



INDIAN STATISTICAL INSTITUTE

Theoretical Statistics and Mathematics Unit, Kolkata

SEMINAR

Date: April 29, 2024

Time: 04:15 PM

VENUE:

L- infinity

(5th Floor, A.N. Kolmogorov Bhavan), ISI Kolkata

TITLE:

**Network Cross-Validation and Model Selection via
Subsampling**

SPEAKER:

Sayan Chakrabarty

University of Illinois, Urbana-Champaign

ABSTRACT:

Complex and larger networks are becoming increasingly prevalent in scientific applications in various domains. Although a large number of network models and methodology exist for such networks, cross-validation on networks remains challenging due to the unique structure of network data. The current paper proposes a general cross-validation procedure based on subsampling for networks. The proposed algorithm divides the original network in multiple subnetworks with a shared overlap part, producing a training set consisting of the subnetworks and a test set with the node pairs between the subnetworks. This train-test split provides the basis for a network cross-validation procedure that can be applied on a wide range of model selection and parameter tuning problems for network data. The method is computationally efficient for large networks as it uses smaller subnetworks for the training step. The theoretical justification of the method is shown for model selection in blockmodels and selecting the latent dimension of random dot product graphs. Numerical results indicate that the proposed algorithm accurately performs model selection and parameter tuning on many simulated and real networks from a diverse set of models. They also suggest that the method is computationally faster than existing methods for network cross-validation.

ALL ARE CORDIALLY INVITED