

Chapter XVII

Network Querying Techniques for PPI Network Comparison

Valeria Fionda

Università della Calabria, Italy

Luigi Palopoli

Università della Calabria, Italy

ABSTRACT

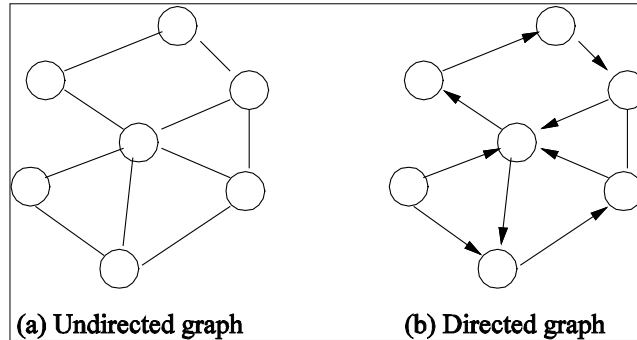
The aim of this chapter is that of analyzing and comparing network querying techniques as applied to protein interaction networks. In the last few years, several automatic tools supporting knowledge discovery from available biological interaction data have been developed. In particular, network querying tools search a whole biological network to identify conserved occurrences of a query network module. The goal of such techniques is that of transferring biological knowledge. Indeed, the query subnetwork generally encodes a well-characterized functional module, and its occurrences in the queried network probably denote that this function is featured by the associated organism. The proposed analysis is intended to be useful to understand problems and research issues, state of the art and opportunities for researchers working in this research area.

INTRODUCTION

Biological data about molecular interactions are growing quickly. This fast growth has rendered the design and the development of automatic tools necessary for retrieving interesting information and discovering new knowledge. Interaction data are usually represented through the *biological networks* model. The comparison of such biological networks across species or different conditions is useful to understand the mechanisms underlying life processes (Zhang, 2008). Generally, biological networks are represented as graphs that can be fed as input to techniques suitable for topological and functional comparison. Such techniques analyze those networks by exploiting specialized algorithms and methodologies in order to infer new information about cellular activity and evolutive processes of the species.

A graph is a set of objects called *nodes* or *vertices* and connected by links called *edges*. More formally, a graph is an ordered pair $G = (V, E)$, in which V is the set of nodes and E is the set of edges, so that the elements

Figure 1. (a) An example of undirected graph and (b) an example of directed graph



from E are pairs of elements from V . In an *undirected* graph, an edge linking nodes A and B can be traversed in both directions. In a *directed* graph, each edge is intended to be traversable in just one direction.

Different types of graphs are used to represent different types of biological networks. In fact, several kinds of biological networks have been defined, each of which stores interaction information related to specific entities or molecules. Main types of networks thereof are: transcriptional regulatory networks, signal transduction networks, metabolic networks and protein interaction network (or PIN). In the *transcriptional regulatory networks* the nodes of the graph represent genes and the edges are directed. An edge connects a source gene to a target gene if the source gene produces an RNA or protein molecule that functions as a transcriptional activator or inhibitor of the target gene. If the gene is an activator, then it is the source of a positive regulatory connection; if it is an inhibitor, then it is the source of a negative regulatory connection. In the *signal transduction networks*, the graph vertices represent proteins and the edges are directed. This type of network stores information about the processes by which a cell converts one kind of signal or stimulus into another. In particular, the signal transduction corresponds to the relaying of molecular signals or physical ones (for example, sensory stimuli) from a cell's exterior to its intracellular response mechanisms. In the *Metabolic networks*, the nodes represent metabolites and the edges are directed. These networks store the set of metabolic and physical processes of the cell and comprise the chemical reactions underlying the metabolism as well as the regulatory interactions that guide these reactions. In the *Protein Interaction Networks*, instead, the nodes represent proteins and the edges are undirected. They store information about the set of interactions between pairs of proteins in a proteome.

In the last few years, due to the large amount of experimentally discovered interaction data, many databases have been devised and made free accessible online. These databases allow storing and retrieving molecular interaction information (Kanehisa, 2000; Salwinski, 2004; Chaturyamontri, 2006). Many of them allow the user to search both online for interaction data and to download database files containing all stored interaction information.

Moreover, several automatic tools supporting knowledge discovery from available interaction data have been developed. Among them, the most related to the topic of this chapter are those tools designed to compare biological networks (see also, in this volume, the chapter *Discovering Interaction Motifs from Protein Interaction Networks*).

To illustrate, there are several ways to compare biological networks, but three of them, *network alignment*, *network integration* and *network querying*, can be regarded as the most significant ones (Sharan, 2006). Network alignment is the process of globally comparing two or more networks of the same type in different species, identifying similarity and dissimilarity regions. Network alignment is commonly applied to detect, across species, conserved sub-networks which are likely to represent common functional modules (Singh, 2007). Network integration is the process of combining several networks of the same species and different kinds, to study their

21 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/network-querying-techniques-ppi-network/5571

Related Content

Dynamics of Protein-Protein Interaction Network in Plasmodium Falciparum

Smita Mohanty, Shashi Bhushan Pandit and Narayanaswamy Srinivasan (2009). *Biological Data Mining in Protein Interaction Networks* (pp. 257-284).

www.irma-international.org/chapter/dynamics-protein-protein-interaction-network/5569

A Particle Swarm Optimization based Hybrid Recommendation System

Rabi Narayan Behera and Sujata Dash (2016). *International Journal of Knowledge Discovery in Bioinformatics* (pp. 1-10).

www.irma-international.org/article/a-particle-swarm-optimization-based-hybrid-recommendation-system/172002

Statistical Analysis for Radiologists' Interpretations Variability in Mammograms

Ahmad Taher Azar (2012). *International Journal of Systems Biology and Biomedical Technologies* (pp. 28-46).

www.irma-international.org/article/statistical-analysis-radiologists-interpretations-variability/75152

Dynamic Analysis of the Possible Effects of Leptin in Some Metabolic Disorders in Obesity

Alejandro Talamino and Laura M. Roa Romero (2012). *International Journal of Systems Biology and Biomedical Technologies* (pp. 1-15).

www.irma-international.org/article/dynamic-analysis-possible-effects-leptin/75150

Performance Assessment of Learning Algorithms on Multi-Domain Data Sets

Amit Kumar and Bikash Kanti Sarkar (2018). *International Journal of Knowledge Discovery in Bioinformatics* (pp. 27-41).

www.irma-international.org/article/performance-assessment-of-learning-algorithms-on-multi-domain-data-sets/202362