



# Privacy

EECS 195

Spring 2019

Zhou Li



# Objectives

- Define privacy and fundamental computer-related privacy challenges
- Privacy principles and laws
- Inference attacks and new tracking technologies
- Email privacy
- Privacy concerns in emerging technologies
- Communicating Anonymously



# What Is Privacy?

- Privacy is **the right to control who knows certain aspects about you, your communications, and your activities**
- Types of data many people consider private:
  - Identity
  - Finances
  - Health
  - Biometrics
  - Privileged communications
  - **Location data**



# More on privacy

- Privacy is subjective
  - What one person considers private is that person's decision
  - There is no universal standard
- Privacy depends on context
  - Probably people prefer to stay private after working hours
  - Culture is also a big factor
- Privacy and confidentiality
  - Confidentiality protects what one person considers private
- Privacy can have a cost
  - Might limit benefit one user can get, causing inconvenience



# Computer-Related Privacy Problems

- Data collection
  - Huge numbers of records can be collected by computer
- Notice and consent
  - Notice of collection and consent to allow collection of data are foundations of privacy
  - But it is often impossible to know what is being collected
  - Not all companies explicitly show notice and consent
- Control and ownership of data
  - Once a user gives consents, the data is out of that user's control
  - It may be held indefinitely or shared with other entities



# Privacy Principles and Policies

- Fair Information Practices
  - Advise to Secretary of US department of health, education and welfare on privacy issues (1973)
- US Privacy Laws
- Non-U.S. Privacy Principles
- GDPR
- CCPA



# Fair Information Practices

- Data should be obtained lawfully and fairly
- Data should be relevant to their purposes, accurate, complete, and up to date
- The purposes for which data will be used should be identified and that data destroyed if no longer necessary for that purpose
- Use for purposes other than those specified is authorized only with consent of data subject or by authority of law
- Procedures to guard against loss, corruption, destruction, or misuse of data should be established
- It should be possible to acquire information about the collection, storage, and use of personal data systems
- The data subjects normally have a right to access and challenge data relating to them
- A data controller should be designated and accountable for complying with the measures to effect these principles



# U.S. Privacy Laws

- **The 1974 Privacy Act** embodies most of the principles above but applies only to data collected by the **U.S. government**
- Other federal privacy laws (focusing on individual data types):
  - **HIPAA** (healthcare data)
  - GLBA (financial data)
  - COPPA (children's web access)
  - FERPA (student records)
- State privacy law varies widely



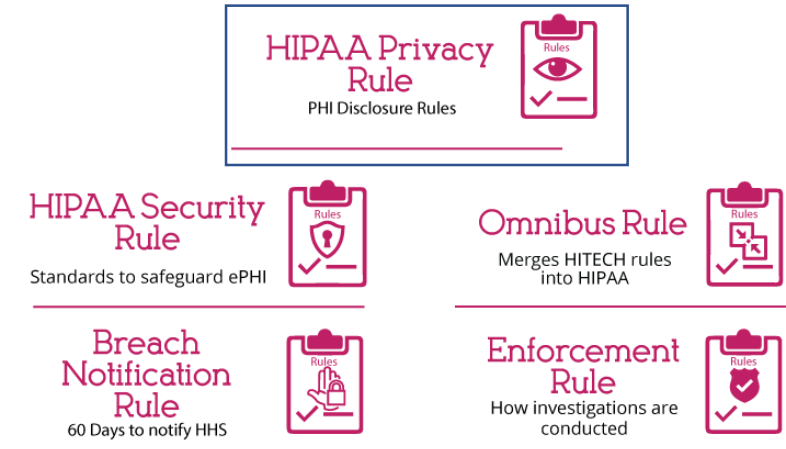


# HIPAA

- Health Insurance Portability & Accountable Act (1996)
- Privacy Rule is only one part of the Act
  - Addresses the use and disclosure of individuals' health information called "**Protected Health Information (PHI)**"
  - Permits important uses of information, while protecting the privacy of people who seek healthcare
- A study on HIPAA's impact shows
  - Data transfer was more explicit
  - Consumers still had little control over data disclosure and dissemination
  - Statements were longer, harder to understand



