Community Detection with Graph Neural Network using Markov Stability

Shunjie Yuan
School of Cyber Engineering
Xidian University
Xi'an, China
shunjiey@foxmail.com

Chao Wang
School of Cyber Engineering
Xidian University
Xi'an, China
wangchao@xidian.com

Qi Jiang
School of Cyber Engineering
Xidian University
Xi'an, China
jiangqixdu@xidian.edu.cn

Jianfeng Ma
School of Cyber Engineering
Xidian University
Xi'an, China
jfma@mail.xidian.edu.cn

2022 International Conference on Artificial Intelligence in Information and Communication (ICAIIC) | 978-1-6654-5818-4/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICAIIC54071.2022.9722614

***Abstract*—Community detection is a fundamental task in network analysis. With the recent development of deep learning, some community detection methods related to deep learning have been proposed. However, these methods still face limitations with respect to accuracy and runtime. In this paper, we propose a graph neural network (GNN) based overlapping community detection method CDMG from the perspective of optimizing Markov Stability, which is a statistical property of the Markov process quantifying the quality of a community partition. Specifically, we train a graph neural network to generate the node embedding defined as the community affiliation weight matrix that denotes the strength of nodes' membership in communities while maximizing the Markov Stability. Then the community affiliation weight matrix is converted to a community affiliation matrix representing the community partition. Experiments on several real-world networks demonstrate the superiority of CDMG compared to other representative community detection algorithms. Additionally, since Markov Stability relies on a time parameter Markov Time, we observe that there exists a Markov Time threshold for a network. When using the Markov Time near the threshold, CDMG can produce a better community partition with much higher accuracy.**

***Index Terms*—Complex Network, Community Detection, Graph Neural Network, Markov Stability**

I. Introduction

Many real-world systems can be represented as complex networks, such as the Internet, neural system, and transportation system. Community detection is fundamental research in network analysis because many further studies rely on community structure. For example, researchers elucidate the relationship between the structure of neuronal networks and the functional dynamics that they implement in the network of Caenorhab-ditis elegans connectome [1]. Because of the importance of community detection, it has attracted a great deal of attention from researchers and numerous algorithms have been proposed.

In recent years, Graph Neural Networks (GNNs) have become a new research hotspot, which is a powerful tool to deal with graph-structured data with deep learning algorithms. Research on GNNs generally can be divided into two categories, spectral-based and spatial-based approaches. Spectral-based methods define graph convolution operations as filters on the frequency domain of a graph. Because GNNs can reveal the higher-order structural information based on the non-linear

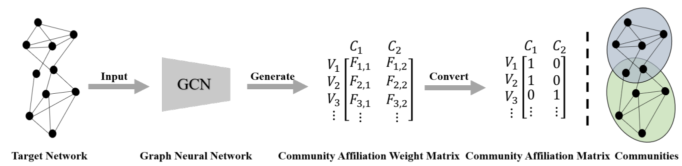

Fig. 1. The overview of the proposed algorithm CDMG. A graph neural network is trained to generate the community affiliation weight matrix that represents the strength of nodes affiliating to communities. Then the community affiliation weight matrix is converted to a community affiliation matrix that denotes the community partition.

feature aggregation and the information propagation across the network, which has been widely applied in several network analysis tasks, such as link prediction, node classification, and community detection [2]–[6]. And our approach is motivated by the success of these works that will be discussed in detail next.

The structure of a network can affect the dynamical behavior that takes place on the network in terms of its high-connectivity within communities. On the other hand, the dynamics can reveal features of the network structure. Based on this idea, Markov Stability [7]–[9], a statistical property of the Markov dynamic, is deployed to measure the quality of community structure in this work. A large Markov Stability always corresponds to a robust community structure, which indicates that a random walker is difficult to escape the communities [7]. For instance, we discuss the influence of community structure on Markov Stability in the well-studied network Karate that has two communities [10]. As we destroy the community structure of 10 percent of nodes each time in Fig. 2 (a), the Markov Stability decreases gradually. And as shown in Fig. 2 (b), the ground-truth partition corresponds to the largest Markov Stability. Not only Markov Stability can indicate the quality of community structure but also the community detection methods based on Markov Stability can provide multiple results compared with other methods as Markov Stability relies on Markov Time.

