UNIVERSITY of GUELPH

COLLEGE of ENGINEERING AND PHYSICAL SCIENCES

SCHOOL OF COMPUTER SCIENCE

Qualifying Examination

Tuesday July 26, 2022 at 9:30am via Zoom

Wenjing Zhang

Information-theoretic Privacy-utility Trade-offs Enhancement for Sequential Data Release

Chair: Dr. Fei Song
Advisor: Dr. Xiaodong Lin
Advisory: Dr. Lei Lei [SoE]
Non-Advisory: Dr. Charlie Obimbo
Non-Advisory: Dr. Ahmed Refaey Hussein [SoE]

Abstract:

Sequential data, such as mobility traces, DNA sequences, web browsing histories, and sensor data, is widely used in numerous real-world applications, including location-based services and genome usage studies. However, direct analysis of the original sequential data can cause severe privacy leakage in people's sensitive information, such as personal habits, medication history, and social relationships. Therefore, it is essential to protect an individual's sequential data against inference attacks from a malicious adversary by perturbing the original data while maintaining certain data utility. The objective of this research proposal is to explore the privacy-utility trade-offs for analyzing two types of sequential data, i.e., mobility data and genome data, from an information-theoretic perspective.

In particular, we present rigorous privacy definitions for sequential data based on information-theoretic metrics, formulate and solve the optimization problems to find the optimal privacy-utility trade-offs, and derive efficient privacy-preserving data release mechanisms via machine learning techniques. The optimal trade-offs are obtained through a deep reinforcement learning approach with known data priors and derived based on a data-driven approach with unknown data priors due to only limited data samples being accessible.