Networks and Models
for the Integrated Analysis of Multi Omics data

Sun Kim

Abstract:
These days, genome-wide measurements of genetic and epigenetics events, a.k.a omics data, are routinely produced; epigenetics is control mechanisms of genetics events as epi- means ‘on’ or ‘upon’. As a result, a huge amount of omics data measured from different genetic and epigenetic events are available. For example, the amount of data at The Cancer Genome Atlas (TCGA) alone exceeds 2.5 peta byte as of October 2016. Unfortunately, the dimensions of omics data is huge, typically tens to hundreds or even millions of thousands while the number of samples are limited typically a few to thousands. Thus mining genetic and epigenetic data measured in different phenotype conditions is a very challenging problem, that is, small data sets on extremely high dimensions. Furthermore, all genetic and epigenetic events are inter-related. Thus it is necessary to perform integrated analysis of omics data sets of different types, which is even more challenging. To address these technical challenges, the bioinformatics community has used virtually all known network based analysis techniques, including recently developed deep neural networks. My group has been trying the network based integrated analysis of omics data at three different levels. First, we have been investigating on computational methods for associating different genetic and epigenetic events, which can be viewed as methods for defining edges in the network. Second, we have been developing mining sub-networks on the phenotype and time dimensions. Third, we have recently begun to investigate on the use of deep learning techniques for the integrated analysis of omics data. An important goal of our research is to combine network analysis and deep learning techniques to construct models or draw maps of cancer cells at multiple levels such as genomic mutations, gene activation/suppressions, epigenetic events including DNA methylation, histone modifications, and miRNA interference, biological pathways, and finally at the whole cell level including tumor heterogeneity and clonal evolution.

Short Bio:
Dr. Sun Kim is a professor and the director of Bioinformatics Institute of Seoul National University. Visit https://cci.drexel.edu/ieeebibm/bibm2016/SunKim_CV_Abstract.pdf for his short CV.