Beyond modularity: density generalized block modeling

J. Reichardt[1], D. R. White[2,3]

[1]Institute for Theoretical Physics, University of Würzburg, Germany,
[2]Dept. of, Antropology, Inst. of Math. Behavioral Sciences, University of California, Irvine, USA
[3]Santa Fe Institute, Hyde Park Road, Santa Fe, NM, USA

Classification
Major categories: Network Theory
Minor categories: Block modeling, functional roles in network

Introduction

The analysis of the structural and statistical properties of complex networks is one of the major foci of complex systems science at the moment. In the context of social networks, the idea that the pattern of connectivity is related to the function of an agent in the network is known as playing a "role" or assuming a "position". Complex systems science has endorsed this idea. By investigating data from a wide range of sources encompassing the life sciences, ecology, information and social sciences as well as economics, researchers have shown that this **intimate relation between topology and function** indeed exists. Hence, understanding the topology of a network is a first step in understanding the function and eventually the dynamics of any network.

Of particular interest in recent years has been the possible decomposition of networks into largely *independent* sub-parts or modules often called "communities". As a community, one generally understands a group of nodes that is densely connected internally but sparsely connected externally. To sociologists the concept of community is known as "cohesive subgroup", but the recent advancements have generalized its applicability much beyond sociology. However, the concept of roles in networks is much wider than mere cohesiveness as it specifically **focuses on the *inter-dependencies* between groups of nodes**. Community structure, emphasizing the absence of dependencies between groups of nodes is only one special case. See Fig 1 for possible network topologies.

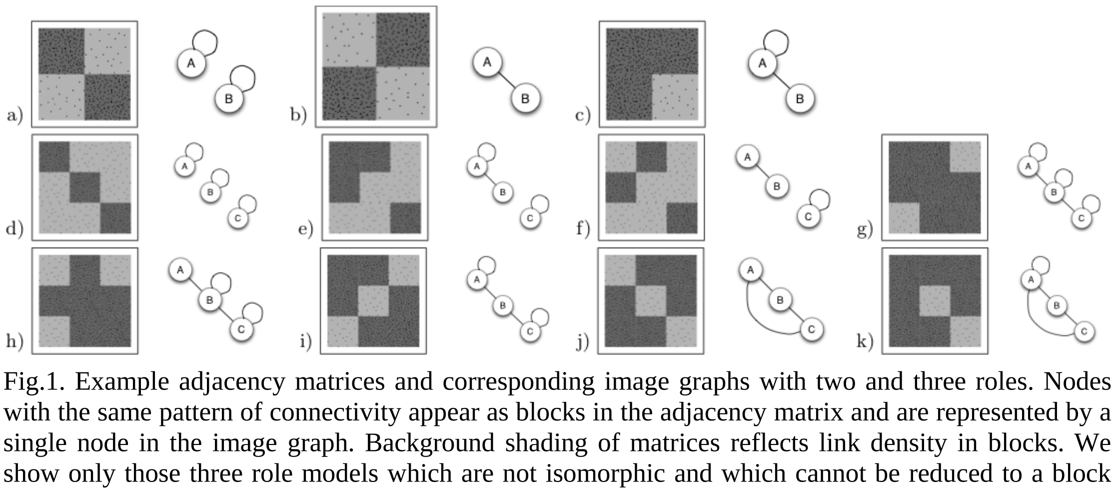

Fig.1. Example adjacency matrices and corresponding image graphs with two and three roles. Nodes with the same pattern of connectivity appear as blocks in the adjacency matrix and are represented by a single node in the image graph. Background shading of matrices reflects link density in blocks. We show only those three role models which are not isomorphic and which cannot be reduced to a block model of two roles only. The two-role-models can be understood as a) modular structure, where nodes connect primarily to nodes of the same role, b) bipartition, with connections primarily between nodes of different type and c) a core-periphery structure with nodes of type A (the core) connecting preferentially among themselves and to nodes of type B (the periphery). The three role models can be seen as combinations of these three basic structures plus the possibility of having intermediates.

Results

We present a framework for automated **discovery of blocking and hence functional classes of nodes** in complex networks [1]. These classes are represented by the nodes of an image graph (“block model") depicting the main patterns of connectivity and thus functional roles in the network. Using a first principles approach, we derive a measure for the fit of a network to any given image graph allowing objective hypothesis testing. From the properties of an optimal fit, we derive how to find the best fitting image graph directly from the network and present a criterion to avoid overfitting. The method can handle both two-mode and one-mode data, directed and undirected as well as weighted networks and allows for different types of links to be dealt with simultaneously. It is non-parametric and computationally efficient. The concepts of structural equivalence and modularity [2] are found as special cases of our approach. We apply our method to networks from economics and biology and show that generalized block structures explain the observed topologies much better than modular structure alone. Further, the classification of nodes into functional roles according to blocks is more robust to noisy data sets and hence particularly valuable for the biosciences where such data frequently occur.

Discussion

The topology of real world networks is very rich. Decomposing then into largely independent subgroups of nodes is often unjustified and oversimplifying the true topology. With the method presented we show how this problem can be overcome in a systematic manner thus opening the way to a better understanding of the topology in networks in terms of functional roles of nodes. The framework we present allows for a wide range of generalizations and applications.

Acknowledgement

References

1. J. Reichardt and D. R. White, *Role models for complex networks*, Eur. Phys. J. B **60**, 217-224 (2007)
1. M.E.J. Newman and M. Girvan, *Finding and evaluating community structure in networks*, Phys. Rev. E **69**, 066133 (2007)