Privacy and Anonymity

Steven M Bellovin

smb+6184@cs. . .

Course Outline

- Legal framework (US and European)
- Data mining and databases
- Anonymous commerce (digital cash)
- Anonymous use of the Internet (onion routing, anonymous browsing, P3P)
- Traffic analysis
- Biometrics and authentication
- Policy and national security considerations
- Possible occasional interruptions to the schedule to read about and discuss privacy-related current events

Class Structure and Grading

- Weekly paper write-ups (33% of the grade)
- Student presentations of assigned reading material (22%)
- A 5–10 page midterm paper (10%, due March 10)
- Class participation (5%)
- A 10–15 page final paper (20%), plus
- In-class presentation (10%) of that paper, during one of the last two class meetings,
- Note: percentages are approximate and subject to change

Write-Ups

- 1-2 page (max!) summary of each reading
- Email me your write-up or notes by the start of each class. (Note: Postscript, PDF or ASCII preferred; Word discouraged. . .)
- Write-ups turned in up to one week late accepted at 50% off
- The summary is just that: a summary of the basic ideas within the reading, in your own words. To be blunt, your write-up should show me that you understood what you read.
- The summary *must* include a discussion of the significance of the reading

Class Presentations

- Sign up for a presentation the week before
- You *don't* need to use slides, Powerpoint, etc.
- We'll be using the conference room computers (Linux and Windows) and projector, though you can use your own laptop if you prefer.
- Assume that the others have read the paper.
- Present the highlights, present the significance, and lead a discussion of the paper.
- 10–15 minute summary of the work. (Remember that your audience will have read the work.)
- 5-10 minutes of your reaction, evaluation, lessons, etc.
- What is important about that work? Why did I assign it?
- 10-15 minutes of class discussion, led by the presenter

The Papers

Think of them as take-home essay exams where you pick the question. They are not expected to be original research.

The easiest approach is to take one topic that I've covered, read the listed papers carefully *plus others you find yourself*, and produce a survey paper on that subject. Another approach is to do a detailed case study of some issue or implemented solution. Feel free to suggest your own topics, too

You must confirm the selection with me.

A good bibliography and good citations are crucial.

The written paper is due when the final for this class would be held.

Bibliographies

- A form of proper attribution
- Where to go for more information
- Sometime a way to judge the credibility of the statement

Resources

- There is wisdom in the world not known to Google
- There are rooms and even whole buildings on this campus known as “libraries”
- The CU library network also has a lot of electronic works

Presenting Your Own Paper

- 10–12 minute presentation of the work
- 5 minutes of class discussion
- Bonus credit for those who present in the first session, since they have one less week to work on it

Contacting Me

- Email is by far the best way to reach me outside of office hours
- For email pertaining to this class, use smb+6184@cs. . .
- I travel a fair amount, and rarely check voice mail when on the road
- If you need to see me and can't make it during my office hours, email me to set up an appointment
- Exceptions to my office hours will be posted on my web page: `http://www.cs.columbia.edu/~smb`

Discussing Legal Works

- This is not a law class
- We are not interested in the legal minutae or in a critique of the legal reasoning
- We are interested in the broad foundations of legal decisions, their effects, and on how these decisions and laws interact with technology

Definitions

anonymity The condition of an identity being unknown or concealed. (RFC 4949)

privacy The right of an entity (normally a person), acting in its own behalf, to determine the degree to which it will interact with its environment, including the degree to which the entity is willing to share its personal information with others. (RFC 4949)

pseudonym A fictitious name, especially a pen name. (The Free Dictionary)

Biases

Many of the assigned readings assume that privacy is a benefit, and should be protected to the extent possible.

"You have zero privacy anyway. Get over it".
—Scott McNealy
CEO, Sun Microsystems

I'm an unabashed privacy advocate. But the course is not about *promoting* privacy, and I don't expect you to agree with me. The course is about understanding privacy, its context, how to achieve it if you want it, and what the drivers are against privacy.

Other Issues?

The U.S. System of Government (Oversimplified)

- 50 states, with their own constitutions and laws
- The federal government, with its constitution and law (see `http://www.usconstitution.net/const.html` for the federal constitution)
- The federal government's powers over state affairs is limited, at least in theory
- All 50 states have an independently-elected governor and legislature
- All states have an independent judiciary

Rights Under the U.S. Constitution

- Most rights are not mentioned in the original constitution; they're in the first 10 amendments ("The Bill of Rights")
- Courts can declare a law unconstitutional, in which case it's not enforceable
- (That power is *not* mentioned in the constitution; it was invented as a political maneuver in 1803, in the case Marbury v. Madison, 5 US 137)
- *District Court* rulings apply to that case only, *Appeals Court* rulings are binding within their region of the country; Supreme Court rulings are binding nationwide

Scope of Rulings

- Courts almost always follow precedent
- Courts only rule on what they have to
- Courts don’t rule on abstract principles, only actual cases
- Rulings frequently limited by facts in the particular case
- You must have “standing” to sue

Example: in re Boucher

- In *in re Boucher* (2007 WL 4246473, Nov. 2007): someone arrested by a customs agent for possession of child pornography on his laptop
- Customs agent saw some child porn — but a later forensic analysis found that the drive was encrypted. Could Boucher be compelled to disclose the key? The magistrate said “no”, because of Boucher’s right not to incriminate himself.
- Decision was by a magistrate: no binding precedent set
- Because of the location of the search — the border — and the fact that Boucher showed he had access, the judge’s reasoning may not stand up
- This case may be unimportant, and the ruling reversed based on the facts of the case — but the *reasoning* may be cited elsewhere

