Structure Learning With Latent Variables

**SPEAKER**  
**Dr Vincent TAN Yan Fu**  
Scientist  
Institute for Infocomm Research  
and  
Adjunct Assistant Professor  
Department of Electrical and Computer Engineering  
National University of Singapore  
Singapore

**DATE**  
18 January 2013 (Friday)

**TIME**  
11:00 am - 12:00 noon  
(Refreshments will be served at 10:45 am)

**VENUE**  
CS Seminar Room, Y6405, 6th Floor  
Yellow Zone, Academic 1  
City University of Hong Kong  
83 Tat Chee Avenue  
Kowloon Tong

**ABSTRACT**

Learning the structure among a large number of variables from high-dimensional data is a fundamental task in many scientific domains. The structure between these variables can be captured diagrammatically by the use of graphical models. We discuss the learning of tree and latent tree models, the latter characterized by subset of variables that are hidden. For trees without latent variables, we derive the error exponents for learning the structure and we also find the extremal structures in terms of the error exponents. Assuming the parameters of the model are kept fixed, it is shown that among all undirected trees, the star has the smallest error exponent (meaning that it is the hardest to learn) and the Markov chain has the highest error exponent. For latent tree models, statistically consistent and low computational and sample complexity algorithms are derived to infer the latent tree structure without any knowledge of the number of hidden variables and their connections to the observed ones. Experiments on real datasets such as the monthly stock returns of the companies in the S&P 100 index reveal fascinating latent dependency structure.

**BIOGRAPHY**

Vincent Y. F. Tan did the Electrical and Information Sciences Tripos (EIST) at the University of Cambridge and obtained a B.A. and an M.Eng degree in 2005. He received a Ph.D. degree in Electrical Engineering and Computer Science from MIT in 2011. After which, he was a postdoctoral researcher at University of Wisconsin-Madison. He is now a scientist at the Institute for Infocomm Research (I2R), Singapore and an adjunct assistant professor at the Department of Electrical and Computer Engineering at the National University of Singapore. He held two summer research internships at Microsoft Research during his Ph.D. studies. His research interests include learning and inference in graphical models, statistical signal processing and network information theory. Vincent received the Charles Lamb prize, a Cambridge University Engineering Department prize awarded to the student who demonstrates the greatest proficiency in the EIST. He also received the MIT-EECS Jin-Au Kong outstanding doctoral thesis prize and the A*STAR Philip Yeo prize for outstanding achievements in research.

All are welcome!