Introduction

Nicola Zannone
Outline

Defining Privacy

Privacy Threats

Applications
- Location-based services
- RFID
- Social Networks

Privacy Enhancing Technologies
Defining Privacy

What is Privacy?

Give your own definition of privacy

Compare with your neighbor and discuss
Defining Privacy

- Abstract and subjective concept, hard to define
  - Dependent on cultural issues, study discipline, stakeholder, context
- Popular definitions:
  - “The right to be let alone” (Brandeis & Warren, 1890)
    - Focus on freedom from intrusion
  - “The right of the individual to decide what information about himself should be communicated to others and under what circumstances” (Westin, 1967)
    - Focus on control
    - Focus on information self-determination
  - “The freedom from unreasonable constraints on the construction of one’s own identity” (Agre, 1999)
    - Focus on autonomy
Privacy Paradigms

Privacy as Confidentiality

“the right to be let alone”

Privacy as Control

“the right of the individual to decide what information about himself should be communicated to others and under what circumstances”

“the freedom from unreasonable constraints on the construction of one’s own identity”

Privacy as Practice
What Privacy is About

- Privacy concerns personal information
- Privacy is more than confidentiality
  - Freedom from intrusion
  - Control of information about oneself
  - Autonomy
Privacy Threats

“On the Internet, nobody knows you’re a dog.”

July 5, 1993
Privacy Threats

“On the Internet, nobody knows you’re a dog.”

July 5, 1993
Privacy Threats

How much privacy do we have left?

Make your own list of privacy threats

Compare with your neighbor and discuss
Privacy Threats

Data Life Cycle

- Create
- Store
- Use
- Archive
- Share
- Destroy
Privacy Threats

Taxonomy of Privacy Threats (Solove)

- Information collection
  - surveillance
  - interrogation

- Information processing
  - aggregation
  - identification
  - insecurity
  - secondary usage
  - exclusion

- Information dissemination
  - breach of confidentiality
  - disclosure
  - exposure
  - increased accessibility
  - blackmail
  - appropriation
  - distortion

- Invasion
  - intrusion
  - decision interference

Secondary Usage

Use information for a purpose other than the one for which it was obtained

- Use of customer information for marketing
- Sale (or trade) of consumer information to other businesses
- Government agencies’ use of consumer database
Aggregation

Combining and comparing information from more than one database

- Why aggregation is a problem?
  - Data are already in the system
  - Individual pieces of data may be not very telling
  - But when combined together, they may reveal more information about a person
Aggregation

Aggregated information about a person is often used to judge her
- Sharing of government agencies’ databases to detect frauds
- Insurance companies can make decisions based on information of diseases, lifestyles, etc.
- Companies can reject job applications using information from social network websites
- Aggregations of credit reports used to evaluate a person’s financial reputation for granting loans
Video Surveillance

CCTV monitoring at public places

- airports
- malls
- ...

Other Threats

- Invisible Information Gathering
  - Satellite surveillance
  - Loyalty cards (e.g., supermarket club cards)
  - Web-tracking data; cookies
  - ISP monitoring
    - Your ISP “knows” every site you visit
    - Google stores your search history

- Profiling
  - Use of customer preferences to predict behaviors of people
  - Customer profiles used to determine their propensity toward a product/service
  - Government agencies create descriptions of possible terrorists

- Identity Theft

- Discrimination
  - Unfair or unequal treatment of people belonging to a certain category or to a minority, without regards to individual merit
How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did

Every time you go shopping, you share intimate details about your consumption patterns with retailers. And many of those retailers are studying those details to figure out what you like, what you need, and which coupons are most likely to make you happy. Target, for example, has figured out how to data-mine its way into your womb, to figure out whether you have a baby on the way long before you need to start buying diapers.

Charles Duhigg outlines in the New York Times how Target tries to hook parents-to-be at that crucial moment before they turn into rampant — and loyal — buyers of all things pastel, plastic, and miniature. He talked to Target statistician Andrew Pole — before Target freaked out and cut off all communications — about the clues to a customer’s impending bundle of joy. Target assigns every customer a Guest ID number, tied to their credit card, name, or email address that becomes a bucket that stores a history of everything they’ve bought and any demographic information Target has collected from them or bought from other sources. Using that, Pole looked at historical buying data for all the ladies who had signed up for Target baby registries in the past. From the NYT:

[Pole] ran test after test, analyzing the data, and before long
Privacy Threats

Why now?

Computers not needed for the invasion of privacy

Digitization of information make it possible to

- collect and store massive amount of personal information
- correlate information from different sources
- access information on the network locally or remotely

Computers simply make new threats possible and old threats more powerful!!
Applications

Location-based services

RFID

Social networks
How many ways can we be located today?
How many ways can we be located today?

- cell phone (turned on?)
- laptop
- credit card at the gas station
- bank card in the ATM machine
- driving through a monitored intersection
- security camera at the supermarket
- scan badge to enter a building
- ...

