CS208: Applied Privacy for Data Science
Spring 2019

Syllabus

Course Website: http://seas.harvard.edu/~salil/cs208/
Time & Place: Mondays & Fridays 10:30-11:45am, Maxwell Dworkin 221

Course Staff

Instructors:
James Honaker (he/him/his)
Maxwell Dworkin 135

Salil Vadhan (he/him/his)
Maxwell Dworkin 337

Teaching Fellow:
Jayshree Sarathy (she/her/hers)
Maxwell Dworkin 334

Overview

Data scientists, including industry analysts, scientific researchers and data-driven policy makers, often want to analyze data that contains sensitive personal information that must remain private. However, common techniques for data sharing that attempt to preserve privacy either bring great privacy risks or great loss of information. Moreover, the increasing ability of big data, ubiquitous sensors, and social media to record lives in detail brings new ethical responsibilities to safeguard privacy.

The traditional approach to protecting privacy when sharing data is to remove "personally identifying information," but it is now known that this approach does not work, because seemingly innocuous information is often sufficient to uniquely identify individuals. A long literature has shown that anonymization techniques for data releases are generally open to reidentification attacks. Indeed, there have been many high-profile examples in which individuals in supposedly anonymized datasets were re-identified by linking the remaining fields with other, publicly available datasets. Aggregated information can reduce but not prevent this risk, while also reducing the utility of the data to researchers.

This class will provide an overview of the risks of private data leakage in data science applications and a firm foundation in how to measure and protect against these risks using the framework of differential privacy, together with a hands-on examination of how to build algorithms and software to preserve privacy, including a review of the deployed solutions in industry and government.

Differential privacy, deriving from roots in cryptography, is a formal, mathematical conception of privacy preservation. It guarantees that any released statistical result does not reveal information about any single individual. That is, the distribution of answers one would get with differentially private algorithms from a dataset that does not include myself must be indistinguishable from the distribution of answers where I have added my own information.

Using differential privacy enables us to provide wide access to statistical information from a privacy sensitive dataset without worries of individual-level information being leaked inadvertently or due to an adversarial attack. There is now both a rich theoretical literature on differential privacy and numerous efforts to bring differential privacy closer to practice, including large-scale deployments by Google, Apple, Microsoft and the US Census Bureau. This course will set out a foundation in the underlying theory of differential privacy, and then consider the practical elements of implementing and deploying privacy-preserving techniques for data analysis.

Format and Goals

The class will typically alternate between lecture/discussion meetings, which will focus on learning the fundamentals and discussion of important issues, and practicum sessions where there will be some lecture, some demonstration code, and some hands-on computer work. Homeworks will typically involve some analytical/mathematical work to learn techniques, and increasingly as the term progresses, hands-on data immersive coding tasks to test and experiment with approaches to privacy preservation within the context of real datasets and data science questions.

The main components of the course are as follows:

- **Class Participation:** Attendance is mandatory, as our meetings will be highly interactive, especially our practicum meetings. For some class meetings, we will provide material for you to read and possibly comment on in advance. Participation also includes your engagement in section and office hours and on Piazza.
- **Problem Sets:** There will be problems sets due approximately every other week, typically Tuesdays at 11:59pm. These will be progressive, and require reuse of previous solutions, so it is important both to keep up on the problems, review feedback to submissions, and

organize and document previous submitted code so that it can be reused. You will have 6 late days for the semester.

- **Final Project:** You will do a final project on a topic of your choosing. Projects can be done individually or in pairs, with groups of three allowed for ambitious projects. You can do a project that is experimental, or involves system-building, or is theoretical. The project should provide good opportunities to connect the course material to your other interests and get some exposure to the frontier of research in differential privacy. The project will involve submitting topic ideas for feedback (due with problem set 1, Feb. 19, and revised as part of problem set 2, due March 12), a detailed project proposal (due April 9), a written paper (draft due in reading period, final version in exam period), and a project presentation (in exam period). We will post more details about the final project, including some directions to look for topics, early in the course.

