Proceedings of the 2022 IEEE 25th International Conference on Computer Supported Cooperative Work in Design

SARNMF: A Community Detection Method for Attributed Networks

Junwei Cheng[1], Weisheng Li[1], Kunlin Han[3], Yong Tang[1], Chaobo He[1,2,*], Nini Zhang[1,*]
[1]School of Computer Science, South China Normal University, Guangzhou, China
[2]School of Information Science and Technology, Zhongkai University of Agriculture and Engineering, Guangzhou, China
[3]Computer Science Department, University of Southern California, Los Angeles, USA
Email:[jung, liws, ytang, zhangnini]@m.scnu.edu.cn,
hechaobo@foxmail.com, kunlinha@usc.edu

2022 IEEE 25th International Conference on Computer Supported Cooperative Work in Design (CSCWD) | 978-1-6654-0527-0/22/$31.00 ©2022 IEEE | DOI: 10.1109/CSCWD54268.2022.9776178

***Abstract*—Community detection is one of the hottest research topics in attributed networks analysis. Nonnegative matrix factorization (NMF) is widely used in community detection of attributed networks because of its high interpretability and extensibility. However, the existing NMF based methods still encounter some obstacles which affect the performance of community detection. Firstly, it is impossible to solve the problem of sparse semantic description. Besides, these methods cannot integrate the heterogeneity of topology structure and nodes attributes. Obviously, these methods cannot accurately identify community structure and assign specific semantic descriptions to each community. To overcome the aforementioned problems, we propose a novel method which combines graph neural networks with weighted-traction regularization. Moreover, we use graph neural networks to discover the semantic characteristics between adjacent nodes which can alleviate the problem of sparse semantic description. Furthermore, the regularizer we proposed can improve the performance of community detection in attributed networks. Experiments on some real attributed networks show that the method we proposed not only is better than some representative related methods but also can assign specific semantic descriptions to each community at the same time.**

***Keywords*—Community detection; Attributed networks; Nonnegative matrix factorization; Graph neural networks.**

I. Introduction

Attributed networks are ubiquitous and widely characterize many complex networks. For example, in co-author networks [4], each node corresponds to an author and it has its own attributes, such as research interest, keywords of paper and personal information. Also each edge denotes the co-authoring relationship. Interestingly, there are many valuable research topics in attributed networks analysis. Community detection is one of them which boosts diversified practical applications, such as friend recommendation [23], advertising [9] and virus traceability [25].

Great efforts have been devoted to community detection in attributed networks, such as deep learning based methods [13], modularity based methods [21] and NMF based methods [11]. Extensive research demonstrates that NMF based methods has achieved significant advantages compared with other methods, because of its high interpretability and extensibility (see [7]). However, designing effective and efficient methods for community detection is still non-trivial for two reasons. First of all, these methods ignore the coupling relationship between adjacent nodes, so it is impossible to solve the problem of sparse semantic description. Secondly, they cannot integrate the heterogeneity of topology structure and nodes attributes, making it hard to deal with the coupling relationship between adjacency matrix and attribute similarity matrix. All of them affect the recognition of community structure and cannot assign accurate semantic descriptions to each community.

To overcome the aforementioned problems, we present a novel method that combines graph neural networks with weighted-traction regularization (namely SARNMF). It consists of two major components (shown in Fig. 1). The first part is attributed network embedding. It is obtained by using graph neural networks. The second part is joint nonnegative matrix factorization with weighted-traction regularization. And then, we derive iterative rules to optimize our method. In conclusion, the work of this paper can be summarized as

- We make use of the semantic characteristics of adjacent nodes to get the attributed network embedding by using graph neural networks, so that we can solve the problem of sparse semantic description.
- We propose a novel strategy, which includes functionality of sampling, aggregation and weighted-traction regularization, namely SARNMF, to improve the performance of joint community detection of both adjacency matrix and attribute similarity matrix.
- We conduct some experiments on six real attributed networks, and the results show that SARNMF performs better than some representative related methods. In addition, our method is used to analyze the SCHOLAT Social Network dataset, and the results show that the method we proposed can accurately assign specific semantic descriptions to each community.

The rest of the paper is organized as follows. Section 2 introduces the related work of others in attributed networks. Section 3 introduces notations and the procedure of SARNMF. Section 4 presents the experimental results on real attributed networks. Finally, we draw conclusion in Section 5.

