Rong Shen
M.S.
Electrical Engineering
Major Professor: Zhengdao Wang

Graphical Model and Algorithm for Detecting DNA Copy Number Variation

Abstract:
Next-generation sequencing (NGS) has revolutionized the detection of structural variation in genome. Among NGS strategies, reading depth is widely used and paramorphism information contained inside is in general ignored. We develop an algorithm that can fully exploit both reading depth and paramorphism information. We embed mutation procedure in our system model for estimating prior likelihood of single nucleotide base. Hidden Markov model is used to connect single base into segments and belief propagation algorithm is performed for the optimal solution of the HMM model. Simulations show promising results in detecting important types of structural variation. We have applied the algorithm on the maize B73 and MO17 genome data and compared the results with those obtained from arrayCGH method based micro-array data. Inconsistency between the two sets of data is discussed.