CoCo Seminar Series Fall 2020

Optimal Multi-Source Forecasting and Clique Relaxations in Social Networks

Dr. Zeynep Ertem
Assistant Professor, Systems Science and Industrial Engineering, Binghamton University

Wednesday November 4, 2020 11:00am-12:00pm
Online (Zoom meeting link available on http://coco.binghamton.edu/)

This talk consists of two parts. First, I will describe a new optimization-based forecasting approach applied to public health challenges. Second, I will introduce a novel clique relaxation formulation based on clustering coefficients to find cohesive subgroups in networks. Finally, I will show a detailed complexity analysis of a special case of this clique relaxation problem. Healthcare systems are one of the largest cost items for many countries. For example, in the US 17% of the GDP is spent on healthcare, yet there is abundant evidence that the system has many inefficiencies. In this talk, I will talk about how an operations research-based approach can help improve efficiency in public health leveraging electronic medical health records. Specifically, I will describe a new forecasting algorithm that uses multiple data sources to timely and accurately predict the flu epidemic in the US. This hierarchical framework uses multi-linear regression to combine forecasts from multiple data sources and greedy optimization with forward-selection to sequentially choose the most predictive combinations of data sources. We show that the systematic integration of complementary data sources can substantially improve forecast accuracy over single data sources. Using multiple data sources, this method achieves 15% more accuracy than a baseline model that only uses one data source while forecasting the Center for Disease Control and Prevention (CDC) influenza-like-illness reports. Furthermore, using this framework, I show that out of the more than 600 flu-related data sources some of the best predictions come from electronic health records. Clique relaxations are used in classical models of cohesive subgroups in social network analysis. Clustering coefficient was introduced more recently as a structural feature characterizing small-world networks. Noting that cohesive subgroups tend to have high clustering coefficients, here I introduce a new clique relaxation, α-cluster, defined by enforcing a lower bound on the clustering coefficient in the corresponding induced subgraph. I consider two variations of the clustering coefficient, namely, the local and global clustering coefficient. I analyze certain structural properties of α-clusters and I develop mathematical optimization models for determining α-clusters of the largest size in a network and apply these models to several real-life social networks. In addition, I develop a novel network clustering algorithm based on local α-clusters.

Zeynep Ertem received her Ph.D. in Industrial and Systems Engineering at Texas A&M University in 2015. She held positions at Texas A&M, UT Austin and USC before joining Binghamton in Fall 2020. Her research interests include mathematical modeling and optimization in healthcare systems, including preparedness and responses for infectious diseases. For more information, contact Hiroki Sayama (sayama@binghamton.edu). http://coco.binghamton.edu/