A cross entropy Monte Carlo approach to aggregating results from microRNA and other genomic experiments

Shili Lin
Department of Statistics, Ohio State University

Abstract

MicroRNAs (or miRs) are noncoding RNAs whose role is to repress translation by regulating gene expression through binding to Mrna targets. There are computational algorithms for miR target predictions, but their results vary. Thus, it would be useful to consolidate these results to have a greater degree of certainty before engaging in costly experiments. To this end, a cross-entropy Monte Carlo (CEMC) method is explored for solving this combinatorial optimization problem. In essence, CEMC turns the optimization problem into a problem of estimating rare probabilities, for which an iterative importance sampling technique is utilized to slowly tighten the "net" to place most of the weight on the optimal solution and its neighbors. The results demonstrate that our aggregation method can be a useful tool for short listing genes for downstream experiments. In addition to microRNA, our CEMC rank aggregation method is applicable to a wider class of problems, including aggregation of lists of differentially expressed genes from different mRNA microarray experiments and aggregation of predictions of peptides from mass spectrometry data.

This is joint work with Mr. Jie Ding.