Private Data Analysis

Crypto Internship Proposal

Snips ..

Spring 2017

Motivation

Many companies and organisations have realised the large potential of data analysis, and the widespread use of mobile phones have made it easier than ever to extract and record highly relevant user data, based for instance on location and behavioural patterns. For example, using the GPS unit in modern mobile phones, app companies can easily build local datasets on phones which correlate the profile of the user with the places he or she visits; then, by aggregating these datasets, the company can make place recommendations to other users with similar profiles. However, due to the sensitive nature of such data, users are reasonably reluctant to share it out of a fear of exposure.

Fortunately, modern cryptography has several solutions aiming at solving this problem, including *secure computation*. However, the application of these to big data problems seen in for instance start-up companies is still somewhat limited, and for efficiency would typically require a degree of adaptation of the techniques from both worlds.

Objectives

- Explore intersection between secure computation and big data analysis, with a focus on efficient algorithms for analysing real-world datasets (including identifying relevant and realistic use-cases)
- Implement and evaluate prototypes using existing frameworks and techniques
- Document findings through blog post, technical report, or research paper

Skills

- Suitable for Math/CS/Engineering student with interest in cryptography and machine learning
- Prior experience in big data a plus (including sketching data structures and streaming algorithms)
- Prior programming experience a plus (expect to do some prototyping)

Format

- 4-6 month full-time internship position at Snips in collaboration with university professor
- Suitable for pursuing engineering career or research in machine learning/applied cryptography

Company

Snips is young start-up located in the centre of Paris. Our primary focus is on data science and machine learning, but since we encourage the principle of *Privacy by Design*, we also have a growing team working on privacy enhancing technologies, such as secure computation, homomorphic encryption, and differential privacy. We are currently around 40 employees, including data scientists, software engineers, designers, and cryptographers.