Stochastic Block Models with Multiple Continuous Attributes

Natalie Stanley, Thomas Bonacci, Roland Kwitt, Marc Niethammer, Peter J. Mucha

ABSTRACT

Stochastic block models (SBMs) are probabilistic models for community structure in networks in which nodes within a community are assumed to be connected to nodes within and between communities in a uniform, characteristic way. Typically, only the adjacency matrix is used to perform SBM parameter inference. In this paper, we consider circumstances in which nodes have an associated vector of continuous attributes. The model assumes that the attributes associated with nodes in a network's community can be described by a particular multivariate Gaussian model. Moreover, in this augmented SBM, the objective is to learn the SBM and multivariate gaussian parameters describing each community. While there are recent examples in the literature that combine connectivity and attribute information, we are to our knowledge the first to consider the effect of multiple continuous attributes. We highlight the usefulness of our model for two two network prediction tasks: collaborative filtering and link prediction. As a result of fitting this attributed stochastic block model, one can predict the attribute vector or connectivity patterns for a new node in the event of the complementary source of information (connectivity or attributes, respectively). While our approach is the first stochastic block model (to our knowledge) to include multiple continuous attributes, we highlight the ability of our approach in two tasks: link prediction in collaborative filtering. In this work, we demonstrate the usefulness of our model in two different types of biological networks.

First, we consider a network between a set of subjects encoding the similarity in the bacterial species counts in their microbiomes. For each subject, we consider attributes, such as, BMI, nationality, and age. Second, We consider a protein interaction network encoding regulatory relationships, augmented with experimental data about function and abundance.

Data for tasks 1 and 2: https://www.nature.com/articles/ncomms5344 http://pubs.acs.org/doi/full/10.1021/pr401258d

WOODSTOCK '97 El Paso, Texas USA

DOI: 10.475/123_4

Keywords

Stochastic Block Model, Networks, Community Detection, Attributes, Image Analysis

1. INTRODUCTION

1.1 Network data and node attributes

Uncovering patterns in network data is a common pursuit across a range of fields, such as in biology [10], medicine [8] and computational social science [7]. A powerful way to analyze mesoscale structural organization within a network is with community structure [14, 11]. In this pursuit, the objective is to identify cohesive groups of nodes with a high density of within-group connections and few between-group connections. Numerous approaches exist to accomplish this task [14, 11], but typically only the adjacency matrix conveying the wiring of the graph is taken into account. In certain applications, each node in a network is equipped with additional information (or particular attributes) that was not implicitly taken into account in the construction of the network. For example, one could consider a collection of attributes measured for individuals in a social network (i.e., age, income, level of education). In this work, we seek to incorporate these attributes in the community detection task, such that communities are able to effectively combine both sources of information to understand the organization of components in a system.

Further motivation for the incorporation of node attributes in community detection is also inspired by detectability of communities, as well as possible impact in the image analysis community. The detectability limit in networks is a theoretical notion of the relationship between the number of within-community connections to the number of between-community connections that allows for algorithms to accurately detect community structure [12, 5]. In circumstances where the number of within-community connections is sparse, we propose that the incorporation of attribute information will make communities easier to detect. Further, graph-based image segmentation has shown promise [15, 2], however the probabilistic notion of communities in this context is not well explored.

1.2 Related work

The incorporation of node attributes in network analysis tasks is not well explored, particularly in the context of probabilistic network models. In a modularity based approach, the authors of [3] modified the classical modularity

quality function to incorporate node attribute information. Two probabilistic network and attribute models have also been proposed by [16, 13]. In [16], the authors developed a generative model for attributed network data based on node affiliation. However, in this approach only binary attribute data are considered. In [13], the authors developed a probabilistic framework for incorporating metadata or attributes in community detection, however, the model takes only a single metadata value for each node into account. We seek to extend these existing approaches to reflect continuous attribute data and particularly multidimensional attribute data drawn from a multivariate Gaussian distribution.

