Statistical Sciences
UNIVERSITY OF TORONTO

SEMINAR

October 23, 2014 at 3:30pm
Sidney Smith Hall, Room 1083
Refreshments will be served at 3:15pm

Private analysis of social networks: Differentially private β-model and synthetic graphs

Aleksandra Slavkovic, Pennsylvania State University

Increasing volumes of personal and sensitive data are collected and archived by health networks, government agencies, search engines, social networking websites, and other organizations. The social networks are a prominent source of data for researchers in economics, epidemiology and sociology and have sparked a flurry of research in statistical methodology for network analysis. While the social benefits of analyzing these data are significant, their release can be devastating to the privacy of individuals and organizations. In this talk, we give a brief overview of challenges associated with protecting confidential data, and the problem of releasing summary statistics of graphs needed to build statistical models for networks while preserving privacy of individual relations. Our goal is to ensure that a user is able to perform standard statistical inference with the released data while satisfying a rigorous definition of privacy risk. We present an algorithm for releasing degree sequences of undirected graphs under the framework of differential privacy. The algorithm is designed to provide utility for inference in random graph models whose sufficient statistics are functions of degree sequences. Specifically, we focus on the tasks of existence of maximum likelihood estimates, parameter estimation and goodness-of-fit testing for the beta model of random graphs. We show the usefulness of our algorithm by evaluating it empirically on simulated and real-life datasets. As the released degree sequence is graphical, our algorithm can also be used to release synthetic graphs under the beta model. Time permitting, we will discuss extensions to more general exponential random graph models (ERGMs).

(This is joint work with Vishesh Karwa)