홍익대학교
HONGIK UNIVERSITY

Network Security

<CH 5>

Youn Kyu Lee
Hongik University

Authentication

Part#1

Access Control

- **Authentication:** Are you who you say you are?
  - Determine whether access is allowed
  - Authenticate human to machine
  - Or authenticate machine to machine
- **Authorization:** Are you allowed to do that?
  - Once you have access, what can you do?
  - Enforces limits on actions
- Access Control = Authentication + Authorization

  (but “access control” often used as synonym for authorization)

Are You Who You Say You Are?

- Authentication based on
  - Something you **know**
    (or something you've forgotten)
    - For example, a password
  - Something you **have**
    (or something you had once)
    - For example, a smartcard
  - Something you **are**
    (or something you once were)
    - For example, your fingerprint

Something You Know

- Passwords, Passphrase, Secret Code, …
- Lots of things act as passwords!
  - PIN, Social security number, Mother's maiden name, Date of birth, Name of your pet, etc.
- Passwords: one of the biggest practical problems facing security engineers today

Trouble with Passwords

- Users do NOT select passwords at random
  (humans are incapable of securely storing or
  memorizing high-quality cryptographic keys)
  and attackers know that ..
  eg. 8 characters password *vs.* 8x8 bits crypto key
- Users REUSE passwords at multiple sites and attackers know that .. (and try to crack a poorly protected site to obtain them and …)
- Users often write down passwords on a paper, which can be stolen or be seen by others
- Users are vulnerable to social engineering attacks

Then, Why Passwords?

- Why is “something you know” more popular than “something you have” and “something you are”?
- **Cost**: passwords are ‘almost’ free
- **Convenience**: easier for admin to reset password than to issue a new thumb

Password Experiment

- Three groups of users : each group advised to select passwords as follows
  - **Group A:** At least 6 chars, 1 non-letter
  - **Group B:** Password based on passphrase
    (eg. It's 12 noon and I am hungry -> I'S12&IAH)
  - **Group C:** 8 random characters(alpha or numeric)
- Results
  - **Group A:** About 30% of passwords easy to crack
  - **Group B:** About 10% cracked
    - Passwords easy to remember
  - **Group C:** About 10% cracked
    - Passwords hard to remember

Password Experiment

- In each case, 1/3rd did not comply
  - And about 1/3rd of those easy to crack!
- Assigned passwords sometimes best
- If passwords not assigned, best advice is…
  - Choose passwords based on passphrase
  - Use password cracking tool to test for weak passwords
    (*Hashcat*, *John The Ripper*, …)
- Require periodic password changes? -- Good?
  - hiJude01, hiJude02, hiJude03, … for monthly changes
  - NIST(2017) : recommend long passphrase only changed on compromise
  - Microsoft(since 2019) : not impose password-expiration policy

Attacks on Passwords

- Attacker could…
  - Target one particular account
  - Target any account belonging to a specific target
  - Target any account on a target system
  - Target any account on any system
- Common attack path
  - outsider → normal user → administrator
  - May only require **one** weak password!
  - Typical way of penetration done by APT(Advanced Persistent Threat) attack

Password Retry

- Suppose system locks after 3 bad passwords. How long should it lock?
  - 5 seconds
  - 5 minutes
  - Until SA restores service
- Make lock time increase (exponentially) as a function of the number of wrong password trials

Password File

- Bad to store 'plaintext' passwords in a file
- But we need to verify passwords
- Cryptographic solution: **hash** the passwords
  - Store $y = h(password)$
  - Can verify entered password by hashing
  - Even with "password file," attacker does not obtain passwords easily due to h()'s one-way property
- But attacker can try a *forward search*
  - Guess x and check whether $y = h(x)$

Dictionary Attack

- Attacker can pre-computes $h(x)$ for all x in **dictionary** of common passwords
- Suppose Trudy gets access to password file containing hashed passwords
  - She only needs to compare hashes to her pre-computed dictionary
  - After one-time work, actual attack is trivial
  - *LinkedIn* passwords encrypted by SHA1 and leaked in 2012

    → only took 3 days to discover 90% of passwords
- Can we prevent this attack? Or at least make attacker's job more difficult?

