

Decomposition of Biological Networks using Betweenness Centrality

John W. Pinney¹, Glenn A. McConkey¹, David R. Westhead¹

Keywords: biological networks, community structure, systems biology

Abstract

The biological sciences are making increasing use of network data in the production and testing of hypotheses, from the simple networks of physical interactions between proteins or correlated expression patterns of genes up to the complex network descriptions of cellular processes such as apoptosis. Until recently, analysis of biological networks has focused on the generation of hierarchical clusters of nodes: groups of interacting proteins or coexpressed genes [1]. However, the realization that biological entities are frequently involved in more than one process has led to criticism of standard clustering algorithms that assign a single grouping for each node. These algorithms have additional problems in that they are usually unable to give much insight into the higher-level biochemical network descriptions that are starting to become available via the Systems Biology Markup Language (SBML) [2] and related efforts.

Here we present an alternative approach to network decomposition that generates hierarchical clusters of relationships, rather than entity nodes. This algorithm is developed from the betweenness centrality decomposition method of Girvan and Newman [3] for identifying community structure within networks. It allows biological entities to be considered as participating in many different processes simultaneously, and hence more naturally extracts biologically meaningful modules from an overall network. Our approach is equally well suited to higher-level and simple networks, and we present examples from protein-protein interaction, gene coexpression and metabolic datasets along with an analysis of a recent network model of apoptosis.

References

- [3] Girvan, M. and Newman, M.E.J. 2002. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences USA* 99:7821-7826.
- [2] Hucka, M. *et al.* 2003. The systems biology markup language (SBML): a medium for representation and exchange of biological network models. *Bioinformatics* 19:524-531
- [1] Rives, A.W. and Galitski, T. 2003. Modular organization of cellular networks. *Proceedings of the National Academy of Sciences USA* 100:1128-1133.

¹ Faculty of Biological Sciences, University of Leeds, LS2 9JT, Leeds, UK, E-mail: john@bioinformatics.leeds.ac.uk