SECURITY IN COMPUTING, FIFTH EDITION

Chapter 9: Privacy

https://www.collective-evolution.com/2013/06/13/7-powerful-ways-to-maintain-your-privacy-and-integrity-online/

Chapter 9 Objectives

- Define privacy and fundamental computer-related privacy challenges
- Privacy principles and laws
- Privacy precautions for web surfing
- Spyware
- Email privacy
- Privacy concerns in emerging technologies

What Is Privacy?

- Privacy is the right to control who knows certain aspects about you
  - Such as your communications, activities
- Privacy is an aspect of computer security
  - Part of confidentiality
- Privacy may conflict with other aspects of security
  - Such as availability
  - Example: refusing to reveal personal data to a shop
    - may prevent you from receiving a frequent-shopper discount

What Is Privacy?

- Privacy is considered a human right
    - Cultural and historical roots may define to what extent privacy is deserved
- Privacy issues existed before computers
    - However, computers changed access to data
        - High-speed processing, data storage and transmission capabilities
        - Computers enable data collection and correlation
            - affect privacy

https://www.collective-evolution.com/2013/06/13/7-powerful-ways-to-maintain-your-privacy-and-integrity-online/

What Is Privacy?

- Privacy is a broad topic
- We will discuss privacy issues inextricably linked to computer security
  - Examine the meaning of information privacy
  - Revisit identification and authentication
    - two aspects of computing that have significant privacy implications
  - Discuss how privacy relates to the Internet
    - specifically in email and web access
  - Investigate emerging computer-based technologies
    - for which privacy is important

PRIVACY CONCEPTS

Aspects of Information Privacy

- Information privacy has three aspects:
  - Sensitive data
  - Affected parties
  - Controlled disclosure
- Similar to the three elements of access control:
  - Subject
  - Object
  - Access rights

Controlled Disclosure

- Privacy is the right to control who knows certain aspects about you
  - Such as your communications, activities
- You voluntarily choose who can know which things about you
- Example: people may ask you your address
  - You *decide* if to whether to give it or not

Controlled Disclosure

- Privacy is something over which you can have considerable influence.
  - However, not a complete control
- Once you give your info to a person/system, your control is diminished
  - Depends on what the system does with that information
  - you are ceding it to someone or something else
- You have to trust the person or system to comply with your privacy wishes
  - whether you state those wishes explicitly or not

Sensitive Data

- Some information is usually considered sensitive, such as financial status, certain health data, etc.
- Some data items sensitivity varies:
  - Some people find it more sensitive than others
- In some cases public interest outweights person’s right to security
  - For example, healthcare professionals are frequently required to report instances of highly communicable/deadly diseases
    - even if the stricken person does not want it to be made public
-

https://backup-guard.com/blog/7-surefire-ways-to-protect-sensitive-data/

Sensitive Data

- Types of data many people consider private:
  - Identity
    - name, identifying information
  - Finances
    - credit rating and status, bank details, tax info, etc.
  - Health
    - medical conditions, drug use
    - DNA, genetic predisposition to illnesses
  - Biometrics
    - physical characteristics, fingerprints

https://backup-guard.com/blog/7-surefire-ways-to-protect-sensitive-data/

Sensitive Data

- Types of data many people consider private (cont.):
  - Privileged communications
    - with accountants, doctors, counselors, clergy, etc.
  - Location data
    - general travel plans, current location, travel patterns
  - Digital Footprint
    - email, telephone calls, spam, instant messages, tweets
    - social networking history
  - Opinions, preferences, and membership
    - voting records, etc.

https://backup-guard.com/blog/7-surefire-ways-to-protect-sensitive-data/

Sensitive Data

- Privacy depends on context
  - Example: A famous athlete results may be public
    - whereas you might not want everyone to know how poorly you finished in your last athletic event
- Culture also influences what people consider sensitive
  - for example, discussing salary information may be appropriate/permissible in one culture but not in another

https://backup-guard.com/blog/7-surefire-ways-to-protect-sensitive-data/

Affected Subjects

- We distinguish between subject and owner:
  - Subject: person or entity being described by the data
  - Owner: person or entity that holds the data
- Subject may be a person or an organization
- Companies may have sensitive data:
  - Product plans, customer list, product profitability, etc.
- Hospitals and schools need to protect data:
  - Patient or student information
- Other info may be protected:
  - Negative news, legal decisions, diplomatic matters, etc.

What is Privacy?

