Syllabus and course outline for Security and Privacy for Big Data, Fall 2017.

Instruction to select the course: ECEN 4010/5060 by Prof. Yanmin Gong

Syllabus - ECEN 5060: Machine Learning & Information Security and Privacy — Fall 2017
Instructor: Dr. Yanmin Gong
E-mail: yanmin.gong@okstate.edu
Credits: 3
Class Hours: TR 3:30 — 4:45 pm, CLB 101
Office Hours: Tuesday, Thursday 2:30 — 3:30 pm or by appointment

Course Objective. The purpose of this course is to introduce you to the emerging trends and issues in security and privacy in current web and mobile services. You will learn how to identify security and privacy issues to be tackled, how to choose and apply appropriate approaches, tools and technology, and how to evaluate the security and privacy of a solution. The course begins with an introduction to current security and privacy challenges in big data applications and continues by covering a number of seminal papers in a wide range of security areas.

Prerequisites: The course is open to interested graduate and senior undergraduate students with proper backgrounds in probability and programming experience (no requirement on what language you want to use).

Textbook: This course does not have any official textbook. Most of the course readings will come from seminal papers in the field. Recommended readings: 1. Introduction to Modern Cryptography. 2. Algorithmic Foundations of Differential Privacy.

Grading:
Homework 20%, Exam 50%, Project Proposal 10%, Project Presentation 10%, Project Report 10%, Class Participation 10%
100-90 A, 89-80 B, 79-70 C, 69-60 D, below 60 F
Course Outline (tentative)

Week #1 - Introduction and overview

- Topics: Syllabus. Security and privacy challenges in big data applications.

Week #2 – Cryptography and Private Key (Symmetric) Cryptography

- Topics: Random Oracle model, secret key crypto, Message Authentication Codes (MAC’s), hash function.

Week #3 – Public-Key (Asymmetric) Cryptography

- Topics: Key management, authentication, secret sharing, commitment, zero-knowledge proof.

Week #4 – Advanced topics in cryptography

- Topics: Zero-knowledge proof, homomorphic encryption, secure multi-party computation, witness encryption, functional encryption, identity-based encryption

Week #5 – Anonymity and attacks against anonymity

- Topics: k anonymity, adversarial background knowledge, simulatability, l-diversity, t-closeness.

Week #6 - Differential privacy

- Topics: Differential privacy, Laplace Mechanism and Composability.
- Reading 1: "Calibrating Noise to Sensitivity in Private Date Analysis" C. Dwork, F. McSherry, K. Nissim, A.Smith TCC 2006
- Reading 2: "Composition Attacks and Auxiliary Information in Data Privacy" S. Ganta, S. Kasiviswanathan, A. Smith KDD 2008

Week #7 - Differential privacy (cont.)

- Topics: Exponential and median mechanisms.
- Reading: "Mechanism design via differential privacy", F. McSherry, K. Talwar FOCS 2008
Week #8 - Differential Privacy (cont.)

- Topics: Smooth Sensitivity and Sample Aggregate Framework.
- Reading: "Smooth Sensitivity and Sampling in Private Data Analysis", K. Nissim, S. Raskhodnikova, A. Smith STOC 2007

Week #9 - Keeping track of rare events

- Topics: The sparse vector technique and private query release.
- Reading: "The algorithmic foundations of Differential Privacy", Roth, Dwork (sections 3.6, 4.2)

Week #10 - Real world use cases and applications

- Topics: Privacy in graphs, releasing census data.
- Reading 1: "Publishing Graph Degree Distribution with Node Differential Privacy", Wei-Yen Day et al, SIGMOD 2016
- Reading 3: "Formal privacy protection for data products combining individual and employer frames" Haney et al

Week #11 - Variations of differential privacy

- Topics: Location Privacy and geo-indistinguishability, pan-privacy, membership privacy, Pufferfish Privacy, crowd-blending privacy.
- Reading 1: "Pufferfish: A Framework for Mathematical Privacy Definitions", D. Kifer, A. Machanavajjhala TODS 2013 (sections 7,8,10 optional)

Week #12 - Privacy and machine learning

- Topics: Classification and Regression under Differential Privacy.
- Reading 2: "Deep learning with Differential Privacy", Abadi et al

Week #13 - Privacy and machine learning (cont.)

- Topics: Model Inversion: Privacy attacks on Learning.
- Reading 3: "Stealing machine learning models via prediction apis" Tamer, Usenix Security 2016

**Week #14 - Efficiency, utility, privacy trade-offs**

- Topics: Data analytics over encrypted data.
  - Reading 1: "Private Data Analytics on Biomedical Sensing Data via Distributed Computation", Gong et al TCBB 2016

**Week #15 - Course wrap-up and review**

- Topics: Overflow from previous topics and review for comprehensive final exam.

**Week #16 - Final Exam week**

- Events: Comprehensive final exam.