Privacy Engineering

Instructor

Daniel Aranki, Department of Electrical Engineering and Computer Sciences, UC Berkeley.

Biography

Daniel Aranki received a PhD in computer science from UC Berkeley in 2017. He received a BSc in computer engineering from the Department of Electrical Engineering at Technion—Israel Institute of Technology, Haifa, Israel, in 2011. Between 2007 and 2011, he worked in the Mobile Wireless Group at Intel Corporation, Haifa, Israel. During his time there, he worked on WiFi receiver design, design and verification flow automation, and WiFi system architecture design. He is the executive director of the Berkeley Telemonitoring Project. His research interests include machine learning, statistical analysis, privacy, information disclosure, and health telemonitoring.

Course Description

With the rise and advancement of artificial intelligence (AI) and machine learning (ML), we are facing new privacy challenges at an ever-growing pace. In this introductory technical course, we survey privacy mechanisms that are applicable to systems engineering. In particular, we focus on the inference threat that is arising because of the aforementioned advancements in AI and ML.

To situate the course in the bigger context, we first briefly discuss the history of privacy and compare two major examples of general legal frameworks for privacy from the United States and the European Union. As a segue to the technical part of the course, we then survey two design frameworks of privacy that may be used to guide the design of privacy-aware information systems.

Finally, we survey a number of threat-specific technical privacy frameworks and discuss their applicability in different settings. Namely, we start by discussing the origins of statistical study of privacy with randomized response. We then discuss anonymization and confidentiality techniques including k-anonymity, l-diversity, t-closeness, and delta-presence. Afterward, we discuss the hardness of absolute protection under these models. From there, we survey semantic privacy models including differential privacy, privacy protection against “honest but curious” agents, and private disclosure of information. We motivate many of the aforementioned models by health care applications.

The course overviews a broad number of paradigms of privacy from a technical point of view. The course is designed to assist system engineers and information systems professionals in getting familiar with the subject of privacy engineering and train them in implementing those mechanisms. In addition, the course is designed to coach those professionals to critically think about the strengths and weaknesses of the different privacy paradigms. These skills are important for cybersecurity professionals and enable them to effectively incorporate privacy-awareness in the design phase of their products.

Prerequisites

This is a graduate-level course, admission to a graduate degree is a prerequisite. Undergraduate students need the instructor's permission.

Relevant Background

Although not required as a prerequisite to this course, knowledge in linear algebra can benefit students in learning and using MATLAB. In addition, knowledge in statistics, probability theory, and/or information theory is relevant to this course. (The basic necessary background in these topics will be covered in this class.)

Course Objectives

By the end of this course, students will be able to

- Describe the different technical paradigms of privacy that are applicable for systems engineering
- Critique the strengths and weaknesses of the different privacy paradigms
- Implement such privacy paradigms, and embed them in information systems during the design process and the implementation phase
- Stay updated about the state of the art in the field of privacy engineering

Course Deliverables

- Weekly Homework
- Participation
- 3 x Labs
- Implementation Project
- Exam

Collaboration Policy and Academic Honesty

We encourage studying in groups of two to four people. This applies to working on homework, discussing labs and projects, and studying for the exam. However, students must always adhere to the UC Berkeley Code of Conduct (http://sa.berkeley.edu/code-of-conduct) and the UC Berkeley Honor Code (https://teaching.berkeley.edu/berkeley-honor-code). In particular, all materials that are turned in for credit or evaluation must be written solely by the submitting student or group. Similarly, you may consult books, publications, or online resources to help you study. In the end, you must always credit and acknowledge all consulted sources in your submission (including other persons, books, resources, etc.).

General Grading Philosophy

The course will be graded on an absolute scale, and the grades will not be fitted to a specific curve. This is a graduate-level course, and we trust that different students will have varying levels of interests in the different subjects in the course. As such, the grading scheme is designed to acknowledge this intellectual diversity.

Late Submission Policy

Solutions of homeworks and labs will be discussed during the lectures and discussions of the course. Therefore, any assignment that is submitted after the deadline will be returned without grading and will receive a grade of zero.

Homework

Weekly homework will be assigned on readings and topics discussed during class lectures. Homework should be submitted individually.

Homeworks are due before the beginning of the discussion of every week. Each student will get two homework drops without penalty. Please note that a late submission is considered a drop.

