Data Privacy

Fall 2015

Course outline

Instructor: Periklis A. Papakonstantinou, e-mail: periklis.research@gmail.com, 1WP-1072

Teaching Assistant: Hafiz Asif, e-mail: hafiz.asif@rutgers.edu, 1WP-1023A

Lectures: Monday 6-9pm @ 1WP-502; office hours: Monday 4.30pm

Emailing: Make sure that your emails have in their subject "data privacy" or "privacy class"

Textbooks and readings: The main source is the text "*Algorithmic Foundations of Differential Privacy*", by Dwork and Roth. We shall also rely on readings from Sovolone's monograph "*Nothing to Hide: The False Tradeoff between Privacy and Security*", and Nissenbaum's "*Privacy in Context: Technology, Policy, and the Integrity of Social Life*" and complement this with various philosophical and legal essays. For the short introduction to cryptography we will use extracts from "Introduction to Modern Cryptography" by Katz and Lindell. *In every class there will be a detailed, carefully chosen list of chapters and sections that cover everything discussed in class.*

Web-page: TBA

Course description, objectives & prerequisites

Privacy is a rather controversial and truly interdisciplinary issue. It brings together philosophy and law, mathematics and computer science, and information sciences. Every class with content must have a focus point (otherwise it's just abstract content-less "discussions"). This class will touch all these areas with our focus *on the foundations of privacy*. That is, much of our effort will be in conceptualizing in the form of *formal definitions* "what does privacy mean", and also giving solutions in the form of formal *algorithmic constructions that provably work*.

This class is a modern, mathematically rigorous introduction to the topic. It puts forth both the basics and the state-of-the-art in the field, and is designed for a diverse audience. Ideally, courses in probability theory or theoretical statistics, combinatorics, theory of algorithms, mathematics of computation, and theory of databases will be helpful to the student. *However, the design of the course assumes only minimal prerequisites.* The only actual requirement is some level of mathematical maturity (equivalent to a reasonably good freshman undergraduate). Every formal knowledge required in these fields will be covered in class. The assumption is that the student will put effort in self-studying from the presented material in case she/he is missing any or all of the prerequisites.

Topics

- Notions of privacy
- Regulations and legal framework (focus on the US system)
- Formal notions of privacy – focus on k-anonymity & differential privacy
- Differentially private databases: constructions and mechanisms
- Impossibility results
- Introduction to modern cryptography I: computing together with untrusted parties
- Introduction to modern cryptography II: storing/retrieving information in untrusted databases

*** For BUS see the accompanying document**

Grading

Here is the breakdown of the final grade.

- **Scribe notes : 20% – also preparing readings before each class**
- **4 Quizzes : 20% (5% each)**
- **3 Assignments : 30% (equally weighted)**
- **Term project & oral presentation : 30%** – at the level of a published research paper
- * **Bonus points:** adds up to 5% from BUS (see BUS document)

Preparing the reports, collaboration, missed assignments/tests & re-marking requests

- The assignments should be done individually by each student. You are not only allowed but also encouraged to form study groups. Your assignment report must be prepared solely by you (avoid plagiarism).
  What type of collaboration is not considered plagiarism: during your meetings to collaborate for an assignment (i) no electronic collaboration is allowed (you can only meet in person), (ii) you should not discuss the very details of the solutions, and (iii) you are not allowed to take any transcript out of your meeting; i.e. you cannot take with you any notes or any form of electronic record. Then, you let at least one hour pass in between this meeting and you starting preparing your report.
- *Every* construction and *every* question that asks for a bound *must be accompanied with a proof of correctness or a proof of the bound etc.* Such a proof corresponds to $\approx$ **60%** of the mark.
- No late assignments accepted. If there is an acceptable and well-documented reason the instructor will arrange for redistribution of marks.