Privacy
CS 236
On-Line MS Program
Networks and Systems Security
Peter Reiher

Privacy

- Data privacy issues
- Network privacy issues
- Some privacy solutions

What Is Privacy?

- The ability to keep certain information secret
- Usually one's own information
- But also information that is "in your custody"
- Includes ongoing information about what you're doing

Privacy and Computers

- Much sensitive information currently kept on computers
  - Which are increasingly networked
- Often stored in large databases
  - Huge repositories of privacy time bombs
- We don't know where our information is

Privacy and Our Network Operations

- Lots of stuff goes on over the Internet
  - Banking and other commerce
  - Health care
  - Romance and sex
  - Family issues
  - Personal identity information
- We used to regard this stuff as private
  - Is it private any more?

Threat to Computer Privacy

- Cleartext transmission of data
- Poor security allows remote users to access our data
- Sites we visit save information on us
  - Multiple sites can combine information
- Governmental snooping
- Location privacy
- Insider threats in various places

Some Specific Privacy Problems

- Poorly secured databases that are remotely accessible
  - Or are stored on hackable computers
- Data mining by companies we interact with
- Eavesdropping on network communications by governments
- Insiders improperly accessing information
- Cell phone/mobile computer-based location tracking

Do Users Care About Privacy?

- Evidence suggests yes, but . . .
- Not necessarily in the way researchers think
  - E.g., data suggests teenagers aren't worried about privacy from hackers
  - They worry about privacy from their parents
- One must consider the actual privacy goals of users in protecting privacy

Data Privacy Issues

- My data is stored somewhere
  - Can I control who can use it/see it?
- Can I even know who's got it?
- How do I protect a set of private data?
  - While still allowing some use?
- Will data mining divulge data "through the back door"?

Privacy of Personal Data

- Who owns data about you?
- What if it's really personal data?
  - Social security number, DoB, your DNA record?
- What if it's data someone gathered about you?
  - Your Google history or shopping records
  - Does it matter how they got it?

Protecting Data Sets

- If my company has (legitimately) a bunch of personal data,
- What can I/should I do to protect it?
  - Given that I probably also need to use it?
- If I fail, how do I know that?
  - And what remedies do I have?

Options for Protecting Data

- Careful system design
- Limited access to the database
  - Networked or otherwise
- Full logging and careful auditing
- Store only encrypted data
  - But what about when it must be used?
  - Key issues
  - Steganography

Data Mining and Privacy

- Data mining allows users to extract models from databases
  - Based on aggregated information
- Often data mining allowed when direct extraction isn't
- Unless handled carefully, attackers can use mining to deduce record values

An Example of the Problem

- Netflix released a large database of user rankings of films
  - Anonymized, but each user had one random identity
- Clever researchers correlated the database with IMDB rankings
  - Which weren't anonymized
  - Allowed them to match IMDB names to Netflix random identities

Insider Threats and Privacy

- Often insiders need access to private data
  - Under some circumstances
- But they might abuse that access
- How can we determine when they misbehave?
- What can we do?

Local Examples

- Over 120 UCLA medical center employees improperly viewed celebrities' medical records
  - Between 2004-2006
- Two accidental postings of private UCLA medical data in 2011
- UCLA is far from the only offender

Encryption and Privacy

- Properly encrypted data can only be read by those who have the key
  - In most cases
  - And assuming proper cryptography is hazardous
- So why isn't keeping data encrypted the privacy solution?

Problems With Data Encryption for Privacy

- Who's got the key?
- How well have they protected the key?
- If I'm not storing my data, how sure am I that encryption was applied?
- How can the data be used when encrypted?
  - If I decrypt for use, what then?

A Recent Case

- Yahoo lost 450,000 user IDs and passwords in July 2012
  - The passwords weren't encrypted
  - Much less salted
- Password file clearly wasn't well protected, either
- Who else is storing your personal data unencrypted?