- Privacy is controlled disclosure of info
- After disclosing something, subject cedes control to the receiver
- Privacy has a cost
  - May conflict with availability of data
    - Also aspect of security

Computer-Related Privacy Problems

- Privacy issues existed before computers
  - However, computers changed access to data
  - Computers and networks have affected the feasibility, speed, and reach of some data
    - Including unwanted disclosures
  - Search engines enable finding one item out of billions
    - the equivalent of finding one sheet of paper out of a warehouse full of boxes of papers
  - Networks openness and technology portability greatly increase risk of disclosures affecting privacy
    - such as laptops, tablets, cell phones, etc.

Dimensions of privacy

- Rezgui et al [2003] defines eight dimensions of privacy:
  - *Information collection*: Data are collected only with knowledge and explicit consent.
  - *Information usage*: Data are used only for certain specified purposes.
  - *Information retention*: Data are retained for only a set period of time.
  - *Information disclosure*: Data are disclosed to only an authorized set of people.

Dimensions of privacy

- Rezgui et al [2003] defines eight dimensions of privacy:
  - *Information security*: Appropriate mechanisms are used to ensure the protection of the data.
  - *Access control*: All modes of access to all forms of collected data are controlled.
  - *Monitoring*: Logs are maintained showing all accesses to data.
  - *Policy changes*: Less restrictive policies are never applied after-the-fact to already obtained data.

Computer-Related Privacy Problems

- Computer usage raises certain privacy issues, including:
  - Data Collection, notice and consent, control and ownership of data

Computer-Related Privacy Problems

- Data collection
  - Advances in computer storage make it possible to hold and manipulate huge numbers of records
    - those advances continue to evolve
  - Example:
    - Google's stored data estimated in multiple petabyte (10^{15}) range
      - accounting for 0.01 percent of the world's total energy usage
    - Facebook servers process ~2.4 billion pieces of content daily
    - Amazon's servers have more than 17 million monthly visitors
  - Data seems not to be thrown away
    - Just moved to slower secondary media or more storage bought

https://www.ibm.com/analytics/hadoop/big-data-analytics

Computer-Related Privacy Problems

- Notice and consent
  - Notice of collection and consent to allow collection of data are foundations of privacy
  - With modern data collection, it is often impossible to know what is being collected
- Entry into a website may require acknowledgment of “terms of use,”
  - describe what is collected and why
    - recourse if you prefer not to have something collected
  - Studies show people do not read them before accepting

https://www.theguardian.com/technology/2017/mar/03/terms-of-service-online-contracts-fine-print

Computer-Related Privacy Problems

- Control and ownership of data
    - Once a user consents to provide data, the data is out of that user’s control.
    - It may be held indefinitely or shared with other entities.
        - For example, merchants may sell your data
    - You have little control over dissemination of your data
        - Disseminated data are almost impossible to get back
    - Example: electronic posting
        - someone may copy it before you delete it
            - It may never be deleted

Computer-Related Privacy Problems

- Control and ownership of data (cont.)
  - Someone may post embarrassing information about you
    - You may want it removed
  - European Union try to enforce “the right to be forgotten”
    - To allow removal of old embarrassing information
    - To exercise the right to be forgotten and request removal from a search engine, one must complete a form through the search engine's website

Computer-Related Privacy Problems

- Control and ownership of data (cont.)
  - Legal ownership is different in different countries
  - In the European Union, subjects own their data
    - Must give permission before it can be used in a variety of ways
  - In the United States, the data’s holder is the owner
    - => letting copies escape to someone else is a problem

Fair Information Practices

- In 1973 a report was created to advise the government about privacy issues
  - Led by Willis Ware
  - The report proposes a set of practices

Fair Information Practices

- Data should be obtained lawfully and fairly
- Data should be relevant to their purposes, accurate, complete, and up to date
- Purposes for data usage should be identified
  - Destroy data if no longer needed for that purpose
- Use for purposes other than those specified is authorized only with consent of data subject
  - or by authority of law

Fair Information Practices (cont.)