New stuff to add: peel, clauset controversy, tiago and dhric

1.3 Stochastic Block Models

To accomplish our objective of integrating node attributes in a probabilistic framework, we will model connectivity in a network with the widely-used stochastic block model (SBM) [9]. This model assumes that edges within a community are connected within and between communities in a characteristic or probabilistic way. To fit this model to network data, the objective is to partition the nodes into communities, such that these assignments maximize the likelihood of the model according to the observed edges. In this inference problem for a network with N nodes and K communities, a $K \times K$ probability matrix, **theta** and a N-length vector, $\mathbf{z}$ are learned. The matrx θ described the probability of connections within and between communities, while $\mathbf{z}$ gives the node-to-community assignments. While this modeling framework is well-studied, recent attention hs focused on the ability to integrate extra information (attributes or metadata), into the inference problem and whether it is appropriate to do so.

1.4 Contributions and Paper Outline

We summarize our contributions to the problem of finding communities in attributed network data as follows.

- We developed a probabilistic model for node-to-community structure that uses adjacency matrix connectivity and continuous node attributes to determine a community assignment for the nodes. Our model is the first to our knowledge to allow for the augmentation of a classic stochastic block model with multiple continuous attributes.
- We provide details for the inference technique that can be used to effectively estimate model parameters.
- We demonstrate that by fitting an attributed stochastic block model allows for good performance in link prediction and collaborative filtering results. In particular, we demonstrate these tasks on two different biological examples.

The paper is organized as follows. In section 2, we describe the model and inference methods for estimating parameters. In section 3, we perform experiments on synthetic networks. Section 4 shows how the method can be applied in image segmentation tasks. Finally in section 5 we discuss the method, its limitations and future work.

2. MODEL

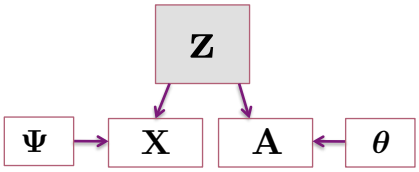

Figure 1: **Modeling community membership in terms of attributes and connectivity**. Node-to-community assignments specified by $\mathbf{z}$ are determined in terms of adjacency matrix information, $\mathbf{A}$ and attribute matrix information, $\mathbf{A}$. $\mathbf{A}$ and $\mathbf{X}$ are assumed by be generated from a stochastic block model and a mixture of multivariate gaussian distributions, parameterized by $\boldsymbol{\theta}$ and $\boldsymbol{\Psi}$, respectively.

2.1 Objective

We seek to incorporate both connectivity ($\mathbf{A}$) and attribute information ($\mathbf{X}$) to infer node-to-community assignments, $\mathbf{Z}$. Note that for a network with N nodes, K communities and p measured attributes, $\mathbf{A}$, $\mathbf{X}$, and $\mathbf{Z}$ have dimensions $N \times N$, $N \times p$ and $N \times K$, respectively. In particular, $\mathbf{Z}$ is a binary indicator matrix, where entry z_{ic} is 1 if and only if node i belongs to community c. We also define $\mathbf{z}$ to be the N-dimensional vector of node-to-community assignments. Following the assumption of [16], we assume that connectivity and attributes are conditionally independent, given the community membership label. The graphical model for the relationship between node-to-community labels, connectivity and attribute information is shown in Fig. 1.

To infer the $\mathbf{Z}$ that best explains the data, we adopt a likelihood maximization approach. That is, we seek to find the partition of nodes to communities that best describes the observed connectivity and attribute information. Given the conditional independence assumption of $\mathbf{X}$ and $\mathbf{A}$, we can express the log likelihood of the data, $\mathcal{L}$ as the sum of connectivity and attribute log likelihoods, $\mathcal{L}_A$ and $\mathcal{L}_X$, respectively, as

$$\mathcal{L} = \mathcal{L}_A + \mathcal{L}_X \ . \tag{1}$$

This likelihood reflects the joint distribution of the adjacency matrix, $\mathbf{A}$, the attribute matrix, $\mathbf{X}$, and the matrix of node-to-community indicators, $\mathbf{Z}$; formally, we have

$$\mathcal{L} = p(\mathbf{A}, \mathbf{X}, \mathbf{Z}) \ . \tag{2}$$

Given that $\mathbf{Z}$ is a latent variable that we are trying to infer, we can approach the problem using the expectation maximization (EM) algorithm [6]. By doing this, we will alternate between estimating the posterior probability that a node i has community label c, or

$$p(z_{ic} = 1 \mid \mathbf{X}, \mathbf{A}) \tag{3}$$

and estimates for $\boldsymbol{\theta}, \boldsymbol{\Psi}$, i.e., the model parameters specifying the adjacency and attribute matrices, respectively.