Labs

In addition to homework, there will be labs throughout the course that will demonstrate some of the concepts taught in this course. Labs should be submitted in groups of three or four.

Implementation Project

Students must work on an implementation project that will require several weeks of programming. Each project should be submitted by a team of three or four.

Exam

There will be one examination that will review the material covered in the course and evaluate the student's understanding. During the time of the exam, collaboration and consultation with other people is prohibited. You may, however, consult with course materials (homework, labs, readings, etc.) and your own written notes.

Participation

Participation and taking an active part in every aspect of the course are key to internalizing the material of the course. Participation includes, but is not limited to, i) active participation in discussions, ii) discussing assignments with other students, and/or iii) activity in the class forum (by asking questions and/or contributing to answering other students' questions).

Disability Accommodation

If you need disability-related accommodations in this class, if you have emergency medical information you wish to share with me, or if you need special arrangements in case the building must be evacuated, please inform me as soon as possible.

Readings

There is no single textbook reading for this course; the recommended weekly readings will include book chapters, published articles, reports, and statutes. See the detailed syllabus below for week-by-week reading assignments.

List of Topics by Week

We will cover a variety of topics in this course. The following syllabus is detailed by week. Please read the required material before viewing the asynchronous videos of each week.

Week 1

Readings:

- United States Congress. (1974). Privacy Act of 1974.
- European Parliament and Council of the European Union. (2016). The General Data Protection Regulation (EU) 2016/679.
- Miller, B., Huang, L., Joseph, A. D., & Tygar, J. D. (2014). I know why you went to the clinic: Risks and realization of https traffic analysis. In *International Symposium on Privacy Enhancing Technologies Symposium* (pp. 143-163). Springer, Cham.
- Wright, C. V., Ballard, L., Monrose, F., & Masson, G. M. (2007). Language identification of encrypted voip traffic: Alejandra y roberto or alice and bob? In *USENIX Security Symposium* (Vol. 3, pp. 43-54).
- Federal Trade Commission. (1998). Privacy online: A report to Congress. *Washington, DC, June*, 10–11.

Summary:

Overview of class. Administrative matters. Course objectives. A brief historical overview of privacy. General privacy regulatory frameworks in the United States and the European Union. Fair information practices by the Federal Trade Commission. Examples of privacy attacks. The privacy landscape. Why isn't encryption enough? The inference threat.

Assignments and Submissions:

- Homework 1 assigned.
- Lab 1 assigned.

Week 2

Readings:

- Cavoukian, A. (2012). Privacy by design: origins, meaning, and prospects for assuring privacy and trust in the information era. In *Privacy protection measures and technologies in business organizations: aspects and standards* (pp. 170-208). IGI Global.
- Langheinrich, M. (2001). Privacy by design—principles of privacy-aware ubiquitous systems. In *International conference on Ubiquitous Computing* (pp. 273-291). Springer, Berlin, Heidelberg.

- Schaar, P. (2010). Privacy by design. *Identity in the Information Society*, *3*(2), 267-274.
- From: Feller, W. (1968). *An introduction to probability theory and its applications, volume 1 (3rd ed.)*. New York: Wiley.
    - You may find the section, Note on the Use of the Book, useful.
    - Introduction.
    - Chapter I, The Sample Space.
    - Chapter V, Conditional Probability; Stochastic Independence, Sections 1–4.

Summary:

Introduction to an engineering-oriented general-purpose privacy framework: privacy by design. The principles that drive privacy by design: proactive not reactive; privacy as the default setting; privacy embedded into design; positive sum (full functionality); end-to-end (life cycle) security; visibility and transparency; and respect for user privacy (user-centric design).

Review of probability theory, statistical theory, and information theory principles. Axioms of probability. Events and probability spaces. Statistical independence. Conditional probability. Bayes' law.

Assignments and Submissions:

- Homework 1 due.
- Homework 2 assigned.

Week 3

Readings:

- From: Feller, W. (1968). *An introduction to probability theory and its applications, volume 1 (3rd ed.)*. New York: Wiley.
    - Chapter IX, Random Variables; Expectation, Sections 1–5 and 9.
    - Chapter VI, The Binomial and the Poisson Distributions.

