Computer Science Principles

ONLINE PRIVACY

Announcements

Read Chapter 2 of "Blown to Bits: Your Life, Liberty, and Happiness after the Digital Explosion"

Acknowledgement: These slides are extended versions of slides prepared by Prof. Arthur Lee and Tony Mione for earlier CSE 101 classes. Some slides are based on Prof. Kevin McDonald at SBU CSE 101 lecture notes and the "Blown to Bits" book by Hal Abelson, Ken Ledeen, and Harry Lewis.

Leaving digital footprints

Do you know that:

- Every time you visit a website, you leave a trace? Your IP address (your computer's Internet address) is transmitted to the webserver you are visiting.
- Your cell phone company knows exactly where you have been and when?
- Your smart phone records your GPS location in every photo you take?
- Your credit card company knows not only how much you spent, but what you buy?
- Google logs every one of your searches in its databases?
- Online advertisers can track your browsing habits?

Standing on lettuce

- In 2012 there was a Burger King employee who anonymously posted the photo on the right to 4Chan.org
- Within 15 minutes of the post, 4Chan users extracted the GPS information from the photo and someone contacted the press and another person left a message on Burger King's website
- The next day the shift manager at the area Burger King store was shown the photo. He knew immediately who it was, and the employee was fired that day.

- The concept of "online anonymity" is largely an illusion

Mobile geo-tagging

Basically, your smart phone is keeping track of you wherever you go

Any time you tweet on Twitter or make a Facebook post or snap a photo you may be telling the world where you are

Here's a thought: every time a parent snaps a photo of his child – at school, home, the playground – they are making a digital record of every place they regularly takes their children, and on what days and times

If those photos are posted on the Web without stripping the EXIF data (which has the GPS coordinates and other info), this could be an invitation for kidnappers or other criminals to target that family's kids

Mobile geo-tagging

Are there other situations where having GPS coordinates on the web would not be a good idea?

- A jogger who runs the same route on a regular schedule and posts her jogging updates on Facebook.
- A family posting their whereabouts on the Web when they are on vacation. This could be an invitation for a burglary.
- Some mobile apps track your GPS position if you give them permission. Would you want unknown data brokers buying and selling your movement? What could they do with such information?

Digital footprints

Geo-tags give one example of **digital footprints** or **information footprints**

What are some ways you personally generate a digital footprint?

- Intentional posting/online activity: social networks, restaurant reviews, product reviews
- Information generated, recorded, and/or transmitted by devices: GPS, voice commands
- Metadata associated with intentional online activities: IP addresses and other info collected by websites
- Data collected (and potentially shared) by other entities

Have you ever searched for your name in a search engine, like Google? What did you find?

Big Brother is watching

In his novel 1984, the author George Orwell described a futuristic world where the government (Big Brother) is constantly spying on its people and trying to control them

Not only do governments collect a lot of data about people, but so do companies

Considering this, what are some examples of digital and nondigital activities you engage in on a daily basis?

- What kind of data does that activity generate?
- What could other people do with that information?
- What could they figure out about you?
- Is that something you want them to know about you?

Big Brother is watching

In the 2013 Edward Snowden leak, it was revealed that since 2007, the US National Security Agency's (NSA) PRISM program has been collecting Internet communication directly from companies, including Microsoft, Yahoo!, Google, Facebook, YouTube, Skype, AOL, and Apple

What kinds of information could be collected from these sources?

- Email, chats, photos, videos, file transfers, video conferencing, social networking details, login times/places

Big Brother is watching

So, your information footprint consists of much more than just the information you intentionally share online

Metadata: "data about data" – additional information that apps and devices exchange along with content, often without the user's knowledge (GPS data, IP addresses and other information we have been talking about)

In June 2013, Snowden leaked that the US NSA was collecting metadata on millions of phone calls (length of calls, originating phone numbers and numbers called, but not the calls themselves)

What would be the arguments, pro and con for this kind of bulk data capture by governments?

It’s not just you

How can other people contribute to your digital footprint?

- Posts and photo tags
- Incidental and deliberate recordings
- Business records
- Public records
- Any kind of records!

Can any of this data be controlled? If so, how? If not, why not?