2.2 Attribute Likelihood

For a network with K communities, we assume that each particular community i has an associated p-dimensional mean $\boldsymbol{\mu}_i$ and $p \times p$ covariance matrix, $\boldsymbol{\Sigma}_i$. Note that these parameters uniquely identify a p-dimensional multivariate Gaussian

distribution. To specify this model for all K communities, we define the parameter $\mathbf{\Psi} = \{\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \ldots \boldsymbol{\mu}_k, \mathbf{\Sigma}_1, \mathbf{\Sigma}_2, \ldots \mathbf{\Sigma}_K\}$. Then, the probability of the i-th row of $\mathbf{X}$, denoted as $\mathbf{x}_i$, giving the values of the p attributes for node i, can be modeled as

$$p(\mathbf{x}_i \mid \mathbf{\Psi}) = \sum_{c=1}^{K} \pi_c p(\mathbf{x}_i \mid \boldsymbol{\mu}_c, \mathbf{\Sigma}_c) \ . \tag{4}$$

Here, $p(\mathbf{x}_i \mid \boldsymbol{\mu}_c, \mathbf{\Sigma}_c)$ is the probability density function for the multivariate Gaussian and π_c is the probability that a node is assigned to community c.

2.3 Adjacency Matrix Likelihood

For the adjacency matrix, $\mathbf{A}$ and the $K \times K$ matrix of stochastic block model parameters, $\boldsymbol{\theta}$, the probability of observing the connectivity pattern of node i, or the i-th row in the adjacency matrix, denoted as $\mathbf{a}_i$, is given by

$$\begin{aligned} \log p(\mathbf{a}_i \mid \boldsymbol{Z}, \boldsymbol{\theta}) = \sum_{c_1=1}^{k} z_{i,c_1} [\sum_{c_2=1}^{k} \sum_{j \in \mathcal{N}(i)} z_{j,c_2} \log(\boldsymbol{\theta}_{c_1,c_2}) \\ - \sum_{j \notin \mathcal{N}(i)} z_{j,c_2} \log(1 - \boldsymbol{\theta}_{c_1,c_2})] \end{aligned} \tag{5}$$

2.4 Inference

To use EM to maximize the likelihood of the data, we break the process into the E-step and M-Step.

E-Step. During the E-step, we use the current value of learned model parameters, $\boldsymbol{\theta}$ and $\mathbf{\Psi}$ to compute the posterior, given in Eq. (3). The posterior, or expectation $\mathbb{E}(z_{ic})$, of node i belonging to community c, is given by

$$\begin{aligned} \mathbb{E}(z_{ic}) &= p(z_{ic} = 1 \mid \mathbf{x}_i, \mathbf{a}_i) \\ &= \frac{p(\mathbf{x}_i, \mathbf{a}_i, z_{ic})}{p(\mathbf{x_i}, \mathbf{a_i})} \\ &= \frac{p(\mathbf{x_i} \mid z_{ic}) p(\mathbf{a}_i \mid z_{ic}) \pi_c}{\sum_{c=1}^{K} p(\mathbf{x_i} \mid z_{ic}) p(\mathbf{a}_i \mid z_{ic}) \pi_c} \ . \end{aligned} \tag{6}$$

M-Step. In the M-step, we can compute updates for $\boldsymbol{\theta}$ and $\mathbf{\Psi}$ using this expectation.

Since, the attributes follow a Gaussian mixture model, it can be shown that the update for the mean vector describing community c, $\boldsymbol{\mu}_c$, can be computed as

$$\boldsymbol{\mu}_c = \frac{\sum_{i=1}^{N} \mathbb{E}(z_{ic}) \mathbf{x}_i}{\sum_{i=1}^{N} \mathbb{E}(z_{ic})} \ . \tag{7}$$

Similarly, the update for the covariance matrix describing a community, $\mathbf{\Sigma}_c$, is computed as

$$\mathbf{\Sigma}_c = \frac{\sum_{i=1}^{N} \mathbb{E}(ic)(\mathbf{x}_i - \boldsymbol{\mu}_c)(\mathbf{x}_i - \boldsymbol{\mu}_c)^T}{\sum_{i=1}^{N} \mathbb{E}(ic)} \ . \tag{8}$$

