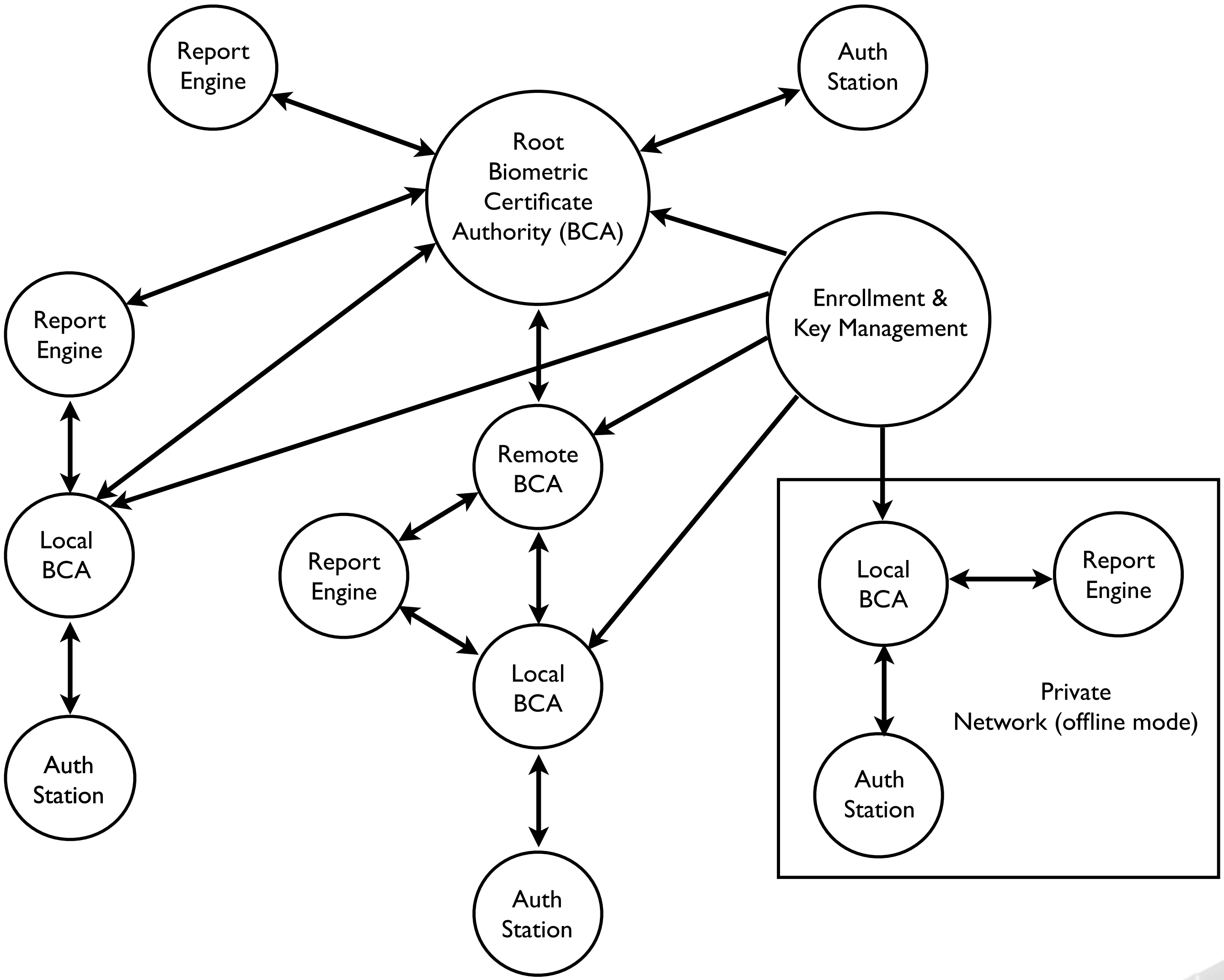
- Scheirer and Boult 2009[1]
  - Let B be a revocable biotoken. A bipartite biotoken B_p is a transformation $bb_{j,k}$ of user j's k^{th} instance of B. Any bipartite biotoken $B_{p,k}$ can match any revocable biotoken B_k for the same user.
  - $bb_{j,k}$ must allow the embedding of some data d into B_p
    - $bb_{j,k}(w_{j,k}, T_k, d)$
  - If $B_{p,k}$ and B_k match, d is released