What role is there, if any, for government to play in controlling how your personal information is shared?

It’s ironic that the same entity that makes and enforces privacy laws also collects vast amounts of private data on the people it represents

Internet Vigilantism

In 2005 in Korea a woman's dog pooped in a Korean subway and she refused to clean up the mess despite offers for help

- The incident was captured on a fellow passenger's phone and posted online

This information was shared around the world and she became known as 개똥녀 or "Dog Poop Girl" in English.

The woman was publicly shamed by thousands of people and she quit her university studies

- Is this an appropriate response for the incident?

Before the digital revolution, it would have been embarrassing for the woman but likely only a small group of people would have known about the incident

Medical records and HIPAA

US Health Information Portability and Accountability Act (HIPPA)

- Federal law that took effect in 2003

Among other things, HIPAA requires the establishment of national standards for electronic health care transactions and the provided security and privacy of health data

It guarantees the right of patients to view their own records, find out who has accessed their health records and know how such records will be shared

However, these rules only apply to electronic medical records used by healthcare professionals – they do not apply to life insurance companies, workers' compensation, auto insurance plans that have health benefits, and others

Medical records and HIPAA

How might these entities (to which HIPAA does not apply) use your personal medical information to their benefit?

Would you withhold sensitive personal and medical information from your doctor if you knew there was a possibility that it might not remain private? Why or why not?

Suppose it someday becomes inexpensive to perform a full genetic profile of people. How could it be used for positive and negative ends?

- Would you want that information recorded about yourself? Why or why not?
- *Gattaca* (1997) – a movie that depicts a dystopian future where genetics is Everything

The illusion of anonymity

There is virtually no anonymity on the Internet

Why would someone want to be anonymous?

In 2008, Gene Cooley of Georgia became the target of a string of “anonymous” posts on the website Topix.com

His attacker (an employee at a local store) made libelous, defamatory posts about him, accusing him of being a drug addict, career criminal, and pedophile

Cooley used a subpoena (court order) to get the IP address of the poster, sued her, and was awarded $404,000 in damages

Read more at: https://abcnews.go.com/Technology/topix-innocent-mans-life-destroyed-anonymous-online-poster/story?id=15963310

The illusion of anonymity

Some details of your computer or mobile device's setup are communicated to your Internet service provider, and often to the site or service you are using

- Even web browser features like "private browsing" are misleading – your IP address and browser-specific information can still be recorded at the other end

Have you ever browsed some items for sale on an online merchant, and then on a completely different website started seeing ads for similar products?

- What does this tell you about your web surfing?
- Would you change your online purchasing habits if you knew your browsing activity would be shared with third parties? Why or why not?

The illusion of anonymity

Here is what Stony Brook University collects when you visit its website:

- The Internet Protocol address of the computer that accessed this Web site
- The type of browser, its version, and the operating system on which that browser is running
- The Web page from which the user accessed the current Web page
- The date and time of the user's request
- The pages that were visited and the amount of time spent at each page

See www.stonybrook.edu/sb/privacy.shtml

The illusion of anonymity

Always assume you have less anonymity, and therefore less privacy, when you're doing something electronically than you would if you were doing it non-electronically

So, if we want to participate in the digital world while protecting our privacy, what might be some "best practices"?

- Only give out as much personal information as you must
- Before you enter any information in a form or allow an app or service to access information about you, ask yourself what they need it for
- What are your options if you don't want to give the information they want?

Why do we give up our privacy?

Generally, because we feel the benefits outweigh the costs

What are some of the benefits?

1. Saving time
2. Saving money
3. Convenience
4. Desire to participate in online groups
5. Desire to attract attention
6. Because you can't live any other way

1. Saving time

In Korea there is the Hi-pass system for making wireless toll payments without having to stop.

- The US has a similar E-Zpass system

But what information do others get in return?

- At the very least you are telling the government when you passed through a certain point on the road

In the US, some states have roads where your E-ZPass ID is read when you enter the road and then you pay a toll when you leave. What are the implications of this?

E-ZPass information has been used in hundreds of court cases, both criminal and civil:

- murder trials, overtime fraud cases, divorce proceedings

2. Saving money

Many stores have point cards or rewards cards

When you scan your rewards card at the register, you give the store permission to keep a record of everything you bought

Why might this be a good thing for the consumer?