- Procedures to guard data should be established
  - against loss, corruption, destruction, or misuse of
- Should be possible to acquire information about collection, storage, 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
  - applies only to data collected by the U.S. government
- Other federal privacy laws:
  - HIPAA (healthcare data) HIPAA online
  - GLBA (financial data)
  - COPPA (children's web access)
  - FERPA (student records)
- State privacy law varies widely
- A list can be found at: United States Privacy Laws

Non-U.S. Privacy Principles

- European Privacy Directive (1995)
  - Applies the Ware Committee's principles to governments and businesses
  - Also provides for extra protection for sensitive data, strong limits on data transfer, and independent oversight to ensure compliance
- A list of other nations' privacy laws can be found at http://www.informationshield.com/intprivacylaws.html

Authentication and Privacy

- Identification: asserting who you are
  - Identities are often well known, predictable/guessable
    - i.e., bank number written on checks, etc.
  - => An impersonator may easily claim to be you
    - by presenting one of your known identifiers
- Authentication: confirms you are who you purport to be
  - Should be reliable

Authentication and Privacy

- Privacy issues occur when confusing authentication and identification
- For example:
  - U.S. social security number was never intended to be an identifier
  - Often serves as an identifier, an authenticator, a database key, or all three
  - => if someone knows your social security number, it may impersonate you

Authentication and Privacy

- Another example:
  - Fraudulent emails may be sent from an email ID
  - But email may be spoofed
    - So wrong person may be suspected

Authentication and Privacy

- **Anonymized Records:**
  - To preserve privacy, researchers often deal with anonymized records
    - records from which identifying information has been removed
  - If those records can be reconnected to the identifying information, privacy suffers
  - Linking a few databases may still reveal private information

Authentication and Privacy

- **Anonymized Records:**
  - Example: A study by Sweeny [2001] showed 87% of US population can be identified by combining 5-digit zip code, gender, and date of birth
  - => it is difficult to anonymize data effectively
  - Many medical records are coded with at least gender and date of birth
    - those records are often thought to be releasable for anonymous research purposes
    - May be a threat to privacy

Privacy-Preserving Data Mining

- Removing identifying information from data doesn’t work
  - Even if the overtly identifying information can be removed, identification from remaining data is often possible

Privacy-Preserving Data Mining

- Data perturbation
  - Data perturbation can limit the privacy risks associated with the data without impacting analysis results
  - **Perturb** the values of the database by a small error:
    - Add a small error to the true values
    - some reported values will be slightly higher than their true values and other reported values will be lower
    - Statistical measures such as sum and mean will be close but not necessarily exact
  - Data mining often focuses on correlation and aggregation
    - both can generally be reliably accomplished with perturbed data

Privacy on the Web

- Users may seem anonymous online:
  - A user can visit websites, send messages, and interact with applications without revealing an identity
- However cookies, adware, spybots, and malicious code may be used to track users
  - Resulting in a largely one-sided anonymity
  - Sophisticated web apps can know a lot about a user
    - but the user knows relatively little about the application

Precautions for Web Surfing

- Governments, companies and people may want to track your activities
  - gather information about you
- Tracking technology includes cookies and web bugs
- These technologies are frequently used to monitor activities without the user’s knowledge

Precautions for Web Surfing

- Cookies
  - Cookies are a way for websites to store data locally on a user’s machine
  - They may contain sensitive personal information, such as credit card numbers
  - A cookie is a text file stored on the user’s computer
    - passed by the user’s browser to the website
      - when the user goes to that site.

Precautions for Web Surfing

- Cookies
  - Each cookie file consists of a pair of data items sent to web browser by the visited website: a key and a value
  - Key is the URL of the site establishing the cookie
  - Value includes:
    - data and expiration date
    - path and domain of the server it is delivered to
      - Supposed to protect against a site accessing another one’s cookies

Precautions for Web Surfing

- Cookies (cont.)
  - Cookies may contain sensitive information
    - Credit card numbers, name and address of user, etc.
  - Sensitive information should be encrypted
    - or otherwise protected in the cookie
  - It is up to the site to define or determine what kind of protection it applies to its cookies
  - The user never knows if or how data are protected

