

CS Seminar

A Genome Compression Algorithm Supporting Manipulation

Dr Liqing Zhang
Assistant Professor
Department of Computer Science
Virginia Tech, USA

Date:
June 7, 2010
Monday
4:00 pm

Venue:
Room 328
Chow Yei Ching Building
The University of Hong Kong

Abstract:

With the advent of the thousand dollar genome, one can anticipate the need to store, communicate, and manipulate many human genomes. Data compression methods have been developed to store and communicate genomes efficiently. Unfortunately, these methods do not support efficient manipulation (e.g., subsequence retrieval) of the compressed genome. We develop a data compression scheme that achieves both efficient storage and efficient sequence manipulation. We demonstrate the practicality of the method on two databases of genomes, one for the human mitochondrion and one for the H3N2 virus. In both cases, we achieve high compression ratios and $O(\log n)$ subsequence retrieval times.

About the Speaker:

Dr. Zhang completed her Ph.D. in molecular evolution and population genetics at the University of California, Irvine. Since 2004, she has been a tenure-track faculty in the Department of Computer Science at Virginia Tech. She is also a faculty in the GBCB (Genetics, Bioinformatics and Computational Biology) graduate program. Her research interests include evolutionary and population genomics, computational biology, and bioinformatics.

All are welcome!
For enquiries, please call 2859 2180 or email enquiry@cs.hku.hk
Department of Computer Science
The University of Hong Kong

