

UAB Grand Challenge - Concept Paper

Name of the UAB Grand Challenge: Network Biology and Predictive Modeling in the Context of Big Data

Principal Point of Contact's Name, Title, e-mail address, and phone number:

Dr. Shahid Mukhtar, Assistant Professor, smukhtar@uab.edu, 205-934-8335

Description of the problem to be addressed, including its importance to the state of Alabama and how it is generalizable to other states, the U.S. and the world

Researchers, teachers, clinicians, and health care providers at the University of Alabama at Birmingham (UAB), across the state of Alabama and in the US are generating data in an exponential fashion. For instance, a biological system constitutes a network of interconnected cellular components. The structural, physical, and functional connections among these components are termed edges, while the molecules involved in the interactions are referred to as nodes. The widespread utilization of genome-wide or “-omics” approaches that seek to integrate large-scale data such as genomes, proteomes, transcriptomes, interactomes, and others into descriptive models, helps elucidate the molecular functions at the systems. Given the integrative nature of these datasets, the question concerns here is how these omics-based datasets interact with each layer and how to interpret the outcome holistically.

Networks consist of systems' components, referred to as 'nodes', and interactions between them, termed 'edges'. The understanding of how networks function in the schema of an entire system is fueling the development of network based approaches. Such a framework is pertinent to evaluation of diverse networks including social media-fueled networks, telecommunication networks, banks transaction networks of businesses, information flow on the internet and biological networks. It is becoming increasingly evident that distinctive networks in nature and technology, from the microscale of cells to the macroscale of ecosystems and the global community landscape of social networks, often exhibit universal, and highly reproducible topological characteristics. Elucidating these physical characteristics, as well as functional interaction properties of a particular system.

Considering that diverse networks share similar organizational landscapes, and rate of information flow through a network is dependent on the connectivity of its components, several parameters of centrality measurements may act as indicators of important nodes in a system. Moreover, network analysis has been a recent focus in biological sciences due to its ability to synthesize global visualizations of cellular processes and predict functions based on inferences from network properties. In addition, mathematical modeling enhances intuitive predictions leading to prioritize candidate genes and innovative conclusions that can be subsequently addressed experimentally.

Owing to the integrative nature of systems biology and network biology, topological network analyses have been interchangeably used across diverse networks. For instance, network scientists have applied these approaches in ecology/ecosystem, physiology, neural networks, food web, human diseases, and network medicine.

Predictive modeling, machine learning deep learning can further substantiate our understanding between genotype to phenotype relationships. While qualitative phenotypes in every discipline of

biology, medicine and other units are easy to visualize and interpret, quantitative phenotypes are merging as new era of systems biology. For instance, genetic variants, Genome-Wide Association Studies (GWAS) and their associations to the phenotypes have been improved by machine learning. Linkages of seemingly unrelated diseases with similar set of genes have been only possible through network analysis as well as machine learning approaches. Thus, big data analysis is a grand challenge not only in the state of AL but in the entire world.

Desired outcomes and the conceptualization of the plan of work to achieve them

Big data analyses in the context of network biology, predictive modeling, and machine learning will allow us understand the relationships with different layers of omics-data, genotype to diseases relationships in precision medicine, discovery of important nodes in an ecosystem. Data analytics will significantly improve the healthcare industry in managing and interpreting the data. It can provide a better relationships among researchers, clinicians, physicians and patients in developing precision and personalized medicine. More importantly, social network aspect will allow perfect the business model of physician-patient relationships. This will also allow us organize the future vision.

List of potential team members (individuals and organizations)

Currently, I'm the only member of the team and would like to join a team after the first round of grand applications. As a young Systems Biologist, I have accomplished what many in my profession are striving to achieve: In sum, I have published 35 articles in high ranked journals such as SCIENCE, CELL, NATURE, Nature Communications, PNAS, Current Biology, Molecular Systems Biology, PLoS Pathogens, Cell Host Microbe, Trends in Plant Science, The Plant Journal, Plant Physiology and BMC Genomics. My published work attained an immediate recognition in the field of data science as my findings were breakthrough and served to open new avenues in Systems Biology. In addition, I'm also playing leadership roles in my discipline for the last several years. I'm looking forward to joining a team in big data analytics.