CE879 - Information Security Mng. & Eng.

Lecture 4: Identity Protection

Department of Computer Engineering
Sharif University of Technology
Spring 1404

S4Lab

Acknowledgments: Some of the slides are fully or partially obtained from other sources. A reference is noted on the bottom of each slide to acknowledge the full slide or partial slide content. These slides were initially developed by Seyedeh Atefeh Musavi and Mehdi Kharrazi.

Identity Protection

S4Lab

- What is the identity ?
- What are its usages and importance?
- How are you identified

What is the identity?

Identity

S4Lab

- What does your identity consist of?
  - Given to you:
    - Name? DoB? National ID?
    - Other examples?

Image from *The New Yorker* cartoon by Peter Steiner, 1993.

[Image: The New Yorker cartoon by Peter Steiner, 1993.]

When a concrete identity is given

S4Lab

- For online banking or eCommerce, many types of info is needed:
  - Email, address, phone number, or age.
- In eCommerce scenarios the payment process can be painful.
  - Entering a 16+ digits credit card number manually is painful on a touchscreen.
- Concrete identities:
  - Services such as PayPal, Amazon Payments or Google Wallet.
    - Tokenizing sensible credentials such as the payment details.
  - State-provided OpenID in the citizen’s ID card in Lithuania for election.

[LeBlanc, J., & Messerschmidt, T., Identity and data security for web development: Best practices. " O'Reilly Media, Inc.”, 2016]

Identity

S4Lab

- What does your identity consist of?
  - Given to you:
    - Name? DoB? National ID?
  - Provided by you:
    - Personal Knowledge
    - Your face, voice, fingerprint
    - your behavior
    - etc.

[Image: The New Yorker cartoon by Peter Steiner, 1993.]

Personal knowledge repo

S4Lab

- Social network apps are repositories for knowledge about people.
- Social identity came up with the rise of social networks
  - A very moderate form of identity, people tend to share quite casually.
  - Profile information often concentrate on social connections, interests and hobbies.
- Facebook or Google+ allow the user to quickly access other services using their already populated profiles.
- Leveraging social identity is completely valid and even encouraged for services such as games, media consumption and of course social networks.

[LeBlanc, J., & Messerschmidt, T., Identity and data security for web development: Best practices. " O'Reilly Media, Inc.", 2016]

How much do you share?

S4Lab

- A good example is use of your location as your ID.
- Which granularity is required depends on the application.
- Your postal address (when purchasing a good).
- Your BLE proximity to a beacon for proximity marketing (up to 100 meters)
- Your GPS position for outdoor positioning services (2-5 meter accuracy).
- Your proximity by a Wi-Fi RTT for indoor positioning services(1-2 meter accuracy).

[Image: https://ukdiss.com/examples/mobile-application-ibeacons.php]

How true/fake would it be?

S4Lab

- IDs are some thing which always have associations with you.
  - How much this association should regards to your real character?
- Real IDs
  - Your true name: University services (e.g. here!)
  - A chosen username: no need to be true.
- Pseudonym with some information about the subject : Car tags (for ordinary observer)
- Pure pseudonym : a random cookie in your browser.
- Fully Anonymous: is this an ID?!

AOL case

- In 2006, AOL released 20 million search queries for 650.000 users.
- (Pseudo)-Anonymized by removing AOL id and IP address.
- Easily de-anonymized in a couple of days by looking at queries.
- Thelma Arnold's identity was betrayed by AOL records of her Web searches, like ones for her dog, Dudley, who clearly has a problem.

[Image: https://www.nytimes.com/2006/08/09/technology/09aol.html]

http://benjamin-nguyen.fr/CAPPRIS/Castelluccia-anonymisation.pdf

PII

S4Lab

- Personally Identifiable Information
- Information that can be used to distinguish or trace an individual's identity, either alone or when combined with other personal or identifying information that is linked or linkable to a specific individual.
- Different interpretation and legislation in different countries
- Is your IP a PII?