Applications
Location-based services
Applications
Location-based services

Location Based Services

- location-based traffic monitoring and emergency services
  - e-Call, traffic congestion control
- location/service finder:
  - where is the nearest restaurant, gas station,...
  - variable pricing applications
  - congestion pricing
  - pay-as-you-drive

- social applications
  - Geotagged Twitter
  - Google Latitude
    (currently integrated in Google Maps)
Applications
Location-based services

Why is this a problem?

➤ do you want to be seen at certain locations?
   ➤ abortion clinic, AIDS clinic, business competitor, or political headquarters

➤ what can be inferred about a person based on location?
   ➤ any important location...
      ➤ desk in a building
      ➤ home location
      ➤ future locations
   ➤ and even identification!
What is RFID?

- **Radio Frequency Identification**
- A micro-chip transmitting data when exposed to radio waves
- Conceptually similar to bar code
  - Used to identify and track objects of interest
RFID Data Collection Process

RFID Tag → RFID Reader → Middleware → EPC Network → Enterprise System
RFID Applications

- Toll payment transponder
- Libraries
- Passports
- Human implantation for medical purposes
2006: President of Colombia agreed to require Colombian citizens to be implanted with RFID chips before they could gain entry into the US for seasonal work.

2008: UK jails considering RFID implants for prisoners.

2015: Tenants of a hi-tech office block in Sweden implanted with RFID chips in order to gain access to the office and operate photocopiers.
RFID – Privacy Issues

- Traceability
- Disclosure of embarrassing information
- Discrimination
- Target by muggers
- …
Social Networks

- Many people publish their personal information on social networks
  - Facebook, MySpace, LinkedIn, etc.

Privacy paradox

There is only one thing in the world worse than being Facebook stalked, and that is not being Facebook stalked.

Atwan & Lushing, 2008
Social Networks – Privacy Issues

- Used by professors, parents, and employers
- Used by law enforcement and college administrators for disciplinary action
- Used for surveillance and data mining
- Used by merchants for marketing
- Use by other members for criminal purposes (e.g., stalking, invasion of privacy)
Social Networks – Privacy Issues

Mafia mobster arrested after girlfriend posts picture of herself in Spain on Facebook

By NICK PISA
Last updated at 4:24 PM on 25th August 2011

One of Italy’s most-wanted Mafia mobsters has been arrested after his blundering girlfriend posted pictures of them on Facebook.

Salvatore D’Avino, 39, was traced by detectives after they spotted the snaps of his heavily pregnant girlfriend Brada Hint in front of sign giving the name and location of a Spanish restaurant.

Italian police alerted the Spanish authorities who traced the couple to their hideout in Marbella on the Costa del Sol and arrested D’Avino.
Scott Torgeson, a radio show host in Columbus, OH, was allegedly fired for making an insulting tweet about former NFL star and current ESPN analyst Desmond Howard. From Radio Ink Magazine:

Over a week ago Torgeson tweeted, “I wish Desmond Howard would get fired or die so I can watch Gameday again.” Today he was officially fired by the station.
How can we protect privacy?

Make your own list of privacy technologies

Compare with your neighbor and discuss
Recall Privacy Paradigms

Privacy as Confidentiality

"the right to be let alone"

Privacy as Control

"the right of the individual to decide what information about himself should be communicated to others and under what circumstances"

"the freedom from unreasonable constraints on the construction of one’s own identity"

Privacy as Practice
Scenario

queries

results
Privacy as Control: Assumptions

- Collection and processing of personal information is useful and necessary
  - Search engines: historical search data helps improve search algorithms
  - Hospital: health records
- Organizations have an interest in protecting user privacy
  - Organization trusted to enforce user privacy preferences
- Privacy problems arise when personal information is misused
Privacy as Control: Threats

- Database of queries can be breached
  - information made public
  - information abused by malicious insider and sold for profit
- Query data used for illegitimate purposes (secondary usage)
- Query data shared with other parties without user consent (e.g., advertisers)
- Time queries are stored (“right to be forgotten”)
Privacy as Control: Goals

- Focus on:
  - control over personal data
  - compliance with data protection regulations
- Providing individuals with means to control use of their information
  - Informed consent, privacy settings
- Providing organizations with means to
  - define and enforce security and privacy policy
  - prevent/detect misuse of personal information
Privacy Settings

- Allow users to specify their privacy preferences
  - privacy preference languages, e.g. APPEL, XPref
- Make it easier for users to configure their privacy settings
  - default suites of privacy settings
  - privacy wizards that automate configuration of settings

**Note:** Enforcement of privacy settings is done by the organization
Purpose-based Access Control

- Regulate access to data
- Privacy-aware Access Control languages
  - Specify which action a user can perform on a given objects
  - Specify the allowed usage of data
- Policy enforcement
  - Ensure the purpose of data access is complaint with intended purpose