II. Related Work

Community detection is an effective method to analyze attributed networks. It aims to discover a group of nodes that

* Corresponding author.

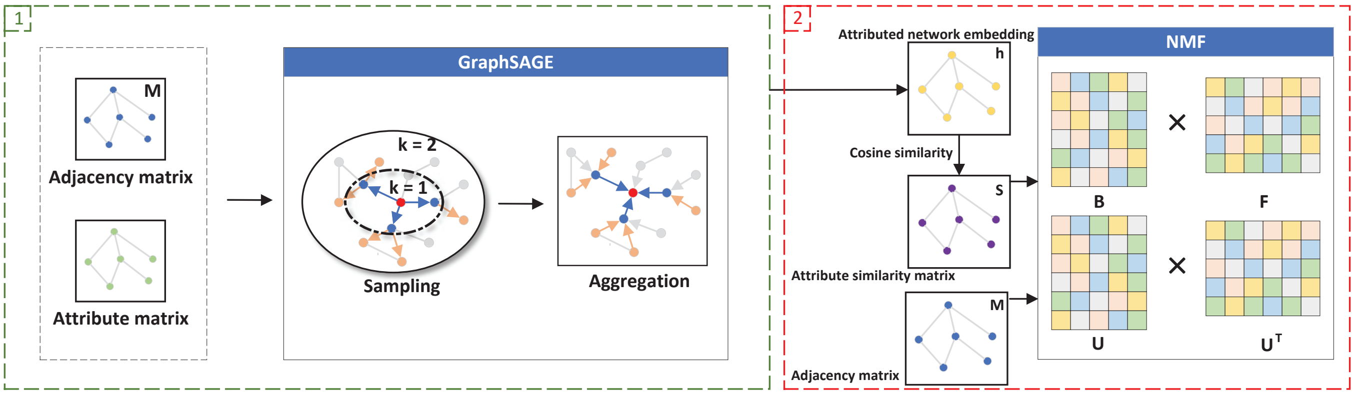

Fig. 1. The method of SARNMF, which consists of two major components: attributed network embedding by using GraphSAGE and joint nonnegative matrix factorization with weighted-traction regularization. Firstly, we input adjacency matrix and attribute matrix to GraphSAGE and obtain the attributed network embedding. And then we process it further by using cosine similarity. Finally, we input the processing result together with adjacency matrix to NMF and we can infer the communities from **U**.

are well connected inside and weakly connected outside [16] [15]. At present, community detection methods of attributed networks can be classified into three categories (see [3]), including (1) early fusion methods [18] [5], (2) simultaneous fusion methods [12] [19] and (3) late fusion methods [14] [1]. Early fusion and simultaneous fusion ideas are widely used because of their high flexibility and extensibility.

Especially, Chang et al. [2] proposed a model, namely Factorized Similarity Learning, to integrate the link, node attributed and user supervision into an uniform framework, which helps to compute the large scale of data efficiently. Huang et al. [8] proposed JWNMF method based on graph clustering, which integrates topology structure and nodes attributes by using a new collective nonnegative matrix factorization method. Pei et al. [18] proposed a nonnegative matrix tri-factorization based clustering framework with three types of graph regularization, which combines the topology and attributes. Guo et al. [5] argues that feature values and linkages provide useful information from different perspectives but they are not consistent. Therefore, they proposed CFOND, a consensus factorization based framework to ensure that the cluster structures are consistent across information. Wang et al. [22] proposed SCI, which improves upon community detection and provides a semantic interpretation to the resultant network communities. Qin et al. [20] proposed a community detection method that controls the contribution of attributed with respect to the identified mismatch degree between the topological and attributed information. Jin et al. [10] designed a new community detection method. It contains substantial explanation to discover latent relationship between network communities and content clusters. Xu et al. [24] proposed a method, namely FTAI, with flexible, robust and suitable for community detection in social networks.

Although great achievements have been made by applying those methods, these methods still encounter some obstacles, which lead to the failure to achieve the desired effect of community detection. The main cause of this problem is that they treat all of nodes attributes as isolated, rather than as a whole. Therefore, it is hard to avoid losing the relationship between the two respective semantic characteristics of adjacent nodes when dealing with the attribute matrix. Besides, these methods cannot integrate the heterogeneity of topology structure and nodes attributes. These problems hinder the improvement of performance on community detection in attributed networks. In this paper, we design a novel method, namely SARNMF, to attack these two problems.

