Biometrics

- Measurement of some biological characteristic
- “Something-you-are” authentication factor
- Essentially how people authenticate one another
Biometrics

• Some attractive features
  • Minimal user effort
  • Nothing extra to carry or remember
    • Hard to lose!
  • Can’t be shared (usually)

• Some drawbacks
  • Not always accurate
  • Work poorly for some people
  • Security challenges (to be discussed)

Here are some examples…
Face recognition

• Pros:
  • Very intuitive
  • Can use ordinary camera
    • Or one on mobile device

• Cons:
  • Poor accuracy
    • (Purported improving rapidly)
  • Not terribly secret
Fingerprint

• Pros:
  • Lots of experience in law enforcement
  • Belief in uniqueness
    • (We’ll talk about this…)

• Cons:
  • Social stigma
  • Spoofable
Iris

• Pros:
  • Extremely accurate
    • Estimated 250 bits of entropy in iriscode!
    • Can in principle yield a cryptographic-strength key
  • Non-invasive
    • Note: not retina!

• Cons:
  • Requires special camera
  • Very sensitive to lighting conditions
  • People confused about difference between the iris and retina…
  • Literature historically gives poor statistical characterization
    • (New model, Feb. '16: http://www.cl.cam.ac.uk/~jgd1000/
Other types

- Some less common ones:
  - Hand geometry
  - Retina
  - Keystroke dynamics
  - Gait
  - Pulse
  ...We won’t discuss these today...

- Even less common:
  - Ear recognition
  - Body odor
  ...We won’t discuss and will try to forget...
And also...

**Forget Fingerprint: Car Seat IDs Driver’s Rear End**

*Wall Street Journal*

**360 pressure-sensing disks**

- $900 each
- 2% false rejection rate

Shigeomi Koshimizu shows his seat sensor. Tokyo, has developed an ultra-sensitive sheet that sometime down the line could make the contours of a driver’s rear end an integral part of a car's security system.
Many current uses of biometrics

• Law enforcement
  • E.g., FBI database of fingerprints

• Government services
  • E.g., delivery of welfare / social services

• Traveler authentication
  • E.g., passports, TSA Pre

• Unlocking your mobile phone
  • E.g., iPhone

• Securing national treasures
  • E.g…
How does biometric authentication work?
Two goals for biometrics

• **Identification**
  • Goal: Learn a person’s identity
    • E.g., identify criminal from fingerprint or DNA at crime scene

• **Authentication**
  • Goal: Determine whether claimed identity is correct
    • E.g., this is really Woody
Registration

Alice

Template
Template is stored

Stored in, e.g.,
• mobile device
• database
• smartcard
Authentication

It’s Alice!!!
Match is “fuzzy”

• Every time a biometric is presented, it looks slightly different

• E.g., fingerprint:
  • Rotation
  • Pressure
  • Angle of presentation
  • Chapping (NYC winters)

• And it may not work for everyone
  • E.g., people with small fingers, bricklayers
Key concepts

• **False acceptance rate (FAR)**
  
  …or “fraud rate”
  
  • Probability that wrong biometric or forgery (e.g., fingerprint) is accepted

• **False rejection rate (FRR)**
  
  …or “insult rate”
  
  • Probability that valid user is rejected

• U.K. banks set an FAR of 1%, insult rate of 0.01% [R. Anderson, *Security Engineering*]
  
  • …showing the emphasis on convenience over security

• iPhone TouchID has claimed (2013) an FAR of 0.002%
  
  • So in this setting, fingerprint is far from *unique*
Big security architecture questions

• Where is the template stored?
• How is the template protected?
• Where is the match performed?

Security is important because...
Revocation of biometrics is hard

You have only so many fingers…

First password

Second password
Biometric secrecy
Classical biometric authentication

Is it Woody? Yes, it’s Woody!
Classical biometric authentication

Is it Woody? Yes, it’s Woody!
Classical biometric authentication

Hello,
Mr. Woody Allen
In these scenarios, biometric data need not be kept secret

- Spoofing is difficult with human oversight
- Indeed, your face is public anyway
- (Assuming, of course, that passport is not a forgery)

But what happens when...
A human-guided process
Becomes automated?
Secrecy of biometric data is now more important to security

• **Reason 1:** Automation will mean relaxation of human oversight
  - More opportunity for spoofing
  - Holding up photos instead of presenting faces, fake fingerprints, etc.

Schiphol airport: Iris scanning
Secrecy of biometric data is now more important to security

- **Reason 2:** On-device and remote authentication
Attacks
Some attacks

“My voice is my passport”
Fake fingerprints
Eyeballs in a bag
Spoofing / cloning

- Apple TouchID
- Chaos Computer Club hack (Starbug)
  - A week or two after TouchID release
  - Moderately sophisticated attack converts fingerprint photo to wood glue prosthetic
Spoofing: Gummy fingers

Making an Artificial Finger from Residual Fingerprint

Materials

A photosensitive coated Printed Circuit Board (PCB)
“10K” by Sanhayato Co., Ltd.