Summary:

Review of probability theory, statistical theory, and information theory principles. Axioms of probability. Events and probability spaces. Statistical independence. Conditional probability. Bayes' law. Random variables. Probability distributions. Expectation. Variance. Entropy. Supervised learning. Unsupervised learning. Cross-validation. Confusion matrix.

Assignments and Submissions:

- Homework 2 due.
- Homework 3 assigned.

Week 4

Readings:

- Warner, S. L. (1965). Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, *60*(309), 63-69.

Summary:

Reviewing the beginning of the study of privacy in statistics: randomized response. Personal-attribute-disclosure protection while maintaining utility for population estimates. Background and problem statement. Proposed technique. Formal analysis. Extracting the maximum likelihood estimate. Estimate confidence intervals. Examples: sensitivity of estimates and designing a survey with confidence interval requirements.

Assignments and Submissions:

- Homework 3 due.
- Homework 4 assigned.
- Forming project groups.

Week 5

Readings:

- Sweeney, L. (2002). *k*-anonymity: A model for protecting privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, *10*(05), 557-570.

Summary:

Anonymization: Protecting identity disclosure through *k*-anonymity. Background and motivation. Example linking attack. Defining the *k*-anonymity privacy model. Good practices: randomized order, complementary releases, and temporal dependencies.

Assignments and Submissions:

- Homework 4 due.
- Homework 5 assigned.
- Lab 1 due.
- Lab 2 assigned.

Week 6

Readings:

- Machanavajjhala, A., Gehrke, J., Kifer, D., & Venkitasubramaniam, M. (2006). *l*-Diversity: Privacy Beyond *k*-Anonymity. In *22nd International Conference on Data Engineering (ICDE'06)(ICDE)* (p. 24). IEEE.

Summary:

More on anonymization: on the relationship between identity disclosure and attribute disclosure. Protecting attribute disclosure through *l*-diversity. Motivation and intuition. Attacks on *k*-anonymity. Formal treatment of attribute disclosure. The *l*-diversity privacy model. Recursive *l*-diversity. Entropy *l*-diversity. Properties of *l*-diversity.

Assignments and Submissions:

- Homework 5 due.
- Homework 6 assigned.

Week 7

Readings:

- Li, N., Li, T., & Venkatasubramanian, S. (2007). *t*-closeness: Privacy beyond *k*-anonymity and *l*-diversity. In *Data Engineering, 2007. ICDE 2007. IEEE 23rd International Conference on* (pp. 106-115). IEEE.

Summary:

More on anonymization: problems with *l*-diversity. Protection of attribute disclosure through *t*-closeness. Motivation, intuition, and comparison to *l*-diversity. Definition of *t*-closeness. Metrics. Definition of earth mover's distance (EMD). Calculating EMD for numerical and categorical attributes.

Assignments and Submissions:

- Homework 6 due.
- Homework 7 assigned.
- Project proposals due.

Week 8

Readings:

- LeFevre, K., DeWitt, D. J., & Ramakrishnan, R. (2005). Incognito: Efficient full-domain *k*-anonymity. In *Proceedings of the 2005 ACM SIGMOD international conference on Management of data* (pp. 49-60). ACM.

Summary:

Implementing algorithms that achieve the anonymization privacy models discussed so far. Suppression and generalization. Global and local recoding. Generalization hierarchies. The importance of the generalization and subset properties. Describing the algorithm Incognito. Proving correctness of Incognito. Analysis of Incognito. Demonstration of Incognito.

Assignments and Submissions:

- Homework 7 due.
- Homework 8 assigned.

Week 9

Readings:

- Bayardo, R. J., & Agrawal, R. (2005). Data privacy through optimal *k*-anonymization. In *Data Engineering, 2005. ICDE 2005. Proceedings. 21st International Conference on* (pp. 217-228). IEEE.
- LeFevre, K., DeWitt, D. J., & Ramakrishnan, R. (2006). Mondrian multidimensional *k*-anonymity. In *Data Engineering, 2006. ICDE'06. Proceedings of the 22nd International Conference on* (pp. 25-25). IEEE.

Summary:

Implementing algorithms that achieve the anonymization privacy models discussed so far. Pruning generalizations. Value reordering. Single- and multi-dimensional cuts. Describing the algorithms *k*-Optimize and Mondrian. Proving correctness of *k*-Optimize and Mondrian. Analysis of *k*-Optimize and Mondrian. Demonstration of *k*-Optimize and Mondrian. Comparing Incognito, *k*-Optimize and Mondrian.