III. METHODOLOGY

In this section, we firstly give an overview of method and introduce notations. And then, the procedures of SARNMF method and optimization are presented.

A. Overview and notations

As represented in Fig. 1, SARNMF consists of two major components which are attributed network embedding and joint nonnegative matrix factorization. Focusing on the first component, we use the graph neural networks to discover the coupling relationship between the attribute of adjacent nodes. Also it solves the problem of sparse semantic description. In the second component, we decompose attribute similarity matrix and adjacency matrix jointly by using weighted-traction regularization. Actually, the regularizer we proposed can integrate the heterogeneity of topology structure and nodes attributes. All in all, the method we proposed overcome these problems and improves the performance of community detection.

Given an undirected attributed network $G = (V, E, \mathbf{A})$ over a set of n nodes $V = \{v_1, v_2, ..., v_n\}$, a set of e edges $E = \{(v_i, v_j)\}$ and $\mathbf{A} \in \mathbb{R}^{n \times d}$, which denotes the nodes attribute matrix with d being the dimensionality of the nodes attributes. The adjacency matrix $\mathbf{M} = (m_{ij})_{n \times n}$ whose element $m_{ij} = 1$ if there is an edge between nodes i and j, and 0 otherwise. The attribute similarity matrix $\mathbf{S} = (s_{ij})_{n \times n}$ whose element s_{ij} is the jth attribute similarity of the ith node. Regularizer is constructed by weighted matrix $\mathbf{X}_{n \times n}$ and traction matrix

$\mathbf{T}_{k\times n}$. The embedding of node v is denoted as h_v and its adjacent node set is denoted as N_v.

B. Attribute network embedding

Graph neural networks is widely used in graph mining due to its excellent performance and better interpretability. Hamilton [6] proposed GraphSAGE based on inductive learning to generate attribute embedding by sampling and aggregating the semantic characteristics of adjacent nodes. In this paper, we refer to this method to complete the first part of our method. The core steps of this model are sampling and aggregation. In terms of sampling, we use random walk method to visit the corresponding adjacent nodes, in which hyperparameter k represents the maximum number of layers visited. In terms of aggregation, we use mean aggregator to aggregate adjacent semantic descriptions. The mean aggregator can be written as:

$$h_v^t = \sigma(W \cdot MEAN(\{h_v^{t-1}\} \cup \{h_u^{t-1}, \forall u \in N_v\})) \tag{1}$$

where W is the parameter vector of layer.

C. Community detection model

Given G, NMF approximates the $\mathbf{M}$ by the product of two nonnegative low-rank matrices. Among them, the first matrix corresponds to the common basis matrix and the second matrix is a feature matrix. The objective function can be written as:

$$\min_{\mathbf{U}\geq 0} ||\mathbf{M} - \mathbf{U}\mathbf{U}^T||_F^2 \tag{2}$$

where $\mathbf{U}$ is communities membership obtained by $\mathbf{M}$ decomposition.

To solve the problem of sparse semantic descriptions, we get the attributed network embedding by using GraphSAGE algorithm. These embeddings not only contain nodes characteristics but also the information of adjacent nodes. Thus, it is easy to calculate the cosine similarity between each two nodes by using attributed network embedding. The calculation result forms the $\mathbf{S}$. Then, the decomposition formula, which is based on attribute similarity, can be written as:

$$\min_{\mathbf{B}\geq 0, \mathbf{F}\geq 0} \alpha||\mathbf{S} - \mathbf{B}\mathbf{F}||_F^2 \tag{3}$$

where α controls the relevant importance of attribute similarity matrix. $\mathbf{B}$ and $\mathbf{F}$ are bias matrix and mapping matrix respectively. And they are obtained by decomposition of $\mathbf{S}$.

