

When: Friday 15:00 – 16:00

Where: ETB 1003

Speaker: **Prof. Anirban Bhattacharya**

Assistant Professor
Department of Statistics
Texas A&M University



Title: **Community Detection in Networks: A Probabilistic Approach**

Date: **02-17-2017**

Abstract: A fundamental problem in network analysis is clustering the nodes into groups that share a similar connectivity pattern. Existing algorithms for community detection assume the knowledge of the number of clusters or estimate it *a priori* using various selection criteria and subsequently estimate the community structure. Ignoring the uncertainty in the first stage may lead to erroneous clustering, particularly when the community structure is vague. We instead propose a coherent probabilistic framework (MFM-SBM) for simultaneous estimation of the number of communities and the community structure. An efficient Markov chain Monte Carlo (MCMC) algorithm is proposed which obviates the need to perform reversible jump MCMC on the number of clusters. The methodology is shown to outperform recently developed community detection algorithms in a variety of synthetic data examples and in benchmark real-datasets. We derive non-asymptotic bounds on the marginal posterior probability of the true configuration, and subsequently use it to prove a clustering consistency result. The methodology is demonstrated on discovering activation centers in brain connectivity networks obtained from the human connectome project.

(Joint work with Junxian Geng and Debdeep Pati)

Biography: Bhattacharya holds a Ph.D. in statistics from Duke University. He is an assistant professor at Texas A&M Statistics. His research interests include: latent variable models for multivariate categorical and count data, Bayesian variable selection in linear and non-linear models, probabilistic models for analysis of network data, trade-offs between computational and theoretical complexity in big data problems, shrinkage priors in high dimensions.