

Abstract

Several approaches have been used to address the problem of predicting protein-protein interactions (PPI). Some of them are based on biological features extracted from a protein sequence (such as, amino acid composition, GO terms, etc.); others use relational and structural features extracted from the PPI network which can be represented as a graph. Qi *et al.*^[3] divide the protein interaction prediction task into three sub-tasks: (1) prediction of physical (or actual) interaction among proteins, (2) prediction of proteins belonging to the same complex and (3) prediction of proteins belonging to the same pathway. They apply several feature classifiers on the prediction tasks considered. Licamele & Getoor^[2] combined the link structure of the PPI graph with the information about proteins in order to predict the interactions in a yeast dataset, gathered from several databases.

In related work, but a completely different application domain, Hsu *et al.*^[1] address the problem of collaboratively recommending friends for a person, based on structural features extracted from a given social network graph. Noticing the similarity between the friends recommendation problem and the protein-protein interaction prediction problem (i.e., proteins can be associated with users and interactions can be regarded as friendship relationships), in our work, we explore the approach used in Hsu *et al.*^[1] in the context of a protein “friends” recommendation. We identify nine structural features described by Hsu *et al.*^[1] based on the PPI graph and use them to learn classifiers for predicting new interactions. Two datasets containing *Saccharomyces cerevisiae* PPI are used to test the proposed approach. Both these datasets were assembled from the Database of Interacting Proteins (DIP). We assembled the first data set directly from DIP in April 2006, while the second data set has been used in previous studies by Qi *et al.*^[3]. Several classifiers are trained using the structural features that we extract from the interactions graph.

The results show good performance (in terms of accuracy and AUC score), proving that the structural features are highly predictive with respect to PPI. We also compare our results with the results of two existing approaches despite the fact that we use fewer features in our prediction problem when compared with the existing approaches. We discover that we have comparable results (slightly higher accuracy and slightly lower AUC score) when compared with the results of the approach by Licamele & Getoor^[2]. We also notice that we have a significantly better AUC score when compared with the AUC score of the approach by Qi *et al.*^[3].

References

- [1] Hsu, W.H., King, A.L., Paradesi, M.S.R., Pydimarri, T., Weninger, T. Collaborative and Structural Recommendation of Friends using Weblog-based Social Network Analysis, *Proc. of Computational Approaches to Analyzing Weblogs - AAAI 2006 Technical Report SS-06-03*, 55-60.
- [2] Licamele, K., Getoor, L. Predicting Protein-Protein Interactions Using Relational Features, *Proc. of ICML Workshop on Statistical Network Analysis 2006*.
- [3] Qi, Y., Bar-Joseph, Z., Klein-Seetharaman, J., Evaluation of different biological data and computational classification methods for use in protein interaction prediction. *Proteins: Structure, Function, and Bioinformatics*, Volume 63, Issue 3, 2006, 490-500.