

SCHOOL OF COMPUTER ENGINEERING

Proposed Research Projects With Scholarship Support

Supervisor: A/P Kwoh Chee Keong, Division of Information System
Co-Supervisor: Dr Li Xiaoli, Adjunct Ast Prof / Principal investigator, Knowledge Discovery
Department Institute for Infocomm Research (I2R) xlli@i2r.a-star.edu.sg

Research Area: Bioinformatics; Data-Mining; Knowledge Discovery

Project Title: Core-Attachment based Mining for Protein Complexes & Small-molecule Interactions

Abstract ;

In this project, we will investigate and develop core-attachment based methods for protein complexes mining. Protein complexes often correspond to dense subgraphs in interactions networks, biologically a protein complex, like a community, often has a core and attachments. Our objective is to formulate the protein complexes mining algorithm that model this inherent organization which has not been previously done.

Aims:

- To design efficient algorithms to detect protein complexes in PPI networks with modelling of their inherent modularity (core and attachment structure).
- To predict protein-small-molecule interactions, using a PU learning based technique with our protein complexes discovered from PPI networks.

Significance:

Protein–protein interactions have a key role in most biological processes, and offer attractive opportunities for therapeutic intervention. Many proteins are functional only after they are assembled into a protein complex. Experimental detection of protein complexes and developing small-molecules that modulate protein–protein interactions is expensive, time consuming and difficult. Thus, effective and efficient computational methods to find these biologically important complexes; and rapid and direct analysis of the specific interaction of small-molecules to proteins could be of substantial importance to the discovery and screening for new drug molecules. As clinical trials are most expensive part of the pipeline, if failure can be predicted before this point, it saves time and money. Hence it is of utmost importance to identify lead compounds in the early stages of drug discovery that will be most likely to succeed, which boils down to predicting interactions, or rather, non-interactions.

Although many approaches have been proposed to detect dense subgraphs as protein complexes, all these methods ignore the inherent organization of protein complexes (*e.g.*, protein complex has a core-attachment structure). Methods that can provide insights to their organization are greatly desired. Currently, computational methods for predicting protein-small-molecule interactions are mainly based on the 3D structures of proteins and small-molecules. Due to the limited data for protein 3D structures, novel methods should be devised to predict protein and small-molecule interactions, without relying on 3D structural information.

Requirements:

Applicants must have a strong background in Computer Science and/or closely related areas (e.g. Mathematics, Computer Science, Bioinformatics, Statistics and Physics) and excellent skills in both written and spoken English, as the working language of the Faculty is English.