In this paper, we propose a new community detection algorithm named CDMG (Community Detection based on Markov Stability and Graph Neural Network) to detect communities with GNNs from the perspective of optimizing Markov

(a) (b)

Fig. 2. The influence of community structure on Markov Stability in Karate. As we destroy the community structure step by step, the Markov Stability decreases gradually. And the ground-truth partition corresponds to the highest Markov Stability. Detecting communities with our method is a reverse process of the above.

Stability. Specifically, we use Markov Stability as the loss function which measures the quality of a community partition, and we train a graph neural network to generate the node embedding defined as the community affiliation weight matrix that represents the strength of nodes affiliating to communities while minimizing the loss function. The community affiliation weight matrix then is converted to a community affiliation matrix, a binary matrix representing the community partition. The overview of CDMG is shown in Fig. 1.

The main contributions are listed as follows:

- We propose a GNN based method CDMG for overlapping community detection from the perspective of optimizing Markov Stability.
- We conduct experiments on four real-world networks and the results demonstrate that CDMG outperforms other established methods in most cases.
- We discuss the influence of the time parameter Markov Time on the performance of CDMG and find out that there is a Markov Time threshold for a network. When Markov Time t is around the threshold, CDMG can result in better community partition with high accuracy.

The rest of the paper is organized as follows. In section II, we introduce some related work in community detection. Section III contains the explicit details about Markov Stability, Graph Neural Network, and the proposed method CDMG based on them. Section IV provides a thorough evaluation of our method and shows its superior performance compared with other representative algorithms on several real-world networks. Section V concludes our work.

II. Related work

Many community detection methods have been proposed from different perspectives. One direction to uncover community structure is to optimize the measures that quantify the quality of community structure like modularity [5], [11]. Another direction to detect communities is to infer the relationship between vertices and communities based on nonnegative matrix factorization [12]–[15]. In this section, we mainly focus on some deep learning based methods. These methods can be broadly divided into methods based on certain neural networks like Generative Adversarial Networks (GANs) [16], GNNs [5], [6], and Attention Model [17] and methods based on Graph Representation with Cluster algorithms [18]–[22].

In [16], a generative adversarial network for community detection, CommunityGAN, is designed, which includes a generator that tries to generate a vertex subset with a high probability to be a clique, and a discriminator that tries to discriminate the clique from the generator. The output of CommunityGAN is the node embedding which is also the community affiliation weight matrix representing the community partition. Recently, several GNN based community detection methods have been proposed. For example, Tsitsulin et al. train a single-layer graph neural network to generate the community affiliation weight matrix while minimizing the loss function, the reformulated modularity composed of the community affiliation weight matrix [5]. Shchur et al. present the NOCD model that generates the community affiliation weight matrix with a graph neural network, where the balanced negative log-likelihood of the Bernoulli–Poisson (BP) model composed of the community affiliation weight matrix is used as the loss function [6]. The core idea of BP model is that the more communities two nodes are in common, the more likely they are to be connected by an edge. As we can see, the key point of community detection with GNNs is to build an efficient model, where community structure information such as community affiliation weight matrix can be integrated into the loss function. Moreover, the Attention Model has also been introduced into the community detection field. Lobov et al. propose a new model which is based on the Transformer model [23]. Specifically, they utilize the encoder part of the Transformer to transform the Bethe Hessian embeddings and produce the probability of each cluster for each node while optimizing the soft modularity loss function [17].

On the other hand, several graph representational learning algorithms have been designed. Given any graph, the graph representational algorithms can learn a low-dimensional vector for each vertex, which can then be used for various network analysis tasks, such as community detection. DeepWalk [19] utilizes random walk to generate node sequences and adopt Skip-Gram to learn node embedding, which preserves second-order proximity. Node2Vec [20] is an extended version of DeepWalk where it deploys a biased random walk to generate node sequences. LINE [21] preserves both the first-order and the second-order proximity while learning node embedding. MNMF [22] is an NMF-based representation learning model, which preserves both the microscopic structure (first and second-order proximities) and mesoscopic community structure. ComE [24] is a framework that jointly solves community detection, community embedding, and node embedding together, and it adopts multivariate Gaussian distributions to represent communities based on the output node embedding.

III. Methodology

Given an undirected network $G = (V, E)$, where V is a set of nodes and $|V| = N$, E is a set of edges among nodes. The goal of community detection is to assign nodes into K communities. Such assignment can be represented as a community affiliation weight matrix $F \in \{x \mid 0 \leq x \leq 1\}^{N\times K}$. From this perspective detecting communities boils down to inferring the community affiliation weight matrix F when given the target network G.