Most existing methods cannot effectively integrate the heterogeneity of topology structure and nodes attributes. Therefore, we propose a regularizer including $\mathbf{M}$ and $\mathbf{S}$ to cope with the problem. Our goal is to make $\mathbf{U}^T$ as close as possible to $\mathbf{F}$. Then, the regularizer can be written as:

$$\min_{\mathbf{X}\geq 0, \mathbf{T}\geq 0} ||\mathbf{X}\mathbf{F}^T - \mathbf{U}||_F^2 + ||\mathbf{F}^T\mathbf{T} - \mathbf{X}||_F^2 \tag{4}$$

As we have described above, we get the complete loss function where integrating the Eq. (2) to Eq. (4) as follows:

$$\begin{aligned} \min \mathcal{L}(\mathbf{U},\mathbf{B},\mathbf{F},\mathbf{X},\mathbf{T}) = ||\mathbf{M} - \mathbf{U}\mathbf{U}^T||_F^2 + \alpha||\mathbf{S} - \mathbf{B}\mathbf{F}||_F^2 + \\ ||\mathbf{X}\mathbf{F}^T - \mathbf{U}||_F^2 + ||\mathbf{F}^T\mathbf{T} - \mathbf{X}||_F^2 \\ s.t. \quad \mathbf{U}\geq 0, \mathbf{B}\geq 0, \mathbf{F}\geq 0, \mathbf{X}\geq 0, \mathbf{T}\geq 0 \end{aligned} \tag{5}$$

D. Model optimization

We can derive the Eq. (5) as:

$$\begin{aligned} \mathcal{L} = tr(\mathbf{M}^T\mathbf{M} - \mathbf{M}^T\mathbf{U}\mathbf{U}^T - \mathbf{U}\mathbf{U}^T\mathbf{M} + \mathbf{U}\mathbf{U}^T\mathbf{U}\mathbf{U}^T) + \\ \alpha \times tr(\mathbf{S}^T\mathbf{S} - \mathbf{S}^T\mathbf{B}\mathbf{F} - \mathbf{F}^T\mathbf{B}^T\mathbf{S} + \mathbf{F}^T\mathbf{B}^T\mathbf{B}\mathbf{F}) + \\ tr(\mathbf{F}\mathbf{X}^T\mathbf{X}\mathbf{F}^T - \mathbf{F}\mathbf{X}^T\mathbf{U} - \mathbf{U}^T\mathbf{X}\mathbf{F}^T + \mathbf{U}^T\mathbf{U}) + \\ tr(\mathbf{T}^T\mathbf{F}\mathbf{F}^T\mathbf{T} - \mathbf{T}^T\mathbf{F}\mathbf{X} - \mathbf{X}^T\mathbf{F}^T\mathbf{T} + \mathbf{X}^T\mathbf{X}) \end{aligned} \tag{6}$$

Thus, the partial derivative of $\mathbf{U}, \mathbf{B}, \mathbf{F}, \mathbf{X}$ and $\mathbf{T}$ are deduced as:

$$\frac{\partial \mathcal{L}}{\partial \mathbf{U}} = -2\mathbf{M}^T\mathbf{U} - 2\mathbf{M}\mathbf{U} + 4\mathbf{U}\mathbf{U}^T\mathbf{U} - 2\mathbf{X}\mathbf{F}^T + 2\mathbf{U} \tag{7}$$

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \mathbf{F}} = -2\alpha\mathbf{B}^T\mathbf{S} + 2\alpha\mathbf{B}^T\mathbf{B}\mathbf{F} + 2\mathbf{F}\mathbf{X}^T\mathbf{X} - \\ 2\mathbf{U}^T\mathbf{X} + 2\mathbf{T}\mathbf{T}^T\mathbf{F} - 2\mathbf{T}\mathbf{X}^T \end{aligned} \tag{8}$$

$$\frac{\partial \mathcal{L}}{\partial \mathbf{B}} = -2\alpha\mathbf{S}\mathbf{F}^T + 2\alpha\mathbf{B}\mathbf{F}\mathbf{F}^T \tag{9}$$

$$\frac{\partial \mathcal{L}}{\partial \mathbf{X}} = -2\mathbf{T}^T\mathbf{F} - 2\mathbf{U}\mathbf{F} + 2\mathbf{X}\mathbf{F}^T\mathbf{F} + 2\mathbf{X} \tag{10}$$

and

$$\frac{\partial \mathcal{L}}{\partial \mathbf{T}} = 2\mathbf{F}\mathbf{F}^T\mathbf{T} - 2\mathbf{F}\mathbf{X} \tag{11}$$

