Lecture 1 (slides of interest)

Balachander Krishnamurthy

AT&T Labs–Research

http://www.research.att.com/~bala/papers

What the course is *not* about

Course URL: `https://www.cs.columbia.edu/~bala/s14`

- Data privacy
- NSA or other government/legal issues
- Any ATT products

Disclaimer: Everything that I say in class are my opinions and not my employer's

Course: high-level view and pre-requisites

- Learn about Internet privacy
- Read several technical papers
- Write code to help improve understanding of privacy as it relates to OSNs
- Write a potentially submission-quality paper
- Pre-requisites:

  Coding ability
  Comfort with reading technical publications
  Strong desire to participate actively in class

What the course is about

- State of the art awareness about the privacy problem on the Internet
- With a focus on OSNs
- Understanding the entities involved in the privacy issue and how they act
- Detecting privacy leakages
- Tools for privacy protection
- Less about filling your head with ideas and more about pointing various streams of possibilities

What you will do

- Some homework (alas!)
- Reading papers and summarizing/presenting them
- Coding (if you don't say *cool!*, then you are in the wrong class)
- Writing a paper possibly of submission quality (not as hard!)
- My overall assessment of you will be based on active class participation
- Skipping classes will impact your grade

Grade breakdown

- 10% class attendance and active participation
  (subjective and can influence overall grade!)
- 20% homework
- 45% on group project
- 25% on writeup: 12 page potentially submission-quality paper

Homework, project component deadlines

Homework: 1: 1/29, 2: 2/12, 3: 3/5 (due BEFORE classes start on those days)
Project: Proposal: 2/26, design doc/progress report: 3/26
Presentations: 4/16, 4/23
Final paper: 4/30

- No extensions
- As adjunct instructor (have a full time job) time: limited, schedule: inflexible
- Do not ask for extensions
- If you have a medical or personal emergency, I'll refer you to the Dean's office and accept their guidance
- Anticipate problems and adjust your schedule; make backups
- Did I mention: No extensions?

Academic honesty

Please see: http://www.cs.columbia.edu/education/honesty

Discussion is ok but copying is $lame$

I view your improperly helping someone else to cheat as tantamount to cheating.

I look for privacy leakages for a living; so chances of you escaping scrutiny is low.

If you write code as a group save your svn logs.

I intend to police closely: there are several technical and non-technical mechanisms available to me.

Cheating is unfair to the vast majority of students who work hard for their grade. Impact on cheaters will not be pretty.

Citation rules

If it is not your idea, then cite the source.

Say how you build on it or differ from it.

Wikipedia is not a source; secondary citations are problematic.

See https://owl.english.purdue.edu/owl/section/2/8/

Material

Lecture 1 uses material from 'A guide to privacy law' Robert Rothman, lecture by Joss Wright at Oxford, a National Academies Press report entitled 'Engaging privacy and information technology in a digital age' edited by Waldo, Lin, and Millett.

Privacy in simpler terms

I get to decide what information about me that you get to see/hear

- when
- in what context
- for how long
- for further dissemination (or not)

etc. At its root, there is an element of *control*. If you do not control information about yourself, then who does?

Why let others define you?

Viewing privacy from different angles

- Consumer
- Onine/offline views
- Researchers views

[MFSV]: Pew report: Digital Footprints `http://www.pewinternet.org/Reports/2007/Digital-Footprints/1-Summary-of-Findings.aspx`

Contextual privacy [Nis11]

- Gathering and sharing information must be specific to that context
- Looking up information in a library is different from searching online
- Even though the end goal of the user may be the same
- Incentives vary: online, advertising becomes a factor; hard to be sure of absence of bias
- Libraries protect your book checkout history but what about online searches?
- "Rather than look for similarity of action, examine similarity of function/purpose"
- Contexts should constrain how information flows

[Nis11]: Helen Nissenbaum, A Contextual Approach to Privacy Online

Summary of the privacy problem

- Perception/expectations vary: beyond just cultural differences
- User's expectations vary: libertarian to paranoid
- Majority's perception: often situational and sadly often too late
- Disconnect between offline and online: most are savvy offline but fail online
- Key differences: hidden parties, speed, extent, and duration of spread
- Absence of understanding of monetary value of data
- Misalignment of incentives across the entities

A brief introduction to OSNs

- Facebook, LinkedIn, Twitter, Digg, Renren, Weibo, Flickr, Line, WeChat, SnapChat
- MySpace, Orkut, LiveJournal, Gowalla, Dodgeball (last 2 are dead)
- Others?
- Anyone on any unusual OSNs? (Path!)
- Is YouTube an OSN?

Homework 1– DUE Next week beginning of class!

Q1: Two paragraphs on your views on privacy both offline and online with a focus on the differences between the two along any axes that you can think of – there are no right or wrong answers but am looking for breadth of views.

Q2: 1 para each: What do you think are the differences between
(a) privacy and security (b) privacy and anonymity

Q3: What steps/tools/techniques do you use to protect your online privacy?

Q4: optional/extra credit: Why I am not using Piazza for this course

Papers to read this week
[Nis11]: A Contextual Approach to Privacy Online
[Kri10]: I know What you will do next summer
[MFSV]: Digital Footprints

for possible short presentation in next class

Next class

- Duality of privacy and security
- Privacy and security
- The role of usability
- Crypto and its failure
- Privacy and anonymity
- Anonymization techniques
- Anonymization myths
- Why hasn't duality yielded more?

References

[Kri10] Balachander Krishnamurthy. I know what you will do next summer. *ACM SIGCOMM CCR*, 40(5), 2010. http://www.research.att.com/~bala/papers/ccr10-priv.pdf.

[MFSV] Mary Madden, Susannah Fox, Aaron Smith, and Jessica Vitak . Digital footprints. http://www.pewinternet.org/Reports/2007/Digital-Footprints/1-Summary-of-Findings.aspx.

[Nis11] Helen Nissenbaum. A contextual approach to privacy online. *Daedalus the Journal of the American Academy of Arts & Sciences*, 140(4):32–48, Fall 2011. http://www.amacad.org/publications/daedalus/11_fall_nissenbaum.pdf.