Hindawi
Complexity
Volume 2020, Article ID 5712815, 15 pages
https://doi.org/10.1155/2020/5712815

Research Article

A Unified Bayesian Model for Generalized Community Detection in Attribute Networks

Qiang Tian,[1] Wenjun Wang,[1] Yingjie Xie,[1] Huaming Wu,[2] Pengfei Jiao,[3] and Lin Pan[4]

[1]*College of Intelligence and Computing, Tianjin University, Tianjin, China*
[2]*Center of Applied Mathematics, Tianjin University, Tianjin, China*
[3]*Center of Biosafety Research and Strategy, Law School, Tianjin University, Tianjin, China*
[4]*School of Marine Science and Technology, Tianjin University, Tianjin, China*

Correspondence should be addressed to Lin Pan; linpan@tju.edu.cn

Received 18 June 2020; Revised 7 August 2020; Accepted 18 August 2020; Published 29 August 2020

Academic Editor: Eric Campos

Identification of community structures and the underlying semantic characteristics of communities are essential tasks in complex network analysis. However, most methods proposed so far are typically only applicable to assortative community structures, that is, more links within communities and fewer links between different communities, which ignore the rich diversity of community regularities in real networks. In addition, the node attributes that provide rich semantics information of communities and networks can facilitate in-depth community detection of structural information. In this paper, we propose a novel unified Bayesian generative model to detect generalized communities and provide semantic descriptions simultaneously by combining network topology and node attributes. The proposed model is composed of two closely correlated parts by a transition matrix; we first apply the concept of a mixture model to describe network regularities and then adjust the classic Latent Dirichlet Allocation (LDA) topic model to identify community semantically. Thus, the model can detect broad types of network structure regularities, including assortative structures, disassortative structures, and mixture structures and provide multiple semantic descriptions for the communities. To optimize the objective function of the model, we use an effective Gibbs sampling algorithm. Experiments on a number of synthetic and real networks show that our model has superior performance compared with some baselines on community detection.

1. Introduction

With the advent of the era of big data and the diverse channels for acquiring data, we have obtained a large amount of data from complex systems in the real world [1]. In particular, we can obtain not only diversified entities in complex systems but also a variety of related descriptions (attributes) of them. Attributed complex networks are usually used to analyze and study these data [2, 3]. Taking social systems as an example, nodes denote individuals and edges represent interactions between them. At the same time, individuals have personal information about gender, age, country, job, race, and so on, which represent their unique attributes. The sufficient and effective application of structural and attribute information is of great value for complex network analysis.

At present, exploring the structural regularities and functions of the network is a significant part of complex network analysis [4]. One of the most essential tasks is community detection. It is believed that nodes within the same community typically have similar structural characteristics and properties. The detection of communities or modules in a network is conducive to understanding organization rules of complex networks, exploring latent patterns, and predicting the behavior of complex systems. A number of successful community detection approaches have been proposed, which fall into different categories, such as hierarchical clustering algorithms [5, 6], modularity optimized approaches [7], statistical inference [8–10], spectral algorithms [11–14], generative model [15–18], and Markov dynamic algorithms [19–21]. For review, the readers can refer to [22].

However, most conventional community detection methods only consider the network structure but ignore the attributes of nodes. In fact, the attributes of nodes help to improve the performance of community detection, because nodes with similar attributes tend to belong to the same community [23, 24]. Different from network structures that specify node connectivity, node attributes provide the semantics of nodes and underlying network [15]. Therefore, when the nodes in the network are divided into different communities, the node attributes in the same community can reveal the community semantics, which is somewhat similar to the Latent Dirichlet Model (LDA). Thus, the missing structural information can be supplemented and more in-depth community detection can be carried out when semantic information and structural information are used complementarily. Recently, some methods have also been proposed to combine the attributes and structural information for better community detection. They include heuristic-based methods [3, 25] and probabilistic inference-based methods [26, 27]. In addition to obtaining better results of community detection, the node attributes also provide semantic descriptions of the communities. These descriptions help to reveal why certain nodes are divided into a group and understand the functions of communities. Therefore, detecting communities and identifying the underlying semantics of communities make complex network analysis full of significance. Some methods have been developed in [15, 28].

Most methods that have been proposed for community detection are typically only appropriate for assortative community structure; i.e., the nodes within a community are densely connected [22, 29]. They usually assume that such certain structural regularity exists in the target network. However, the assumption may not always correspond to the true intrinsic structure of the network, which limits the applicability of the existing methods. Beyond that, there are other types of important structural regularities in the real networks and the networks may contain multiple structures simultaneously, for example, disassortative structure (bipartite structure) [30], i.e., a kind of structure pattern in which most of the edges are across different communities and mixture structure, i.e., a kind of structure contains both assortative and disassortative structures, and so on. Due to the rich diversity of community regularities in real-world networks, there may be several unknown types of structures in the networks. Therefore, it is urgent to propose some methods to adapt to the realistic situations and to carry out generalized community detection. So, in this paper, we called these assortative and disassortative structures in the complex networks as generalized communities similar to [30]. Some methods [4] have been proposed to detect generalized communities in complex networks.

In particular, although node attributes may carry essential semantic information of communities, there are few ways to detect generalized communities, that is, detecting broad types of network structural regularities and combining network structures and attributes. Chen et al. [31] developed a Bayesian nonparametric attribute (BNPA) model and explored various types of network structures, but the model did not provide multiple semantic descriptions of the communities.

As a result, considering the rich diversity of community regularities in real networks, nodes attributes can not only improve the quality of generalized community detection but also identify the latent semantic characteristics of communities, identify the generalized communities, and provide semantic descriptions, which are worth studying in the complex network analysis. All the above methods neglect solving this twofold problem. Instead, we propose a unified generative model to detect communities in a wide variety of network structures without any prior knowledge of the certain type of intrinsic regularities in the networks. We also derive the semantic descriptions of the communities by combining the network structure and attributes at the same time. Our model is composed of two closely related parts by a probability transition matrix. The first is the topology part in which communities are described based on a mixture model, assuming that nodes in the same groups have similar link patterns (no matter whether there are more links within the communities or between communities). The second is the attribute part, in which semantic information is identified by the classic topic model (LDA) [23]. We assume that each community has several topics; i.e., the distribution of topics exists in each community. A probability transition matrix is used to reveal the potential corrections between topics and communities. It can handle the problem that the topics from attributes and the communities from networks are not well matched. We finally use a Gibbs sampling algorithm to optimize the objective function. Extensive experiments on a number of synthetic and real networks have shown that our model performs better than some baselines on community detection.

In summary, the contributions of this paper are as follows:

(i) As we know, it is the first time we propose the generalized community in the attribute networks, in which the nodes have some link patterns with others and semantic similarity in the network

(ii) We propose a unified generation model to analyze the attribute networks and detect the generalized community structure as well as its semantic description; it can describe the internal relationship between topological structure and node attribute of the network

(iii) We also develop an effective Gibbs sampling algorithm and experiments show its better performance compared with some baselines

2. Related Work

To explore the network structural regularities, some methods for detecting generalized communities have been proposed. Recently, node attributes have attracted extensive attention in the complex network analysis.

Newman and Leicht [30] developed a mixture model to explore the network structure with only links. In this method, the nodes with the same link patterns were divided

into the same groups. It modeled the relationships between communities and nodes. The probability that a node was connected to other nodes in the network was related to the community to which the node belonged. Closely connected nodes may not belong to the same community. Thus, a broad of structural signatures could be explored without any prior assumptions about the structure of the network. Hua-Wei et al. [4] focused on identifying the intrinsic structural rules in networks. In this model, the nodes within the same groups had a similar link preference to other groups. A block matrix was defined to denote the probability that the randomly selected edge linked two distinct groups. It could detect broad types of structural regularities by modeling network structures.

There were several methods for content analysis, such as Latent Dirichlet Model (LDA) [23]. The method focused on node attributes and identified the set of nodes whose attributes were similar. Several community detection approaches combining network topologies and node attributes have also been proposed. Some methods only used node attributes to improve the performance of community detection, while others provided the semantic descriptions of communities. Ruan et al. [25] proposed a method for determining the strength of the edges between nodes using content information, which is also applicable to graph clustering. Yang et al. [27] used a discriminative model that combines node attributes and network topologies to detect communities. However, this method focused on community detection without describing the relevant attributes of each community. It did not provide a semantic description of the community. Pool et al. [28] proposed a heuristic method to detect communities by optimizing the community scores. This heuristic method reported too many relatively small communities, some of which had only two or three nodes. Chakraborty and Sycara [32] developed a model based on nonnegative matrix trifactorization method to detect communities via modeling network structure and contents. However, this method mainly used additional attributed information to identify communities and failed to infer the relationship between communities and attributes. Chen et al. [31] developed a Bayesian nonparametric attribute (BNPA) model to explore structural regularities in networks. This model combined network structures and node attributes for community detection and assumed that network structures and node attributes shared the same community memberships; i.e., attribute clusters and network communities were the same. However, attributes and community structures may not always align at all; they could not give multiple semantic descriptions of communities. Wang et al. [33] proposed a model that combined network topology and node semantic information to identify communities. It integrated topology-based community memberships and node-attributes-based community attributes (or semantics) in the framework of nonnegative matrix factorization. The model was based on two important observations: if the community memberships of two nodes are similar, they will have a high probability to produce adjacent edges, and if their attributes are related to the underlying community attributes, they will likely be in the same community. The use of node contents improved the result of community detection and provided a semantic description to the resultant network communities. He et al. [15] introduced a generative model consisting of two parts, one for communities and the other for semantics, exploring the network structure and interpreting the functional modules semantically. The method was only applicable to the network with assortative structures and failed to detect generalized community. More discussions on attribute networks can be found in related surveys by Bothorel et al. [34] and Chunaev [35].

