Mining Large-scale Graph Data
EECS 598-008 – Special Topics

Lecture: T/Th 1:30-3:00pm @ EECS 1311
Discussion: Th 3-4pm @ FXB 1012

Instructor: Danai Koutra

Overview:

Graphs naturally represent information ranging from links between webpages to friendships in social networks, to collaborations between coauthors and connections between neurons in our brains. These graphs often span billions of nodes and interactions between them. Within this deluge of interconnected data, how can we extract useful knowledge, understand the underlying processes, make interesting discoveries, and contribute to decision-making?

This course will cover recent methods and algorithms for analyzing large-scale graphs, as well as applications in various domains (e.g., neuroscience, web science, social science, computer networks). The focus will be on scalable and practical methods, and students will have the chance to analyze large-scale datasets. The topics that we will cover include clustering and community detection, recommendation systems, similarity analysis, deep learning, summarization, and anomaly detection in the graph setting.
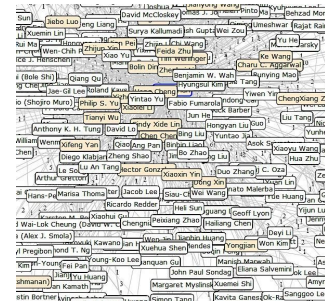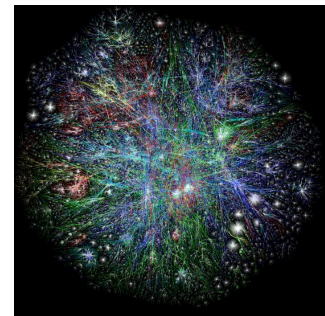
Logistics:

This course will be based on readings, and will have a mix of lectures and student presentations. We will read a selection of recent papers that have appeared in top data mining conferences and journals. The coursework will consist of 2-3 short assignments that will familiarize the students with the challenges of large-scale network analysis, as well as a semester-long group project that will be related to topics discussed in class and selected by students.

Prerequisites:

Students are expected to (1) have basic knowledge of linear algebra, (2) be familiar with probability theory/statistics, and (3) have good programming skills (e.g., Python, JAVA, Matlab, R). Some familiarity with machine learning will be helpful, but it is not required.

Website: http://web.eecs.umich.edu/~dkoutra/courses/W18_598