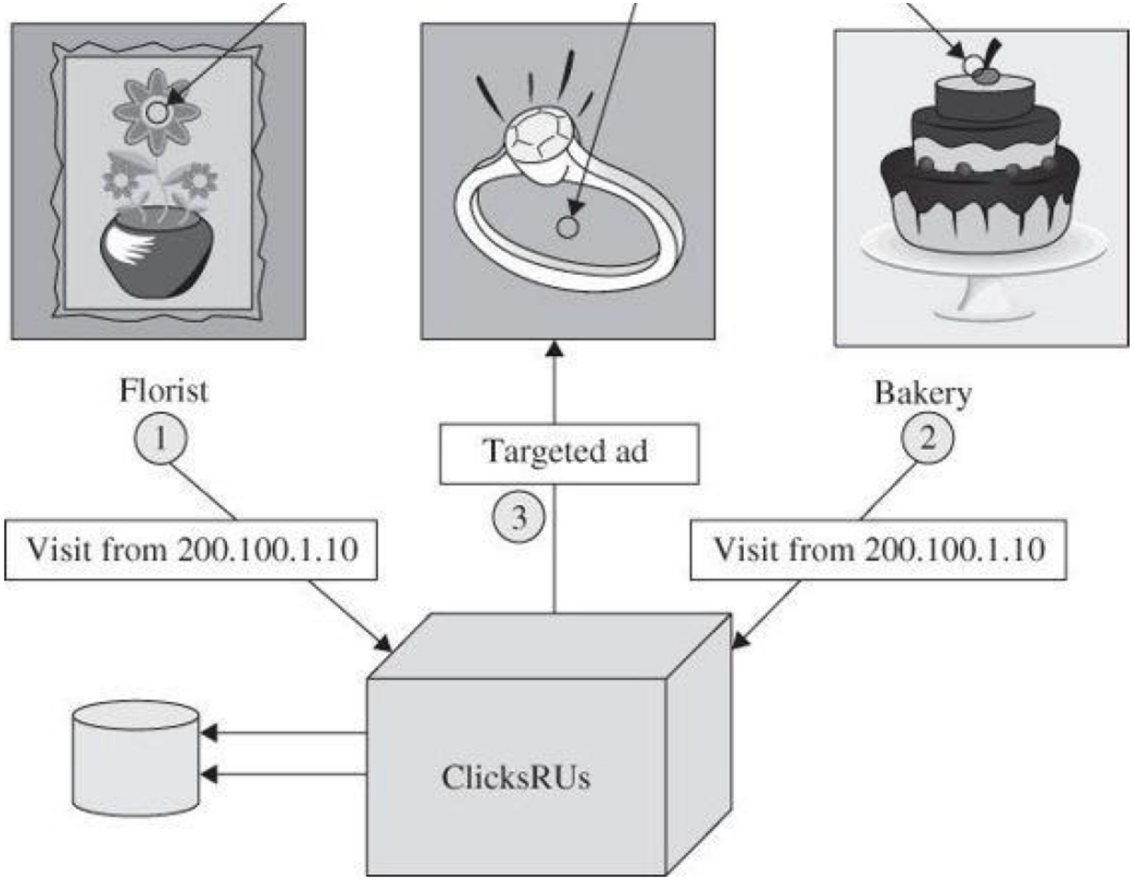
Precautions for Web Surfing

- Third-party tracking cookies
  - A web page can contain cookies for organizations
    - Called third-party cookies
  - Some companies specialize in tracking users by having numerous popular sites place their cookies in users' browsers
    - The third-party tracking firm receives reports from individual sites and correlates the data to provide predictive intelligence
  - This tracking information is used for online profiling, which is generally used for targeted advertising

Precautions for Web Surfing

- Web bugs
  - Cookies store and return data but cause no action
  - A web bug is more active than a cookie
  - Has the ability to immediately send information about user behavior to advertising services
  - Can also invoke more bugs and hence more code