Constitutional Roots

- Limits on government actions; does not apply to private behavior
- The word “privacy” is not mentioned in the U.S. Constitution
- A right to privacy is inferred from other provisions
- First Amendment: anonymous speech has a long history; right to receive information also protected
- Fourth Amendment: warrants and probable cause required for searches. Held in 1967 to encompass wiretaps
- Fifth Amendment: note distinction between testimonial evidence and physical evidence
- Ninth Amendment: states that other rights exist
- Fourteenth Amendment extended certain rights to the states

Legal Research

- `http://www.findlaw.com` — statutes, cases, etc., under the “For legal professionals” tab
- `http://www.law.cornell.edu` — similar
- LexisNexis, via CU library — comprehensive, annotated, complex
- `http://www.lib.uchicago.edu/~llou/mpoctalk.html` — research guide
- `http://thomas.loc.gov` — good for bills currently before Congress
- Many, many more

Legal Notation: Court Cases

United States v Morris (1991, CA2 NY) 928 F2d 504

United States v Morris Common name; lists parties to the case

1991 Year of the decision

CA2 NY Court of Appeals, Second Circuit, New York

928 F.2d 504 Volume 928, page 504, of the Federal Reporter, Second Series (yes, cases are cited by page number!)

See `http://www.faqs.org/faqs/law/research/part1/` for more details

Similar formats are used even when referring to cases from other countries: Semayne's Case, 5 C. Rep. 91a, 77 Eng. Rep. 194 (K.B. 1603).

Legal Notation: U.S. Code

18 USC 1030 (Computer Fraud and Abuse Act)

18 Title 18 of the U.S. Code (Crimes and Criminal Procedure)

USC U.S. Code or other body of laws or regulations ("CFR" is the Code of Federal Regulations)

1030 Section 1030

Computer Fraud and Abuse Act The common name. Strictly speaking, it applies to the *Public Law* that created that section (or made major amendments to it)

The U.S. Code is the orderly compilation; sections are created and amended by Public Laws: P.L. 98-473, Title II, Ch XXI, §2102(a), 98 Stat. 2190.

Some Current Events

- Warrantless NSA surveillance of Americans
  `http://www.eff.org/issues/nsa-spying`
- Justice Dept. subpoenas massive amounts of Google data:
  `http://www.informationweek.com/showArticle.jhtml?articleID=184417576`
- Cell phone calling records for sale:
  `http://www.washingtonpost.com/wp-dyn/content/article/2005/07/07/AR2005070701862_pf.html`

Privacy Challenges in Authentication Systems

Context

- National Research Council report
- Committe composed of computer scientists (including me), lawyers, human factors experts, biometrics experts
- Report is *not* the result of original research by the commitee
- This chapter provides a legal and societal framework for privacy

Privacy Impact of the Decision to Authenticate

- Registration may require disclosure of personal facts or information
- Authenticating may cause creation of records
- Details of an event can augement these records
- Even without personally identifiable information, a dossier can be compiled
- Others may have access to this information

Access Control Systems

- What sort of access is mediated, and when? Are people aware of the borders or the access control? Differs in physical and online worlds.
- Computer technology reduces cost of record keeping
- Computer technology facilitates linkage
- Computer technology enables covert identification on a large scale (think red light cameras)

Legal Foundations of Privacy

- Common law: “[T]he house of every one is to him as his castle and fortress.” Semayne’s Case, 5 C. Rep. 91a, 77 Eng. Rep. 194 (K.B. 1603)
- Doesn’t work as well in today’s interconnected world
- Types of privacy:
  - Bodily integrity
  - Decisional privacy
  - Information privacy
  - Communications privacy

Common Law Roots

- Tort law provides (some) recourse for private misbehavior
- Prosser (1960) defined four separate privacy torts:
  - Intrusion upon seclusion (including private affairs)
  - Public disclosure of private facts
  - Publishing objectionable, false information
  - Misappropriation of name or likeness
- Useful (in this report) as a reminder of what people expect to be protected

Statutory Protections

- Federal Trade Commission has the power to enforce voluntary privacy statements
- Piecemeal statutes protect some privacy, often in response to market failure or narrow court rulings.
- Often regulate private sector behavior; sometimes constrain government
- 11 different federal laws contrain private sector; much personal information is freely available for sale, including to the government
- Personal information voluntarily given to businesses not protected by the Fourth Amendment

Fair Information Practices

- First “code of fair information practices” developed in 1973 at HEW
- Basic rules for minimizing information collection, ensuring due process, protection against secret collection, provide security, ensure accountability
- Emphasize individual knowledge and consent
- Principles are broadly accepted, but individual principles not implemented uniformly

Fair Information Principles and Practices

- Collection limitation
- Data quality
- Purpose specification
- Use limitation
- Security
- Openness/notice
- Individual participation
- Accountability

Privacy of Communications

- Law recognizes need for privacy in new forms of communication
- Need is both value-driven and pragmatic
- Balance between privacy and law enforcement needs
- Protection for new communication mechanisms not as strong as for voice

Conclusions

- Authentication can affect all forms of privacy
- Authentication systems should not infringe on autonomy and legal rights. They should recognize the need for multiple identities
- Fair information practices should be followed