

Two Sample Test of Networks - the Connection Between Manifold Learning and Statistics

Date: 10 March 2021 (Wednesday)

Time: 11:00am - 12:00pm



Seminar link: <https://cityu.zoom.us/j/97935827375>

ABSTRACT

Network data is a major type of object data that has been widely collected or derived from numerous sources such as brain imaging. Such data may contain numeric, topological, and geometrical information, and are thus necessarily considered on manifold for appropriate machine learning and statistical analysis. The development of statistical methodologies for network data is challenging and currently at its very early stage; for instance, the non-Euclidean counterpart of basic two-sample tests for network data is scarce in literature. In this study, a novel framework is presented for two independent sample comparison of networks. Specifically, an approximation distance metric to quotient Euclidean distance is proposed, and then combined with network spectral distance to quantify the local and global dissimilarity of networks simultaneously. A permutational non-Euclidean analysis of variance is adapted to the proposed distance metric for the comparison of two independent groups of networks. Comprehensive simulation studies and real applications are conducted to demonstrate the superior performance of our method over other alternatives. The asymptotic properties of the proposed test are investigated and its high-dimensional extension is discussed as well.

Enquiries: hkids@cityu.edu.hk



Dr Hongyu MIAO GUEST SPEAKER'S PROFILE

Dr Hongyu Miao received his Bachelor (1999) and MS degrees (2002) in Engineering Mechanics from Tsinghua University, and PhD degree in Mechanical Engineering (2007) and MS degree in Biostatistics (2011) from University of Rochester. He joined the Department of Biostatistics and Computational Biology at University of Rochester in 2006. Dr. Miao is currently tenured Associate Professor in the Department of Biostatistics and Data Science at the University of Texas Health Science Center at Houston, School of Public Health. His research interests include statistical learning, network analysis, functional data, and big complex data with applications in clinical trials, connected health, systems biology, infectious diseases, and neural development and disorders. He has published 70+ peer-reviewed journal articles, including the most prestigious statistical and machine learning journals like JASA, Annals of Statistics, SIAM Reviews, and IEEE Trans. PAMI. He led and has been leading multiple NIH/NSF funded projects, served on multiple international/national conference committees, NIH/NSF grant review panels, and academic journal's editorial boards. He is also the Secretary of Houston Chapter of the American Statistical Association (HACASA), and currently serves as the Director of Data Science Education Program as well as the Director of Center for Biostatistics Collaboration and Data services at UT School of Public Health. He has previously (co-)mentored 5 postdoctoral trainees and 20+ graduate students.

All are welcome