Web bugs Example



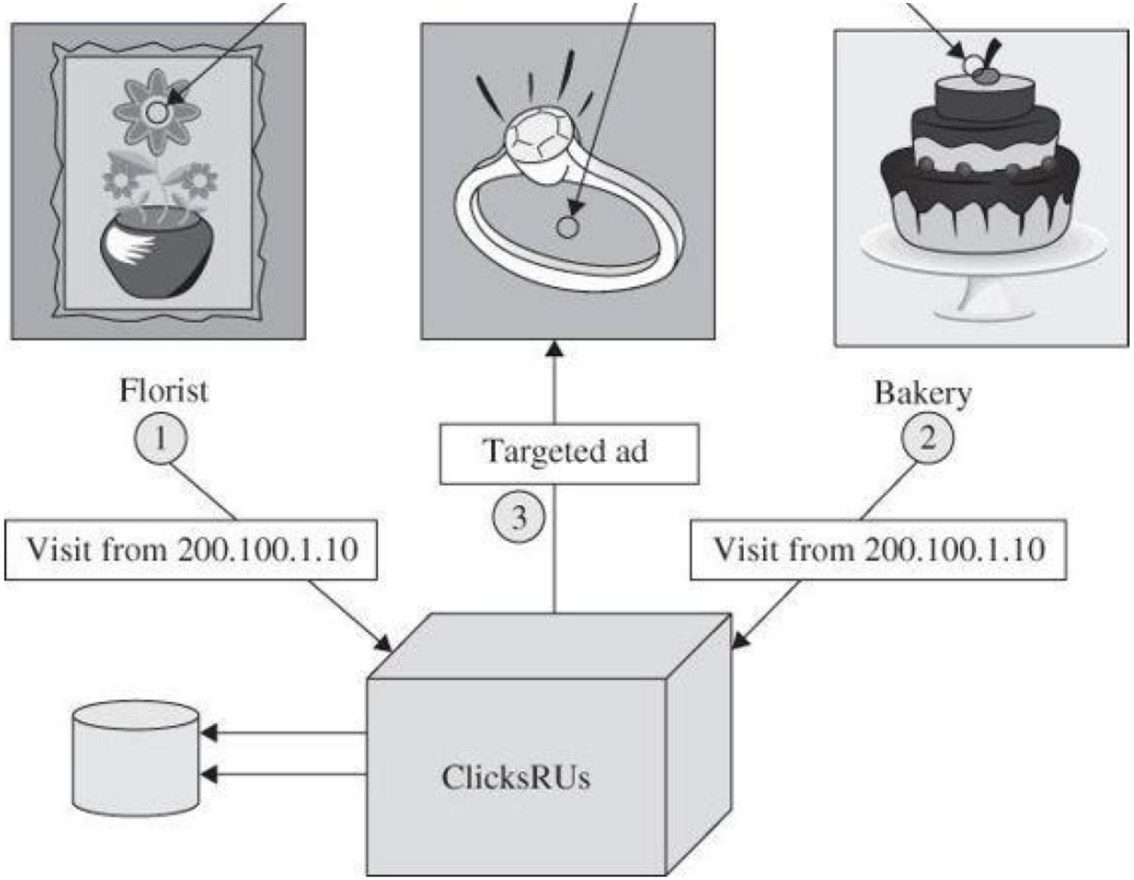
Web bugs Example

- Florist might subscribe to a web tracking service, ‘ClicksRUs’
- The florist includes a web bug in its web image
- When user loads that page, his details are sent to ClicksRUs
  - which then installs a cookie
- User next goes to a bakery’s site that also subscribes to tracking with ClicksRUs
  - the new page will also have a ClicksRUs web bug.

Web bugs Example (cont.)

- This time, ClicksRUs retrieves its old cookie
  - finds that you were last at the florist’s site
    - records the coincidence of these two firms.
- ClicksRUs correlates these data points
- ClicksRUs can inform the florist and the bakery that they have common customers
  - might develop a joint marketing approach.

Web bugs Example



Spyware

- Spyware is code designed to spy on a user, collecting data
- General spyware:
  - Advertising applications, identity theft
- Hijackers:
  - Hijack existing programs and use them for different purposes
    - such as reconfiguring file sharing software to share sensitive information

Spyware

- Adware
  - Displays selected advertisements in pop-up windows or the main browser window
  - Often installed in a misleading way as part of other software packages

Internet Privacy

- **When you make your Facebook profile private, your old posts and photos become private as well**
  - True
  - False

Internet Privacy

- **When you make your Facebook profile private, your old posts and photos become private as well**
  - True
  - [✓ False]
    - If you created public posts on Facebook prior to setting your security to private, then those old posts will still be visible

https://www.pbs.org/newshour/nation/quiz-how-much-do-you-know-about-your-internet-privacy

Internet Privacy

- **When you delete a tweet, it is automatically wiped from the Internet**
    - True
    - False

https://www.pbs.org/newshour/nation/quiz-how-much-do-you-know-about-your-internet-privacy

Internet Privacy

- **When you delete a tweet, it is automatically wiped from the Internet**
  - True
  - [✓ False]
    - Even if you delete tweets, Google and other search engines cache search results from twitter
      - => Deleted tweets are still searchable for a while
    - Retweets of the deleted tweet are deleted
      - But retweets with comments are *not* deleted

https://www.pbs.org/newshour/nation/quiz-how-much-do-you-know-about-your-internet-privacy

Internet Privacy

- **Apps you download can track your online movements and then sell that data to other companies**
  - True
  - False

https://www.pbs.org/newshour/nation/quiz-how-much-do-you-know-about-your-internet-privacy

Internet Privacy

- **Apps you download can track your online movements and then sell that data to other companies**
  - [✓ True]
    - Data brokers buy information from different apps, and aggregate the data with data from stores’ loyalty programs
      - Merging online and offline data is valuable for targeted advertising
  - False

https://www.pbs.org/newshour/nation/quiz-how-much-do-you-know-about-your-internet-privacy