Salt and/or Pepper

- Hash password with **salt**
- Choose random salt s and compute

$$y = h(\text{password}, s)$$

  and store (s,y) in the password file
- Note: The salt s is not secret
- Easy to verify salted password
- But Trudy must re-compute dictionary hashes *for each user*
  - Lots more work for Trudy!

Salt and/or Pepper

- Hash password with **pepper**
- Choose a *common* pepper p and compute $y = h(password, p)$ for *all users* and store all $(y)s$ in the password file and p somewhere else(eg. in a configuration file)
- Note: The pepper p is secret
- Easy to verify peppered password(once p is retrieved)
- But Trudy must
  - guess the pepper OR
  - retrieve it in another way

Number of Hash Iterations

- Another way to increase security of passwords

```
while iteration_counter > 0:
    hash = sha512(hash)
    decrement iteration_counter
```

- For a normal user, calculation will be a bit longer (still, order of milliseconds for a reasonable `iteration_counter` value(100 or so))
- But Trudy(not knowing any password) must do
  - millions of attempts
  - take additional hours/days to retrieve passwords

Other Password Issues

- Reset done via mostly SMS messages(secret code)
  - SIM swapping
  - Malware stealing SMS, Stolen/lost phone
- Failure to change default passwords
  - Wireless Router, Set-top Box, …
  - IoT devices with no I/O(keyboard, ..)
- Attacker-Installed/Spoofed WiFi Access Points
  - showing bogus Website(made exactly alike to the real one) to collect passwords given to it

Other Password Issues

- Social engineering
  - using phishing SMS, emails, phone calls, …
- Error logs may contain "almost" passwords
- Bugs, keystroke logging, spyware, etc.

- *Credential stuffing*
  : use leaked(via various ways) passwords to try to penetrate the most protected site without causing an alarm (it will work if passwords were reused)

Other Password Issues

- Modern Web browsers can cache passwords

  : browsers keep asking “save id/passwords” and autocomplete them using previously saved ones

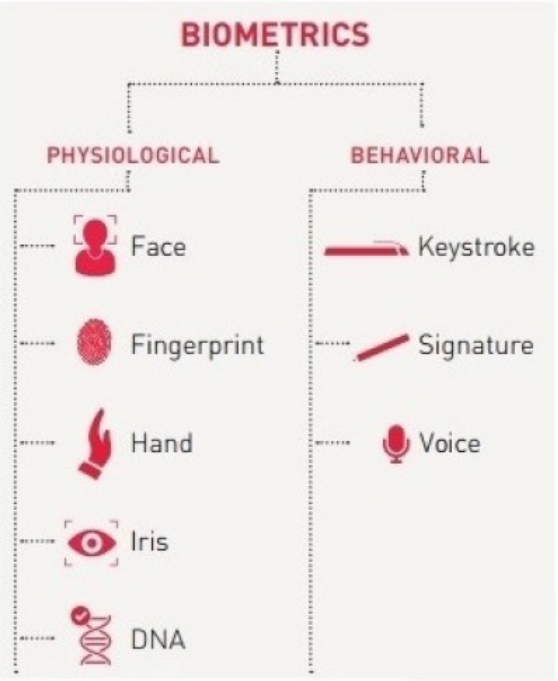
- Password managers(typically, as a browser add-on)

  : need to memorize a *master password*, which encrypts all the individual site passwords

Authentication

Part#2

Biometrics



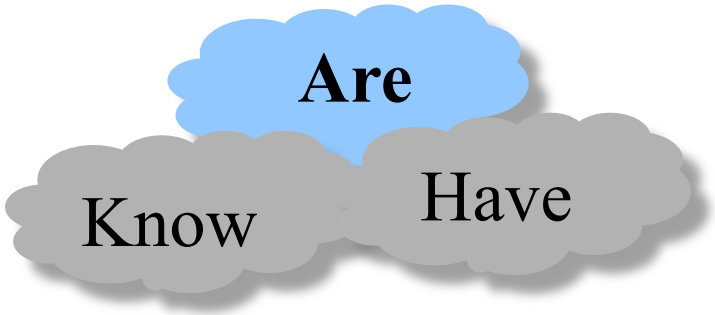
Source: Thalesgroup Biometrics in 2021
https://www.thalesgroup.com/en/markets/digital-identity-and-security/government/inspired/biometrics

Something You Are

-Biometric

- **"You are your key"** — Schneier

- Examples
  - Fingerprint
  - Iris
  - Handwritten signature
  - Facial recognition
  - Voice recognition
  - Gait (walking) recognition
  - …



Why Biometrics?