3. Model Formulation

In this section, we give a formal description of the proposed model, i.e., Generalized Semantic Community (GSC) identification, with the purpose of generalized community detection and semantic identification in the networks.

3.1. Notations. We define an attributed network G with N nodes and M attributes as an $N \times N$ adjacency matrix A and an $N \times M$ attributes matrix X. All the nodes and attributes are denoted as $V = (v_1, v_2, \ldots, v_N)$ and $W = (\omega_1, \omega_2, \ldots, \omega_M)$ in the network. In the adjacency matrix A, $a_{ij} = 1$ if there is an edge from node v_i to node v_j; otherwise, $a_{ij} = 0$. In the attributes matrix X, $x_{it} = 1$ if node i has the t-th attributes ω_t; otherwise, $x_{it} = 0$. Our model is specified by three types of quantities:

(i) Observed quantities: the number of groups K, the number of nodes N, the number of attributes M, the adjacency matrix A, and the attribute matrix X

(ii) Latent quantities: group labels z, where z_i denotes the community membership of node v_i, and the content memberships g, where g_{it} denotes the topic labels of the node v_i's t-th attribute

(iii) Model parameters: $\pi = (\pi_r)_{1\times K}$, where π_r is the fraction of nodes in community r; $\theta = (\theta_{rj})_{K\times N}$, where θ_{rj} is the probability that a certain node in community r connects to node v_j; $\eta = (\eta_{rs})_{K\times K}$, where $\eta_{rs} = p(g_{it} = s \mid z_i = r)$ is the probability that node v_i is in the s-th content cluster given that the community label is r; $\phi = (\phi_{st})_{K\times M}$, where $\phi_{st} = p(x_{it} = 1 \mid g_{it} = s)$ is the probability that the s-th topic generates t-th attributes of node v_i

Table 1 shows the notations of the parameters.

3.2. Problem Definition. Considering the rich diversity of community regularities in real networks, encoding network structure and node attributes simultaneously, and providing the semantic descriptions of the resultant network communities are still the problems that are worth studying in the community detection. However, most existing methods tend to ignore certain aspects of the problems that remain the challenges of current community detection. Given an attributed network, the goal of handling these problems is twofold:

Table 1: Notations.

Type	Symbol	Description
Observed quantities	A	Adjacency matrix
	X	Attribute matrix
	K	Number of communities and topics
Latent quantities	z_i	Community assignment of node v_i
	g_{it}	Topic labels of the node v_i's t-th attribute
Model parameters	π_r	The fraction of nodes in community r
	θ_{rj}	The probability that a certain node in community r connects to node v_j
	η_{rs}	The probability that node v_i is in the s-th content cluster given that the community label is r
	ϕ_{st}	The probability that the s-th topic generates t-th attributes of node v_i
Hyperparameters	$\alpha, \beta, \gamma, \xi$	Acting as noninformation prior of corresponding model parameters with prior distribution

(i) How to divide the nodes into communities and content clusters no matter what kind of network structural regularity the network is?

(ii) How to identify the correlations between communities and attribute topics to provide the best semantic descriptions of communities?

So the problem can be formalized as, given the adjacency matrix A and attributes matrix X as well as the number of communities K, our goal is to obtain the community assignment z_i for each node i and the topic distribution of the communities.

3.3. Model Definition. To achieve the objective, we define a unified Bayesian probabilistic generative model to handle topologies and node attributes at the same time. Our goal is to divide the nodes in networks with extensive structural regularities into K communities and K content clusters, respectively, by using adjacency matrix A as well as attributes matrix X. To model network structure, we assume that the nodes in the same groups have similar link patterns; i.e., the probability of a node connecting to other nodes in the network is the link tendency between the community to which the node belongs and the rest of nodes. We also take a modified LDA model for node attributes. A transition matrix is used to jointly model network structures and node attributes, which connects network communities and attribute topics. To be specific, a community may be characterized by multiple topics, and the topic of each node attribute is derived from the topic distribution of the community to which each node belongs. Then, by extracting the latent correlation between network communities and attributes clusters, multiple semantic interpretations can be provided for each community. Figure 1 shows a graphical representation of this model, and the generation process is as follows:

(1) Sample $\pi \sim \text{Dirichlet}(\alpha)$

(2) For each community $r \in \{1, 2, \ldots, K\}$

(a) Sample $\theta_r \sim \text{Dirichlet}(\beta)$

(b) Sample $\eta_r \sim \text{Dirichlet}(\gamma)$

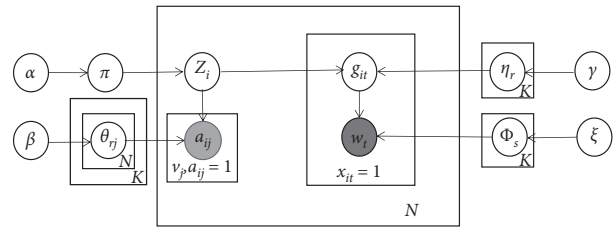

Figure 1: The graphical representation of model.

(3) For each topic s in K topics

(a) Sample $\phi_s \sim \text{Dirichlet}(\xi)$

(4) For each new node v_i, $i \in \{1, 2, \ldots, N\}$

(a) Sample a latent group assignment $z_i \sim \text{Multinomial}(\pi)$

(b) For each node v_j with $v_i \neq v_j$:

(i) Sample edge $a_{ij} \sim \text{Multinomial}(\theta_{z_i})$

(c) For each of the t-th attribute ω_t with $x_{it} = 1$:

(i) Sample $g_{\text{it}} \sim \text{Multinomial}(\eta_{z_i})$

(ii) Sample attribute $\omega_t \sim \text{Multinomial}(\phi_{g_{it}})$.

3.3.1. Generating Model Parameters. We introduce a Bayesian treatment into the model generation process. After the number of communities K is given, model parameters are treated as random variables; we generate model parameters $\pi = (\pi_r)_{1\times K}$, $\theta = (\theta_{rj})_{K\times N}$, $\eta = (\eta_{rs})_{K\times K}$, and $\phi = (\phi_{st})_{K\times M}$, respectively, by the Dirichlet distribution. The parameters are generated based on some hyperparameters, denoted as a K-dimensional vector $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_r, \ldots, \alpha_K)$, an N-dimensional vector $\beta = (\beta_1, \beta_2, \ldots, \beta_j, \ldots, \beta_N)$, a K-dimensional vector $\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_s, \ldots, \gamma_K)$, and an M-dimensional vector $\xi = (\xi_1, \xi_2, \ldots, \xi_t, \ldots, \xi_M)$. The generative process is as follows.

We use Dirichlet distributions to generate the following model parameters, respectively:

$$\begin{aligned}\pi &= (\pi_1, \pi_2, \ldots, \pi_r, \ldots, \pi_K),\\ \theta &= (\theta_1, \theta_2, \ldots, \theta_r, \ldots, \theta_K),\\ \eta &= (\eta_1, \eta_2, \ldots, \eta_r, \ldots, \eta_K),\\ \phi &= (\phi_1, \phi_2, \ldots, \phi_s, \ldots, \phi_K),\end{aligned} \tag{1}$$

$$\begin{aligned}p(\pi \mid \alpha) &= \frac{\Gamma\left(\sum_{r=1}^{K}\alpha_r\right)}{\prod_{r=1}^{K}\Gamma(\alpha_r)}\prod_{r=1}^{K}\pi_r^{\alpha_r-1},\\ p\left(\theta_r \mid \beta\right) &= \frac{\Gamma\left(\sum_{j=1}^{N}\beta_j\right)}{\prod_{j=1}^{N}\Gamma\left(\beta_j\right)}\prod_{j=1}^{N}\theta_r^{\beta_j-1},\\ p\left(\eta_r \mid \gamma\right) &= \frac{\Gamma\left(\sum_{s=1}^{K}\gamma_s\right)}{\prod_{s=1}^{K}\Gamma(\gamma_s)}\prod_{s=1}^{K}\eta_r^{\gamma_s-1},\\ p\left(\phi_s \mid \xi\right) &= \frac{\Gamma\left(\sum_{t=1}^{M}\xi_t\right)}{\prod_{t=1}^{M}\Gamma(\xi_t)}\prod_{t=1}^{M}\phi_s^{\xi_t-1},\end{aligned} \tag{2}$$

where $\Gamma(\bullet)$ represents a Gamma function. All the communities share the same β, and all the topics share the same γ and ξ.

