Econometrics Seminar
May 28, 2024

Shujie Ma
UCR, Department of Statistics

Title: Privacy-Preserving Community Detection for Locally Distributed Multiple Networks

Abstract: In this talk, I will introduce a new efficient and scalable consensus community detection approach and distributed learning algorithm in a multi-layer stochastic block model using locally stored network data with privacy-preserving. Specifically, we develop a spectral clustering-based algorithm named ppDSC. To reduce the bias incurred by the randomized response (RR) mechanism for achieving differential privacy, we develop a two-step bias adjustment procedure. To reduce the communication cost encountered in distributed learning, we perform the eigen-decomposition locally and then aggregate the local eigenvectors using an orthogonal Procrustes transformation. We establish a novel bound on the misclassification rate of ppDSC. The new bound reveals the asymmetric roles of the two edge-flipping probabilities of the RR in the misclassification rate. Through the bound, we can also find the optimal choices for the flipping probabilities given a fixed privacy budget. Moreover, we show that ppDSC enjoys the same statistical error rate as its centralized counterpart, when the number of machines satisfies a polynomial order with the sample size on each local machine and the effective heterogeneity is well controlled.