- More secure replacement for passwords
- Cheap and reliable biometrics needed
  - Today, an active area of research
- Biometrics **are** widely used in security today
  - Iris for extremely accurate authentication
  - Palm print for secure entry
  - Fingerprint to unlock car door, etc.
- But with some limitations
  - Non-negligible false-alarm rates in hostile environments
  - High cost for enrollment, usage, and maintenance
  - Application limited for crime scene forensics

Historical Biometrics

- Handwritten Signatures
  - Rare for signatures disputed in court cases, as the context mostly makes it clear who did what
  - So it actually works fairly well in practice, but it certainly is not bulletproof
    - tablet-based signature recognition systems(from early 1990's) record not just the shape of the curve but also its dynamics(the velocity of the hand, where the pen was lifted off …)
    - and they compare captured signatures against ones enrolled previously
    - these systems show error rates of several %s, at best
  - Works OK as a deterrent to attackers

Historical Biometrics

- Face Recognition
  - Oldest identification mechanism of all
  - Photo ID-based authentication used widely – accurate?
  - Neural network revolution since 2012
    - most accurate algorithm(thanks to CNNs) finds matching entries from galleries of 12 million individuals, with a miss rate approaching 0.1%
    - fastest algorithm needs just ½ second to match one against 10 million other faces
    - some algorithms can match side-view photos to galleries of frontal photos (called *pose invariance)*

Historical Biometrics

- Face Recognition
  - Neural network revolution since 2012
    - iPhone X's dot projector paints one's face with tens of thousands of dots and this hardware-based scheme helps to deal with makeup, some sunglasses and facial hair (and claimed to have error rate of one in a million)
    - needs improvement on how to handle variations in pose, age, illumination and expression
    - regulations against face recognition products vary across countries
- Fingerprints : will be discussed later

Historical Biometrics

- Voice Recognition (a.k.a. speaker recognition)
  - Identifying a speaker from a short utterance

    (≠*speech recognition*, which is for transcribing speech

    and need to ignore speech idiosyncrasies)
  - Varies whether recognition is text-dependent or not, whether environment is noisy, whether operation must be real time, and whether one needs to *verify speakers* or *recognize them* from a large set
  - A clever attack: *breaking up* intercepted messages from victim *into* quarter second segments that are then *cut and pasted* to provide new, deceptive messages
  - "Deepfake" editing software for voice and image morphing

Ideal Biometric

- **Universal** — applies to (almost) everyone
  - In reality, no biometric applies to everyone
- **Distinguishing** — distinguish with certainty
  - In reality, cannot hope for 100% certainty
- **Permanent** — physical characteristic being measured never changes
  - In reality, OK if it to remains valid for long time
- **Collectable** — easy to collect required data
  - Depends on whether subjects are cooperative
- Also, safe, user-friendly, etc., etc.

Biometric Modes

- **Identification**(for forensics) — Who goes there?
  - Compare **one-to-many**
  - Example: The FBI fingerprint database
- **Authentication** — Are you who you say you are?
  - Compare **one-to-one**
  - Example: Thumbprint mouse
- Identification problem is much more difficult
  - More “random” matches and needs more comparisons
  - In many cases, subjects are not cooperative at all
- Here we are mostly interested in authentication

Enrollment vs. Recognition

- Enrollment phase
  - Subject's biometric info put into database
  - Must carefully measure the required info
  - OK if slow and repeated measurement needed
  - Must be very precise
  - May be weak point of many biometric
- Recognition phase
  - Biometric detection, when used in practice
  - Must be quick and simple
  - But must be reasonably accurate (with low error rates)

Cooperative Subjects?

