Scalable Big Graph Data Processing

KISUNG LEE, SCHOOL OF COMPUTER SCIENCE, GEORGIA INSTITUTE OF TECHNOLOGY

The application of graph data analytics is virtually unlimited because graph data are everywhere, from the friendship graphs of social networks to networks of the human brain. Even though graph data analytics is essential for gaining insight into big graphs, large-scale graph processing is complex because of its graph-specific challenges, including complicated correlations among data entities, highly skewed distribution, various graph operations, and the sheer enormity of graph data. This presentation will focus specifically on two new distributed systems for the scalable processing of big graph data. I will first present a graph system that efficiently supports graph pattern query processing (subgraph matching) by scalable graph partitioning and efficient distributed query processing. I will then describe a distributed system for iterative graph computations that can reduce memory requirements for running iterative graph algorithms while ensuring competitive performance. I will conclude the presentation by introducing a set of challenges for developing a general purpose graph analytics system that can support both efficient graph query processing and fast iterative graph computations under one unified system architecture.

www.cse.lsu.edu/research/seminars