3.3.2. Generating Observed and Latent Quantities. At first, we sample the latent community membership z_i for every node v_i from a multinomial distribution independently. It is described as

$$p(z_i = r \mid \pi) = \pi_r, \quad r = 1, 2, \ldots, K. \tag{3}$$

After the latent community membership z_i of nodes v_i is explicit, we generate edge a_{ij} as the following definition:

$$p\left(a_{ij} \mid \theta_{z_i}\right) = \theta_{z_i j}^{a_{ij}}, \tag{4}$$

where θ_{rj} denotes the "preferences" for any node in community r to link to node v_j, regardless of which community that node v_j is in. Nodes in the same community have a common link "preference" without any assumptions about network structure regularities. Thus, generalized communities can be detected. Then, we sample the latent topics membership g_{it} for each attribute ω_t of node v_i from a multinomial distribution independently, defined as

$$p\left(g_{it} = s \mid \eta_{z_i}\right) = \eta_{z_i s}. \tag{5}$$

As η_{rs} denotes the probability that node v_i is in the s-th semantic topic while it is divided into r-th community, that is, η_{rs} provides the transition from communities to topics, the topic assignment and community membership of node do not always match well. This is why the community may have several topics.

We generate attributes ω_t as the following definition:

$$p\left(x_{it} \mid \phi_{g_{it}}\right) = \phi_{g_{it}t}^{x_{it}}. \tag{6}$$

Then, the probability of the network G with N nodes and M attributes is

$$\begin{aligned}p(A, X, z, g, \pi, \theta, \eta, \phi \mid \alpha, \beta, \gamma, \xi) &= p(A \mid z, \theta)p(X \mid g, \phi)p(g \mid z, \eta)p(z \mid \pi)p(\pi \mid \alpha)p(\theta \mid \beta)p(\eta \mid \gamma)p(\phi \mid \xi)\\ &= \prod_{i=1}^{N}\left(\pi_{z_i}\prod_{j=1}^{N}\theta_{z_i,j}^{A_{ij}}\right)\cdot\prod_{i=1}^{N}\prod_{t=1}^{M}\left(\eta_{z_i g_{it}}\right)^{x_{it}}\cdot\prod_{i=1}^{N}\prod_{t=1}^{M}\left(\phi_{g_{it}t}\right)^{x_{it}}\cdot\frac{\Gamma\left(\sum_{r=1}^{K}\alpha_r\right)}{\prod_{r=1}^{K}\Gamma(\alpha_r)}\prod_{r=1}^{K}\pi_r^{\alpha_r-1}\\ &\cdot\prod_{r=1}^{K}\frac{\Gamma\left(\sum_{j=1}^{N}\beta_j\right)}{\prod_{j=1}^{N}\Gamma\left(\beta_j\right)}\prod_{j=1}^{N}\theta_{rj}^{\beta_j-1}\cdot\prod_{r=1}^{K}\frac{\Gamma\left(\sum_{s=1}^{K}\gamma_s\right)}{\prod_{s=1}^{K}\Gamma(\gamma_s)}\prod_{s=1}^{K}\eta_{rs}^{\gamma_s-1}\cdot\prod_{s=1}^{K}\frac{\Gamma\left(\sum_{t=1}^{M}\xi_t\right)}{\prod_{t=1}^{M}\Gamma(\xi_t)}\prod_{t=1}^{M}\phi_{st}^{\xi_t-1}.\end{aligned} \tag{7}$$

It is subject to $\sum_{r=1}^{K}\pi_r = 1$, $\sum_{j=1}^{N}\theta_{rj} = 1$, $\sum_{s=1}^{K}\eta_{rs} = 1$, and $\sum_{t=1}^{M}\phi_{st} = 1$.

4. Model Optimization

To exactly infer that the latent variables z and g are intractable, we use Gibbs sampling [36] and slice sampling [37] to sample the latent variables z and g and hyperparameters (α, β, and γ), respectively.

4.1. Inference. Because the Dirichlet and Multinomial distributions are conjugate, equation (2) can be simplified as

$$
\begin{aligned}
p(A, X, z, g \mid \alpha, \beta, \gamma, \xi) &= p(A \mid z, \beta) \cdot p(z \mid \alpha) \cdot p(X \mid g, \xi) \cdot p(g \mid z, \gamma) \\
&= \prod_{r=1}^{K} \frac{\Gamma\left(\sum_{j=1}^{N} \beta_j\right)}{\prod_{j=1}^{N} \Gamma\left(\beta_j\right)} \frac{\prod_{j=1}^{N} \Gamma\left(m_r^j + \beta_j\right)}{\Gamma\left(\sum_{j=1}^{N} \left(m_r^j + \beta_j\right)\right)} \cdot \frac{\Gamma\left(\sum_{r=1}^{K} \alpha_r\right)}{\prod_{r=1}^{K} \Gamma\left(\alpha_r\right)} \frac{\prod_{r=1}^{K} \Gamma\left(n_r + \alpha_r\right)}{\Gamma\left(\sum_{r=1}^{K} \left(n_r + \alpha_r\right)\right)} \\
&\cdot \prod_{r=1}^{K} \frac{\Gamma\left(\sum_{s=1}^{K} \gamma_s\right)}{\prod_{s=1}^{K} \Gamma\left(\gamma_s\right)} \frac{\prod_{s=1}^{K} \Gamma\left(N_r^s + \gamma_s\right)}{\Gamma\left(\sum_{s=1}^{K} \left(N_r^s + \gamma_s\right)\right)} \cdot \prod_{s=1}^{K} \frac{\Gamma\left(\sum_{t=1}^{M} \xi_t\right)}{\prod_{t=1}^{M} \Gamma\left(\xi_t\right)} \frac{\prod_{t=1}^{M} \Gamma\left(M_s^t + \xi_t\right)}{\Gamma\left(\sum_{t=1}^{M} \left(M_s^t + \xi_t\right)\right)},
\end{aligned} \tag{8}
$$

with

$$
\begin{aligned}
p(z \mid \alpha) &= \int p(z \mid \pi) p(\pi \mid \alpha) \mathrm{d}\pi, \\
p(g \mid z, \gamma) &= \int p(g \mid z, \eta) p(\eta \mid \gamma) \mathrm{d}\eta, \\
p(A \mid z, \beta) &= \int p(A \mid z, \theta) p(\theta \mid \beta) \mathrm{d}\theta, \\
p(X \mid g, \xi) &= \int p(X \mid g, \phi) p(\phi \mid \xi) \mathrm{d}\phi,
\end{aligned} \tag{9}
$$

where m_r^j denotes the number of outlinks whose tail nodes belong to r and whose head node is v_j; n_r denotes the number of nodes in community r; M_s^t denotes the number of ω_t which is generated by topic s; and N_r^s denotes the total number of topics s generated by community r.

The inference process is in Algorithm 1.

4.1.1. Sampling z. For each node v_i, given the community assignment for all other nodes, the community probability of the node z_i choosing community r is

$$
\begin{aligned}
p\left(z_i = r \mid z_i, g, A, X\right) \propto & \prod_{j=1}^{L_i} \frac{m_{r,i}^j + \beta_j}{m_{r,i} + \sum_{j=1}^{N} \beta_j + j - 1} \cdot \frac{n_r + \alpha_r}{N + \sum_{r=1}^{K} \alpha_r} \\
& \cdot \prod_{s \in g_i} \prod_{t \in m_i, g_{it} = s} \left(M_{s,i}^t + \xi_t\right) \prod_{j=1}^{N(is)} \frac{1}{M_{s,i} + \sum_{t=1}^{M} \xi_t + j - 1} \\
& \cdot \prod_{s \in g_i} \prod_{j=1}^{N(is)} \left(N_{r,i}^s + \gamma_s + j - 1\right) \prod_{j=1}^{m_i} \frac{1}{N_{r,i} + \sum_{s=1}^{K} \gamma_s + j - 1},
\end{aligned} \tag{10}
$$

where L_i denotes the outlinks of nodes v_i; $m_{r,i}$ denotes the number of outlinks from community r except node v_i; $m_{r,i}^j$ denotes the number of outlinks from community r except edges a_{ij}; n_r denotes the number of nodes in community r; N is total number of nodes; g_i denotes the topic labels of the attributes of v_i; m_i denotes the attributes of v_i; g_{it} denotes the topic of v_i's t-th attribute; $M_{s,i}^t$ denotes the number of node attributes whose topic is s except v_i's attribute ω_t; $M_{s,i}$ denotes the number of nodes' attributes whose topic is s except the attributes of v_i; $N_{r,i}^s$ denotes the total number of topics s generated by community r except v_i; and $N(\mathrm{is})$ denotes the attributes of v_i whose topic is s.

