

Winter School 2010

The Chinese University of Hong Kong

January 24-30, 2010

Room 1009, 10/F., William M.W. Mong Engineering Bldg., CUHK

Algorithms for Biclustering Problems

By

Prof. Lusheng Wang

Professor, Department of Computer Science

City University of Hong Kong

Abstract:

One of the main goals in the analysis of microarray data is to identify groups of genes and groups of experimental conditions (including environments, individuals, and tissues) that exhibit similar expression patterns. This is the so-called biclustering problem. In this talk, we consider two variations of the biclustering problem: the consensus submatrix problem and the bottleneck submatrix problem. The input of the problems contains an $m \times n$ matrix A and integers l and k . The consensus submatrix problem is to find an $l \times k$ submatrix with $l < m$ and $k < n$ and a consensus vector such that the sum of distances between the rows in the submatrix and the consensus vector is minimized. The bottleneck submatrix problem is to find an $l \times k$ submatrix with $l < m$ and $k < n$, an integer d and a center vector such that the distance between every row in the submatrix and the vector is at most d and d is minimized. We show that both problems are NP-hard and give randomized approximation algorithms for special cases of the two problems. Using standard techniques, we can derandomize the algorithms to get polynomial time approximation schemes for the two problems. To the best of our knowledge, this is the first time that approximation algorithms with guaranteed ratios are presented for microarray data analysis.

Biography:

Lusheng Wang received the PhD degree from McMaster University, Hamilton, Ontario, Canada, in 1995. Currently, he is a professor in the Department of Computer Science, City university of Hong Kong. His research interests include algorithms, bioinformatics, and computational biology.