European Court of Justice rules IP addresses are personal data

Court finds details may be 'personal' if website can identify people using ISP data

Wed, Oct 19, 2016, 16:51

[https://www.irishtimes.com/business/technology/european-court-of-justice-rules-ip-addresses-are-personal-data-1.2835704]

[https://www.gsa.gov/reference/gsa-privacy-program/rules-and-policies-protecting-pii-privacy-act]

Identity usages

S4Lab

- Usage:
  - History —> User Profile
    - Search history, background image, etc.

User Profile

S4Lab

- Enhancing user experience by utilizing identity :
  - Login doesn’t necessarily have to be the first point of contact for users and often harms the conversion process of turning visitors into users by forcing them to register or login.
  - Leveraging existing profiles, such as a user’s social identity, can help easing the way once the user did decide to register by pre-populating profile information and therefore lowering the amount of information the user has to type in manually.

[LeBlanc, J., & Messerschmidt, T., Identity and data security for web development: Best practices. " O'Reilly Media, Inc.”, 2016]

Identity usages

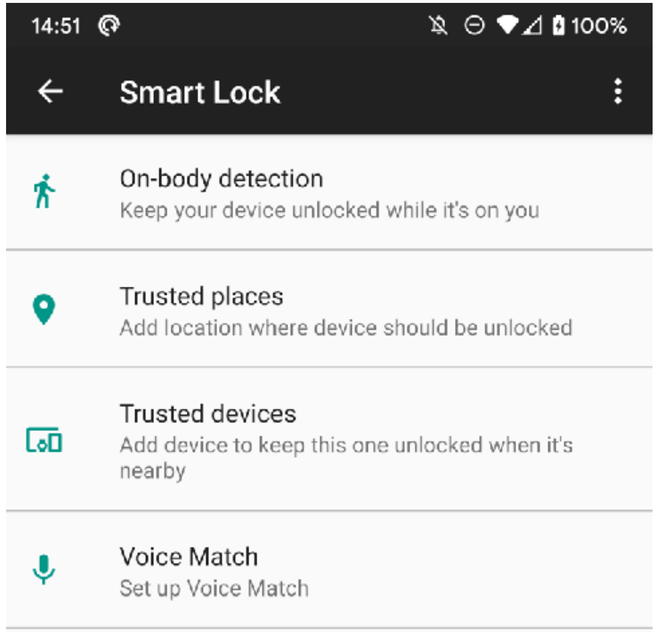
S4Lab

- Usage:
  - History —> User Profile
    - Search history, background image, etc.
  - **Authentication —> Access/Service**
    - **Bank account access, Online shopping**

An example for Access/Service

S4Lab

- Introducing the concept of Trust Zones.
- Google introduced this concept for Android as a feature known as Smart Lock:
  - On-body detection
  - Trusted places
  - Trusted devices
  - Trusted faces
  - Voice match



[Image: https://www.androidpolice.com/2019/09/04/trusted-face-smart-unlock-method-has-been-removed-from-android-devices/]

[LeBlanc, J., & Messerschmidt, T., Identity and data security for web development: Best practices. " O'Reilly Media, Inc.", 2016]

Identity usages

S4Lab

- Usage:
  - History —> User Profile
    - Search history, background image, etc.
  - Authentication —> Access/Service
    - Bank account access, Online shopping
  - **Anonymity —> John/Jane Doe**
    - **Anonymous web surfing, Whistleblower**

What is anonymization?

S4Lab

- Any technology which removes personally identifiable information (PII).
- You should have control over your identity through:
  - Your Internet activities.
  - Your other data may become available online by other parties.
- OK, So let’s remove me (my PIIs) from every DBs.
- Not possible why?
  - I would need them later myself.
  - The data market needs your data!