4.1.2. Sampling g_i. For node v_i in community r, given the topic assignment for all the attributes except the attribute ω_t, the topic probability of the attribute ω_t choosing topic s is

$$
\begin{aligned}
p\left(g_{it} = s \mid g_{it}, z, A, X\right) \propto & \left(M_{s,it}^t + \xi_t\right) \cdot \frac{1}{M_{s,it} + \sum_{t=1}^{M} \xi_t} \\
& \cdot \left(N_{r,it}^s + \gamma_s\right) \cdot \frac{1}{N_{s,it} + \sum_{s=1}^{K} \gamma_s},
\end{aligned} \tag{11}
$$

where $M_{s,it}^t$ denotes the number of ω_t whose topic is s except v_i's attribute ω_t; $M_{s,it}$ denotes the number of all the attributes whose topic is s except v_i's attribute ω_t; $N_{r,it}^s$ denotes the number of nodes' attributes whose topic is s and whose nodes belong to community r except v_i's attribute ω_t; and $N_{r,it}$ denotes the number of node attributes that belong to community r except v_i's attribute ω_t.

4.2. GSC Models. Our model can also only handle edges or nodes' attributes in the networks.

Require: adjacency matrix A, attributes matrix X, iterations T, and specified group number K
Ensure: group assignment z
0: initialize α, β, γ, ξ, set n_r, m_r, m_r^j, M_s, M_s^t, N_r, and N_r^s to 0
Initialize each node's latent community label z_i
(1)//sampling z, g_i, α, β, γ, and ξ
(2)for $te = 1$ to T do
(3) **for** $i = 1$ to N do
(4) //get the current community assignment of node v_i
(5) update n_r, m_r, m_r^j, M_s, M_s^t, N_r, and N_r^s
(6) **for** $k = 1$ to K do
(7) compute probability $p(z_i = k)$ according to equation (8)
(8) **end for**
(9) Gibbs sampling for z and obtain $z_i = r$
(10) update n_r, m_r, m_r^j, M_s, M_s^t, N_r, and N_r^s
(11) **for** $t = 1$ to M do
(12) //get the current topic assignment of attribute ω_t
(13) update M_s, M_s^t, N_k, and N_k^s
(14) **for** $s = 1 = 1$ to K do
(15) compute probability $p(g_{it} = s)$ according to equation (9)
(16) **end for**
(17) Gibbs sampling for g_i and obtain $g_{it} = s$
(18) update M_s, M_s^t, N_k, and N_k^s
(19) **end for**
(20) **end for**
(21) slice sampling for α, β, γ, and ξ in (0, 1)
(22)**end for**

Algorithm 1: Inference for GSC.

4.2.1. GSC-Link. The probability of only considering the links can be written as

$$p(A, z \mid \alpha, \beta) = p(A \mid z, \beta) \cdot p(z \mid \alpha). \tag{12}$$

The community probability of node i choosing community k is

$$p\left(z_i = r \mid z_i, A\right) \propto \prod_{j=1}^{L_i} \frac{m_{r,i}^j + \beta_j}{m_{r,i} + \sum_{j=1}^{N} \beta_j + j - 1} \cdot \frac{n_r + \alpha_r}{N + \sum_{r=1}^{K} \alpha_r}. \tag{13}$$

4.2.2. GSC-Attr. The probability of only considering the attributes can be written as

$$p(X, z, g \mid \alpha, \gamma, \xi) = p(X \mid g, \xi) \cdot p(g \mid z, \gamma) \cdot p(z \mid \alpha). \tag{14}$$

The community probability of node i choosing community k is

$$\begin{aligned} p\left(z_i = r \mid z_i, g, X\right) \propto & \cdot \prod_{s \in g_i} \prod_{t \in m_i, g_{it} = s} \left(m_{s,i}^t + \xi_t\right) \prod_{j=1}^{N(is)} \frac{1}{m_{s,i} + \sum_{t=1}^{M} \xi_t + j - 1} \\ & \cdot \prod_{s \in g_i} \prod_{j=1}^{N(is)} \left(N_{k,i}^s + \gamma_s + j - 1\right) \prod_{j=1}^{m_i} \frac{1}{N_{k,i} + \sum_{s=1}^{K} \gamma_s + j - 1}. \end{aligned} \tag{15}$$

The topic probability of the attribute ω_t choosing topic s is the same as GSC.

5. Experiments and Analysis

Firstly, we experiment on three different synthetic networks with different structure regularities (i.e., assortative, disassortative, and mixture structures) to evaluate the quality of community detection and analyze the superiority of modeling on the network with a rich diversity of structures. Then, we assess the interpretability of communities in an online music system. Finally, we evaluate on real networks and do a comparison with state-of-the-art methods.

As the ground truth of communities in the networks is known, we use the following Normalized Mutual Information (NMI) [38] to compare all the methods:

$$NMI(G, G') = \frac{2MI(G, G')}{H(G) + H(G')}, \quad (16)$$

where $G = (G_1, G_2, \ldots, G_k)$ is the ground truth of communities in the network, and $G' = (G'_1, G'_2, \ldots, G'_k)$ is the community identified by the method. $H(G)$ and $H(G')$ are the entropies of G and G', respectively, and $MI(G, G')$ denotes the mutual information between them. The higher NMI is, the better the result is.

To describe parameter estimation in GSC more adequately, we describe the changing trend of likelihood function with the number of iterations in Figure 2(a), and each curve in Figure 2(b) shows the changes of the log-likelihood of Cora with one of four hyperparameters when other hyperparameters are determined by slice sampling. It can be seen that the log-likelihood of GSC quickly converges at about 150th iteration. The log-likelihood probability is less sensitive to α, β, and ξ while γ made a big difference.

5.1. Experiment on Synthetic Networks with Different Structure Regularities. Firstly, we conduct experiments on synthetic networks to evaluate the quality of community detection. Then, we assess on real networks and do a comparison with state-of-the-art methods.

The first synthetic network is a random network in Newman's method [15]. The network consists of 128 nodes divided into 4 disjoint communities with $z_{in} + z_{out} = 16$. As $\rho(= z_{in}/32) > \rho(= z_{out}/96)$, z_{in} (the edges linking to nodes within community) is much larger than z_{out} (the edges linking to nodes in other communities). For every node v_i, we generate a $4h$-dimensional binary attribute (i.e., x_i) to divide the nodes of 4 content clusters with $h_{in} + h_{out} = 16$. In this paper, h_{in} denotes the number of attributes for every node v_i with $x_{it} = 1$ associated with its community and h_{out} (noisy attribute) denotes the number of attributes for every node v_i with $x_{it} = 1$ corresponding to the other communities. In particular, we generate the $(s - 1 \times h + 1)$-th to $(s \times h)$-th attributes for each node in the s-th cluster by a binomial distribution with mean $\rho_{in} = h_{in}/h$ and generate the remaining attributes by the binomial distribution with mean $\rho_{out} = h_{out}/(3h)$.

We set $h = 50$ and consider that the topologies and contents share the same membership. The node attributes' matrix and the community attributes' matrix are shown in Figure 3. We first set $z_{out} = 8$ and change h_{out} from 0 to 12 with an increment of 1. We adapt GSC-link using network topology alone as the baseline method. Other comparison methods are NEMBP [15] and SCI [33], which use both network topologies and attributes. As shown in Figure 4(a), our method can use the complementary structural information in node attributes to improve the quality of community detection when $h_{out} < 12$. Even when $h_{out} = 12$, the cluster structures of node attributes disappear; our model GSC can get better results than baseline method GSC-link. Then we set $h_{out} = 8$ and change z_{out} from 0 to 9 with an increment of 1. As shown in Figure 4(b), our method also can perform better than GSC-attr. In general, the proposed method can get better results of community detection by using topology and content information.

The second synthetic network is Newman's model [30] of 108 nodes. It consists of 8 keystone nodes without community labels and other nodes link to them according to their community membership. The remaining 100 nodes are equally divided into 4 groups, and the edges between these nodes are randomly linked, with the mean degree of every node being 10. The keystone nodes are {101, 102, 103, 104}, {103, 104, 105, 106}, {105, 106, 107, 108}, {101, 102, 107, 108}.

In particular, each community has a unique signature set of keystones, and only the link pattern to keystones can identify the community; thus the structure of this network is neither assortative nor disassortative.

