Dynamic network analysis reveals a protein's role in aging and a person's risk of mental health problems

Tijana Milenkovic, Department of Computer Science and Engineering, University of Notre Dame

Networks (or graphs) are powerful models of complex systems in various domains, from biological cells to societies to the Internet. With increasing availability of dynamic (temporal) real-world networks, how to efficiently study these data? One can model a dynamic network as a single aggregate static network, or as a series of time-specific snapshots, each being an aggregate static network over the corresponding time window. Then, one can use established methods for static analysis on the resulting aggregate network(s). However, this loses valuable temporal information either completely, or at the interface between different snapshots, respectively.

To overcome these drawbacks, we have developed novel approaches for studying a dynamic network more explicitly, by capturing inter-snapshot relationships. For example, some of our approaches are based on well-established graphlets (subgraphs, Lego-like building blocks of complex networks), which have been proven in numerous contexts in static network research. These contexts include our past measure of (static) graphlet degree centrality (GDC), which quantifies the position (centrality, importance) of a node in a (static) network based on how many graphlets of each type the node participates in. That is, the larger and denser the extended network neighborhood of a node, the more central the node. In our work that this talk will focus on, we have developed new theory to allow for graphlet-based analyses of dynamic networks. That is, our dynamic graphlets, an approach called graphlet orbit transitions (GoT), and our dynamic version of GDC allow for quantifying the change in the network position (i.e., extended neighborhood) of a node with time. We have demonstrated the superiority of the graphlet-based approaches for dynamic network analysis over existing approaches in several tasks, such as classification, alignment, or clustering of dynamic networks from various domains.

In this talk, we will demonstrate usefulness of our dynamic network analysis approaches in two tasks: studying the role of a protein in the aging process based on its position in a dynamic molecular network, and analyzing an individual's mental health based on their position in a dynamic social network, as follows.

First, incidence of many complex diseases, such as cancer, Alzheimer's disease, and even COVID-19 increases with age. Understanding the molecular mechanisms behind the aging process, including identification of human genes (i.e., their protein products) implicated in aging, is important for treating such aging-related diseases. However, wet lab experimental analyses of human aging are hard due to long human life span and ethical constraints. Computational identification (i.e., prediction) of aging-related genes via machine learning from human -omics data can fill in this gap. In this context, we have integrated aging-specific gene expression data with context-unspecific protein-protein interaction (PPI) network data to infer a dynamic aging-specific PPI subnetwork. Then, we have developed a machine learning model that when applied to the dynamic subnetwork to analyze how genes' PPIs change with age yields extremely high aging-related prediction performance, with F-score of 91.4%, while its best static counterpart yields F-score of "only" 74.3%. So, our predictive model could guide with high confidence the discovery of novel aging-related gene candidates for future wet lab validation.

Second, mental disorders such as depression and anxiety are public health issues. Early interventions can significantly reduce risk of developing mental disorders. Yet, most people do not seek treatments due to a lack of awareness of their disorders. A way to raise awareness is to develop computational approaches for predicting whether and when an individual will become at risk of a mental disorder. Innovative technologies such as wearable sensors can provide a wealth of data relevant to mental health. In this context, we have leveraged rich longitudinal data from the recent NetHealth study containing individuals' social interaction data collected via smartohones, health-related behavioral data (physical activity and sleep duration) collected via Fitbit devices, and a variety of individuals' trait data (including mental health) collected via surveys. We have modeled the NetHealth data as a dynamic network and developed a machine learning model for predicting one's likelihood of being depressed or anxious based on how their position in the dynamic network changes over time. We have found that our dynamic network-based predictive model significantly outperforms its static network counterparts as well as its fairly comparable non-network counterpart that has traditionally been used for mental health prediction.

Hulovatyy, Chen, and Milenkovic, Exploring the structure and function of temporal networks with dynamic graphlets, *Bioinformatics*, 2015.

Li and Milenkovic, Improving supervised prediction of aging-related genes via dynamic network analysis, *arXiv:2005.03659*, 2020.

Liu, Hachen, Lizardo, Poellabauer, Striegel, and Milenkovic, The power of dynamic social networks to predict individuals' mental health, in *Proceedings of the Pacific Symposium on Biocomputing (PSB)* 2020.