



# Professor Jiawei HAN

Abel Bliss Professor of Computer Science  
University of Illinois at Urbana-Champaign



## Construction, Exploration and Mining of Information Networks

### **Abstract:**

People and informational objects are interconnected, forming gigantic, interconnected, integrated information networks. By structuring these data objects into multiple types, such networks become semi-structured heterogeneous information networks. Most real world applications that handle big data, including interconnected social media and social networks, medical information systems, online e-commerce systems, or database systems, can be structured into typed, semi-structured, heterogeneous information networks. For example, in a medical care network, objects of multiple types, such as patients, doctors, diseases, medication, and links such as visits, diagnosis, and treatments are intertwined together, providing rich information and forming heterogeneous information networks. Effective construction, exploration and analysis of large-scale heterogeneous information networks poses an interesting but critical challenge.

In this talk, we first present a set of data mining scenarios in heterogeneous social and information networks and show that mining typed, heterogeneous networks is a new and promising research frontier in data mining research. Departing from many existing network models that view data as homogeneous graphs or networks, the semi-structured heterogeneous information network model leverages the rich semantics of typed nodes and links in a network and can uncover surprisingly rich knowledge from interconnected data. This heterogeneous network modeling will lead to the discovery of a set of new principles and methodologies for mining and exploring interconnected data, such as rank-based clustering and classification, meta path-based similarity search, and meta path-based link/relationship prediction. Then we discuss our recent progress on construction of quality semi-structured heterogeneous information networks from unstructured data. We will also point out some promising research directions in this domain.

### **About the speaker:**

Jiawei Han is the Abel Bliss Professor of Computer Science at University of Illinois at Urbana-Champaign. He has been researching into data mining, information network analysis, database systems, and data warehousing, with over 600 journal and conference publications. He has chaired or served on many program committees of international conferences, including PC co-chair for KDD, SDM, and ICDM conferences, and Americas Coordinator for VLDB conferences. He also served as the founding Editor-In-Chief of ACM Transactions on Knowledge Discovery from Data and as the Director of Information Network Academic Research Center supported by U.S. Army Research Lab. He is a Fellow of ACM and Fellow of IEEE, and received 2004 ACM SIGKDD Innovations Award, 2005 IEEE Computer Society Technical Achievement Award, 2009 IEEE Computer Society Wallace McDowell Award, and 2011 Daniel C. Drucker Eminent Faculty Award at UIUC. His book "Data Mining: Concepts and Techniques" has been used popularly as a textbook worldwide.

**Date:** 19 June 2014 (Thursday)  
**Time:** 4:30 p.m. – 5:30 p.m.  
**Venue:** Room 513, William M. W. Mong Engineering Building

**\*\* ALL ARE WELCOME \*\***

