香港浸會大學 HONG KONG BAPTIST UNIVERSITY

DEPARTMENT OF COMPUTER SCIENCE 計算機科學系

DEPARTMENT OF COMPUTER SCIENCE

SEMINAR

2024 SERIES

Fill a Gap in Differential Privacy: A Binary Data Scenario

DATE & TIME

22 MAR 2024 (FRI) 9:30 – 10:30 AM

ONLINE VIA ZOOM

DR. PAN LI

Associate Professor
Department of Electrical, Computer, and Systems Engineering
Case Western Reserve University

ABSTRACT

Data sharing and releasing are undoubtedly critical for building a data-driven future. However, data sovereignty, regulations, and privacy concerns may prevent data holders from sharing their data, and hence significantly hinder the development of data-driven applications. Since introduced in 2006, differential privacy (DP) has been employed as a de facto standard for privacy-preserving data extraction from statistical databases. A few classic DP mechanisms have been proposed for continuous- and scalar-data queries. However, how to query binary- and matrix-valued data from a database in a differentially private manner has been a very important yet challenging problem and rarely studied, resulting in a big gap in DP theories. Particularly, in many real-world applications, binary- and matrix-valued data are ubiquitous and indispensable, e.g., structures of social networks, "yes" or "no" answers in questionnaires, and point mutations in genomic data, where privacy concerns may arise under a variety of circumstances.

In this talk, I will introduce a new exclusive or (XOR) mechanism utilizing a matrix-valued Bernoulli distribution that addresses the DP of binary- and matrix-valued data and fills this gap. I will theoretically analyse the privacy and utility guarantee of the proposed mechanism. Then, I will discuss how to calibrate the parameters in the matrix-valued Bernoulli distribution. Additionally, I will describe an Exact Hamiltonian Monte Carlo based sampling scheme to efficiently generate samples from this distribution. At the end, I will present experiment results in practical case studies demonstrating that the XOR mechanism notably outperforms state-of-the-art DP methods in terms of utility (e.g., classification accuracy), and even achieves comparable utility to the non-private mechanisms.

SPEAKER'S BIOGRAPHY

REGISTER NOW

Enquiries: 3411-2385 **Email:** comp@comp.hkbu.edu.hk **Website:** https://bit.ly/bucs-events