We anticipate placing roughly equal weight on each of the above three elements in determining final grades.

By the end of the course, we hope that you will all be able to:
- Identify and demonstrate risks to privacy in data science settings,
- Correctly match differential privacy technology with an application,
- Safely implement privacy solutions, and experimentally validate the performance and utility of algorithms.
- Understand differential privacy at a level sufficient to engage in research about best practices in implementation, apply the material in practice, and/or connect it to other areas,
- Formulate and carry out an interesting, short-term independent research project, and present the work in both written and oral form.

Prerequisites

Basic probability, algorithms, and programming at the level of CS109/AC209. STAT110 and CS124 should also be sufficient preparation.

Diversity and Inclusion

We would like to create a learning environment in our class that supports a diversity of thoughts, perspectives and experiences, and honors your identities (including race, gender, class, sexuality, socioeconomic status, religion, ability, etc.). We (like many people) are still in the process of learning about diverse perspectives and identities. If something was said in class (by anyone) that made you feel uncomfortable, please talk to us about it. If you feel like your performance in the class is being impacted by your experiences outside of class, please don't hesitate to come and talk with us. As a participant in course discussions, you should also strive to be open-minded and respectful of your classmates. (Statement modified from one by Dr. Monica Linden at Brown University.)

Health Accommodations

If you have a physical or mental health condition that affects your learning or classroom experience, please let us know as soon as possible so that we can do our best to support your learning (at minimum, providing all of the accommodations listed in your AEO letter if you have one). (Statement adapted from one by Prof. Krzysztof Gajos.)

Support Structures

Everyone can benefit from support during challenging times. Not only are we happy to listen and make accommodations with deadlines as needed, we can also refer you to additional support structures on campus, including, but not limited to, the below.

- Bureau of Study Counsel
- InTouch
- Counseling and Mental Health Services, 617-495-2042
- Let's Talk
- Room 13, 617-495-4969

Collaboration Policy

Students are encouraged to discuss the course material and the homework problems with each other in small groups (2-3 people). Discussion of homework problems may include brainstorming and talking through possible solutions, but should not include one person telling the others how to solve the problem. In addition, each person must write up their solutions independently, and these write-ups should not be checked against each other or passed around. While working on your problem sets, you should not refer to existing solutions, whether from other students, past offerings of this course, materials available on the internet, or elsewhere. All sources of ideas, including the names of any collaborators, must be listed on your submitted homework.

In general, we expect all students to abide by the Harvard College Honor Code. We view us all (teaching staff and students) as engaged in a *shared mission* of learning and discovery, not an adversarial process. The assignments we give and the rules we set for them (such as the collaboration policy) are designed with the aim of maximizing what you take away from the course. We trust that you will follow these rules, as doing so will maximize your own learning (and thus performance on exams) and will maintain a positive educational environment for everyone in the class. We welcome and will solicit feedback from you about what more we can do to support your learning.

Topics to be Covered

- Privacy Attacks on “De-Identified” Data and Statistical Data Releases
    - Reidentification attacks
    - Reconstruction attacks
    - Membership attacks
- Foundations of differential privacy
    - Definition and interpretation
    - Basic mechanisms (Laplace, Gaussian, randomized response, histograms, exponential)
    - Composition of differential privacy
    - Survey of known algorithms and experimental validation
- Implementing (centralized) differential privacy
    - US Census Bureau’s (planned) deployment
    - Privacy budgeting
    - Interactive query interfaces
    - Differentially private programming platforms
- The local model of differential privacy
    - Basic theory and mechanisms
    - Randomized response, histograms
    - Comparison with the centralized model
    - Deployments by Google and Apple
- Other possible topics (depending on time and interest)
    - Statistical inference and machine learning under differential privacy
    - Law & policy considerations
    - Tiered access models
    - Side-channel & randomness attacks on implementations
    - Combining differential privacy and cryptography