EMAIL SECURITY

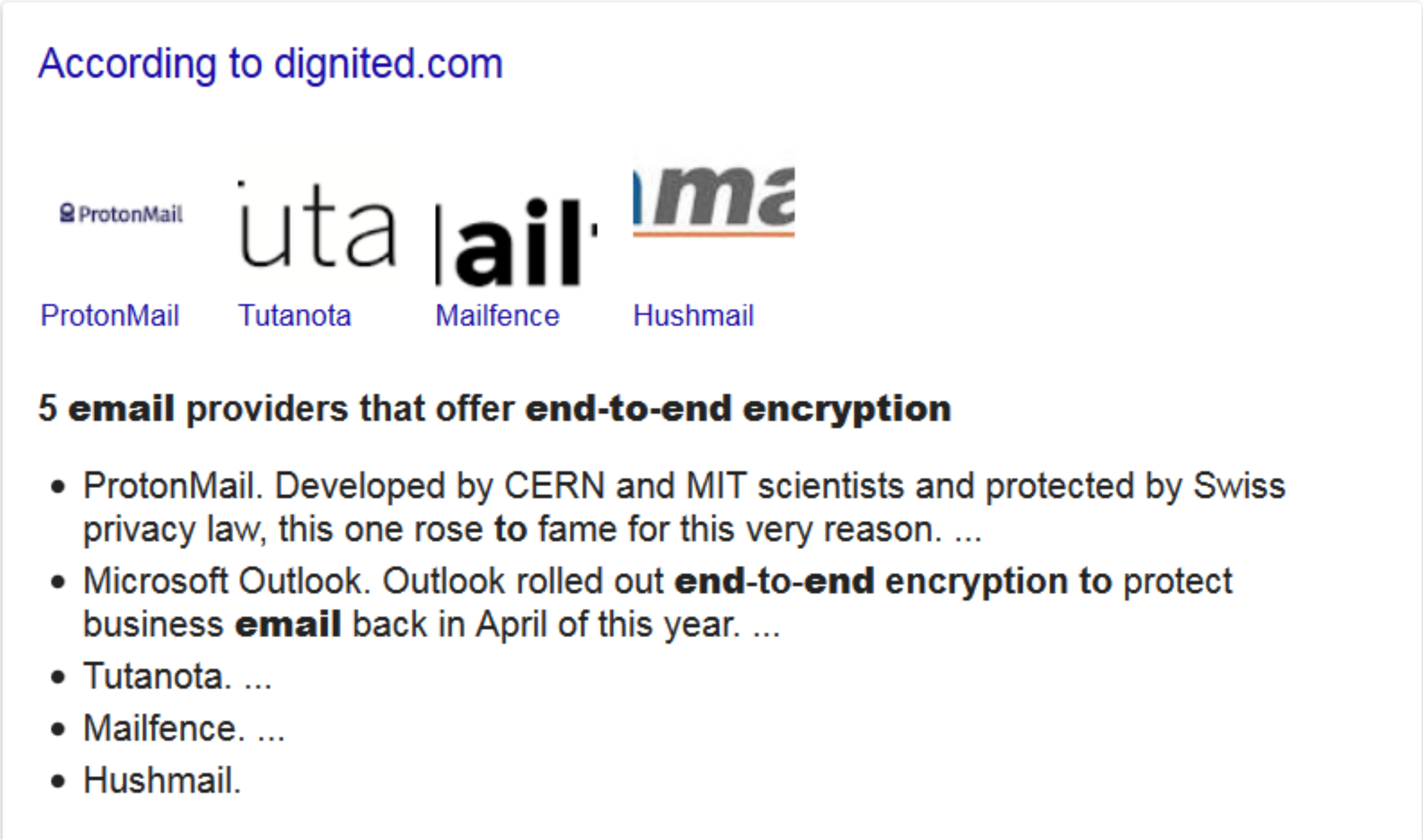
Where Does Email Go?