Solid gelatin sheet “GELATINE LEAF” by MARUHA CORP

320JPY/sheet

200JPY/30grams

Spoofing: Gummy fingers

Residual Fingerprint

- Enhancing
- Capturing
- Image Processing

Fingerprint Image

- Printing
- Mask
- Exposing
- Developing
- Etching
- Mold

Digital Microscope

Adobe Photoshop 6.0

Cyanoacrylate Adhesive

Transparent Film

UV light

Photosensitive Coated PCB

Inkjet Printer

KEYENCE VH6300: 900k pixels

Canon BJ-F800: 1200x600dpi
Spoofing: Gummy fingers

Gelatin Liquid

Drip the liquid onto the mold.

Put this mold into a refrigerator to cool, and then peel carefully.
Spoofing: Gummy fingers

Mold: 70 JPY/piece
(Ten molds can be obtained in the PCB.)

Gummy Finger: 50 JPY/piece
Maybe we want weak security...

Carjackers swipe biometric Merc, plus owner's finger

Sometimes you might not want such great security...
By John Lettice, 4 Apr 2005

A Malaysian businessman has lost a finger to car thieves impatient to get around his Mercedes’ fingerprint security system. Accountant K Kumaran, the BBC reports, had at first been forced to start the S-class Merc, but when the carjackers wanted to start it again without having him along, they chopped off the end of his index finger with a machete.

Although security systems of this sort are typically fitted to high end cars (because of import duties, Kumaran’s car is reported to have been worth $75,000 “second-hand” - under the circumstances, we think we’d have said ‘at resale’), they’re not in essence complicated devices - requiring not only a fingerprint, but also a PIN.
Or perhaps we want liveness detection

• In fingerprint readers
  • Capacitance
  • Color changes
  • Perspiration

• For iris scanning
  • Pupil dilation
But that doesn’t always work either

- Gummy fingers are transparent and thus expose color
- For iris you can check pupil dilation, but…
- Thankfully, there’s a handbook…
Things can get grotesque

“…in countries like South Africa where fingerprints are used to pay pensions, there are persistent tales of ‘Granny’s finger in the pickle jar’ being the most valuable property she bequeathed to her family.” [R. Anderson, *Security Engineering*, Chap. 15]
Deployments and deployment challenges
Touch ID

• Uses secure hardware
  • Introduced in ARM A7 (now A10 in iPhone 7) with “Secure Enclave”
• (Apple Pay also benefitting)
HTC caught storing fingerprints AS WORLD-READABLE CLEARTEXT

Android biometric banks more Fort Nope than Fort Knox.
India’s AADHAAR system

- Holds fingerprints, iris scans, and facial scans of 600+ million people
- Used to deliver subsidies, deliver wages to bank accounts, control fraud, etc.
- Very different security problem than iPhone
  - iPhone generally holds one user's template
  - AADHAAR holds entire country's templates!
- Compromise endangers security of entire country!
- How to protect templates?
Fuzzy cryptography
Protecting big databases of biometric templates?

Ideally, we would **hash** templates.
Two problems

1. Hash easily cracked
   • FAR ➔ Guessing probability 0.002%
   • What is min-entropy?
   • Weaker than three-character password!
     • {a…z} + {A…Z} + {0…9}

2. Hashing won't work anyway!
   • Remember: Biometrics are "noisy"
     • Small reading errors / variations
Secure sketch

[Dodis, Reyzin, Smith '04]

• Let \( w \) be the secret
  • E.g., biometric template
• Let \( w' = w + e \) be a corrupted secret
  • E.g., biometric reading with noise \( e \)
• Secure sketch scheme has two functions:
  • \( \text{SS}(w) \rightarrow s \) (sketch)
  • \( \text{Rec}(w', s) \rightarrow w \) (if \( e \) is small)
• (Informal) security guarantee: \( s \) leaks little information about \( w \)
• Relies on error-correcting code
More formally

• \((M, m, m^*, t)\)-secure sketch has two functions:
  • \(\text{SS}(w) \xrightarrow{} s\) (sketch)
  • \(\text{Rec}(w', s) \xrightarrow{} w\)
• Correctness: \(\text{Rec}(w', s) \xrightarrow{} w\) (if \(\text{dist}(w' - w) \leq t\))
• Security: For any distribution \(W\) over \(M\), we have:
  \[\tilde{H}_\infty(W \mid \text{SS}(W)) \geq m^*\]

where \(\tilde{H}_\infty(A \mid B) \overset{\text{def}}{=} -\log \left( \mathbb{E}_{b \leftarrow B} \left[ \max_a \Pr[A = a \mid B = b] \right] \right)\)

i.e., adversary can recover value of \(W\) with prob. \(\leq 2^{-m^*}\).
Error-correcting codes: Setting

Alice

Bob

m

"Hello!"

"Hello!"

"Jello!"

"Jello!"

Noisy channel
Error-correcting codes: Intuition

Alice

Bob

m

"Hello!"

HHHeellllllooo!!!

NOISE

HHeelll*llioo!!!

