

Computer Science Seminar Series, 2009

National Capital Region

Machine Learning for Genomic Data Analysis

Speaker: Prof. Yue Wang
Department of Electrical and Computer Engineering
Virginia Tech

Friday, November 13, 2009
1:00PM- 2:00PM, NVC 325

Abstract

After a brief review on the research work being conducted in the CBIL at Virginia Tech, this talk will address in silico correction of normal tissue contamination in analyzing gene copy number changes in cancer research. Copy number change is an important form of structural variations in human genomes. Somatic copy number alterations can cause the acquisition of oncogenes and loss of tumor suppressors in tumorigenesis. Recent development of SNP array technology makes it possible to examine copy number changes in a genome-wide scale. However, tumor samples often consist of mixed cancer and normal cells. Such tissue heterogeneity poses as a serious hurdle to analyzing copy number changes and could confound subsequent marker identification and diagnostic classification rooted in specific cells. We report here a statistically-principled in silico approach to correct normal tissue contamination and accordingly recover the true copy number profile in cancer cells. We tested the proposed method on three simulations and one real datasets and obtained highly promising results validated by the ground truth and figure of merit. We expect this newly developed method to be a useful tool in routine copy number analysis.

Biography

Yue Wang received his Ph.D. degree in electrical engineering from University of Maryland Graduate School in 1995. From 1996 to 2003, he was an assistant and later associate professor of electrical engineering at The Catholic University of America. In 2003, he joined Virginia Polytechnic Institute and State University and is currently the Grant A. Dove Professor of electrical and computer engineering. Yue Wang became an elected Fellow of The American Institute for Medical and Biological Engineering (AIMBE) and ISI Highly Cited Researcher by Thomson Scientific in 2004. His research interests focus on statistical pattern recognition, machine learning, signal and image processing, with applications to computational bioinformatics and biomedical imaging for human disease research.