According to the multiplication update rule of Oja [17], we obtain the following update rules for $\mathbf{U}, \mathbf{B}, \mathbf{F}, \mathbf{X}$ and $\mathbf{T}$:

$$\mathbf{U} = \mathbf{U} \odot \frac{\mathbf{M}^T\mathbf{U} + \mathbf{M}\mathbf{U} + \mathbf{X}\mathbf{F}^T}{2\mathbf{U}\mathbf{U}^T\mathbf{U} + \mathbf{U}} \tag{12}$$

$$\mathbf{F} = \mathbf{F} \odot \frac{\alpha\mathbf{B}^T\mathbf{S} + \mathbf{U}^T\mathbf{X} + \mathbf{T}\mathbf{X}^T}{\alpha\mathbf{B}^T\mathbf{B}\mathbf{F} + \mathbf{F}\mathbf{X}^T\mathbf{X} + \mathbf{T}\mathbf{T}^T\mathbf{F}} \tag{13}$$

$$\mathbf{B} = \mathbf{B} \odot \frac{\mathbf{S}\mathbf{F}^T}{\mathbf{B}\mathbf{F}\mathbf{F}^T} \tag{14}$$

$$\mathbf{X} = \mathbf{X} \odot \frac{\mathbf{U}\mathbf{F} + \mathbf{T}^T\mathbf{F}}{\mathbf{X}\mathbf{F}^T\mathbf{F} + \mathbf{X}} \tag{15}$$

$$\mathbf{T} = \mathbf{T} \odot \frac{\mathbf{F}\mathbf{X}}{\mathbf{F}\mathbf{F}^T\mathbf{T}} \tag{16}$$

where $\odot$ denotes element-wise product. The procedure of SARNMF is illustrated in Algorithm 1.

IV. Experimental study

To evaluate the performance of the proposed method, we conduct some experiments on real attributed networks. In this section, we introduce some datasets, evaluation metrics and method settings. And then, we design some experiments to justify the performance of our method. We completed an ablation study to prove the indispensability of graph neural networks for our method. Finally, we apply our method in the SCHOLAT Social Network dataset.

Algorithm 1: SARNMF

Input:

$G = (V, E, \mathbf{A})$:An attributed network;
$\mathbf{S}$: Attribute similarity matrix;
$\mathbf{M}$: Adjacency matrix;
$\mathbf{X}$: Weighted matrix;
$\mathbf{T}$: Traction matrix;
α: Weight for attribute similarity matrix;

Output:

C: Communities in attributed networks.

```
Initialize B, F and U by using random sampling;
Initialize X by using identity matrix;
Initialize T by using 𝟙
while not convergent of algorithm or not reach the
  maximum number of iterations do
    Fix B, F, X, T, update U via Eq. (12);
    Fix B, U, X, T, update F via Eq. (13);
    Fix F, U, X, T, update B via Eq. (14);
    Fix B, F, U, T, update X via Eq. (15);
    Fix B, F, U, X, update T via Eq. (16);
Infer the communities C based on U^T;
return C.
```

A. Datasets and evaluation metrics

We introduce evaluation metrics in this subsection that are widely used to estimate the results of community detection, including the Modularity and F-score. Besides, we use real attributed networks with class labels to validate performance of various methods, including WEBKB (contain Wisconsin, Washington, Texas and Cornell), Cora and Citeseer which are summarized in Table I.

TABLE I
STATISTICS OF ATTRIBUTED NETWORKS, WHERE $|V|$ AND $|E|$ DENOTE THE NUMBER OF VERTICES AND EDGES, RESPECTIVELY.

Dataset	$\|V\|$	$\|E\|$	#Attribute
Cornell	195	304	1703
Texas	187	328	1703
Wisconsin	265	530	1703
Washington	230	446	1703
Cora	2708	5278	1433
Citeseer	3312	4536	3703

Modularity is usually used to measure the strength of the network community structure. And the higher its value shows that the better result of community detection. Modularity is defined as

$$\mathrm{Modularity} = \frac{1}{2\mathrm{e}} \sum_{\mathrm{ij}} (\mathbf{M}_{\mathrm{ij}} - \frac{\mathrm{k_i k_j}}{2\mathrm{e}}) \delta(\mathrm{i}, \mathrm{j}) \tag{17}$$

where e represents the number of edges and $\delta(i, j)$ is 1 if i and j in the same community, 0 otherwise.

