

NUMERICAL ANALYSIS AND SCIENTIFIC COMPUTING SEMINAR

Numerical and Streaming Analyses of Centrality Measures on Graphs

Eisha Nathan
Georgia Institute of Technology

Abstract: Graphs are a natural representation for modeling relationships between entities across many different fields of research. In real-world networks today, new data is constantly being produced, leading to the notion of dynamic graphs. An important query when analyzing graphs is to identify the most important vertices in a graph. Vertex importance is termed as centrality, and centrality scores can be used to provide rankings on the vertices of a graph. Specifically, I focus on Katz Centrality, a linear algebra based metric that measures the affinity between vertices in a graph as a weighted sum of the number of walks between them. Calculating Katz scores involves approximating the solution to a linear system using iterative solvers. Currently, we do not have a sense of how the numerical error from the iterative method in the approximation to the centrality vector affects the identification of the highly ranked vertices. In the first part of the talk, I bridge the two research areas of numerical analysis and network analysis by providing insight into how bounding the error between the approximate and exact solution in the numerical problem affects the correct relative ranking of vertices in the data mining problem. Typically, Katz scores are calculated through linear algebra. In the second section of the talk, I present an new alternative, agglomerative method of calculating personalized Katz scores. The algorithm presented is several orders of magnitude faster than the typical linear algebra approach and is able to return good quality scores of vertices. Finally, updating centrality scores of the vertices in a dynamic graph quickly becomes expensive to recompute from scratch every time the underlying graph is changed, as is the case in evolving networks. I conclude the talk by presenting a new algorithm for updating the Katz scores of vertices in a dynamic graph that is faster than static recomputation while maintaining good quality of the scores returned.

Friday, March 2, 2018, 2:00 pm
Mathematics and Science Center: W301

MATHEMATICS AND COMPUTER SCIENCE
EMORY UNIVERSITY