**Big Graph Data Processing**

**Professor Wook-Shin Han**  
Department of Creative IT Engineering and  
Department of Computer Science and  
Engineering  
POSTECH

---

**Abstract:**

Graphs are used to model many real objects such as biological structures, social networks, chemical compounds, ontology, and web graphs. Many real applications in computer science, bioinformatics, chemistry, physics, healthcare, and geology require efficient and effective management of graph structured data. In this talk, I will first introduce the iGraph framework (VLDB 2010, VLDB 2013) for supporting subgraph indexing and subgraph isomorphism algorithm. Subgraph isomorphism has many important applications such as motif search in bioinformatics, chemical compound search in chemistry, social network search, RDF query processing, and malware detection in computer science. Next, I will introduce a novel subgraph isomorphism algorithm called Turbo$\_${ISO} (SIGMOD 2013), which significantly outperforms the state-of-the-art methods by up to orders of magnitude. I will also introduce TurboGraph (KDD 2013), which is a graph analytics engine that I am currently developing. TurboGraph is the first truly parallel graph engine that exploits 1) full parallelism including multi-core parallelism and FlashSSD IO parallelism and 2) full overlap of CPU processing and I/O processing as much as possible. TurboGraph enables us to process billion-scale graphs on a just single PC, while many existing distributed approaches need hundreds or thousands of machines to achieve similar performance. Thus, for the first time, TurboGraph will serve as a core technology to open trillion-scale graph triangulation methods which was accepted to SIGMOD 2014.

---

**About the Speaker:**

Professor Wook-Shin Han is currently an Associate Professor in the Department of Creative IT Engineering and the Department of Computer Science and Engineering in POSTECH. Before that, he was an Associate Professor in the Department of Computer Science and Engineering in Kyungpook National University. He obtained his Ph.D. from KAIST in 2001 under supervision of Professor Kyu-Young Whang. His primary research efforts for the past 19 years have been devoted to developing new techniques in DBMS "engine research." He has developed an object-relational DBMS supporting multiple language bindings. He has also developed the tight coupling technology of DBMS with IR features. At the IBM Almaden Research Center, he has developed progressive query optimization inside the parallel DB2 as a postdoc. He also invented the new concept of "parallelizing query optimization" for faster query compilation by exploiting the multi-core architecture. Recently, he has developed a framework called iGraph for comparisons of subgraph isomorphism indexing and query processing algorithms as well as an ultra-fast graph analytics engine called TurboGraph. He published at major international journals and conferences, including SIGMOD, VLDB, SIGKDD, ICDE, WWW, IEEE Transactions on Knowledge and Data Engineering, and VLDB Journal. He regularly serves as a PC member for VLDB, SIGMOD, and ICDE. He served as an associate editor for several international journals including IEEE Transactions on Knowledge and Data Engineering. He will serve as an industrial co-chair for ICDE 2015.

---

All are welcome!  
For enquiries, please call 2859 2180 or email enquiry@cs.hku.hk

Department of Computer Science  
The University of Hong Kong