Privacy-enhancing technologies

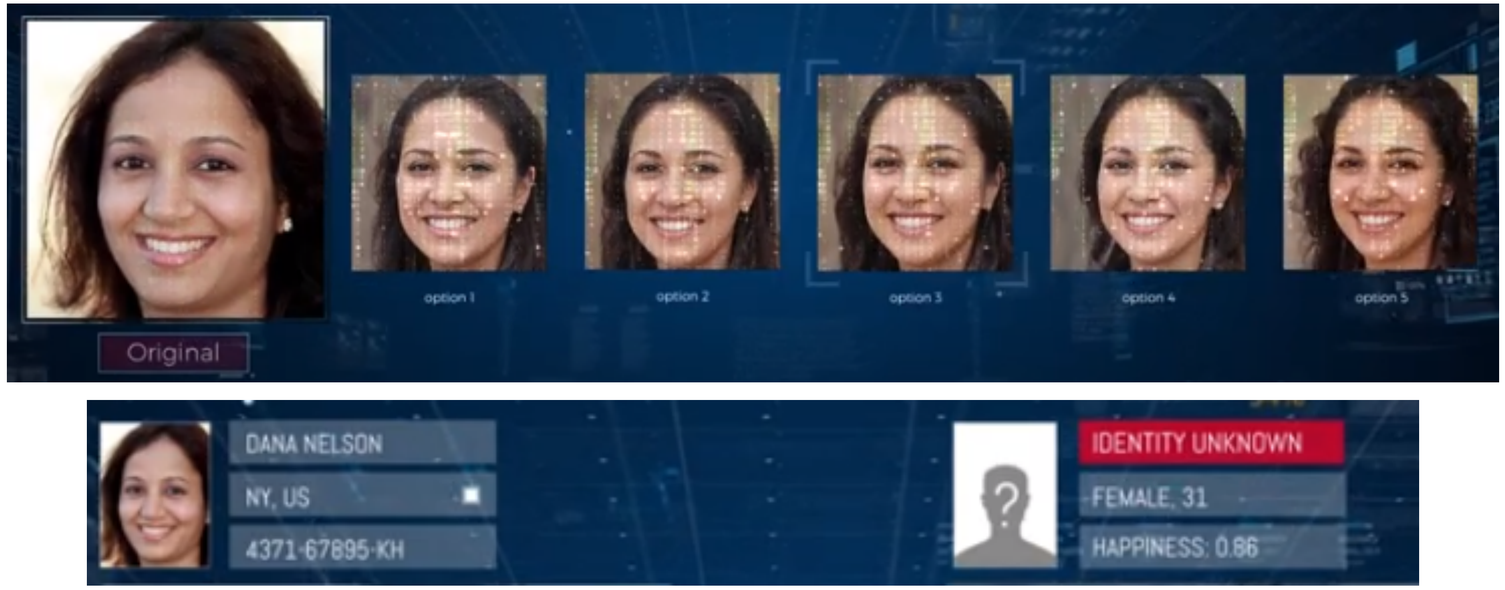
S4Lab

- Best technologies are those which:
  - Convince you/regulations that the records are anonymous.
- Enough?

Privacy-enhancing technologies

S4Lab

- Best technologies are those:
  - Convince you/regulation that the records are anonymous.
  - But the data is still valuable to be analyzed.
  - Example: Maintaining key face attributes.



[Image: https://www.d-id.com/smart-video-anonymization/]

Some Data anonymization methods…

S4Lab

- Random perturbation
  - Input perturbation
  - Output perturbation
- Generalization
  - The data domain has a natural hierarchical structure.
- Suppression

[http://benjamin-nguyen.fr/CAPPRIS/Castelluccia-anonymisation.pdf]

How identification takes place?

Identification process

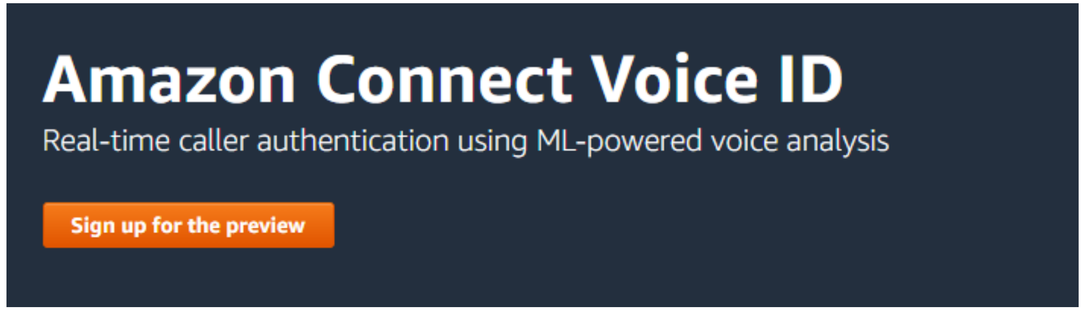
S4Lab

- Proving you are, who you say:
  - Provided to you:
    - 2FA, OTP …
  - Provided by you with explicit consent:
    - Password, Secret knowledge.
  - Provided by you without explicit consent:
    - Your face, your voice, your behavior.

