The Hong Kong University of Science and Technology

Department of Mathematics

PhD THESIS EXAMINATION

Network Data: from analysis to applications

By

Mr. Ting LI

ABSTRACT

In this thesis, we mainly consider two problems about network data.

First, we propose a novel statistic of networks, the normalized clustering coefficient, which is a modified version of the clustering coefficient that is robust to network size, network density and degree heterogeneity under different network generative models. In particular, under the degree corrected block model (DCBM), the "in-out-ratio" could be inferred from the normalized clustering coefficient. Asymptotic properties of the proposed indicator are studied under three popular network generative models. The normalized clustering coefficient can also be used for networks clustering, network sampling as well as dynamic network analysis. Simulations and real data analysis are carried out to demonstrate these applications.

Second, we propose a new algorithm, called weighted inverse Laplacian (WIL), for predicting labels in partially labeled networks. It is a traditional topic to do community detection in networks. However, it is less discussed how to get more accurate predictions if some of the community labels are observed. The idea comes from the first hitting time in random walk, and it also has nice explanations both in information propagation and the regularization framework. By combining two different kinds of normalization, WIL is more flexible and has more tolerance of community imbalance and degree heterogeneity. We also propose a partially labeled degree-corrected block model (pDCBM) to describe the generation of partially labeled networks. We show that WIL ensures the misclassification rate goes to 0 as the number of nodes goes to infinity, and that it can handle situations with greater imbalance than traditional Laplacian methods. WIL outperforms other state-of-the-art methods in most of our simulations and real datasets, especially in unbalanced networks and heterogeneous networks.

Date: 11 July 2019, Thursday

Time: 3:00 p.m.

Venue: Room 3494 (near lifts 25-26)

Thesis Examination Committee:

Chairman	:	**Prof. Theodore Henry King CLARK, ISOM/HKUST**
Thesis Supervisor	:	**Prof. Bingyi JING, MATH/HKUST**
Member	:	**Prof. Can YANG, MATH/HKUST**
Member	:	**Prof. Man Yu WONG, MATH/HKUST**
Member	:	**Prof. Ning CAI, IEDA/HKUST**
External Examiner	:	**Prof. Guangming PAN Division of Mathematical Sciences/Nanyang Technological University**

(Open to all faculty and students)

The student's thesis is now being displayed on the reception counter in the General Administration Office (Room 3461).