To update the parameters of $\boldsymbol{\theta}$, we follow the method in [4] and update the probability of an edge existing between community q and l, given by $\boldsymbol{\theta}_{ql}$ as,

$$\boldsymbol{\theta}_{ql} = \frac{\sum_{i \neq j} \mathbb{E}(z_{iq}) \mathbb{E}(z_{jl}) x_{ij}}{\sum_{i \neq j} \mathbb{E}(z_{iq}) \mathbb{E}(z_{jl})} \tag{9}$$

We continue the process of iterating between the E-step and M-step until the change in the data log-likelihood, $\mathcal{L}$, is below a predefined tolerance threshold.

3. SYNTHETIC DATA RESULTS

We first test the performance of our model and inference procedure on a synthetic example. Here, we generated a network for a stochastic block model with $N = 200$ nodes and $K = 4$ communities . Edges in the adjacency matrix were generated according to a stochastic block model, parameterized as follows:

$$p(A_{ij} = 1) \sim \begin{cases} \text{Bernoulli}(.10), & \text{if } z_i \neq z_j \\ \text{Bernoulli}(.25), & \text{if } z_i = z_j \end{cases} \tag{10}$$

Note that $\mathbf{z}$ is a 200-dimensional vector, where z_i identifies the community label for node i.

Fig. 2(a) shows an example network generated according to this parametrization. Nodes are colored by their community assignment, with orange, green, purple, and magenta corresponding to communities 1-4, respectively. Looking at the network, it is apparent that there is not a strong notion of detectable community structure. That is, nodes are not separated into the desired homogeneous clumps that should be consistent with community assignment. To model attributes, for a community c, we randomly generated an 8-dimensional vector, $\boldsymbol{\mu}_c$, where each entry is from a Gaussian with 0-mean and unit variance. Associated with all c, $c \in \{1, 2, 3, 4\}$ is a 4×4 diagonal covariance, $\mathbf{\Sigma}_c = \text{diag}(1.25)$. Moreover, using the $\boldsymbol{\mu}_c$ and $\mathbf{\Sigma}_c$, a sample attribute vector can be generated. That is, the attribute vector for node i, $\mathbf{x}_i$ is generated as,

$$\mathbf{x}_i \sim \mathcal{N}(\boldsymbol{\mu}_{z_i}, \mathbf{\Sigma}_{z_i}) \tag{11}$$

where $\mathcal{N}(\cdot, \cdot)$ denotes a multivariate Gaussian.

Fig. 2(b) shows a PCA plot of the attribute vectors associated with each node in an example synthetic experiment and hence, each point represents a node. Since, the true dimension of these feature vectors is 8, this plot provides a projection to 2 dimensions that allows for visualization of the relatedness between the nodes, according to the attributes. One can observe that members of community 2 are overall nicely separated in attribute space but members of communities 3 and 5 are especially hard to discern.

4. USING THE FITTED ATTRIBUTED SBM FOR LINK PREDICTION AND COLLABORATIVE FILTERING

One of the benefits of a generative network model is that it can be applied to prediction tasks. Most notably, in the absence of one source of information about a node (connectivity or attributes), the model can be used to predict the complementary information source (attributes or connectivity, respectively). By fitting an attributed SBM, we found that obtain successful performance in two fundamental network prediction tasks, link prediction and collaborative filtering.

In the link prediction problem, when given two node stubs, the objective is to determine whether a link exists between

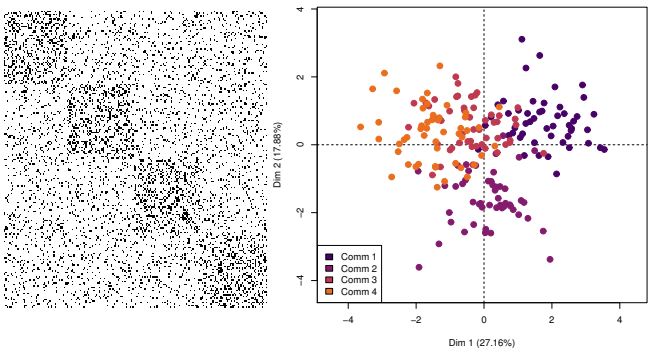

Figure 2: **Synthetic Experiments.** NMI kmeans=0.68, NMI SBM=0.65, NMI attributeSBM=0.83

them. Since we are modeling connectivity with a stochastic block model, we can predict links using the learned parameters. In particular, we highlight how this task can be performed using just the attribute information of the node stubs of interest. In the experiments to follow, we compare to 3 commonly-used link prediction methods. In all of these methods, a score is computed for all pairs of edge candidates and ultimately the top x set of edges with highest weights are kept (where x is some user-defined parameter). Let m and n be a pair of nodes and $\Gamma(m)$ denote the set of neighbors for a node m. Then, under the following 3 link prediction methods, we can calculate the score of the potential link as, Score(m, n)