A. Markov Stability

Markov Stability [7]–[9] is a statistical property of the Markov dynamic, which evaluates the quality of a community partition in terms of the persistence of the Markov dynamics within the communities during the time scale t, which means the larger Markov Stability is, the more unlikely a random walker is to escape the communities within time t. One main advantage of the community detection methods based on Markov Stability is that they can reveal community structure with different Markov Time t because Markov Stability is a time-parametrized function. Markov Time acts as a resolution parameter for community detection [1]. For an undirected and unweighted network G, its topology is encoded in the adjacency matrix $A \in R^{N\times N}$. We define the n-dimensional vector d with components $d_i = \sum_{j=1}^{n} A_{ij}$, the diagonal matrix $D = \operatorname{diag}(d)$ and the total weight of the degrees of the networks is $\mathrm{m} = \sum_{i,j} A_{ij}/2$. Then we define a discrete-time Markov process governed by the following dynamics:

$$P_{t+1} = P_t D^{-1} A \equiv P_t M \tag{1}$$

where P_t is the probability vector and M is the transition matrix. Given a partition H at time t, the Markov Stability is defined as the trace of the clustered autocovariance of the diffusion process:

$$ms(t, H) = \operatorname{trace}\left(H^T \left[\prod P(t) - \pi^T \pi\right] H\right) \tag{2}$$

where $\pi = d^T/2m$ is a unique stationary distribution of the process, $P(t) = M^t$ and $\Pi = \operatorname{diag}(\pi)$. The optimal community partition H corresponds to the maximal Markov Stability. But maximizing (2) is an NP-hard problem with no guarantees of global optimality, the existing methods utilize Louvain or other heuristic methods to optimize (2) [7], [25], but here we adopt a graph neural network.

B. Graph Neural Network

Graph Neural Networks are a class of models that can perform non-linear feature aggregation and information propagation with respect to network structure. For the purpose of this work, we use Graph Convolutional Network (GCN) [26] to output node embedding for vertices. Given the node attributes X, a single-layer GCN can be defined as:

$$F := GCN(A, X) = \delta(\hat{A}XW) \tag{3}$$

where δ is the non-linear activation function, such as ReLU, $\hat{A} = \widetilde{D}^{-1/2}\tilde{A}\widetilde{D}^{-1/2}$ is the normalized adjacency matrix, $\tilde{A} = A + I_N$ is the adjacency matrix with self-loops, and $\widetilde{D}$ is the diagonal degree matrix of $\tilde{A}$. If node attributes X are not available, we can use A as node features.

C. Method

Markov Stability can evaluate the quality of the network structure. The output of GCN can be considered as an embedding of nodes with the aim of preserving the network structure [6]. Therefore, we propose to optimize the output of GCN with Markov Stability. The core idea of our method is to generate a community partition with the maximal Markov Stability by using a graph convolutional network. Specifically, we utilize a 2-layer graph convolutional network to generate the node embedding defined as the community affiliation weight matrix F:

$$F := \mathrm{GCN(A, X)} = \mathrm{ReLU}\left(\hat{A}\,\mathrm{ReLU}\left(\hat{A}XW^1\right)W^2\right) \tag{4}$$

The main difference between our GCN model and the standard GCN is that we introduce normalization after the second graph convolution layer, which leads to noticeable improvements in performance. The Markov Stability is used as the loss function which is defined as:

$$\mathcal{L}(F) = -\operatorname{trace}\left(F^T\left[\Pi P(t) - \pi^T\pi\right]F\right) \tag{5}$$

Different from the Markov Stability mentioned before, here we use the community affiliation weight matrix F to approximate the community affiliation matrix $H \in \{0, 1\}^{N\times K}$ where H_{uv} denotes that node u belonging to community v or not. By minimizing (5), we can find the optimal community affiliation weight matrix F, then we convert it to the community affiliation matrix H, assigning the nodes to the communities with a threshold p. If F_{uc} is bigger than the threshold p, we believe that node u belongs to community c and set H_{uc} to 1 else 0. The threshold p and the Markov Time t are two hyperparameters which will be discussed in next section. To sum up, using GCN and Markov Stability for community detection has several advantages. First, GCN can generate similar community affiliation weight vectors for neighboring nodes, which improves the quality of community detection. Second, the node attributes can be incorporated into the model. Finally, because Markov Stability relies on Markov Time t, our method can provide multiple results when using different Markov Time t. And we design experiments to analyze the influence of Markov Time t on the performance of CDMG in next section.