## Five HIPAA Rules





# Problems with Privacy Laws

- Overlap between target areas of different laws
  - Which law covers privacy protection of student's health center bills paid by credit card? Healthcare law? Credit report law? Educational privacy law?
  - Important to determine which law applies to single piece of data
- Gaps between laws are not always covered
  - New technologies, new devices and new types of data
  - Different state laws interpret the privacy differently



# California Consumer Privacy Act of 2018 (CCPA)

- Originally introduced back in February of 2017, AB 375 was signed into law by California Governor Jerry Brown on **June 28, 2018**.
- A “mini” version of GDPR, because while it has much of the consumer privacy protection of GDPR, it doesn’t have the data security aspects of GDPR.
- CCPA provides consumers with the following five enumerated rights:
  - The right of Californians to know what personal information is being **collected** about them.
  - The right of Californians to know whether their personal information is **sold or disclosed and to whom**.
  - The right of Californians to **say no to the sale** of personal information.
  - The right of Californians to **access their personal information** that a business is holding.
  - The right of Californians to **equal service and price**, even if they exercise their privacy rights.



# Non-U.S. Privacy Principles

- European Privacy Directive (1995)
  - Established because automated systems are more pervasive
  - Applies to governments and businesses
  - Provides extra protection for sensitive data, strong limits on data transfer, and independent oversight to ensure compliance
- A list of other nations' privacy laws
  - Japan, Australia, Canada, ...
  - can be found at <http://www.informationshield.com/intprivacylaws.html>
- Laws could conflict across nations
  - EU law forbids sharing data with companies or governments in countries whose privacy laws are not as strong
  - Though some “safe harbor” principles are agreed as temporary solutions



# General Data Protection Regulation (GDPR)


- \*Enforcement begins **May 25 2018**
  - More recent and more strict
- GDPR provides European individuals (data subjects) the right to:
  - Know what personal data is collected and how it is used
  - Have incorrect personal data updated
  - Have personal data erased / "to be forgotten"
  - Have personal data exported
- Companies have 30 days to comply with the request\*
- Penalties up to **4%** annual revenues or **20M** euro, whichever greater





# GDPR Cookie Consent Notification


A screenshot of the Analytics mania website. The header is dark blue with the 'Analytics mania' logo and social media icons. Below the header is a blurred image with the word 'Respect' visible. A red rectangular box highlights a GDPR cookie consent notification overlay at the bottom of the page. The notification text states: 'We use cookies to offer you a better browsing experience, analyse site traffic, personalise content, and serve targeted ads. Read how we use cookies and how you can control them in our "Cookie Settings". By using our site, you consent to our use of cookies.' At the bottom of the notification, there is a link to 'Cookie Settings' and a blue button labeled 'Accept Cookies'. A close button (X) is also present in the bottom right corner of the notification.

 Analytics mania

f t in

Respect

We use cookies to offer you a better browsing experience, analyse site traffic, personalise content, and serve targeted ads. Read how we use cookies and how you can control them in our "Cookie Settings". By using our site, you consent to our use of cookies.

[Cookie Settings](#) [Accept Cookies](#) 





# Individual Actions to Protect Privacy

- Do things **anonymously**
  - Web anonymity reduces fear of discrimination [MUL99]
  - People researching private matter, such as health issue or sexual orientation, are more likely to seek anonymous source
  - But it'll be revealed when you pay for something (except BitCoin)
- Keep **multiple identities**
  - E.g., bank account numbers, driver license numbers
  - Your identities are numbers **linked to your name**
  - But what if your name is changed?
- Use **pseudonyms**
  - A.k.a, unique identifiers that link records in server's database but **cannot be tracked back to your real identity**