**Note**: No control after disclosure of data
Purpose Control

- Verify whether data have been used for the intended purpose
- Auditing Mechanism
  - log data access and processing operations
  - analyze logs to detect policy violations
Privacy as Control: Main characteristics

- Privacy is defined as the ability to specify acceptable data usage through
  - policies defined by users (settings)
  - policies defined by the organization
- Organizations are trusted to enforce policies
Privacy as Confidentiality: Assumptions

- Lack of transparency and data protection enforcement
  - once data are under the control of an organization it is very difficult to verify how they are actually used
  - abuse of personal information may not be evident to individuals
- Organizations that collect and process user data are not necessarily competent and honest, security is expensive
  - Incentive to collect and use personal data for financial gain (without regard for user privacy)
  - Large number of reported privacy breaches (due to the lack of appropriate security practices)
- Placing trust in organizations makes individual vulnerable
Privacy as Confidentiality: Threats

- Queries themselves are sensitive
- Inferred profiles (aggregated data)
- Linkability of user information across different contexts
- Identifiability: queries are hard to anonymize
- Massive collection of user information is considered in itself a privacy threat
  - allow a variety of privacy violations: discrimination, manipulation, opportunistic abuse
  - information asymmetries reinforce power asymmetries: surveillance society
Privacy as Confidentiality: Goals

- Focus on minimal information disclosure
- Create an individual autonomous sphere free from intrusion
- Disclosure of information is BY DEFAULT prevented, or information is minimally disclosed in a way that cannot be linked back to the individual
  - Individuals may still disclose information voluntarily
Anonymous credentials

- Based on zero-knowledge
- Prover can prove
  - he holds a credential with certain attributes, or
  - any expression on them (simple arithmetic, boolean)
    e.g., age > 18, gender = female
- Verifier gains no more information
- Application: E-cash
Tor ("The Onion Router")

- Anonymous communication technique over a computer network
- Tor directs Internet traffic through an overlay network
- Difficult to trace users’ Internet activities
  - Prevent someone watching Internet connection from learning what websites are visited
  - Prevent websites from learning users’ physical location
Steganography and Covert Communications

- Encryption: hide data content
- Anonymity/unlinkability: hide identities / relations
- Unobservability: hide existence

- Communications:
  - Hide the fact that there is any communications
  - Embed a communication within another
  - Covert channels: hide secrets within public information

- Storage:
  - Hide the existence of files
  - Under coercion one can deny there is any file to decrypt
Privacy Enhancing Technologies
Privacy as Confidentiality

Privacy as Confidentiality: Main Characteristics

- Privacy is defined as properties hard-coded in the technology itself
  - Goal of technology is to ensure privacy properties hold
- Preventing data disclosure
- Minimize the need to trust others for handing sensitive data
Privacy as Practice: Assumptions

- Transparency provides users with an understanding of the system
  - Transparency produces awareness
  - Awareness evokes actions
Privacy as Practice: Threats

- User has no means to know:
  - which data are collected
  - for which purpose his data are used
  - how data are aggregated into profiles
  - which decisions are made based on these profiles
Privacy as Practice: Goals

- Focus on user awareness
- Data collection is made transparent
  - how, why, which data are collected
  - users can intervene in the collection and use of data
- Data processing is made transparent
  - discover undesirable privacy practices of organizations
P3P: Platform for Privacy Preferences

- Allow websites to communicate their privacy policies
- Provides a standard XML format to encode privacy policies
- Help users understand privacy policies

**Note**: No enforcement!
Transparency Tools

- Give users a better understanding of information flow, state and history
  - Become aware of how they (and others) participate in the socio-technical system
  - Verify whether their personal privacy level fits within the current working of the system
  - How they can take actions to change their own behavior as well as the socio-technical environment

- Tools
  - Google dashboard: see which info is associated to your account
  - Facebook: view how others see your profile
Privacy as Practice: Main characteristics

- Main objective is to support users in decision making
- Potential to uncover malicious behavior by organizations
## Summary

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>Confidentiality</th>
<th>Practice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Privacy definition: who fills the term privacy with meaning?</td>
<td>Organization-centric</td>
<td>Techno-centric</td>
<td>User-centric</td>
</tr>
<tr>
<td>Data Lifecycle</td>
<td>Information usage</td>
<td>Information disclosure/collection</td>
<td>Transparency in data disclosure, collection and use</td>
</tr>
<tr>
<td>Trust/adversary model</td>
<td>Organization trusted, individual untrusted</td>
<td>Powerful threat model, rigorous security analysis</td>
<td>No threat model</td>
</tr>
</tbody>
</table>
In this course...

- Focus on privacy as control
- Course Outline
  - Access Control
    - Discretionary Access Control
    - Mandatory Access Control
    - Role-based Access Control
    - Attribute-based Access Control
  - Usage Control
  - Privacy-aware Access Control
  - XACML