

**Stony Brook University
The Graduate School**

Doctoral Defense Announcement

Abstract

Joint Analysis of Gene and Protein Data

By

Chen Ji

Early detection is critical in the successful treatment of life-threatening diseases such as cancer. A vital component of this research is the identification and correlation of disease-related genetic and proteomic biomarkers based on gene micro-array data and proteomic mass spectra data from diseased and control subjects. Such knowledge is crucial in discovering the underlying genetic disease pathways, in drug development and in early diagnosis.

In this work, we first propose a quality control algorithm to improve proteomic data acquisition from mass spectrometer. We then demonstrate a novel variance component approach for biomarker detection and for population homogeneity examination. A major contribution of this thesis is the development of the scoring method that would yield the predictive disease probability rather than the traditional crude binary (yes/no) diagnosis. We present the s-CART and s-RF classifiers – the improved scoring variants of the binary classification and regression tree (CART) and Random Forest (RF) classifiers. Finally, we illustrate the biological and statistical process of integrating the genomic and proteomic data through a human platelet study conducted at the Stony Brook University Medical Center.

Date: August 7, 2007

Time: 11:00 AM

Place: Math 1-122

Program: Applied Mathematics and Statistics

Dissertation Advisor: Wei Zhu