Computer Science Colloquium

Analysis of Massive Information Networks

Philip S. Yu
University of Illinois at Chicago

Abstract: With the ubiquity of information networks and their broad applications, there have been numerous studies on the construction, and mining of information networks in multiple disciplines, including social network analysis, World-Wide Web, database systems, data mining, machine learning, and networked communication and information systems. However, large graphs may often be disk resident, and such graphs cannot be efficiently processed. In this talk, we examine the issues of on-line analytic processing, summarization and indexing of large graphs. Specifically, the problem of connectivity in the context of massive graphs is considered. In many large communication networks, social networks and other graphs, it is desirable to determine the minimum-cut between any pair of nodes. We will discuss how a connectivity index can be developed for a massive-disk resident graph. A sampling based approach is deployed to create compressed representations of the underlying graph. Trade-off between processing efficiency and accuracy will be shown.

Bio: Philip S. Yu is a Professor in the Department of Computer Science at the University of Illinois at Chicago and also holds the Wexler Chair in Information Technology. Before joining UIC, he spent most of his career at IBM Thomas J. Watson Research Center and was manager of the Software Tools and Techniques group. Dr. Yu is a Fellow of the ACM and the IEEE. He served as the Editor-in-Chief of IEEE Transactions on Knowledge and Data Engineering (2001-2004). He is an associate editor of ACM Transactions on Knowledge Discovery from Data and also ACM Transactions of the Internet Technology. He serves on the steering committee of IEEE Int. Conference on Data Mining. He was a member of the IEEE Data Engineering steering committee. Dr. Yu received a Research Contributions Award from IEEE Intl. Conference on Data Mining in 2003. His research interests include data mining, privacy, data stream, and database systems. He has published more than 560 papers in refereed journals and conferences. He holds or has applied for more than 300 US patents. Dr. Yu was an IBM Master Inventor when at IBM.

Friday, October 30, 2009, 3:00 pm
Mathematics and Science Center: W301

Mathematics and Computer Science Emory University