Jaccard: $\text{Score}(m,n) = \frac{\Gamma(m)\cap\Gamma(n)}{\Gamma(m)\cup\Gamma(n)}$

Adamic Adar: $\text{Score}(m,n) = \sum_{c\in(m)\cap\Gamma(n)} \frac{1}{\log|\Gamma(c)|}$

Preferential Attachment: $\text{Score}(m, n = |\Gamma(m)|\times|\Gamma(n)|$

Conversely, the collaborative filtering problem seeks to predict a node's attributes based on its similarity to its neighbors. For some node of interest, we can use our fitted attributed SBM model to predict a node's attributes, given only the information about its connectivity. Formally, for node i, we seek to predict $\mathbf{x}_i$. In the following experiments, we compare our results to two common collaborative filtering approaches. Let $\mathcal{N}^k(m)$ be the set of k-nearest neighbors in the network for node m. Let $\hat{\mathbf{x}}_i$ be the predicted attribute vector for node i and s_{ij} be a similarity measure between nodes i and j.

Neighborhood Avg: $\hat{\mathbf{x}}_i = \frac{1}{|\mathcal{N}^k(i)|}\sum_{j\in\mathcal{N}^k(i)}\mathbf{x}_j$

Weighted Neighborhood Avg: $\hat{\mathbf{x}}_i = \frac{1}{\sum_{jin\mathcal{N}^k(i)} s_{ij}}\sum_{j\in\mathcal{N}^k(i)} s_{ij}\mathbf{x}_j$

We show results for these two tasks in two different biological network examples in section X. In particular, the experiments were designed in the following ways.

4.1 Link Prediction Experiments

For our link prediction (figures x and y), we first choose 25 pairs of nodes that have an edge in the network and 25 pairs that do not. We then test each of these 50 edges in a leave one out manner. That is, for each of the 50 edges, we fit the attributed SBM to the network with the two stubs of the edge left out. We then use the nearest neighbor in attrivute space of each stub as the the input to each of the 3 baseline community detection methods (Jaccard, Adamic Adar, and Preferential Attatchment). To use our attributed SBM in this link prediction task, we also consider the most commonly observed community among the 3 nearest neighbors for the stubs of the edge of interest. Again, using the nearest neighbors (denote n and m of the stubs, then we define the link prediction score for the edge as θ_{z_n,z_m}, or the probability that an edge exists between nodes n and m according to the fitted model.

4.2 Collaborative Filtering Experiments

In collaborative filtering experiments, the objective is to predict the vector of attributes for each node. In our experiments, we used leave-one-out validation to predict the attribute vector for each node. That is, for each node in the network, we created a single node test set. The training set, was then the rest of the network with the node to predict removed. For this single test set node, we identified neighbors it connects to in only connectivity space within the training set. For standard collaborative filtering approaches (Neighborhood average and weighted neighborhood average), the predicted attribute for the test set node is then the specified averaging of the neighbors. To use our model for this task, we first fit the attributed SBM model to the training set. Similar to the standard link prediction approaches, we identify the nearest neighbors for our test node in connectivity space within the training set. We then predict the community membership of our test node to be the most-frequently observed community among its neighbors. Using this community assignment, c, we then predict the attribute vector for our test node to be μ_c, or the mean vector that was learned to describe community c.

5. APPLICATIONS IN BIOLOGICAL NETWORKS

While the motivation for the development of this method was not motivated by problems in biological data, we evaluate the potential to combine similarity or relational information between a set of entities, whether that be proteins, or biological samples, and experimental data and metadata. Our application of this model to biological problems provides a framework to predict attribute or connectivity information about a new observation. Note that we do not intend to suggest any new biological insights, but rather that we can combine two sources of information for prediction tasks and alternative definitions of what constitutes a community in the data.

