

Big Data Analytics 2013



Invited talk

Partitioning & Clustering Big Graphs

George Karypis, *University of Minnesota*

Sparse and unstructured graphs arise in many application domains including data mining, pattern recognition, social network analysis, scientific computing, parallel processing, computer graphics, operations research, optimization, databases, and VLSI design. As the capacity to collect information and to model complex systems has soared, so has the size of the graphs that are being generated and need to be processed and analyzed. This massive growth in scale, combined with an increase in the complexity and diversity of the types of analysis that need to be performed, has created the need to develop frameworks that allow for the efficient processing and analysis of extremely large graphs.

Within the research dedicated to solving these problems, graph partitioning, by enabling intelligent *divide-and-conquer* approaches for solving many graph-based problems, is an important technique for facilitating the development of effective and computationally efficient approaches for processing and analyzing large, sparse, and unstructured graphs. In addition, by enabling the discovery of well-connected subgraphs, graph partitioning has become an essential tool for detecting regularities in large unstructured graphs that can be used to gain insights into the underlying data and direct subsequent analysis.

In this talk we will present an overview of the various partitioning algorithms that exist in the METIS family of multilevel serial and parallel (hyper)graph partitioning algorithms and will present our ongoing work on extending these algorithms to classes of irregular graphs arising in Big Data analysis, to exploit multi-core architectures, and to develop effective multilevel modularity based serial & parallel clustering algorithms.