IV. Evaluation

In this section, we perform a thorough evaluation of CDMG and show its superior performance compared to other competing methods for overlapping community detection. First, we describe the experimental networks, the accuracy metrics, and the comparative methods as well as the parameter setting of CDMG. Then we analyze the results and runtime of these community detection methods and explore the influence of Markov Time on the performance of CDMG especially. The experiments were performed on a computer running Windows Server 2016 with Intel(R) Xeon(R) CPU E5-2620 v4 @ 2.10GHz CPUs, 96GB of RAM.

A. Datasets

We conduct experiments on a variety of real-world networks including LiveJournal, Amazon, YouTube and DBLP. The statistics of these large-scale networks are summarized in Table I. Considering both the training time of these learning methods and the performance of the machine we used, we only sample four subgraphs with 100 ground-truth communities from these large-scale networks. LiveJournal is a free on-line blogging community where user can find friendship and form a group which other people can join. The vertices represent users and edges represent friendship. This subgraph has 1118 vertices and 2047 edges. The network of Amazon is collected by crawling its website, which is based on purchase information of the Amazon website. The vertices represent products, and edges represent those products that are frequently co-purchased. This subgraph has 3225 vertices and 10262 edges. YouTube is a popular American online video-sharing platform that includes a social network where the vertices represent users, and edges represent friendship among users. In the YouTube social network, users form friendships with each other and users can create groups that other users can join. This subgraph has 4890 vertices and 20787 edges. The DBLP computer science bibliography provides open bibliographic information on major computer science journals and proceedings. In the DBLP network, vertices represent authors, and edges represent that the authors have published at least one paper together. This subgraph has 10824 vertices and 38732 edges.

TABLE I
STATISTICS OF SEVERAL LARGE-SCALE NETWORKS.

Network	$\|V\|$	$\|E\|$	$\|C\|$
Amazon	0.34M	0.93M	49K
YouTube	1.10M	3.00M	30K
DBLP	0.43M	1.30M	2.5K
LiveJournal	4.00M	34.9M	310K

B. Metrics

To quantitatively evaluate the performance of these community detection methods, we choose two metrics: overlapping Normalized Mutual Information (NMI) [27] and Omega Index (Ω-Index) [28]. NMI is widely used to measure the performance of a community detection algorithm, which adopts the criterion used in information theory to compare the detected communities and the ground-truth communities. Omega Index is the overlapping version of Adjusted Rand Index, which is based on pairs of nodes in agreement in two partitions.

C. Comparative Methods

Clique Percolation Method (CPM) [29] is a typical overlapping community detection algorithm which assumes that communities consist of overlapping complete subgraph. SymNMF [12] is a general framework for graph clustering, which inherits the advantages of NMF by enforcing nonnegativity on the clustering assignment matrix. NSED [13] is a nonnegative symmetric encoder-decoder approach proposed for community detection. Node2Vec [20] is a node embedding algorithm, which adopts biased random walk and Skip-Gram model to embed vertices. LINE [21] is a node embedding algorithm, which preserves the first-order and the second-order proximity among embeddings. MNMF [22] is an NMF-based node embedding method which considers the microscopic structure (the first-order and second-order proximities) and mesoscopic community structure. Because LINE, Node2Vec and MNMF are representational learning algorithms, we use the K-means algorithm to detect communities based on their output node embeddings. NOCD [6] is a community detection algorithm with a graph neural network, which is defined by Bernoulli-Poisson model. ComE [24] is a framework for community detection, community embedding, and node embeddings, which jointly detects communities and learns the embeddings.

D. Parameter Setting

After obtaining the community affiliation weight matrix F, we need to convert it to the community affiliation matrix H. If F_{uc} is bigger than the threshold p, we believe that node u belongs to community c. The threshold p is defined as $p = \sqrt{-\log(1-\epsilon)}$ where ϵ is the background edge probability $\epsilon = 2|E|/|V|(|V|-1)$. The basic intuition about assigning nodes to communities is that if two nodes belong to the same community, then the probability of having an edge between them through the community should be bigger than the background edge probability [16]. Markov Time t is another hyperparameter that determines the time of the Markov process. In general, we set Markov Time t to 1. We also analyze the influence of Markov Time t on community detection with a certain community number.

E. