

Exploiting Indirect Neighbours and Topological Weight to Predict Protein Function from Protein-Protein Interactions

Hon Nian Chua, Wing-Kin Sung, and Limsoon Wong

School of Computing and Graduate School for Integrated Sciences and Engineering, National University of Singapore,
3 Science Drive 2, Singapore 117543
{g0306417, dcswk, dcswls}@nus.edu.sg

Abstract. Most approaches in predicting protein function from protein-protein interaction data utilize the observation that a protein often share functions with proteins that interacts with it (its level-1 neighbours). However, proteins that interact with the same proteins (i.e. level-2 neighbours) may also have a greater likelihood of sharing similar physical or biochemical characteristics. We speculate that two separate forms of functional association accounts for such a phenomenon, and a protein is likely to share functions with its level-1 and/or level-2 neighbours. We are interested to find out how significant is functional association between level-2 neighbours and how they can be exploited for protein function prediction.

We made a statistical study on recent interaction data and observed that functional association between level-2 neighbours is clearly observable. A substantial number of proteins are observed to share functions with level-2 neighbours but not with level-1 neighbours. We develop an algorithm that predicts the functions of a protein in two steps: (1) assign a weight to each of its level-1 and level-2 neighbours by estimating its functional similarity with the protein using the local topology of the interaction network as well as the reliability of experimental sources; (2) scoring each function based on its weighted frequency in these neighbours. Using leave-one-out cross validation, we compare the performance of our method against that of several other existing approaches and show that our method performs well.