Identification process

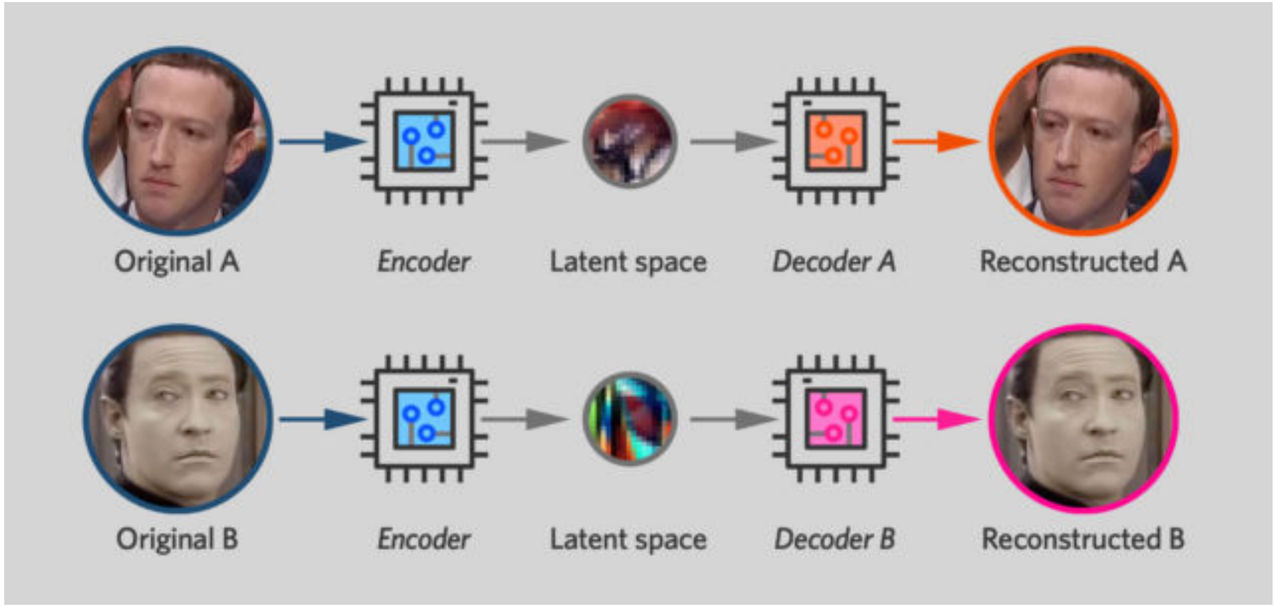
S4Lab

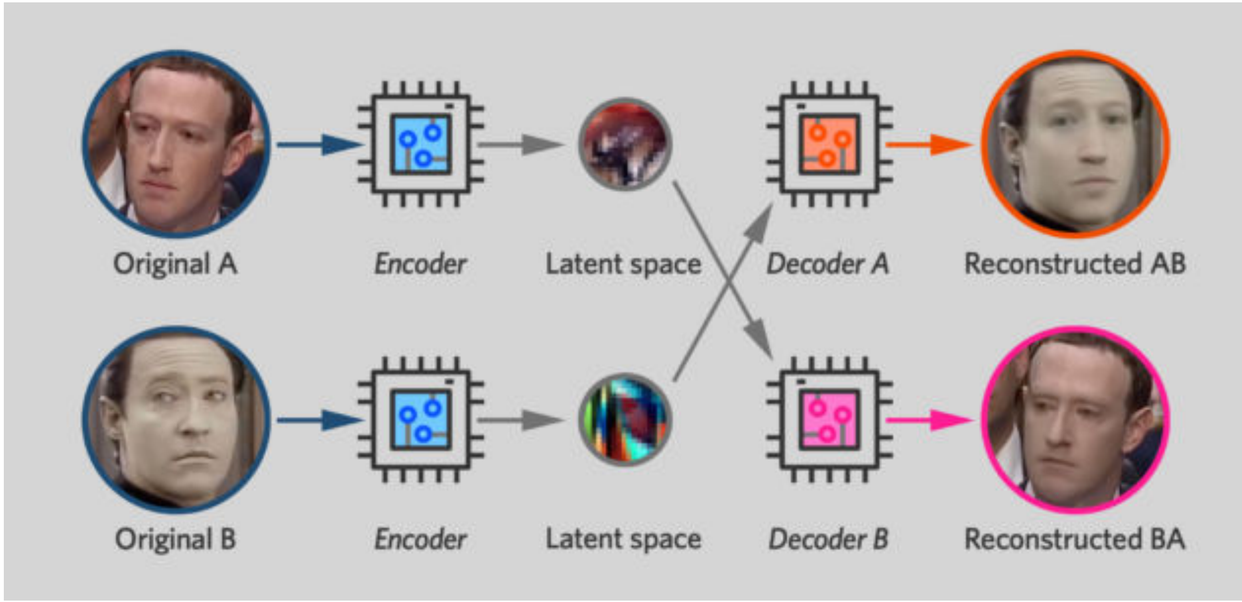
- Proving you are, who you say:
  - Provided to you:
    - 2FA, OTP …
  - Provided by you with explicit consent:
    - Password, Secret knowledge.
  - Provided by you without explicit consent:
    - Your face, your voice, your behavior.
- Proving you are not, whom they say you are:
  - Deep Fake



[Image: https://aws.amazon.com/connect/voice-id/]

Deep fake





- Two auto-encoders.
- One trained by thousands of video frames of you, another by the target.
- The outputs are then swapped.

[https://jeritanyangnikmat.blogspot.com/2019/12/i-created-my-own-deepfakeit-took-two.html]

Deep fake

S4Lab

- So let’s imagine that the court does not accepts media as an acceptable evidence unless it is shown to be authentic by the state-of-the-art deep fake detection tools.
- Is the deep fake’s problem solved?

Deep fake dilemma

S4Lab

- We may have complex deep networks to provide fake detection tools.
- But we still lack sufficient processes to control the disinformation process against an identity.
  - Twitters blue tick?
- So how you –as an individual – protect your self against deep fakes?
  - Recording the original media in every case?
  - Limiting your web images/videos (in number of images or special angles)?
  - Use a ledger technology to be informed or able to show modifications to your media?
  - Use of controlled capture
  - Are these suggestions even possible?

Deep fake dilemma

S4Lab

- So we have some original but fake (none genuine) data.
- New ideas to handle the technology.
  - “don’t panic but prepared” approach.
- What about copy-right of these fake products?
- Can you claim your product (e.g. an animation) has been copied?
- Technologies and startups to provide/verify authentic data:
  - “controlled capture” and “verified-at-capture” approach (e.g. Serelay):