- Authentication — cooperative subjects
- Identification — uncooperative subjects
- For example, face recognition systems
  - used in Las Vegas casinos to detect known cheaters (terrorists in airports, etc.)
  - often do not have ideal enrollment conditions
  - subject will try to confuse recognition phase
- Cooperative subject makes it much easier
  - 1:1 question and yes/no decision
  - subjects are generally cooperative

Biometric Errors

- **False accept(fraud/type1)** vs. **False reject(insult/type2)**
  - False accept — Trudy mis-authenticated as Alice
  - False reject — Alice not authenticated as Alice
- For any biometric, can decrease fraud or insult, but other one will increase
- For example
  - 99% voiceprint match: low false accept, high false reject
  - 30% voiceprint match: high false accept, low false reject
- **Equal error rate:**
  - rate where *false accept rate* = *false reject rate*
  - A way to compare different biometrics

Fingerprints

- Long history of usage in a number of countries
- A systematic studies on fingerprint identification
  - F. Galton's scheme of classifying fingerprint patterns
  - E. Henry's indexing system: assigning 1 bit to whether or not a suspect's 10 fingers had a whorl(a type of circular pattern)
  - Used for essentially 2 purposes: identifying people and crime scene forensics (by police forces)
- Identifying people
  - FBI's NGI(Next Generation Identification) service: identifies about 8 thousand fugitives a month
  - US D. of Homeland Security's IDENT: holds fingerprints on 200 million aliens and matches them against a watch list of bad guys

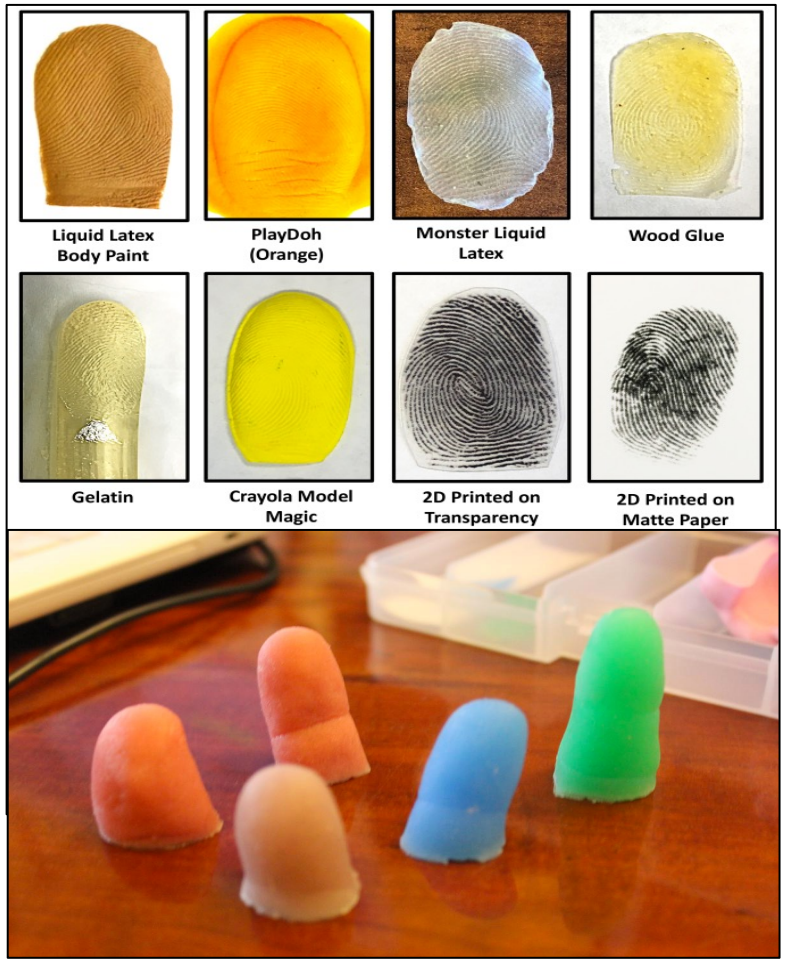
Fingerprints

- Identifying people
  - India's *Aadhaar* : has fingerprints and iris codes of most residents (initially to support welfare payment fraud detection)
  - Equal error rate of typical automatic fingerprint identification systems: slightly below 1% *per finger* (US DHS program requires each visitor's 10-prints!)
  - Other problems: some have damaged fingerprints and rare with no conventional fingerprint patterns at all
  - How to attack?
    - molded fingerprints with cooking gelatin
    - latent fingerprints reactivated(or transferred) using adhesive tape