5.1 Microbiome Subject Similarity Results

Motivation

In the analysis of biological data, it is often useful to cluster subjects based on a set of their measure biological features and to then determine what makes each of the subgroups different. One type of biological data gaining much attention in recent years is metagenomic sequencing data, used to profile the composition of a microbiome. We refer to this as the 'metagenomic profile' and each feature is a count for each bacterial species, also known as operational taxonomic unit (OTU). In their paper (Ref.), the authors conducted a study among subjects across a variety of ethnicities, body mass index (BMI) classifications, and age groups. Additionally, each subject has a count of 130 different bacterial species, obtained through sequencing.

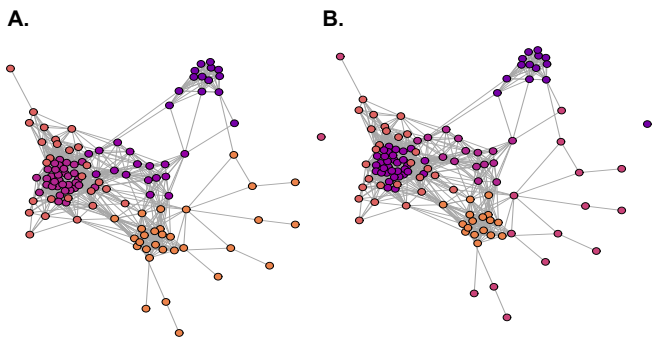

Figure 3: **Visualization of fitting attributed model to microbiome subject similarity network:** Nodes are colored by one of 5 community assignments as predicted with the attributed SBM.

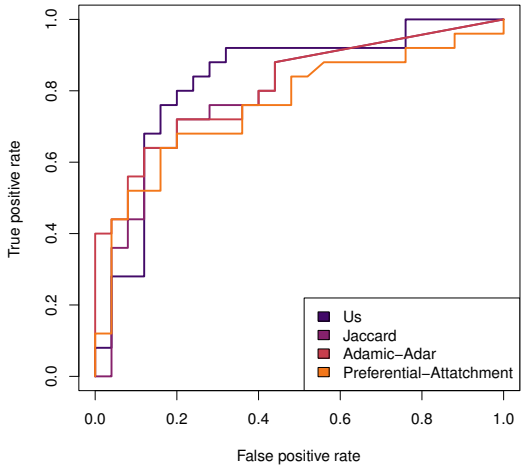

Figure 4: **Link Prediction on the microbiome subject similarity matrix:**

Pre-Processing We extracted a subset of the subjects being from Eastern Europe,Southern Europe ,Scandinavia, and the United States and constructed a between-subject similarity network between only individuals who had a BMI measurement. This resulted in a network of 121 nodes, where each edge is the pearson correlation between their microbial compositions. We then removed all edges in the network with weight (correlation) <0.7.

Constructing Node Attributes We clustered samples (nodes) in this network using the Louvain community detection method into one of six clusters. We then built a random forest model to predict each sample's cluster assignment according to their metagenomic profiles. The attribute for each node was then the probability distribution of belonging to each of the 6 clusters of samples. This example highlights a novel way to combine two sources of information, between subject similarity and extra knowledge about the group memberships of the subjects. In figure 3A-B, we plot this subject similarity network and color the nodes by their community assignments using the classic and attributes stochastic block models, respectively. We also computed the normalized mutual information (NMI) between the node-to-community assignments identified from the partition under the louvain algorithm with the results obtained fitting the class and attributed SBM. Since the attributes of the nodes were the probability of belonging to each of the clusters identified under the Louvain algorithm partition, it was expected to observe a higher similarity between the Louvain partition with the attributed SBM partition than witth the SBM partition. This is exactly what we observed, with 0.78 being the NMI between the louvain algorithm partition and the classic SBM. Using the attributed SBM, we observed an NMI of 0.94 with the Louvain algorithm partition.

Microbiome Link Prediction In figure 4a-b we show the results for the link prediction in the sample similarity network. To perform link prediction, we selected 25 pairs of nodes (i.e. edge stubs) from our 121-node subject network. In our prediction task , we sought to model the ability to use only the attribute information and our fitted model to predict whether there was a link. Repeating this experiment 25 times, we obtained a distribution of area under the curve (AUC) values in figure 4A. and plotted the ROC curves corresponding to the experiment closest to the median for all link prediction methods.

