Computational Mathematics Colloquium

Optimization for scalable graph analytics

Kimon Fountoulakis
University of California, Berkeley

Abstract: Graphs, long popular in computer science and discrete mathematics, have received renewed interest because they provide a useful way to model many types of relational data. In biology, e.g., graphs are routinely used to generate hypotheses for experimental validation; in neuroscience, they are used to study the networks and circuits in the brain; and in social networks, they are used to find common behaviors of users. These modern graph applications require the analysis of large graphs, and this can be computationally expensive. Graph algorithms have been developed to identify and interpret small-scale local structure in large-scale data without the requirement to access all the data. These algorithms have been mainly studied in the field of theoretical computer science in which algorithms are viewed as approximation methods to combinatorial problems.

In our work, we take a step back and we analyze scalable graph clustering methods from data-driven and variational perspectives. These perspectives offer complementary points of view to the theoretical computer science perspective. In particular, we study implicit regularization properties of certain methods, we solve data-driven issues of existing methods, we explicitly show what optimization problems certain graph clustering procedures are solving, we prove that existing optimization methods have better performance and generalize to unweighted graphs, and finally we demonstrate how state-of-the-art methods can be efficiently parallelized for modern multi-core hardware.

Monday, February 5, 2018, 4:00 pm
Mathematics and Science Center: W301

Mathematics and Computer Science
Emory University