At first, we study the influence of noise attributes on community detection. ρa represents the proportion of noisy attributes of each node. We change the probability of noisy attributes ρa from 0 to 1 with an increment of 0.1. The node attributes' matrix is shown in Figure 5. When ρa becomes larger, the attributes associated with each community are blurred and less discriminant information is provided for the network community. As shown in Figure 6(b), we almost divide the nodes into 3 communities while only considering network structure. The result gets better when using node attributes in Figure 6(c). As shown in Figure 7(a), our method outperforms GSC-link (even ρa reaches 0.7) and significantly outperforms SCI and NEMBP. It shows that the quality of identified communities improves combining node attributes and network structures. Our model GSC is able to fully use network structure information even if the information of node attributes is erroneous. As ρa increases beyond 0.7, GSC performs worse. It also reveals that node attributes with terrible quality can lower the result of community detection. Figure 7(a) also shows that NEMBP performs worse than GSC-link when ρa reaches 0.2 and SCI performs always much worse. Figures 6(b) and 6(c) represent the results of GSC and NEMBP, respectively, when ρa is 0.5. It can be concluded from the above analysis that GSC is more capable of identifying the networks with mixed structural regularities than SCI and NEMBP.

In this network, the propensity to link to the unique set of keystone nodes determines the group membership. We change the keystone links of each group to change the network structure by varying the keystone links of each group from 100 to 10 with a decrement of 10. We set the probability of noisy attributes $\rho a = 0.5$. We adapt our model with only attributes as the baseline method and NEMBP for comparison. As can be seen in Figure 7(b), our method is also able to perform well even if the keystone links are only 30. The new model represents strong robustness to the changes of network structure. However, the rambling result of NEMBP indicates that it does not work very well for this type of network.

The third network [31] has both a community and a bipartite structure with 100 nodes and 402 edges as shown in Figure 6(e). The 100 nodes are equally divided into 5 groups,

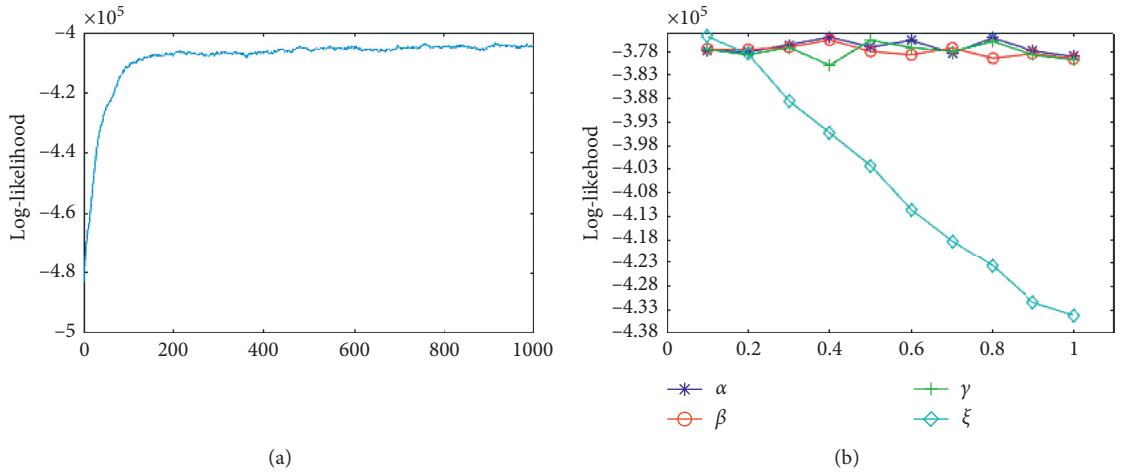

FIGURE 2: (a) Trend of the log-likelihood probability of Cora with iterations. (b) Trend of the log-likelihood probability of Cora with α, β, γ, and ξ, respectively.

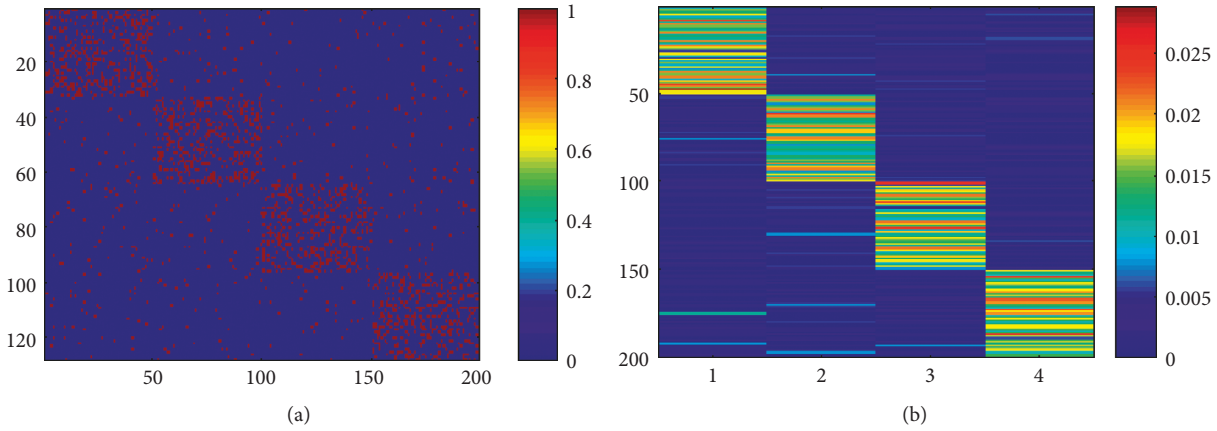

FIGURE 3: (a) The node attributes' matrix. (b) The community attributes' matrix.

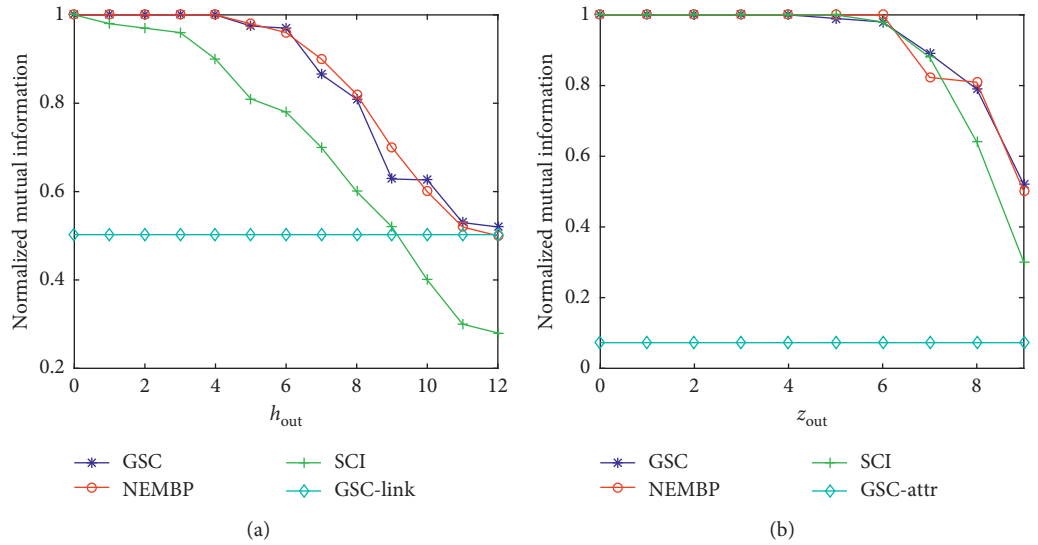

FIGURE 4: The value of NMI of three methods on random networks. (a) h_{out}. (b) z_{out}.

