

SCHOOL OF DATA SCIENCE

SEMINAR SERIES

Are there any community structures in a hypergraph?

Date: 27 February 2019 (Wednesday)

Time: 10:00am to 11:00am

Venue: G5-317 (Lift 2), 5/F, Yeung Kin Man Academic Building (YEUNG),
City University of Hong Kong

Dr. Feng, Yang
Associate Professor of Statistics
Columbia University



Guest Speaker's profile

Yang Feng is an associate professor of statistics at Columbia University. In 2010, he got his Ph.D. in Operations Research & Financial Engineering from Princeton University under the supervision of Professor Jianqing Fan. His current research interest includes network models, high-dimensional statistical learning, and nonparametric methods. He is an associate editor for Journal of Business & Economic Statistics, Statistica Sinica, and Statistical Analysis and Data Mining. His research is partially supported by NSF CAREER Grant DMS-1554804. For more information, please visit <http://www.stat.columbia.edu/~yangfeng/>.

Abstract

Many complex networks in the real world can be formulated as hypergraphs where community detection has been widely used. However, the fundamental question of whether communities exist or not in an observed hypergraph remains unclear. The aim of the work is to tackle this important problem. Specifically, we systematically study when a hypergraph with community structure can be successfully distinguished from its Erdos-Renyi counterpart, and propose concrete test statistics based on hypergraph cycles when the models are distinguishable. For uniform hypergraphs, we show that the success of hypergraph testing highly depends on the order of the average degree as well as the signal to noise ratio. In addition, we obtain asymptotic distributions of the proposed test statistics and analyze their power. Our results are further extended to nonuniform hypergraphs in which a new test involving both edge and hyperedge information is proposed. The novel aspect of our new test is that it is more powerful than the classic test involving only edge information. Simulation and real data analysis support our theoretical findings.