

When: Friday 13:50 – 14:50

Where: ETB 1035

Speaker: Prof. Sing-Hoi Sze

Associate Professor
Department of Computer Science & Engineering
Department of Biochemistry & Biophysics
Texas A&M University



Title: Identifying similar transcripts in a related organism from de Bruijn graphs of RNA-Seq data, with applications to the study of salt and waterlogging tolerance in *Melilotus*

Date: 11-10-2017

Abstract: While a popular strategy in de novo transcriptome assembly algorithms is to assemble the reads by obtaining a de Bruijn graph that represents the transcriptome, an additional step is needed to obtain predicted transcripts from the de Bruijn graph. A similarity search algorithm is then applied to a related organism to obtain information about possible function of these predicted transcripts.

We observe that it is possible to obtain a more complete set of similar transcripts by starting the search from the de Bruijn graph directly. We develop a heuristic extension algorithm to identify paths in the de Bruijn graph that are similar to transcripts from the related organism. We show that our algorithm is able to recover more similar transcripts than existing algorithms, with longer transcripts and a better resolution of isoforms. We construct new RNA-Seq libraries for two *Melilotus* species, and apply our algorithm to study salt and waterlogging tolerance in these two species.

Biography: Dr. Sze is an associate professor of computer science and engineering and of biochemistry & biophysics at Texas A&M University. Before joining Texas A&M, he was a postdoctoral research associate at the University of California, San Diego. He received his Ph.D. in computer science from the University of Southern California. His current research areas include sequence alignment, motif finding, and algorithms for analyzing high-throughput sequencing data.