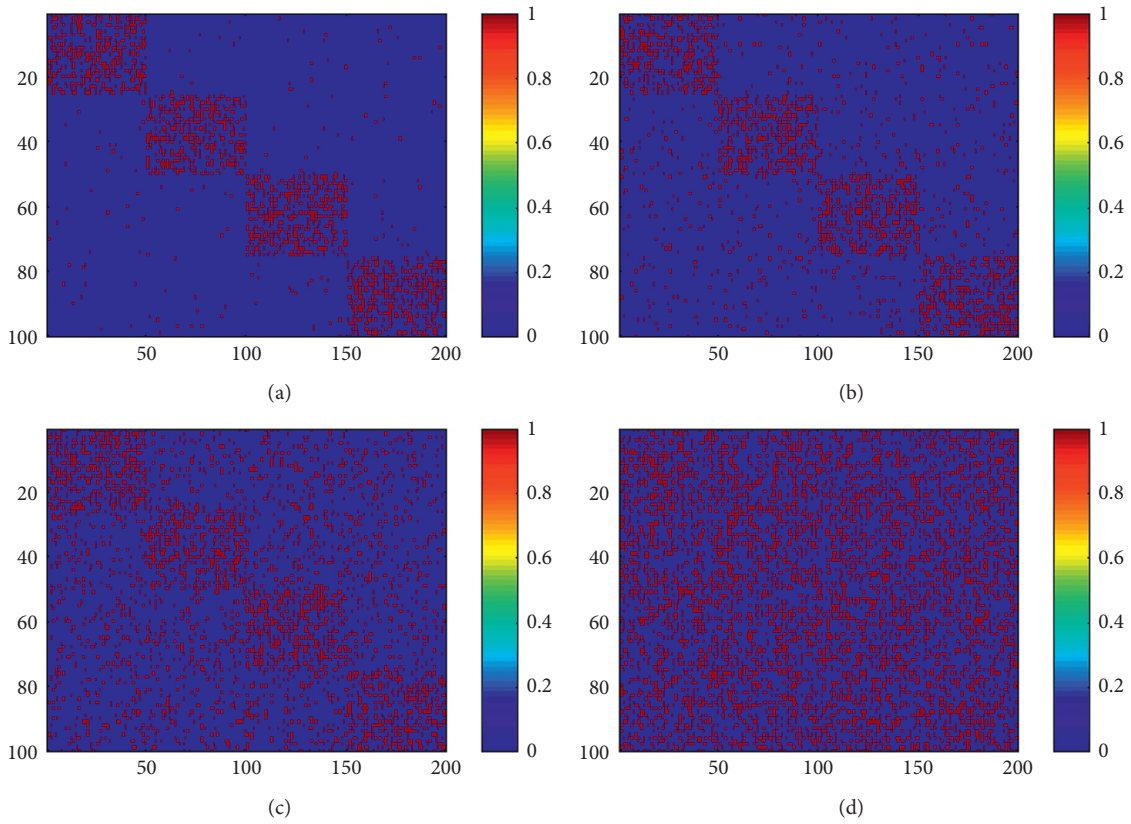

Figure 5: The node attributes' matrix with different ρa. (a) $\rho a = 0.1$. (b) $\rho a = 0.3$. (c) $\rho a = 0.5$. (d) $\rho a = 0.7$.

three of which form an assortative structure, whereas the remaining two form a bipartite structure. For each node v_i, we generate 5×50-dimensional binary attributes; each of the communities and nodes has 50-dimensional relevant attributes. We change the probability of noisy attributes ρa from 0 to 1 with an increment of 0.1. As shown in Figure 7(c), our model always gets better results than NEMBP and SCI. Even when $\rho a = 0.6$, the quality of identified communities is also improved compared with GSC-link, and the NMI is almost 1. Figure 6(f) shows the result of NEMBP when $\rho a = 0.6$. Its performance is much worse than that of GSC.

5.2. Evaluating Efficiency. In this part, we evaluate the efficiency of community detection methods by measuring each method's running time on synthetic networks as we increase the network size. The comparison methods are NEMBP and SCI. The synthetic networks include assortative and disassortative structures. The edges are placed uniformly at random within and between communities in certain numbers. The number of edges within each community is set to 1,200 and the number of edges between a community and the others is set to 600. They form a community structure. The rest of the communities are divided in pairs, the number of edges between two communities in each pair is set to 2,400, and the number of edges between communities in different pairs is set to 1,200. Each pair of groups forms a bipartite structure. The maximum number of nodes in our synthetic network is 7,000, including 12,6000 edges and 700 attributes. We change the scale of the network (Syn-100, Syn-500, Syn-1000, Syn-2000, Syn-3000, Syn-5000, and Syn-7000). The synthetic network of 100 is the third network that we used above. For each synthetic network, we generate 10 K-dimensional binary attributes. We set the ratio of noise attributes to 0.5.

Figure 8 shows the running time of methods versus the network size. Our method is the fastest among the three. When the program runs to convergence, the running time of our method on Syn-7000 is about 5 minutes. For NEMBP, we set the number of iterations in the program to 10; the running time of the program can reach 11 hours even on Syn-2000. The running time of SCI is more than 19 hours.

5.3. A Case Study. In this paper, we use $(\eta_{rs})_{K \cdot K}$ to correlate the communities and attribute topics and evaluate whether it contributes to the descriptions of the communities. We intensively analyze the underlying semantics of communities and provide particular descriptions for some of the communities detected by GSC. Thus, we use the LASTFM dataset, which is a social network from an online music system, that is, Last.fm. It includes 1,892 users and 11,946 attributes of user's favorite music singers and tag assignment. In this network, the ground truth of community partition is unknown, so we decide to detect 38 communities as in [15]. We find that the communities may have one main

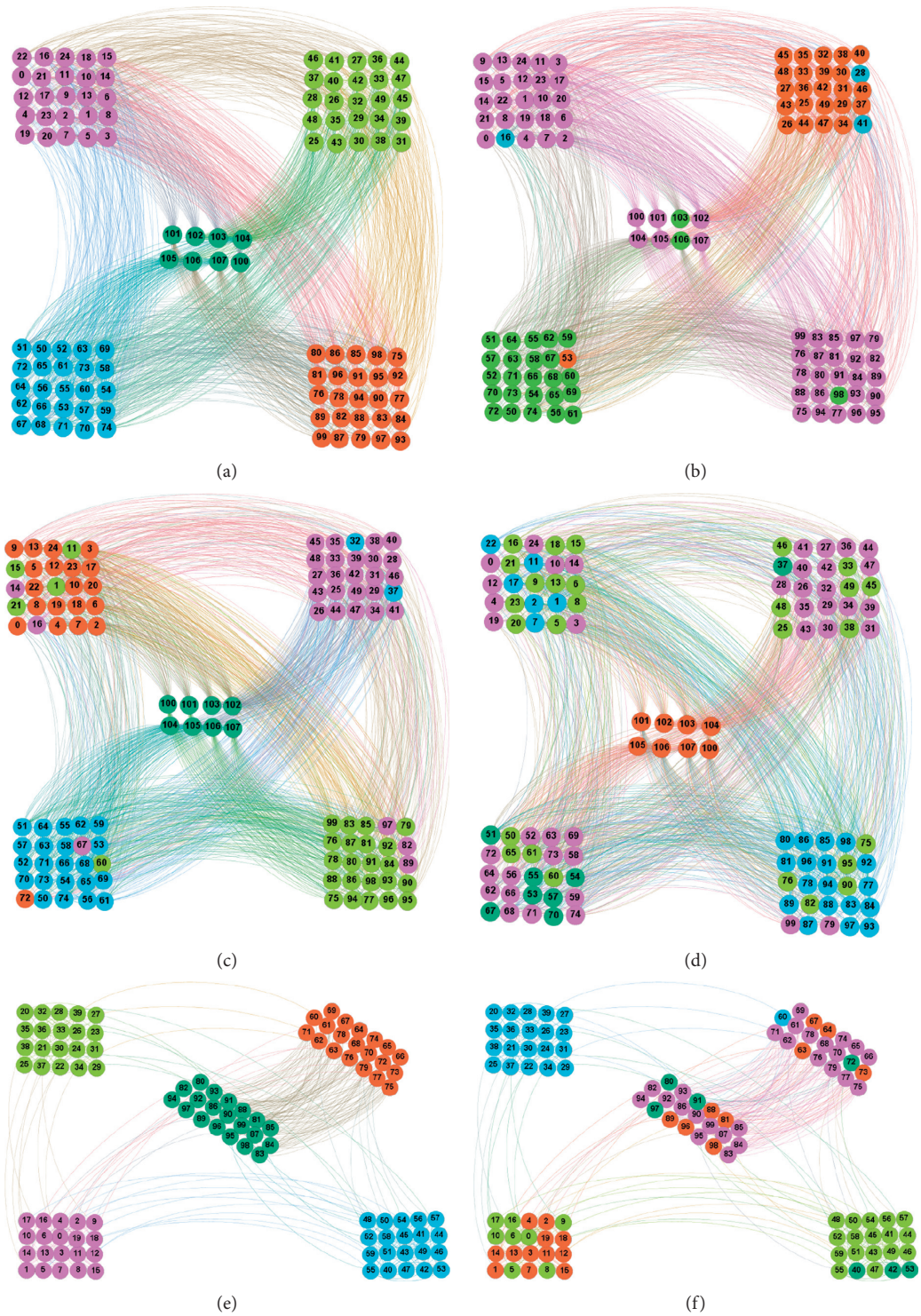

FIGURE 6: Communities detected by GSC and NEMBP models on two synthetic networks. (a) The real community assignment of the synthetic network of 108 nodes. (b) The result of community detection by GSC-link. (c) The result of community detection by GSC, $\rho a = 0.5$. (d) The result of community detection by NEMBP, $\rho a = 0.5$. (e) The real community assignment of the synthetic network of 100 nodes. (f) The result community detection by NEMBP. The nodes of the same communities are in the same colors.

topic or multiple topics; a detailed analysis of the three detected communities with different topics is shown in Figure 9.

