

IDENTIFYING BIOLOGICAL PATHWAYS VIA PHASE DECOMPOSITION AND PROFILE EXTRACTION

Yi Zhang and Zhidong Deng
Department of Computer Science, Tsinghua University
Beijing, 100084, China

Biological processes are always carried out through large numbers of genes (and their products) and these activities are often organized into different cellular pathways: sets of genes that cooperate to finish specific biological functions. Owing to the development of microarray technology and its ability to simultaneously measure the expression of thousands of genes, effective algorithms to reveal biologically significant pathways are possible. However, some open problems such as large amount of noise in microarrays and the fact that most biological processes are overlapping and active only on partial conditions pose great challenges to researchers. In this paper, we proposed a novel approach to identify overlapping pathways via extracting partial expression profiles from coherent cliques of clusters scattered on different conditions. We firstly decompose gene expression data into highly overlapping segments and partition genes into clusters in each segment; then we organize all the resulting clusters as a cluster graph and search coherent cliques of clusters; finally we extract expression profiles from coherent cliques and shape biological pathways as genes consistent with these profiles. We compare our algorithm with several recent models and the experimental results justify the superiorities of our approach: robustly identifying overlapping pathways in arbitrary set of conditions and consequently discovering more biologically significant pathways in terms of enrichment of gene functions.