F-score is usually used to validate agreement of algorithms between the real community and community detected by algorithm, which is defined as

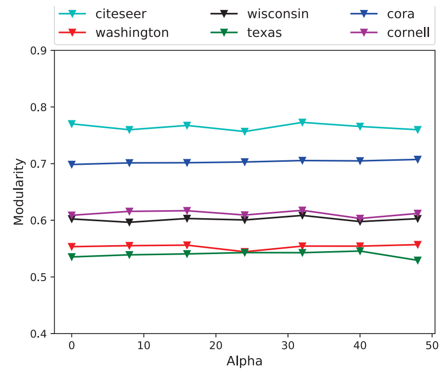

Fig. 2. Parameter α effects, where y-axis denotes Modularity of SARNMF on datasets.

$$\mathrm{F-score} = 2 \times \frac{\mathrm{precision} \times \mathrm{recall}}{\mathrm{precision} + \mathrm{recall}} \tag{18}$$

where the $precision$ and $recall$ are defined as

$$precision = \frac{TP}{TP + FP} \tag{19}$$

and

$$recall = \frac{TP}{TP + FN} \tag{20}$$

where TP is the number of the true positive samples, FP is the number of the false positive samples and FN is the number of false negative samples.

B. Parameter setting and stopping condition

There are three parameters involved in SARNMF: α determines the tradeoff between adjacency matrix and attribute similarity matrix, k determines the range of node attribute sampling and η determines the degree of retention of features. For each dataset, we analyze their characteristics and select suitable parameters for our method, whose details are shown in Table II. Among them, α has a direct impact on the result of community detection. The results on datasets are shown in Fig. 2. From the panel, it is easy to conclude that, as α increases from 0 to 50, the improvement of the SARNMF's Modularity go up and down periodically. Therefore, we only need to select the local optimal value for α. Experiments show that the result of our method converges as the number of iterations increases (shown in Fig. 3). Generally speaking, 100 iterations give a promising result.

TABLE II
HYPERPARAMETER DESCRIPTIONS.

Dataset	GraphSAGE		NMF	
	k	η	α	*epoch*
Cornell	3	50	7	300
Texas	1	64	29	300
Wisconsin	2	40	20	300
Washington	3	60	31	300
Cora	2	64	1	300
Citeseer	2	64	1	300

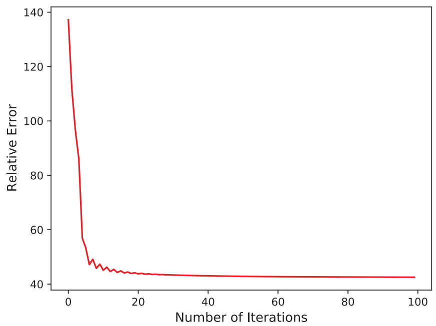

Fig. 3. Convergence analysis of SARNMF.

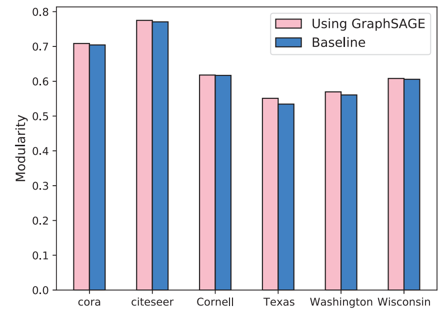

Fig. 4. Performance of methods in Modularity.

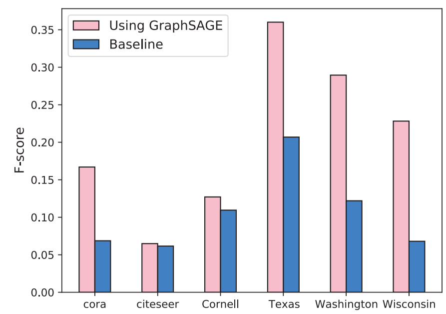

Fig. 5. Performance of methods in F-score.

C. Experiment results

Results of community detection on real attributed networks are listed in Tables III and IV respectively.

As shown in Table III and Table IV, SARNMF outperforms all its competitors on all datasets under Modularity and F-score on attributed networks. After evaluation, we can know that our method average improves by 1.3% of Modularity and 6% of F-score respectively compared with the other method with the best results. Especially, SARNMF increases by 50% of Modularity and 25% of F-score compared with NMF. Likewise, it increases at least by 3.35% of Modularity and 13.58% of F-score compared with other methods. Moreover, the experiment data demonstrates the performance of graph neural networks and weighted-traction regularization.

The running time of SARNMF on Cora, Citeseer, Cornell, Texas, Washington and Wisconsin are 169.93s, 343.39s, 5.79s, 4.54s, 7.92s and 8.18s respectively on a PC with ”RAM: 32G; CPU: Intel I7-6700”.