Corrupted codeword

Codeword

decoding

"Hello!"
Error-correcting code

Key idea: decode / correct by mapping to nearest codeword
Error-correcting code

We can correct up to \( t = \text{floor}[(d - 1) / 2] \) errors

minimum distance \( d \)
(Linear) error-correcting code—some notation

- Alphabet $\Sigma$, a field
- Codebook $C \subset \Sigma^n$
  - Subspace (lattice) of dimension $k$
- Error-correction function $decode: \Sigma^n \rightarrow C \cup \bot$
  - Fixes corrupted codeword or fails and outputs $\bot$
(Linear) error-correcting code

- Alphabet $\Sigma$, a field
- Codebook $C \subset \Sigma$
  - Subspace (lattice) of dimension $k$
- Error-correction function $\text{decode}$: $\Sigma^n \to C \cup \perp$
  - Fixes corrupted codeword or fails and outputs $\perp$

- Message space $M = \Sigma^k$
- Mapping function $\text{map}$: $M \to C$
  - Maps message to codeword
  - Adds redundancy
- Unmapping function $\text{unmap}$: $C \to M$
  - Maps codeword back to message
Example

Minimum distance?

$m = 10$

map

unmap
Example

Binary (10, 2, 5)-code 
\((n, k, d)\)
Secure sketch uses codeword space only (no messages)

- Alphabet $\Sigma$, a field
- Codebook $C \subset \Sigma^n$
  - Subspace (lattice) of dimension $k$
- Error-correction function
  decode: $\Sigma^n \rightarrow C \cup \perp$
  - Fixes corrupted codeword or fails and outputs $\perp$
Secure sketch for Hamming distance

[JW99]
Secure sketch for Hamming distance

\[ JW99 \]
Recovery

$w', w, e$
Recovery
Recovery

decode
Secure sketch functions

- **SS(w)**
  - $c \leftarrow C$
  - $s \leftarrow w - c$

- **Rec(s, w)**
  - $w = \text{decode}(w' - s) + s$

- Suppose code can correct $t$ errors in the worst case

- Observe (if $|e| \leq t$):
  - $\text{decode}(w' - s) + s = \text{decode}(w' - w + c) + s = \text{decode}(c + e) + s = c + s = w$
Why is this secure?

(Assuming uniformly distributed $w$)
Why is this secure?
Given $s$ alone…

there are $|C|$ candidates for $w$!
Applying secure sketches to biometrics

- Server stores $s = SS(w)$ and $H(w)$
  - Given $s$, $H(w)$, hard to recover high-entropy, uniformly distributed $w$
- Treats $w$ like a password

![Diagram: Alice signs $w'$, which is stored in the server. The server returns $H(w)$, which is verified.]
But wait!

1. Most biometrics have low entropy
   • E.g., fingerprints
2. Hamming distance isn't metric used to match!
   • E.g., fingerprints treated as sets of minutiae
   • Possible exception to 1. and 2.: iriscodes

\[ H(w) = \text{Alice} \quad w' \quad \text{Rec} \quad s \quad w = H(w) \]
More constructions

[Dodis, Ostrovsky, Reyzin, Smith '08]

• Secure sketches for
  • Set difference
    • Good for fingerprints?
  • Edit distance
    • Good for password typos?

• Fuzzy extraction
  • Goal: Generate uniformly random key (bitstring) from $w$
  • Basic idea: apply extractor (universal hash) to $w$ after Rec
  • Good for turning biometric into cryptographic key
    • …or would with enough entropy
Another setting
Private face recognition

- Why is confidentiality important?
- Field agents shouldn't learn faces in database.
- FBI shouldn't learn queries involving innocent people.

Facial image captured by FBI field agent

Suspect?

FBI: database of suspects
Private face recognition

Osadchy, Pinkas, Jarrous, Moskovitch '11

• Functionality:
  • Client submits face \( v \) for identification
  • Client learns "match" if \( v \in V \); "no match" otherwise

• Goal:
  • Client learns "match" or "no match"; server learns nothing
  • (Vice versa also possible)

Target face

(blin[ed]) \( v \)

"match" or "no match"

Database \( V \) of suspects' faces
Key idea

• Face encoded as 900-bit vector $v$
  • String of feature, location pairs
  • Feature is binary vector indicating subset of similar entries in "part vocabulary"
  • Location is quantized distance from face center

Parts in vocabulary
Matching algorithm

• Matching now just Hamming distance computation:
  \[ |v - v_i| > T? \]
  over all \( v_i \in V \), where \( T \) is threshold for likely match.
• Privacy-preserving computation achieved using homomorphic encryption and oblivious transfer
• Performance per one-to-one match:
  • Client: 2.8 GHz dual core Pentium D processor and 2GB RAM
  • 213 seconds preprocessing time
  • 0.3 seconds online computation
Research challenge: Fuzzy crypto in practice

- Scads of published papers on fuzzy crypto for biometrics
- Not one rigorous end-to-end conversion of biometric to cryptographic key, i.e.:
  - Statistical characterization of biometric templates on significant population
  - Crisp metric for template matching (suitable for closed-form analysis)
  - Construction of secure sketch
  - Computation of entropy in key yielded by fuzzy extractor