    - With controlled capture, an image, video or audio recording is cryptographically signed, geo-tagged, and timestamped.
    - Verified capture in order to verify quickly, consistently and at scale.
    - The applications should be present at the point of capture.

[https://lab.witness.org/projects/synthetic-media-and-deep-fakes/]

Deep fake attacks

S4Lab

- Deepfake scammer walks off with $25 million (Feb. 2024)
- Significant financial loss suffered by a multinational company's Hong Kong office, amounting to US$25.6 million.
- A phishing message, purportedly from the company’s UK-based chief financial officer, instructing them to execute a secret transaction.
- Simulate a multi-person video conference where all participants (except the victim) were fabricated images of real individuals.
- The scammers were able to convincingly replicate the appearances and voices of targeted individuals using publicly available video and audio footage.

[https://arstechnica.com/information-technology/2024/02/deepfake-scammer-walks-off-with-25-million-in-first-of-its-kind-ai-heist/]

Identity-based abuses/attacks

Identification is a growing concern in IoT

S4Lab

- We'll be authenticating ourselves to things all the time, and they'll be authenticating themselves to us.
  - To know who it should talk to, who it should listen to, and who is allowed to control it.
  - How does your car, know you?
  - Little things like toys and smart light bulbs?
- Not just the humans but machines require reliable IDs and identification.
  - Your thermostat will want to talk to your furnace.
  - Your appliances will want to talk to your electric meter.
  - Your toys will want to talk to each other.
- Attacks will have serious consequences.
  - If I can impersonate you to your devices, I can take advantage of you.
  - This is the identity theft of the future, and it's scary.