- Targeted coupons and special deals for items you purchase on a regular basis could save you money

What is the potential for invasion of privacy?

- What if the stores sells your information to third parties?

Information is valuable

Companies want to collect data about you because it can be very valuable to them for marketing purposes

The “free” online services you use – Kakao, Naver, Facebook, Twitter, Gmail – actually come with a cost: a cost to your privacy and anonymity

- The price you pay is the data you give them

You should generally assume that companies, institutions, and governments are saving and exchanging information about you, both openly and covertly

Pay attention to messages about privacy from companies and institutions

- If you don’t respond, they may share your information by default

Information is valuable

When might you be OK with a company sharing your info?

Which bits of your personal information are more valuable to you?

Which bits do you think advertisers might value most?

Are there pieces of data about you that you would never sell? Why not?

3. Convenience

Sites like Amazon.com and services like Netflix have gotten quite good at making useful recommendations for products and movies to their customers

Why?

- Because whenever you purchase (or even view) an item on Amazon or watch a show on Netflix, a computer system makes a note of it
- We are willing to give up our purchasing history and TV watching habits in the interest of convenience

What's the harm? What might be the harm?

- Have you ever purchased something online or watched a TV program you don't necessarily want the world to know about? Does Amazon know about it?

4. Desire to participate in online groups

People use services like Kakao Talk or Facebook for many useful reasons:

- Keep in touch with friends
- Participate in social groups
- Learn about events

Not using these services can affect our social interactions and daily life

However, companies like Facebook can make money using your information.

- People voluntarily provide tremendous amounts of information on social media
- Companies can sell your information and also use it to better target ads to influence you

Facebook also control services like the News Feed, where they can affect what information you see.

What are some of the risks of a company knowing so much about you?

5. Desire to attract attention

We live in the Age of the Selfie – “look at me and all the cool stuff I’m doing and you’re not!”

Especially prevalent in high school students looking to become “instafamous”

By giving up their privacy in the quest for popularity, teens expose themselves to increased risk for identity theft and cyberbullying

What other risks to safeguarding privacy do you see for those who “post their whole lives” on Facebook, Twitter and Instagram?

- What’s the solution?
- If you must exhibit, you can restrict your profile. But remember that “the Internet never forgets.”

6. Because I have to

In much of modern life it has become impractical to do things in a non-digital way

Consider this: if you were going to buy a TV, laptop, or other expensive item, how likely would you be able to use a check or bring million(s) of won in cash with you?

- Would the business even accept a check?

Many electronic devices now are Internet-capable, if not explicitly Internet-connected all the time: TVs, DVRs, telephones, game consoles, computers, tablets, etc.

- We have quickly changed to an always-connected culture where people use Internet-connected devices without even thinking twice about it. This has become the norm, not the exception.

6. Because I have to

The consequences of this new normal can be pretty serious when company servers are hacked

You can take every precaution there is, but if the companies or government agencies do not safeguard your data, then your privacy is at great risk

What are some major data breaches you have heard of lately?

Some recent major data breaches

2014: hackers steal bank information on 76 million Chase Bank customers: names, addresses, phone numbers and emails

2014: hackers steal information on 150 million eBay users: customer names, encrypted passwords, email addresses, physical addresses, phone numbers and dates of birth

2015: hackers steal information of 80 million health insurance customers from Anthem, including names, dates of birth, Social Security numbers, health care ID numbers, home addresses, email addresses, employment information and income data

2014-2018 hackers stole personal information and credit card information from 383 million Marriott hotel customers.

www.informationisbeautiful.net/visualizations/worlds-biggest-data-breaches-hacks/

6. Because I have to

Just as the credit card has replaced cash, the email has replaced the written letter

Suppose every time you sent a letter to someone, the Korea Post reads your mail and keeps a copy for itself

- You may be unhappy with that, but that's exactly what free email services do so that they can send targeted advertising to your computer screen

What about email you send using your employer's system? Or your school's system?

- Should they be allowed to read it all and retain copies? Why or why not?