TABLE III
PERFORMANCE OF VARIOUS METHODS IN MODULARITY.

Dataset	NMF	FTAI	SCI	SARNMF
cora	0.4427	0.6841	0.7058	**0.7086**
citeseer	0.5282	0.7212	0.7415	**0.7750**
cornell	0.1687	0.6108	0.6093	**0.6179**
texas	0.0502	0.5223	0.5434	**0.5510**
washington	0.2508	0.5377	0.5516	**0.5695**
wisconsin	0.2816	0.5292	0.599	**0.6081**

TABLE IV
PERFORMANCE OF VARIOUS METHODS IN F-SCORE.

Dataset	NMF	FTAI	SCI	SARNMF
cora	0.1458	0.0965	0.0533	**0.1670**
citeseer	0.0467	0.0150	0.0459	**0.0649**
cornell	0.0736	0.1002	0.0982	**0.1272**
texas	0.1110	0.2243	0.1271	**0.3601**
washington	0.1449	0.1247	0.2700	**0.2895**
wisconsin	0.1505	0.1293	0.0876	**0.2283**

D. Ablation study

It is easy to find that our method has obvious advantages in F-score. In order to explain this phenomenon, we make an ablation study to prove that aggregating the underlying characteristics of adjacent nodes can improve the performance of community detection.

We set the same parameters to ensure a single variable, that is, whether the use of GraphSAGE affects the performance of the method. The experimental results are shown in Fig. 4 and Fig. 5 respectively. Interestingly, although the use of GraphSAGE has slightly improved the Modularity, it has greatly improved the F-score. For example, if we use GraphSAGE to calculate $\mathbf{S}$, then the average can improve by 10% and the maximum can improve by 17% in F-score. Therefore, the ablation study demonstrates that retaining the underlying characteristics of adjacent nodes by sampling and aggregating improves the performance of SARNMF.

E. Application

SCHOLAT Social Network [24] is an attributed network derived from SCHOLAT (A Chinese academic social network), which consists of 16007 scholars, 202248 relationships among these scholars, and 25817 unique words as attributes. After analyzing the SCHOLAT Social Network with SARNMF, we discover that the Modularity of SARNMF is 0.7517 and F-score of SARNMF is 0.5796. We randomly select two com-

(a) Semantic topic about scientific research. (b) Semantic topic about computer science.

Fig. 6. Word cloud

munities and count their high-frequency semantic feature as shown in Fig. 6. We can infer that the community represented by Fig. 6(a) is composed of teachers and the community represented by Fig. 6(b) is composed of students majoring in computer science. Based on the results of the aforementioned communities, we can understand the preferences of users in these communities. Therefore, we can provide corresponding information or scholars to users in these communities. The result of this application shows that our method helps to assign specific semantic descriptions to each community.

V. CONCLUSIONS

NMF is widely used in community detection of attributed networks. However, the existing studies still have shortcomings. To overcome these problems, in this study, we propose a novel method namely SARNMF. The advantage of our method is that it can aggregate the semantic characteristics of adjacent nodes and better leverage the attributes of nodes and topology structure to solve their own heterogeneity. The experiment results demonstrate that our method improves the effectiveness of community detection in attributed networks and at the same time assigns specific semantic descriptions to each community.

Although we has improved the performance in community detection of attributed networks by devising SARNMF method, our method has higher time complexity compared with some representative related methods since it has more rules to update. Besides, our method can only be used for static attributed networks analysis. Therefore, we will extend our method to fit in dynamic attributed networks analysis.

ACKNOWLEDGMENT

This work was supported in part by the National Natural Science Foundation of China under Grant 62077045, Grant U1811263 and Grant 61772211, in part by the Humanity and Social Science Youth Foundation of Ministry of Education of China under Grant 19YJCZH049, in part by the Natural Science Foundation of Guangdong Province of China under Grant 2019A1515011292, and in part by the Science and Technology Support Program of Guangzhou City of China under Grant 201905010006.

REFERENCES

[1] Z. Bu, G. Gao, H. Li, and J. Cao. Camas: A cluster-aware multiagent system for attributed graph clustering. *Information Fusion*, 37:10–21, 2017.

