Colloquium

Scalable Big Graph Data Processing

Kisung Lee
Georgia Institute of Technology

Abstract: 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.

Bio: Kisung Lee is a Ph.D. candidate in the School of Computer Science at Georgia Tech. His research interests lie in the intersection of big data systems and distributed computing systems. Kisung has also worked on research problems in spatial data management and social network analytics. He has been a recipient of the best paper awards of IEEE Cloud 2012 and MobiQuitous 2014. Kisung received his B.S. and M.S. degrees in computer science from KAIST and has served as a reviewer for several conferences and journals.

Thursday, February 26, 2015, 4:00 pm
Mathematics and Science Center: W303

Mathematics and Computer Science
Emory University