Fingerprints

- Identifying people
  - How to attack?
    - retrieve other's fingerprints from photos in social media (Mitre project to develop software to harvest people's fingerprints from photos they post on social media)
    - 'masterprint': a fake fingerprint that can be worn on your fingertip and be used to match at least one of the partial prints derived from a typical finger (it works against 6% of users' prints)
    - thermal scanners defeated by rubber molded fingers
    - ultrasonic scanners defeated (accidently) via certain silicon screen protectors (some patterns on the protectors recognized as users' fingerprints)
    - and others(some are scary)…

Fingerprints



Fingerprints

- Crime Scene Forensics
  - Prints found at a crime scene matched against database
  - Forensic error rate has become extremely controversial
    - limit on the size and quality of the image taken
    - several cases convicted of murder mainly due to their crime scene fingerprints later being refuted by other evidences
    - experts are easy to fail esp. on obscured, dirty prints
    - fingerprints can be transferred using adhesive tape, can be framed by another criminal (or by the police)
  - The belief that fingerprint examination has a zero error rate is now widely ridiculed!

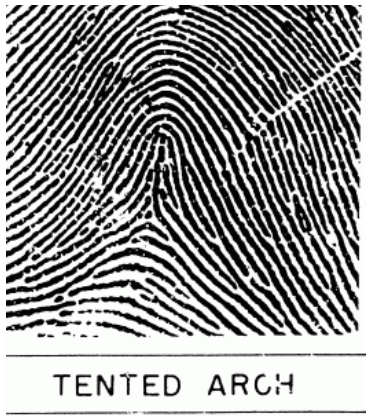
Fingerprint Comparison

- Examples of *loops*, *whorls*, *arches*, and *tents*
- *Minutia* (branches and endpoints of ridges) extracted from these features

Loop(double)