Microbiome Collaborative Filtering Figure X shows the results for the collaborative filtering task on the subject microbiome network

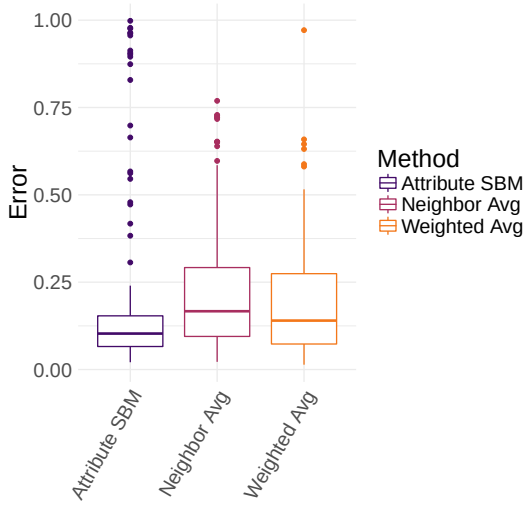

Figure 5: **Collaborative Filtering Accuracy in Microbiome Data**: For each of the 121 nodes, we fit a model to the remaining 120 node network and given the node's closest network neighbors try to compute its attribute value. Error is the 2-norm between true and predicted attributes.

Here, we sought to predict the 6 dimensional attribute vector for each node.

sahfkahfkjh insert more text

5.2 Protein Interaction Network Results

We also apply our attributed SBM approach to the protein interaction network presented in [1]. This network represents interactions between proteins, predicted from the liter-

ature. Associated with each each node (protein), is a vector of experimentally observed modifications resulting from the exposure of cancer cells to a chemotherapeutic drug. Proteins were able to undergo six possible modifications. While communities in this network should reflect functional relatedness among proteins, we also expect that members of a community should share similarities in the observed modification type.

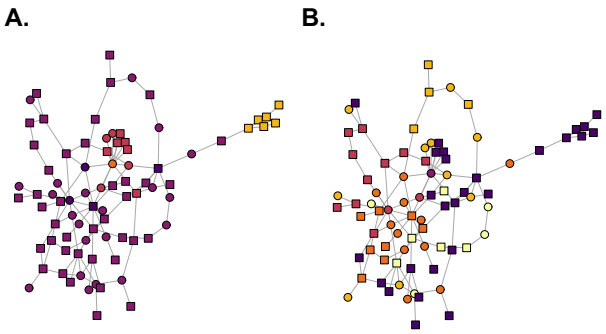

Figure 6: **Protein interaction network**

Data Pre-Processing: We downloaded the network data and the modification information from the supplement of [1].

Constructing Node Attributes: For each node, we constructed its attributes vector as a vector of length 6, where each entry is a binary indicator for which of the 6 modifications was experimentally observed.

Figure 6A-B show the results of fitting a classic SBM and attributed SBM, respectively. The 6 possible modifications exist for 3 biological processes that can either increase or decrease. The node shape reflects whether the modification for a node was an increase (square) or decrease (circle). Nodes are colored by their assignment into 1 of x communities. (mention something about the intuition behind the color groupings).

Next, we studied the entropies of these binary node classifications in each of the communities according to the regular and attributed SBM partitions. The hypothesis was that using the attributed SBM, we should have lower entropy of these two labels within communities because the attribute component of the model should assist in creating communities that are not only spatially relevant but also agree in attributes.