- Janet sends an email to Scott
  - the message is transferred via simple mail transfer protocol (SMTP)
  - The message is the transferred through multiple ISPs and servers
    - before it arrives at Scott’s post office protocol (POP) server
  - Scott receives the email when his email client logs into the POP server on his behalf
  - Any of the servers in this chain of communication can see and keep Janet’s email

Interception of Email

- Email is subject to same interception risks as other web traffic
- Email may be encrypted
  - Populate protocols include S/MIME and PGP encryption
    - Protection is considered end-to-end
      - From client's workstation to recipient workstation

Interception of Email



According to dignited.com

ProtonMail Tutanota Mailfence Hushmail

5 email providers that offer **end-to-end encryption**

- ProtonMail. Developed by CERN and MIT scientists and protected by Swiss privacy law, this one rose **to** fame for this very reason. ...
- Microsoft Outlook. Outlook rolled out **end-to-end encryption to** protect business **email** back in April of this year. ...
- Tutanota. ...
- Mailfence. ...
- Hushmail.

https://www.dignited.com/33439/5-email-providers-that-offer-end-to-end-encryption/

ProtonMail

- End-to-end encrypted email service
- Uses client-side encryption to protect email contents and user data
  - before they are sent to ProtonMail servers
- Service can be accessed through the Tor network
  - Also, through traditional webmail client, iOS or Android apps

Email Privacy

- Ted Talk: Think your email's private? Think again

Google Email

- Google supports TLS (Transport Layer Security) protocol
  - But can only implement this if both sides support it
- According to Google, 40 to 50 percent of emails aren’t encrypted
  - sent between Gmail and other email providers
- Google also offers a chrome extension End-To-End, which provides end-to-end encryption
  - Can encrypt, decrypt, digitally sign, and verify signed messages within the browser
    - using OpenPGP

Monitoring Email

- Some organizations routinely copy all emails sent from their computers
  - Monitoring for inappropriate content
  - Legal affairs

Monitoring Email

- Most employers make employees sign an agreement
  - grants them right to monitor their email and computer usage
    - Signing this agreement normally deprives an employee of any reasonable expectation of privacy
      - No reasonable expectation of privacy in work email

Anonymous or Disappearing Email

- Disposable email addresses from sites like mailinator.com
- Remailers are trusted third parties that replace real addresses with pseudonymous ones
    - to protect identities in correspondence

Anonymous or Disappearing Email

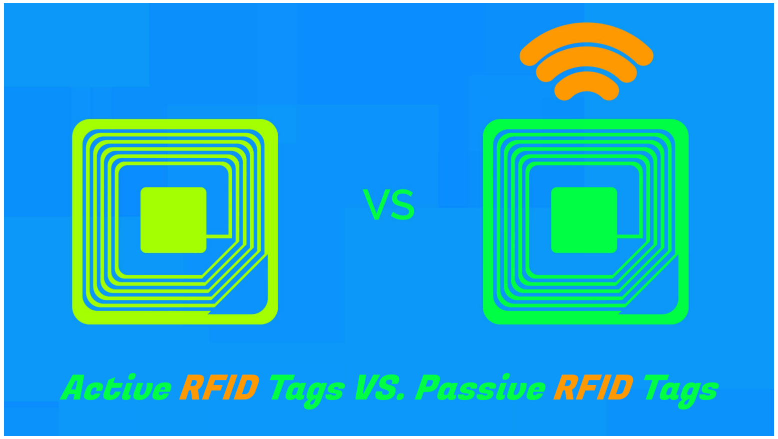
- Multiple remailers can be used in a TOR-like configuration to gain stronger anonymity
- The TOR-like configuration:
  - The sender selects three remailers;
  - he encrypts the message with each of their public keys in succession;
  - he then sends the message through them in the reverse of that order
    - with each one's public key being able to open only one layer of message

Anonymous or Disappearing Email

- Disappearing email
  - Because email travels through so many servers, it cannot be made to truly disappear
  - Messaging services like Snapchat, claim to make messages disappear
    - cannot guarantee that recipients will not be able to save those messages

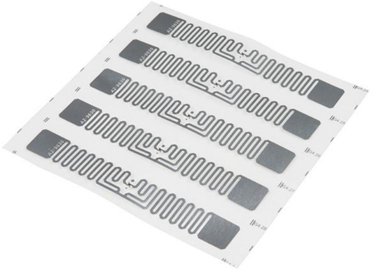
Radio Frequency Identification (RFID)

- RFID tags are small, low-power wireless radio transmitters
- When a tag receives a signal on the correct frequency, it responds with its unique ID number



https://cybra.com/active-rfid-tags-vs-passive-rfid-tags/

Radio Frequency Identification (RFID)



https://www.demco.com/products/Security/Security-Strips-Tags/RFID/ISO-RFID-Tags/_/A-B00186931&PRODUCT=20584890&intcmp=GOOGLE_PRD20584890?gclid=CjwKCAjw39reBRBJEiwAO1m0OTFesaUZQBREMa5E5B129cdi4pjQxoOiTvMGAPxSPDrz3fzenq3n5BoCABIQAvD_BwE&gclsrc=aw.ds
https://www.kr4.us/UHF-RFID-Tag-Set-of-5.html?gclid=CjwKCAjw39reBRBJEiwAO1m0OVX-rpnrH_jbafXA9-e8qsBw2utlw0UyCV7zk8HcatohIQhZyiNk0hoC1zoQAvD_BwE