[Schneier, Bruce. Click here to kill everybody: Security and survival in a hyper-connected world. WW Norton & Company, 2018.]

The identity theft

- AUTHENTICATION IS GETTING HARDER, AND CREDENTIAL STEALING IS GETTING EASIER.
- Credential stealing doesn’t require finding a zero-day or an unpatched vulnerability, plus there’s less chance of discovery, and it gives the attacker more flexibility in technique.

[Schneier, Bruce. Click here to kill everybody: Security and survival in a hyper-connected world. WW Norton & Company, 2018.]

Financial motivations: Tax scam "DirtyDozen"

S4Lab

- The sensitive information included in the Wage and Tax Statement (W-2) has always been appealing to impostors
  - Used an IRS impersonation in 10% of identity deception-based e-mails in this reporting year.
- As a result, valid W-2 forms and standard US Individual Tax Return (1040) forms are available on the dark web at a cost ranging between US $1 and US $52.
- This material, combined with the Social Security Numbers (SSN) and birth dates, which are also available, allows any inexperienced hacker willing to invest an amount of US $1,000 to legally access a United States-based bank account, file a false tax return, claim a refund and cash-out an investment that has doubled or tripled.

[Identity Theft - ENISA Threat Landscape 2020]

Sim jacking

S4Lab

- Or SIM swap scam, Port-Out scam, and SIM splitting

[Image: https://www.wired.com/story/jack-dorsey-twitter-hacked/]

What if your phone dies!

S4Lab

- How the mobile network knows your identity?
  - To send/receive your calls.
- So what if the customer loses the phone/wants to change provider?
  - Get another sim but with similar number.
  - So the attacker buys a cheep sim/phone.
  - Mobile carrier should believe that the attacker is the genuine customer.
  - The attacker knows the required info.
- How?
  - Maybe by open source intelligence (OSINT).
  - Purchase it off the Dark Web.
  - Phish you for it.
  - Or by bribing.
- The goal is to port a telephone number to another device (with different sim card).

[https://www.intrust-it.com/cant-text-or-make-calls-it-could-be-sim-jacking/]

So what?

S4Lab

- Bypass 2FA!
  - Authentications binded to sim cards (sms/call).
- Bank accounts, cryptocurrency wallets, access to emails, high prestige social media account.

[https://vpnoverview.com/internet-safety/cybercrime/identity-theft/]

Process/technology failures

S4Lab

- Two failures; (i) Procedure (ii) Technology.
- Process failure:
  - Allowing cellular network subscribers to easily move between networks and retain their phone numbers.
  - The identity verification step is not always carried out in the most rigorous manner to check the identity.
- Technology failure:
  - When a factor challenge is made with SMS, the sender has no way of knowing if the phone has been simjacked or if another person is using the handset.
  - Although smartphones increasingly have locking methods including various biometric controls, these are all in vain if the SMS is being sent directly to the criminal.
- Reducing crimes such as simjacking is just one part of a wider Customer Identity and Access Management (CIAM) push by a growing number of organizations that includes retailers and banks.

[https://www.vanillaplus.com/2020/04/07/51758-jacked-in-the-box/]

Technology amendment

S4Lab

- Multi-factor authentication (MFA) on all your accounts and use an authenticator app (AA).
- MFA by AA doesn’t send one-time passcodes (OTPs) to you by text or call your phone number to verify your identity.
- Instead, you use a specialized app (there are several good ones from LastPass, Microsoft and Authy) on your smartphone to receive the OTP that you are required to enter when you sign in.
- Authenticator apps are not linked to your SIM.

[https://www.intrust-it.com/cant-text-or-make-calls-it-could-be-sim-jacking/]

Further Reading

S4Lab

- https://theconversation.com/deepfake-videos-could-destroy-trust-in-society-heres-how-to-restore-it-110999
- https://www.cbc.ca/news/business/marketplace-online-prices-profiles-1.4414240