The first example in Figure 9(a) is a community with one main topic. It should be the fans of popular female singers like "Rihanna" and "Britney Spears." Their music are "pop," "rock," and "dance." They are both "female vocalists" and "sexy." As for the community in Figure 9(b), it is a group of fans of "hardcore punk" music. The hardcore punk is

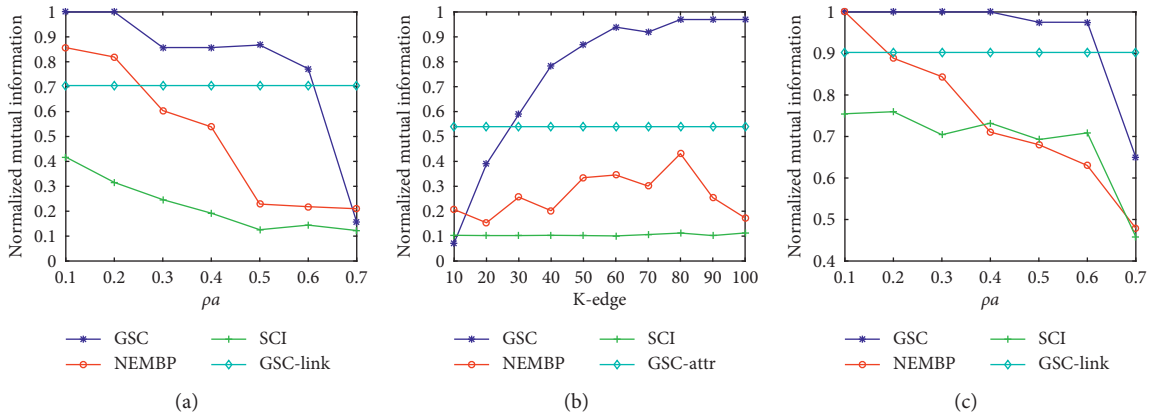

Figure 7: The value of NMI of four methods: (a) on Synthetic 108 with the change of ρa from 0.1 to 0.7, (b) on Synthetic 108 with the change of keystone links from 10 to 100, and (c) on Synthetic 100 with the change of ρa from 0.1 to 0.7.

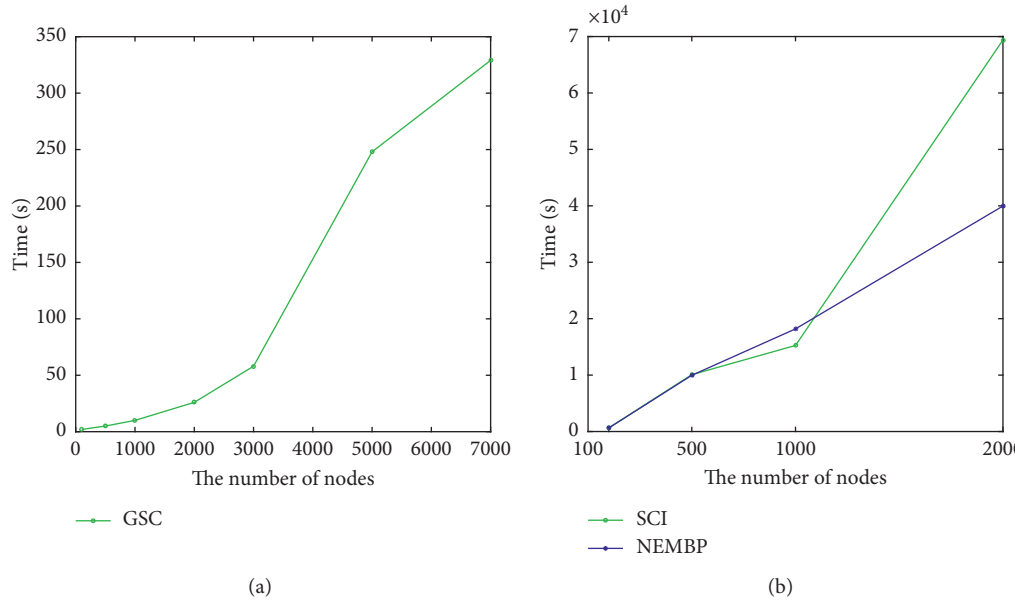

Figure 8: The running time on synthetic networks: (a) GSC on Synthetic networks with the change of nodes from 100 to 7000; (b) NEMBP and SCI on Synthetic networks with the change of nodes from 100 to 2000.

also labeled as hard rock. Glam-sleaze music is a derivative of hard rock and alternative rock coming from a post-punk band. Grunge music is a music genre of indie rock which evolved from hardcore punk. Emotionally-Driven Hardcore Punk (EMO) is an indie rock style, and the Screamo originated from EMO. The last community has two major topics. The communities shown in Figures 9(c) and 9(d) are about the fans of electronic music. One topic is mainly about Electronic Body Music (EBM), which combines elements of industrial music and electronic punk music. The other topic is about IDM. This kind of music was created in the late 80s accompanied by hard edge dance and slow music.

5.4. Experiment on Real Networks. Cora, Citeseer, Terrorist, and Biology are four real networks with both links and contents that we apply in this paper. Cora is a part of Cora citation networks, including 2,708 published articles and 5,429 edges. Each publication is represented by a 1,433-dimensional binary word vector which means the absence or presence of the relating words. The total publications are divided into seven communities. Citeseer is a subset of Citeseer citation networks. It includes 3,312 published articles and 4,732 edges. Each publication is represented by a 3,703-dimensional binary word vector. The total publications are divided into six communities. The Terrorist dataset consists of 1,293 terrorist attacks; each attack is assigned one of 6 labels indicating the type of the attack. Each attack is described by a 106-dimensional binary word vector whose entries indicate the absence or presence of a feature. Biology is a real paper citation network, which is from 435 different biological journals. It contains 10,000 papers connected by links. Each paper is described by a 9,944 0/1-valued keyword vector; two papers are connected if they have a reference relationship. There are 435 nodes representing different

Figure 9: The examples show the word clouds of the main attributes of communities. The sizes of word indicate the probability that they belong to a topic.

biological journals in the network; each paper links to them according to the journal in which it is published. So, the network forms a mixture structure that is similar to the synthetic network of 108. All the papers are split into 435 groups; each group contains papers published in a certain journal. We also use Syn-2000, which includes both community and bipartite structure. The five networks are shown in Table 2.

We compare our GSC model with the methods from three categories: (1) models based on only network structures, that is, GSC-link; (2) models based on only network attributes, such as GSC-attr and LDA; (3) models based on both structures and attributes, such as PCL-DC, NMMA, SCI, and NEMBP.

The results of these models on three networks are shown in Table 3. Our model can use the information of network structure and node attributes simultaneously to identify communities. The model GSC outperforms the other models on Cora and achieves larger NMIs than most of models on Citeseer and Terrorist. The result of GSC is lower than that of NMMA on Citeseer. This is mainly due to the fact that network structures and node attributes are more likely to share the same community memberships. NMMA assumed that attribute clusters and network communities were the same, so it performs better on Citeseer. Sometimes, the community structure is not so obvious when considering only the structural information of the network. The nodes are divided into communities mainly by using their attributes. In this situation, our model can effectively use the information of the attributes. The models based on structure and attributes usually outperform the models with only link or attributes.

Table 2: Statistical characteristics of five networks.

Datasets	N (nodes)	E (edges)	M (attributes)	K	Type
Cora	2708	5429	1433	7	Community
Citeseer	3312	4732	3703	6	Community
Terrorist	1293	3172	106	6	Community
Biology	10000	22662	9944	435	Mixture
Syn-2000	2000	36000	200	20	Mixture

Table 3: NMI (%) of different models on five networks with node attributes.

Models	Datasets					Type
	Cora	Citeseer	Terrorist	Biology	Syn-2000	
GSC-link	16.33	4.696	1.67	26.43	95.86	Link
GSC-attr	28.86	24.31	30.03	4.28	90.26	Attr
LDA	14.61	9.13	**31.95**	5.42	89.60	Attr
PCL-DC	17.54	2.99	5.32	3.29	88.32	Link + attr
NMMA	41.57	**39.95**	25.59	6.86	94.28	Link + attr
SCI	19.26	4.87	8.73	N/A	81.57	Link + attr
NEMBP	44.08	24.27	9.37	N/A	78.68	Link + attr
GSC	**45.96**	25.13	30.45	**29.45**	**99.64**	Link + attr

6. Conclusions

In this paper, we propose a novel Bayesian probability model to detect generalized communities and identify the semantics combining network structures and nodes attributes

and use an efficient Gibbs sampling algorithm to optimize the objective function. Even if the information of node attributes is of poor quality, our method can use the complementary structural information in node attributes to get better results. The model assumes that the network structure and node attributes have different hidden variables and adopts a transition matrix to explore the hidden correlation between communities and topics. Thus, it can provide semantic descriptions of communities to better reveal the characteristics of communities. We evaluate our method on a number of real and synthetic datasets and in a case study. The new method can detect various types of network structures and outperforms several state-of-the-art algorithms.