[2] S. Chang, G. Qi, C. Aggarwal, J. Zhou, M. Wang, and T. Huang. Factorized similarity learning in networks. In *2014 IEEE International Conference on Data Mining*, pages 60–69, 2014.

[3] P. Chunaev. Community detection in node-attributed social networks: A survey. *Computer Science Review*, 37:100286, 2020.

[4] S. Fortunato. Community detection in graphs. *Physics Reports*, 486(3-5), 2009.

[5] T. Guo, S. Pan, X. Zhu, and C. Zhang. Cfond: Consensus factorization for co-clustering networked data. *IEEE Transactions on Knowledge and Data Engineering*, 31(4):706–719, 2019.

[6] W. Hamilton, Z. Ying, and J. Leskovec. Inductive representation learning on large graphs. *Advances in neural information processing systems*, 30, 2017.

[7] C. He, X. Fei, Q. Cheng, H. Li, Z. Hu, and Y. Tang. A survey of community detection in complex networks using nonnegative matrix factorization. *IEEE Transactions on Computational Social Systems*, pages 1–18, 2021.

[8] Z. Huang, Y. Ye, X. Li, L. Feng, and H. Chen. Joint weighted nonnegative matrix factorization for mining attributed graphs. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pages 368–380. Springer, 2017.

[9] Y. Jiang. Advertising target community detection algorithm in wechat network. 2016.

[10] D. Jin, Z. Liu, R. He, X. Wang, and D. He. A robust and strong explanation community detection method for attributed ncetworks. *Chinese Journal of Computers*, 2018.

[11] D. Lee and H. Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*, 401(6755):788–791, 1999.

[12] Z. Li, J. Liu, and K. Wu. A multiobjective evolutionary algorithm based on structural and attribute similarities for community detection in attributed networks. *IEEE transactions on cybernetics*, 48(7):1963–1976, 2017.

[13] F. Liu, S. Xue, J. Wu, C. Zhou, W. Hu, C. Paris, S. Nepal, J. Yang, and P. Yu. Deep learning for community detection: Progress, challenges and opportunities. In *Twenty-Ninth International Joint Conference on Artificial Intelligence and Seventeenth Pacific Rim International Conference on Artificial Intelligence IJCAI-PRICAI-20*, 2020.

[14] M. Newman and A. Clauset. Structure and inference in annotated networks. *Nature Communications*, 7(2-3):11863, 2015.

[15] M. Newman and M. Girvan. Finding and evaluating community structure in networks. *Physical Review E*, 69(2 Pt 2):026113, 2004.

[16] M. Newman, D. Watts, and S. Strogatz. Random graph models of social networks. *Proceedings of the National Academy of Sciences*, 99(suppl 1):2566–2572, 2002.

[17] E. Oja. Principal components, minor components, and linear neural networks. *Neural Networks*, 5(6):927–935, 1992.

[18] Y. Pei, N. Chakraborty, and K. Sycara. Nonnegative matrix tri-factorization with graph regularization for community detection in social networks. *AAAI Press*, 2015.

[19] C. Pizzuti and A. Socievole. Multiobjective optimization and local merge for clustering attributed graphs. *IEEE transactions on cybernetics*, 50(12):4997–5009, 2019.

[20] M. Qin, D. Jin, D. He, B. Gabrys, and K. Musial. Adaptive community detection incorporating topology and content in social networks. volume 161, pages 342–356. Elsevier, 2018.

[21] P. Schuetz and A. Caflisch. Efficient modularity optimization by multistep greedy algorithm and vertex mover refinement. *Physical Review E Statistical Nonlinear & Soft Matter Physics*, 77(2):046112, 2007.

[22] X. Wang, D. Jin, X. Cao, L. Yang, and W. Zhang. Semantic community identification in large attribute networks. 30(1), 2016.

[23] Y. Wang and L. Gao. Social circle-based algorithm for friend recommendation in online social networks. *Chinese Journal of Computers*, pages 1267–1275, 2014.

[24] Q. Xu, L. Qiu, R. Lin, Y. Tang, C. He, and C. Yuan. An improved community detection algorithm via fusing topology and attribute information. In *2021 IEEE 24th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, pages 1069–1074, 2021.

[25] L. Zheng, X. Xu, L. Jia, Z. Lu, and L. Zhang. Malicious url prediction based on community detection. In *2015 International Conference on Cyber Security of Smart Cities, Industrial Control System and Communications (SSIC)*, pages 1–7. IEEE, 2015.