Link Prediction in the Protein interaction network

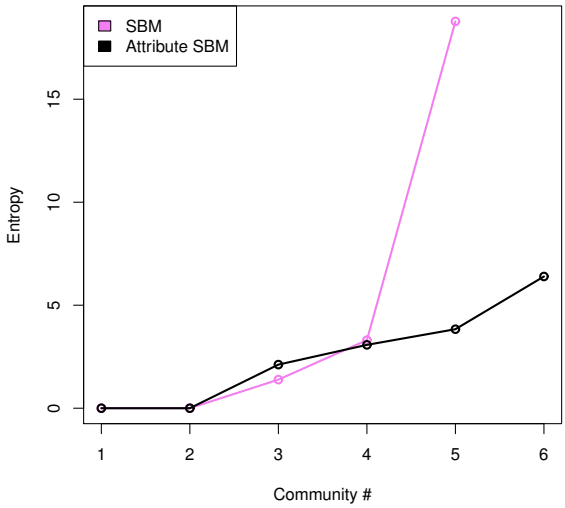

Figure 7: **Community Entropies**

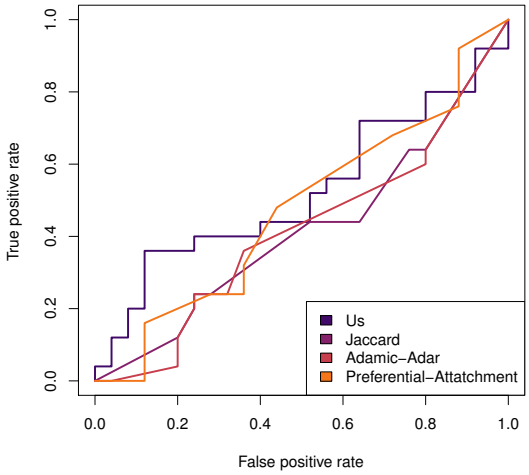

Figure 8: **Community Entropies**

6. DISCUSSION

Detectability problems

7. DISCUSSION

To be filled in.

8. REFERENCES

[1] T. Bonacci, S. Audebert, Stephane, L. Camoin, E. Baudelet, and G. Bidaut. Identification of new mechanisms of cellular responses to chemotherapy by tracking changes in post-translational modifications by ubiquitin and ubiquitin-like proteins.

[2] A. Browet, P.-A. Absil, and P. Van Dooren. Community detection for hierarchical image segmentation. In *Combinatorial Image Analysis*, pages 358–371. Springer, 2011.

[3] D. Combe, C. Largeron, M. Géry, and E. Egyed-Zsigmond. I-louvain: An attributed graph clustering method. In *Advances in Intelligent Data Analysis XIV*, pages 181–192. Springer, 2015.

[4] J.-J. Daudin, F. Picard, and S. Robin. A mixture model for random graphs. *Statistics and computing*, 18(2):173–183, 2008.

[5] A. Decelle, F. Krzakala, C. Moore, and L. Zdeborová. Asymptotic analysis of the stochastic block model for

modular networks and its algorithmic applications. *Physical Review E*, 84(6):066106, 2011.

[6] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society. Series B (methodological)*, pages 1–38, 1977.

[7] D. Greene and P. Cunningham. Producing a unified graph representation from multiple social network views. In *Proceedings of the 5th Annual ACM Web Science Conference*, pages 118–121. ACM, 2013.

[8] J. Guinney, R. Dienstmann, X. Wang, A. de Reyniès, A. Schlicker, C. Soneson, L. Marisa, P. Roepman, G. Nyamundanda, P. Angelino, et al. The consensus molecular subtypes of colorectal cancer. *Nature medicine*, 2015.

[9] P. W. Holland, K. B. Laskey, and S. Leinhardt. Stochastic blockmodels: First steps. *Social networks*, 5(2):109–137, 1983.

[10] D. B. Larremore, A. Clauset, and C. O. Buckee. A network approach to analyzing highly recombinant malaria parasite genes. *PLoS Comput Biol*, 9(10):e1003268, 2013.

[11] J. Leskovec, K. J. Lang, and M. Mahoney. Empirical comparison of algorithms for network community detection. In *Proceedings of the 19th international conference on World wide web*, pages 631–640. ACM, 2010.

[12] R. R. Nadakuditi and M. E. Newman. Graph spectra and the detectability of community structure in networks. *Physical review letters*, 108(18):188701, 2012.

[13] M. Newman and A. Clauset. Structure and inference in annotated networks. *arXiv preprint arXiv:1507.04001*, 2015.

[14] M. A. Porter, J.-P. Onnela, and P. J. Mucha. Communities in networks. *Notices of the AMS*, 56(9):1082–1097, 2009.

[15] J. Shi and J. Malik. Normalized cuts and image segmentation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(8):888–905, 2000.

[16] J. Yang, J. McAuley, and J. Leskovec. Community detection in networks with node attributes. In *Data mining (ICDM), 2013 ieee 13th international conference on*, pages 1151–1156. IEEE, 2013.