Hierarchical models of complex networks

Mikkel N. Schmidt, Morten Mørup, and Tue Herlau
Cognitive Systems, DTU Informatics
Technical University of Denmark

Networks or graphs of relations (edges) between entities (vertices) occur in many areas of science, for example in sociology, representing friendships between people; in economy, representing trade relations; and in biology, representing interactions between proteins. Many of these real-world networks exhibit hierarchical organization, in the sense that the complex structure of edges between vertices can be well described by a latent hierarchical structure underlying the network.

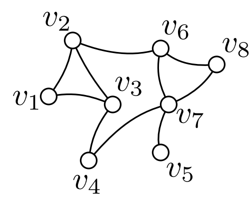

Figure 1: Example of a complex network with $N = 8$ vertices.

We discuss generative nonparametric Bayesian approaches to modeling and infering latent hierarchical structure in complex networks. The models we discuss include existing non-hierarchical networks models as special cases, and thus makes it possible to infer whether or not hierarchical structure is present in a given network.

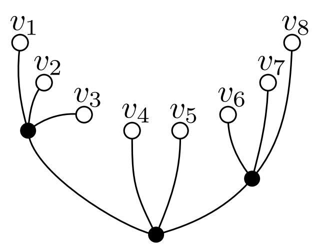

Figure 2: Example of a multifurcating tree with $N = 8$ leaf nodes and $M = 3$ internal nodes. The leaf nodes correspond to vertices in the complex network.

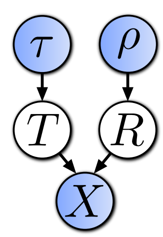

Figure 3: Graphical representation of hierarchical model of complex network.

A general generative model

We represent a complex network with N vertices by an $N \times N$ adjacency matrix X. In the case of an undirected binary network (see Fig. 1) X is a symmetric binary matrix where $X_{ij} = 1$ indicates that an edge exists between node i and j. The following general outline of a probabilistic generative process can be used to characterize a complex network with a hierarchical cluster structure (see also Fig. 3). Several existing hierarchical network models [2, 3, 5] are special cases of this approach.

1. Generate a rooted tree (see Fig. 2) where the leaf nodes corresponds to the vertices in the complex network,
$$T \sim p(T|\tau), \tag{1}$$
where τ are parameters of the prior distribution of the tree. Each internal node in the tree corresponds to a cluster of network vertices.
2. For each internal node r in the tree, generate parameters that govern the probabilities of edges between each of its children.
$$R_r \sim p(R_r|T, \rho), \tag{2}$$
where ρ are parameters of the prior distribution of the edge probabilities.
3. For each pair of vertices i and j in the network, generate an edge with probability governed by the parameters located at the common ancestral nodes in the tree
$$X_{ij} \sim p(X_{ij}|T, R). \tag{3}$$

Inference in such models entails finding the posterior distribution of the tree as well as the edge probability parameters given the observed complex network.

Relation to previous work

The *hierarchical random graph* [2] corresponds to using a flat prior over binary trees (dendrograms) and the the authors "*endow each internal node r of the dendrogram with a probability p_r and then connect each pair of vertices for whom r is the lowest common ancestor independently with probability p_r*" [2].

The approach of *learning annotated hierarchies from relational data* [5] again uses a flat prior over binary trees, and each internal node is assigned a weight variable. A tree-consistent partition is sampled from the tree with probability governed by the weights. For each combination of clusters in this partition, a probability is generated from a Beta distribution, and edges within and between these clusters are generated with these probabilities.

In the *tree-like infinite relational model* [3], in its most simple formulation the prior over trees is constructed such that the partitioning of the vertices on the level below the leaf nodes in the tree are generated from a Chinese restaurant process, and the tree below this level is generated from a flat prior over multifurcating trees. In this model, the likelihood is identical to that of Clauset et al. [2].

Ideas and extensions

The restriction of hierarchical structure to binary trees is unnatural in many applications, and furthermore the number of internal nodes in a binary tree is greater than or equal to the number of internal nodes in a multifurcating tree, leading to an increased computational burden. Although efficient in this sense, the prior in [3] might also not be natural. Another approach is to use an exchangeable process over nested partitionings, such as the *Gibbs fragmentation tree* [4] process, which has an interpretation as a nested Chinese restaurant process.

In existing hierarchical network models, the probability of edges depend only on the parameter located at the nearest common ancestral node in the tree (in [5] as defined by the tree-consistent partition). It is useful to introduce some dependency between the parameters in the tree, for example to force edge probabilities to increase as we move up through the tree, making clusters of nodes higher in the tree more likely to link to each other. Recent studies [6] suggest that this might lead to models with both better interpretation and link prediction performance.

Another idea is to allow the edges to depend not only on the nearest common ancestral node in the tree, but on all common ancestral nodes all the way down to the root. This mechanism also allows clusters higher in tree the to have higher probability of linking: A simple idea is to allow each common ancestral node of vertex i and j to independently generate the edge between the two vertices with probability p_r. The total probability of a link is then

$$p(X_{ij} = 1|T, R) = 1 - (1 - p_{n_0})(1 - p_{n_1}) \cdots (1 - p_{n_L}), \tag{4}$$

where $\{n_0, \ldots, n_L\}$ are the common ancestral nodes of the two vertices (n_0 is the root of the tree). This mechanism can also be used to specify other interesting likelihoods such as mixture models over common ancestral nodes in the tree, akin to the *Bayesian rose tree* model of Blundell et al. [1].

References

[1] Charles Blundell, Yee Whye Teh, and Katherine Heller. Bayesian Rose Trees. In *Uncertainty in Artificial Intelligence, Proceedings of the International Conference on*, 2010.

[2] Aaron Clauset, Cristopher Moore, and M E J Newman. Hierarchical structure and the prediction of missing links in networks. *Nature*, 453(7191):98–101, 2008.

[3] Tue Herlau. Inference in complex networks. Master's thesis, Technical University of Denmark, 2011.

[4] Peter McCullagh, Jim Pitman, and Matthias Winkel. Gibbs fragmentation trees. *Bernoulli*, 14(4):988–1002, 2008.

[5] Daniel M Roy, Charles Kemp, Vikash K Mansinghka, and Joshua B Tenenbaum. Learning annotated hierarchies from relational data. *Advances in Neural Information Processing Systems 19*, 19:1185–1192, 2007.

[6] Morten Mørup, Mikkel N. Schmidt, and Lars Kai Hansen. Infinite multiple membership relational modeling for complex networks. In *Machine Learning for Signal Processing, IEEE International Workshop on (MLSP)*, 2011.