



Bayesian Inference of Interactions in Biological Problems

Jing Maria Zhang

Dissertation Advisor: Professor Jun S. Liu

Recent development of bio-technologies such as microarrays and high-throughput sequencing has greatly accelerated the pace of genetics experimentation and discoveries. As a result, large amounts of high-dimensional genomic data are available in population genetics and medical genetics. With millions of biomarkers, it is a very challenging problem to search for the disease-associated or treatment-associated markers, and infer the complicated interaction (correlation) patterns among these markers.

In this talk, I will address Bayesian inference of interactions in two biological research areas: whole-genome association studies of common diseases, and HIV drug resistance studies, and a new recursive Bayesian algorithm for BN structure.

We have developed a Bayesian model for simultaneously inferring haplotype-blocks and selecting SNPs within blocks that are associated with the disease, either individually, or through epistatic interactions with others. Simulation results show that this approach is uniformly more powerful than other epistasis mapping methods.

We have investigated the HIV drug resistance problem from a new perspective. By probabilistically modeling mutations in the HIV-1 proteases isolated from drug-treated patients, we have derived a statistical procedure that first detects potentially complicated mutation combinations and then infers detailed interacting structures of these mutations.

Finally, the idea of recursively exploring the dependence structure of interactions in the above two research studies can be generalized to infer the structure of graph models.