A Coupled Finite Mixture Model for Transcriptional Module Discovery

Han Li*

The Chinese University of Hong Kong, Hong Kong lihan.sta@gmail.com

Xiaodan Fan

The Chinese University of Hong Kong, Hong Kong xfan@sta.cuhk.edu.hk

Approaches to elucidate complex gene regulatory networks usually rely on the analysis of transcriptional modules (TMs). Two high-throughput technologies, gene expression microarray and Chromatin Immuno-Precipitation on Chip, often provide complementary information for discovering TMs. To efficiently integrate these two data sources, we propose a novel Bayesian model referred to as Coupled Finite Mixture Model (CFMM), which permits a separate clustering for each data source and also explicitly models their dependence. We validate our model in both a synthetic dataset and a real dataset. Our method is shown to find more consensus genes and the resulting TMs have improved biological functional coherence than those inferred by other state-of-the-art methods.

Key Words: Chip-chip data, gene expression, integrative clustering