Whorl

Arch

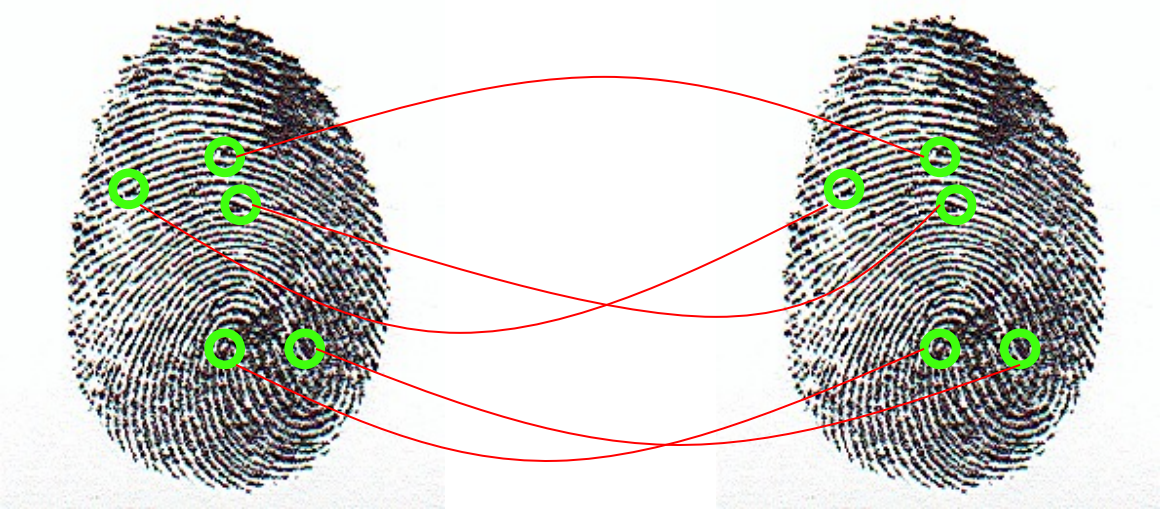


Tent

Fingerprint: Enrollment

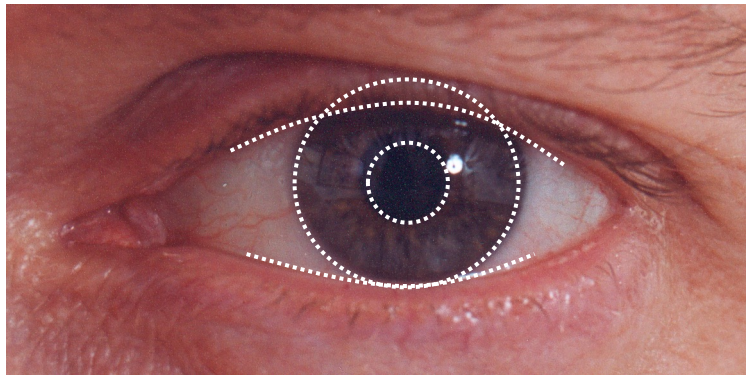
- Capture image of fingerprint
- Enhance image
- Identify points (Minutia)

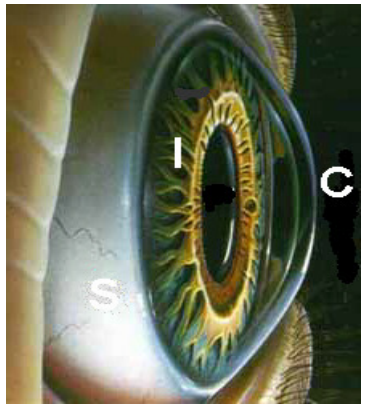
Fingerprint: Recognition



- Extracted points are compared with information st ored in a database
- partial/whole match?
- Aside: Do identical twins' fingerprints differ?

Iris Patterns





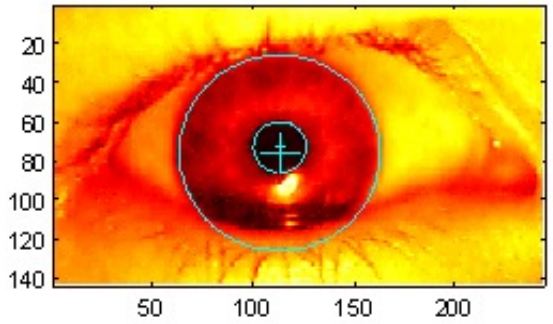
- Iris pattern development is “chaotic”
- Little or no genetic influence
- Different even for identical twins
- Different even for the two eyes of a single in dividual
- Pattern is stable through lifetime

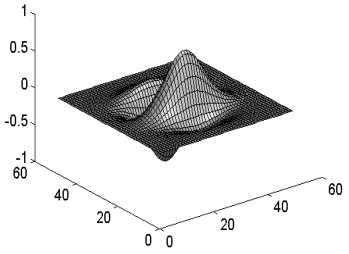
Iris Patterns

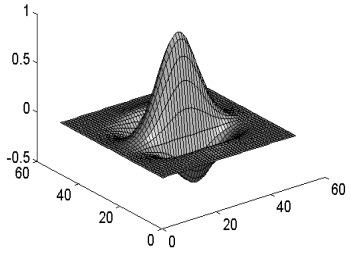
- Modern and innovative way of identifying people
- Has the best error rates of any biometric system when measured under lab conditions
  - False accept rate of (Very close to) zero
  - EER is less than 1/million
  - if one can tolerate a false reject rate of 1 in 10 thousand, the false accept rate would be < 1/trillion
  - In practice, the false reject rate is much higher than this (why? eyelashes, hangovers can cause camera not to see enough the iris) → 4% ~ 6%
- Large-scale deployments in UAE(detecting deportees who return with false papers) and in India(mandatory to billions of people for many purposes)