Radio Frequency Identification (RFID)

- Current uses include:
  - Transit system fare cards
  - Patient records and medical device tracking
  - Sporting event timing
  - Stock or inventory labels
  - Counterfeit detection
  - Passport and identity cards
  - Surgically implanted identity tokens for live stock

Radio Frequency Identification (RFID)

- Privacy concerns:
  - RFID tags become cheaper and more ubiquitous
  - RFID readers are installed in more places
    - => It may become possible to track individuals wherever they go
  - As RFID tags are put on more items, it will become increasingly possible to discern personal information by reading those tags
  - Many RFID are designed to be inexpensive
    - have limited computation power, cannot implement traditional cryptographic protocols

Radio Frequency Identification (RFID)

- RFID tags respond to reader interrogation without alerting their bearers[1]
  - Thus, where read range permits, clandestine scanning of tags is a plausible threat.
- Most RFID tags emit unique identifiers
  - a person carrying an RFID tag broadcasts a fixed serial number to nearby readers,
    - providing a ready vehicle for clandestine physical tracking.
    - Such tracking is possible even if a fixed tag serial number is random and carries no intrinsic data.

[1] RFID Security and Privacy: A Research Survey, A. Juels

Radio Frequency Identification (RFID)

- In addition, attackers can learn personal information about users opinions
  - Study showed information about borrowed library books can be read from RFID tags
    - Impinging on privacy
    - Books choice may be personal

Radio Frequency Identification (RFID)

- The threat to privacy grows when a tag serial number is combined with personal information
  - E.g., when a consumer makes a purchase with a credit card
    - A shop can establish a link between her identity and the serial numbers of the tags on her person.
    - Marketers can then identify and profile the consumer using networks of RFID readers
      - Both inside shops and out

Radio Frequency Identification (RFID)

- Problem of clandestine tracking is not unique to RFID
- It affects many other wireless devices, such as Bluetooth-enabled ones

Radio Frequency Identification (RFID)

- Potential privacy-preserving approaches to RFID’s [Juels 05]:
  - Blasting:
    - disabling a tag
  - Blocking:
    - shielding a tag to block its access by a reader
  - Reprogramming:
    - Such that the tag will emit a different number after a while

Radio Frequency Identification (RFID)

- Potential privacy-preserving approaches to RFID’s [Juels 05] (cont.):
  - Encrypting
    - So the output is selectively available
    - Since RFID processing power is limited, efficient protocols have been suggested
      - Replacing traditional cryptography

Other Emerging Technologies

- Electronic voting
  - Voting should be private but also accurate
    - Collected votes should be authentic
  - Among other issues, research into electronic voting includes privacy concerns
    - such as maintaining privacy of who has voted and who each person voted for
    - In addition, the study emphasized that the public must have confidence in the process
      - otherwise, the public will not trust the outcome

Other Emerging Technologies

- Voice over IP (VoIP) and Skype
  - VoIP adds the possibility of encryption to voice calls
  - it also allows a new set of service providers to track sources and destinations of those calls
    - Even if the voice traffic is solidly encrypted, the source and destination of the phone call will be somewhat exposed
      - through packet headers

Other Emerging Technologies

- Cloud computing
  - Physical location of information in the cloud may have significant effects on privacy/confidentiality protections
  - Cloud data may have more than one legal location at a time
  - Laws could oblige cloud providers to examine user data for evidence of criminal activity
  - Legal uncertainties make it difficult to assess the status of cloud data

Emerging Technologies

- Technologies continue to emerge and mature
- Privacy risks are introduced with them
- Privacy implications should be evaluated for each
  - Follow changes as technology evolves
- Many emerging technologies are developed under financial pressure
  - Rush to market, dealing with privacy issues later
    - Not the recommended way
  - Development approach should consider privacy issues starting from initial design stage

Summary

- What data is considered private is subjective
- Privacy laws vary widely by jurisdiction
- Cookies and web bugs track user behavior across websites
- Spyware can be used to track behavior for targeted advertising
  - or for much more nefarious purposes

Summary

- Email has little privacy protection by default
- Emerging technologies are fraught with privacy uncertainties
  - including both technological and legal issues

Questions?