P67

ReSAPP: Predicting overlapping protein complexes by merging multiple-sampled partitions of proteins

So Kobiki 1 and Osamu Maruyama 2

1 Graduate School of Mathematics, Kyushu University, Japan

2 Institute of Mathematics for Industry, Kyushu University, Japan

Abstract

Many proteins are known to perform their own functions when they form particular groups of proteins, called protein complexes. With the advent of large-scale protein-protein interaction (PPI) studies, it has been a challenging problem in systems biology to predict protein complexes from PPIs. In this paper, we propose a novel method, called ReSAPP (Repeated Simulated Annealing of Partitions of Proteins), which predicts protein complexes from weighted PPIs. ReSAPP, in the first stage, generates multiple (possibly different) partitions of all proteins of given PPIs by repeatedly applying a simulated annealing based optimization algorithm to the PPIs. In the second stage, all different clusters of size two or more in those multiple partitions are merged into a collection of those clusters, which are outputted as predicted protein complexes. In performance comparison of ReSAPP with our previous algorithm, PPSampler2, as well as other various tools, MCL, MCODE, DPClus, CMC, COACH, RRW, NWE, and PP-Sampler1, ReSAPP is shown to outperform the other methods. Furthermore, the value of F-measure of ReSAPP is higher than that of the variant of ReSAPP without merging partitions. Thus, we empirically conclude that the combination of sampling multiple partitions and merging them is effective to predict protein complexes.