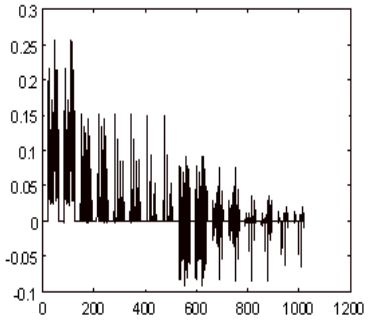
Iris Scan

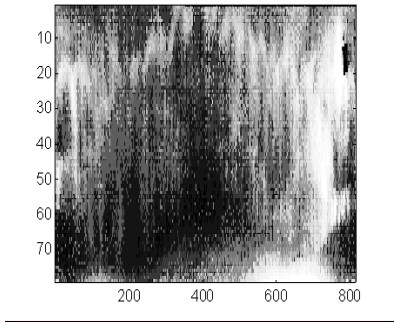
- Scanner locates iris
- Take b/w photo
- Use polar coordinates…
- 2-D wavelet transform
- Get *256 byte iris code*
  (by J. Daugman in 1993)











Measuring Iris Similarity

- Based on Hamming distance
- Define $d(x,y)$ to be
  - # of non match bits / # of bits compared
  - $d(0010,0101) = 3/4$ and $d(101111,101001) = 1/3$
- Compute $d(x,y)$ on 2048-bit iris code
  - Perfect match is $d(x,y) = 0$
  - For same iris, 'expected' distance is 0.08
  - At random, expect distance of 0.50
  - Accept iris scan as match if distance < 0.32

Attack on Iris Scan

- Good **photo** of eye can be scanned
  - Attacker could use photo of eye

- To prevent photo-based attack, scanner could use light to be sure it is a "live" iris OR measure *hippus* – a natural fluctuation in the diameter of the pupil that happens at about 0.5Hz

  (countermeasure: wear a contact lens with the target's iris patterns printed on it)

Equal Error Rate Comparison

- Equal error rate (EER): fraud rate = insult rate
- **Fingerprint** biometric has EER of about < 5%
- In theory, **iris scan** has EER of (very close to) zero
  - But in practice, may be hard to achieve
  - Enrollment phase must be extremely accurate
- Iris codes are currently the most powerful/accurate biometric used in various applications
- Most biometrics much worse than fingerprint!
- Biometrics useful for authentication…
  - …but identification biometrics would need lots of improvement to be useful in real-life cases

Other biometrics

- Typing patterns(keystroke dynamics)
- Vein patterns
- Hand geometry
- Facial thermograms (maps of the surface temperature of the face, from infrared images)
- The shape of the ear
- DNA
  - Valuable tool for crime scene forensics
  - Determining parenthood in child support case
  - But, slow and expensive for real-time applications
  - Privacy problem (reconstruct many info about a person from their DNA sample)

Attacks on Biometric systems

- Hard to tell ‘freshness’ of forensic biometric sample (fingerprint or DNA sample can be ‘planted’)
- Most are vulnerable to suitable ‘recordings’ esp. via unattended operation of biometric auth. devices: voice recognition, iris scanners by photos on a contact lens, molds of fingerprints
- Most are not as accurate for all people: manual workers with damaged fingerprints (some cases hardcore criminals doing this deliberately)

Attacks on Biometric systems

- Most are vulnerable to collusion: Alice let Bob to take a rubber impression of her fingertip and use it to withdraw money from her account. Alice produces a watertight alibi and claims there's a theft and request her money back

Attacks on Biometric systems

- A combination of two or more biometrics will improve either false-accept rate or false-reject rate, while making the other worse: eq. 2 burglar alarms at your home: the prob. of defeating both goes down while the number of false alarms goes up
- 'physiological' biometrics are very hard(if not impossible) to renew / be reissued

Biometrics: The Bottom Line

- Biometrics are hard to forge
- But attacker could
  - Steal Alice's thumb
  - Photocopy Bob's fingerprint, eye, etc.
  - Subvert software, database, "trusted path" …
- And how to revoke a "broken" biometric?

  → a BIG problem (different from a "broken" password/passphrase, which can be "easily" replaced with a new one)
- **Biometrics are not foolproof !**

Q & A

aiclasshongik@gmail.com