1. W. Scheirer and T. Boult, "Bipartite Biotokens: Definition, Implementation, and Analysis," ICB 2009.

Digital Cert. Supporting Biotokens

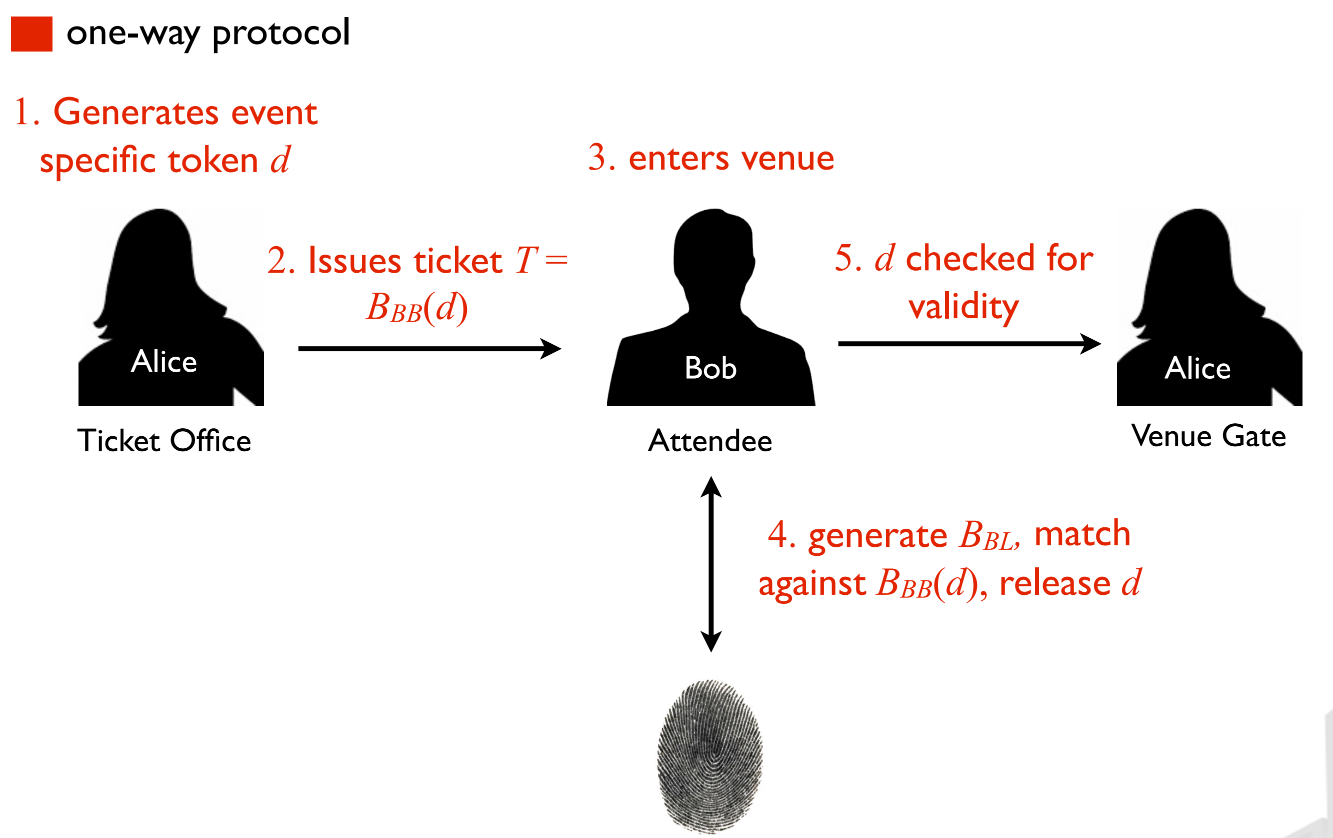


A Biocryptographic Key Infrastructure



Simple Authentication Protocol

- Assume Bob has already enrolled at the ticket company



What does this mean for an event like the Olympics or World Cup?

- Measures Protecting Users
  - The user has control over their biometric data
  - Per event templates from a single base enrollment
  - If a template is stolen, we have a process to revoke and re-issue credentials
- Tighter Event Security
  - Attendee identity assurance

Want to learn more?

- IEEE Transactions on Information Forensics and Security
  http://www.signalprocessingsociety.org/publications/periodicals/forensics
- IEEE Transactions on Pattern Analysis and Machine Intelligence
  http://www.computer.org/portal/web/tpami
- IEEE Workshop on Information Forensics and Security
  http://www.wifs12.org
- IEEE International Conference on Biometrics: Theory, Applications and Systems
  https://sites.google.com/a/nd.edu/btas_2012
- International Conference on Biometrics
  http://atvs.ii.uam.es/icb2013

Thank You!

Questions?