

# Networks in Nature: Dynamics, Evolution, and Modularity

Sumeet Agarwal, Merton College

Doctor of Philosophy, Hilary 2012

## **Abstract**

In this thesis we propose some new approaches to the study of complex networks, and apply them to multiple domains, focusing in particular on protein-protein interaction networks. We begin by examining the roles of individual proteins; specifically, the influential idea of ‘date’ and ‘party’ hubs. It was proposed that party hubs are local coordinators whereas date hubs are global connectors. We show that the observations underlying this proposal appear to have been largely illusory, and that topological properties of hubs do not in general correlate with interactor co-expression, thus undermining the primary basis for the categorisation. However, we find significant correlations between interaction centrality and the functional similarity of the interacting proteins, indicating that it might be useful to conceive of roles for protein-protein interactions, as opposed to individual proteins.

The observation that examining just one or a few network properties can be misleading motivates us to attempt to develop a more holistic methodology for network investigation. A wide variety of diagnostics of network structure exist, but studies typically employ only small, largely arbitrarily selected subsets of these. Here we simultaneously investigate many networks using many diagnostics in a data-driven fashion, and demonstrate how this approach serves to organise both networks and diagnostics, as well as to relate network structure to functionally relevant characteristics in a variety of settings. These include finding fast estimators for the solution of hard graph problems, discovering evolutionarily significant aspects of metabolic networks, detecting structural constraints on particular network types, and constructing summary statistics for efficient model-fitting to networks. We use the last mentioned to suggest that duplication-divergence is a feasible mechanism for protein-protein interaction evolution, and that interactions may rewire faster in yeast than in larger genomes like human and fruit fly.

Our results help to illuminate protein-protein interaction networks in multiple ways, as well as providing some insight into structure-function relationships in other types of networks. We believe the methodology outlined here can serve as a general-purpose, data-driven approach to aid in the understanding of networked systems.