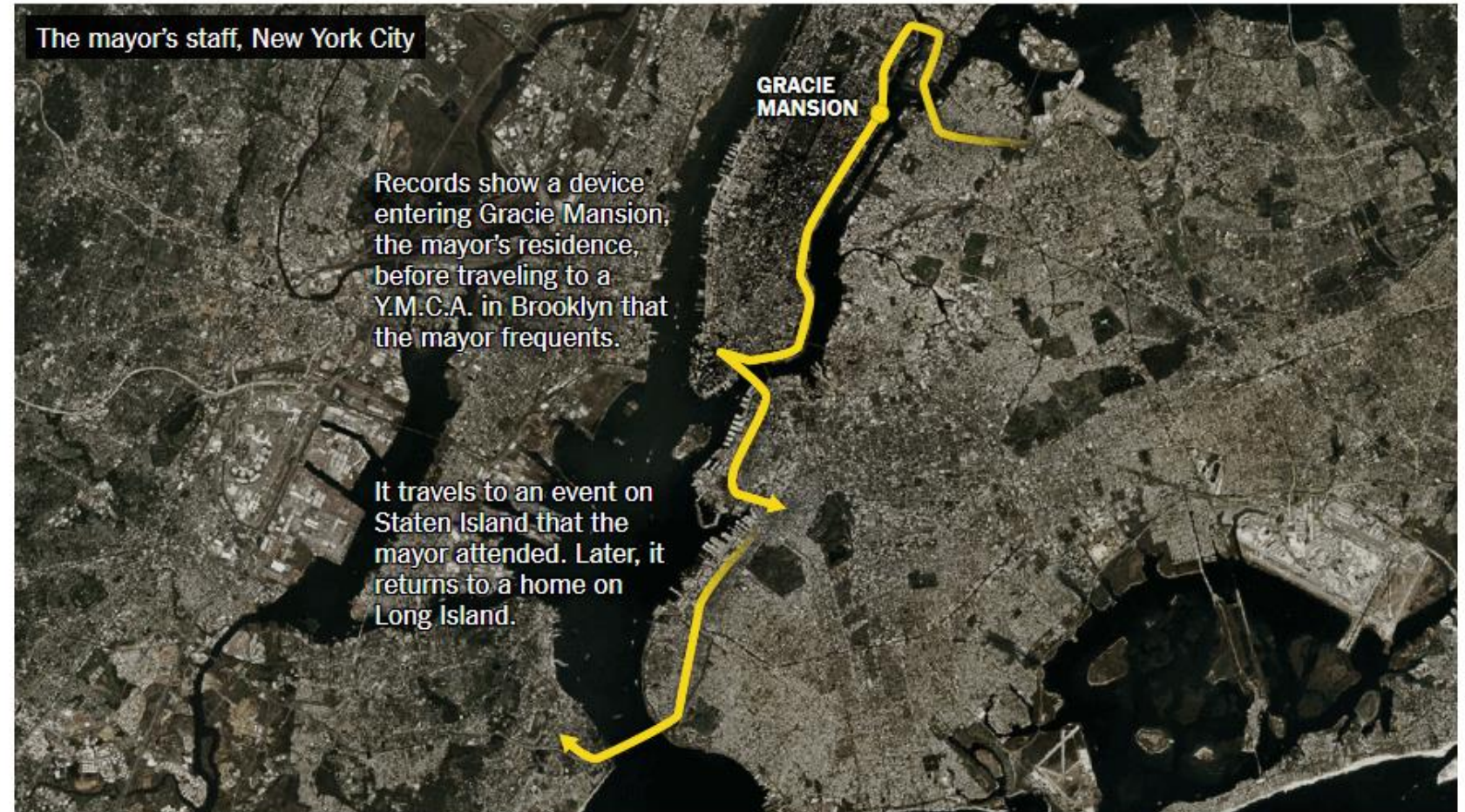
# Record Linkage

- A pseudonym is mapped to a user after authentication
- A user might have many pseudonyms, each associated with some activities
- Though pseudonyms are not supposed to be linked to real identity, **as data accumulation over time, linkage might be possible**
- For example, by collecting your locations multiple times, a mobile app can infer who you are



## Your Apps Know Where You Were Last Night, and They're Not Keeping It Secret

- The identity of a government employee is inferred because his/her multiple locations are related to mayor's events.



By Michael H. Keller | Satellite imagery by Mapbox and DigitalGlobe



# How to address this issue?

- Enforce **privacy-preserving data mining**
  - E.g., government can use it to alleviate people's worry in excessive data collection
- Naïve approach: removing identifying information from data
  - E.g., removing full name from collected data before analysis or release
- The approach doesn't work when statistical inference attack is performed