Assignments and Submissions:

- Homework 8 due.
- Homework 9 assigned.
- Lab 2 due.
- Lab 3 assigned.

Week 10

Readings:

- Nergiz, M. E., Atzori, M., & Clifton, C. (2007). Hiding the presence of individuals from shared databases. In *Proceedings of the 2007 ACM SIGMOD international conference on Management of data* (pp. 665-676). ACM.

Summary:

On confidentiality: comparing to anonymity. Limitations of *k*-anonymity, *l*-diversity, and *t*-closeness for confidentiality. Protecting confidentiality through delta-presence. Motivation. Intuition. Defining the privacy model. Properties of delta-presence. Describing two algorithms to achieve delta-presence: SPALM and MPALM. Proving correctness of SPALM and MPALM. Analyzing SPALM and MPALM. Experiments.

Assignments and Submissions:

- Homework 9 due.
- Midterm exam assigned.

Week 11

Readings:

- Dwork, C. (2006). Differential privacy. In M. Bugliesi et al. (Eds.), *Automata, Languages and Programming.* 33rd International Colloquium. Venice, Italy, July 10–14. *Proceedings*, 4051.
- Dwork, C. (2008). Differential privacy: A survey of results. In *International Conference on Theory and Applications of Models of Computation* (pp. 1-19). Springer, Berlin, Heidelberg.
- Chapters 1–3 from: Dwork, C., & Roth, A. (2014). The algorithmic foundations of differential privacy. *Foundations and Trends® in Theoretical Computer Science*, *9*(3–4), 211-407.

Summary:

Statistical learning threats to data privacy. The utility–privacy trade-off. Problems with pure anonymization. Impossibility of absolute disclosure prevention in statistical databases. Differential privacy: motivation, intuition, definition, and properties. Achieving differential privacy: the randomized response mechanism, the Laplace mechanism, the exponential mechanism.

Assignments and Submissions:

- Midterm exam due.

- Homework 10 assigned.

Week 12

Readings:

- du Pin Calmon, F., & Fawaz, N. (2012). Privacy against statistical inference. In *Communication, Control, and Computing (Allerton), 2012 50th Annual Allerton Conference on* (pp. 1401-1408). IEEE.
- Salamatian, S., Zhang, A., du Pin Calmon, F., Bhamidipati, S., Fawaz, N., Kveton, B., ... & Taft, N. (2013). How to hide the elephant-or the donkey-in the room: Practical privacy against statistical inference for large data. In *GlobalSIP* (pp. 269-272).

Summary:

The model of an honest but curious agent. Extracting utility while protecting against disclosure of undisclosed attributes. The concepts of distortion and cost. Hamming and Euclidean distortions. Log loss cost function: definition and properties. Convex representations of the privacy model. Introducing the “information privacy” model. Comparison to differential privacy. Presenting a practical algorithm to implement the framework through quantization. Experiments.

Assignments and Submissions:

- Homework 10 due.
- Homework 11 assigned.

Week 13

Readings:

- Section 5 of the following book chapter: Aranki, D., Kurillo, G., & Bajcsy, R. (2017). Smartphone Based Real-Time Health Monitoring and Intervention. In *Handbook of Large-Scale Distributed Computing in Smart Healthcare* (pp. 473-514). Springer, Cham.
- Aranki, D., & Bajcsy, R. (2015). Private Disclosure of Information in Health Tele-monitoring. *arXiv preprint arXiv:1504.07313*.

Summary:

Private disclosure of information. Man-in-the-middle (eavesdropper) inference attack threat model. Motivation from health care applications (telehealth). Why isn’t encryption enough? Definition of the private disclosure of information (PDI) privacy model. Intuition of privacy protection through PDI. Analysis of the model. Absolute disclosure prevention (perfect privacy). Experiment.

Assignments and Submissions:

- Homework 11 due.
- Homework 12 assigned.
- Lab 3 due.

Week 14

Readings:

- No readings assigned for this week.

Summary:

Summary of the course. General discussion and reflections. Directions of open research and development. Project presentations.

Assignments and Submissions:

- Homework 12 due.
- Projects due:
    - ~10-minute presentation
    - Project report that includes problem statement, survey of related work, methodology, experimentation, results, and discussion