It is similar to the proposed methods in requiring that the number of communities be provided. This problem is about model selection issue, and we will focus on determining group number automatically in the next step.

Data Availability

The datasets used to support the results of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (61902278), the National Key R&D Program of China (2018YFC0832101), the Livelihood Science and Technology Project of Qingdao (18-6-1-106-nsh), and the National Social Science Foundation of China (15BGL035).

References

[1] C. Mao and W. Xiao, "A comprehensive algorithm for evaluating node influences in social networks based on preference analysis and random walk," *Complexity*, vol. 2018, Article ID 1528341, 16 pages, 2018.

[2] X. Han, D. Chen, and H. Yang, "A semantic community detection algorithm based on quantizing progress," *Complexity*, vol. 2019, Article ID 3475458, 13 pages, 2019.

[3] J. Cheng, X. Su, H. Yang et al., "Neighbor similarity based agglomerative method for community detection in networks," *Complexity*, vol. 2019, Article ID 8292485, 16 pages, 2019.

[4] S. Hua-Wei, C. Xue-Qi, and G. Jia-Feng, "Exploring the structural regularities in networks," *Physical Review E Statistical Nonlinear & Soft Matter Physics*, vol. 84, no. 2, Article ID 056111, 2011.

[5] D. He, X. Yang, Z. Feng, S. Chen, and F. Fogelman-Soulié, "A network embedding-enhanced approach for generalized community detection," in *Proceedings of the International Conference on Knowledge Science, Engineering and Management*, pp. 383–395, Springer, Changchun, China, August 2018.

[6] G. Zhang, D. Jin, J. Gao, P. Jiao, F. Fogelman-Soulié, and X. Huang, "Finding communities with hierarchical semantics by distinguishing general and specialized topics," in *Proceedings of the 27th International Joint Conference on Artificial Intelligence*, pp. 3648–3654, AAAI Press, Stockholm, Sweden, July 2018.

[7] L. Yang, X. Cao, D. He, C. Wang, X. Wang, and W. Zhang, "Modularity based community detection with deep learning," in *Proceedings of the IJCAI*, pp. 2252–2258, New York, NY, USA, July 2016.

[8] B. Karrer and M. E. Newman, "Stochastic blockmodels and community structure in networks," *Physical Review E*, vol. 83, no. 1, Article ID 016107, 2011.

[9] A. Anandkumar, R. Ge, D. Hsu, and S. M. Kakade, "A tensor approach to learning mixed membership community models," *The Journal of Machine Learning Research*, vol. 15, no. 1, pp. 2239–2312, 2014.

[10] D. He, D. Liu, D. Jin, and W. Zhang, "A stochastic model for detecting heterogeneous link communities in complex networks," in *Proceedings of the AAAI*, pp. 130–136, Austin, TX, USA, January 2015.

[11] U.-U. Narantsatsralt and S. Kang, "Social network community detection using agglomerative spectral clustering," *Complexity*, vol. 2017, Article ID 3719428, 10 pages, 2017.

[12] F. Liu, D. Choi, L. Xie, and K. Roeder, "Global spectral clustering in dynamic networks," *Proceedings of the National Academy of Sciences*, vol. 115, no. 5, pp. 927–932, 2018.

[13] F. Krzakala, C. Moore, E. Mossel et al., "Spectral redemption in clustering sparse networks," *Proceedings of the National Academy of Sciences*, vol. 110, no. 52, 2013.

[14] Y. Li, K. He, D. Bindel, and J. E. Hopcroft, "Uncovering the small community structure in large networks: a local spectral approach," in *Proceedings of the 24th International Conference on World Wide Web, International World Wide Web Conferences Steering Committee*, pp. 658–668, Florence, Italy, May 2015.

[15] D. He, Z. Feng, D. Jin, X. Wang, and W. Zhang, "Joint identification of network communities and semantics via integrative modeling of network topologies and node contents," in *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, San Francisco, CA, USA, Febraury 2017.

[16] D. Jin, X. Wang, R. He, D. He, J. Dang, and W. Zhang, "Robust detection of link communities in large social networks by exploiting link semantics," in *Proceedings of the AAAI*, New Orleans, LA, USA, Febraury 2018.

[17] J. Yang, J. McAuley, and J. Leskovec, "Community detection in networks with node attributes," in *Proceedings of the 13th International Conference on Data Mining (ICDM)*, pp. 1151–1156, IEEE, Dallas, TX, USA, December 2013.

[18] J. J. Choong, X. Liu, and T. Murata, "Variational approach for learning community structures," *Complexity*, vol. 2018, Article ID 4867304, 13 pages, 2018.

[19] M. Rosvall, A. V. Esquivel, A. Lancichinetti, J. D. West, and R. Lambiotte, "Memory in network flows and its effects on spreading dynamics and community detection," *Nature Communications*, vol. 5, no. 1, p. 4630, 2014.

[20] G. Rossetti and R. Cazabet, "Community discovery in dynamic networks: a survey," *ACM Computing Surveys (CSUR)*, vol. 51, no. 2, p. 35, 2018.

[21] F. O. Arimoro, H. E. Olisa, U. N. Keke, A. V. Ayanwale, and V. I. Chukwuemeka, "Exploring spatio-temporal patterns of plankton diversity and community structure as correlates of

water quality in a tropical stream," *Acta Ecologica Sinica*, vol. 38, no. 3, pp. 216–223, 2018.
[22] S. Fortunato and D. Hric, "Community detection in networks: a user guide," *Physics Reports*, vol. 659, pp. 1–44, 2016.
[23] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *Journal of Machine Learning Research*, vol. 3, pp. 993–1022, 2003.
[24] J. Xiao, H.-F. Ren, and X.-K. Xu, "Constructing real-life benchmarks for community detection by rewiring edges," *Complexity*, vol. 2020, Article ID 7096230, 16 pages, 2020.
[25] Y. Ruan, D. Fuhry, and S. Parthasarathy, "Efficient community detection in large networks using content and links," in *Proceedings of the International Conference on World Wide Web*, pp. 1089–1098, Rio de Janeiro, Brazil, 2013.
[26] D. Cohn and T. Hofmann, "The missing link—a probabilistic model of document content and hypertext connectivity," in *Proceedings of the International Conference on Neural Information Processing Systems*, pp. 409–415, Denver, CO, USA, 2000.
[27] T. Yang, R. Jin, Y. Chi, and S. Zhu, "Combining link and content for community detection," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 927–936, Paris, France, June 2009.
[28] S. Pool, F. Bonchi, and M. v. Leeuwen, "Description-driven community detection," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 5, no. 2, p. 28, 2014.
[29] S. Fortunato, "Community detection in graphs," *Physics Reports*, vol. 486, no. 3–5, pp. 75–174, 2010.
[30] M. E. J. Newman and E. A. Leicht, "Mixture models and exploratory analysis in networks," *Proceedings of the National Academy of Sciences*, vol. 104, no. 23, pp. 9564–9569, 2007.
[31] Y. Chen, X. Wang, J. Bu, B. Tang, and X. Xiang, "Network structure exploration in networks with node attributes," *Physica A: Statistical Mechanics and Its Applications*, vol. 449, pp. 240–253, 2016.
[32] Y. P. N. Chakraborty and K. Sycara, "Nonnegative matrix trifactorization with graph regularization for community detection in social networks," in *Proceedings of the AAAI*, AAAI Press, Austin, TX, USA, pp. 2083–2089, January 2015.
[33] X. Wang, D. Jin, X. Cao, L. Yang, and W. Zhang, "Semantic community identification in large attribute networks," in *Proceedings of the AAAI*, pp. 265–271, Phoenix, AZ, USA, Febraury 2016.
[34] C. Bothorel, J. D. Cruz, M. Magnani, and B. Micenková, "Clustering attributed graphs: models, measures and methods," *Network Science*, vol. 3, no. 3, pp. 408–444, 2015.
[35] P. Chunaev, "Community detection in node-attributed social networks: a survey," *Computer Science Review*, vol. 37, p. 100286, 2020.
[36] C. Andrieu, N. De Freitas, A. Doucet, and M. I. Jordan, "An introduction to mcmc for machine learning," *Machine Learning*, vol. 50, no. 1-2, pp. 5–43, 2003.
[37] R. M. Neal, "Slice sampling," *The Annals of Statistics*, vol. 31, no. 3, pp. 705–767, 2003.
[38] Z. Chang, X. Yin, C. Jia, and X. Wang, "Mixture models with entropy regularization for community detection in networks," *Physica A: Statistical Mechanics and Its Applications*, vol. 496, pp. 339–350, 2018.