Part of SBU's stated email policy

Stony Brook University email, wherever it may be stored or transmitted, belongs to the University, may be audited by the University at any time, and may be subject to disclosure to a third party, including review by authorized law enforcement personnel. Email accounts are subject to review and disclosure, without notice, when required by law, where a violation of law or University policy may exist, where there is a risk of spoliation, bodily harm, property loss or damage, where the University's mission is jeopardized or during the course of routine system administration.

See it.stonybrook.edu/policies/d106

SBU’s email retention policy

All email is retained within the user's account until it is moved to a trash file ("trash") by the account holder.

Email remaining in trash for 30 days will be purged automatically from trash.

The account holder may immediately purge emails in trash by emptying trash.

Email messages purged from trash will be retained within University files for an additional 30 days after the removal to allow recovery by the email system administrator. After that, purged emails will be expunged automatically from the system.

The evolution of privacy

Concerns about invasion of privacy are nothing new

Lawyers Warren and Brandeis wrote in 1890: “Instantaneous photographs and newspaper enterprise have invaded the sacred precincts of private and domestic life; and numerous mechanical devices threaten to make good the prediction that ‘what is whispered in the closet shall be proclaimed from the house-tops.’”

Even by that time, defamation – malicious, false gossip – was illegal (and still is) in the US

But what about malicious, true gossip? They argued that it could cause mental pain and distress far greater than physical harm. Therefore, people have the right to control what is said about them. They have the “right to be left alone.”

The evolution of privacy

The problem with their position is that it's unenforceable and unconstitutional in the US

- The First Amendment guarantees the right to free speech, even ugly speech

In 1967, author Alan Westin argued that personal privacy needs to be protected not by restricting speech, but by restricting how private information is used: "Privacy is the claim of individuals, groups or institutions to determine for themselves when, how, and to what extent information about them is communicated to others"

This is a change in thinking from "Don't say anything about me at all" to "Don't say anything private about me that I don't want you to share"

Fair Information Practice Principles (FIPP)

In 1973 the US Department of Health, Education, and Welfare attempted to answer this question with its FIPP list for medical data:

1. Openness: there must be no personal data record-keeping systems whose existence is secret
2. Disclosure: there must be a way for a person to find out what information is kept about him and how it's used
3. Secondary Use: there must be a way for a person to prevent other people from accessing one's data without consent
4. Correction: there must be a way for a person to correct or amend one's record
5. Security: the data must be securely held and used only for its intended purposes

Fair Information Practice Principles

These principles were never formally adopted, but they are in use in many governmental and corporate privacy policies

US federal privacy laws are a bit spotty and are a hodgepodge of rules that pertain to specific technologies

Suppose you are a Netflix customer. How would the FIPPs apply to your TV watching habits?

- Openness, Disclosure, Secondary Use, Correction, Security

Let's consider something a little more serious: your bank account data, including bank transactions and ATM usage. How would the FIPPs apply there?

Can privacy be a bad thing?

High privacy standards themselves have a cost

HIPAA and other privacy laws can make medical research hard to conduct

- Access to large datasets about health issues and medical payments can tell us a lot about public health and the cost of medical care
- But HIPAA restricts the release of much of this data

What about financial data? If a bank suspects someone is laundering money or funding terrorists through its accounts, which takes precedence: privacy or security?

Most people don't use encrypted email, but it's readily available. Should the government be given the encryption keys so they can read your email? Why or why not?

There is no escape

Massive data collection has become so deeply engrained in modern society that you would have to live in a cave by yourself to avoid it

In this lesson we didn't even touch on other sources of private and public data that can be used against people:

- Property deeds, business records
- Home electronics: smart electric meters, smart TVs, color laser printers (date, time and printer serial number encoded on print-outs)
- RFID tags embedded in merchandise you purchase
- Surveillance cameras with facial recognition software

There is no escape

One final story: An angry father of a teenage girl went to the manager of a Target store because Target was mailing her coupons for baby clothes and cribs. He angrily asked “Are you trying to encourage her to get pregnant?”

It turns out that the daughter had been buying items at Target that Target’s software determined were typically bought by expectant mothers, and so the software inferred she was pregnant.

As it happens, the software was correct and figured out the girl was pregnant before her father did

Read more at: https://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/#18f9bd106668

Questions?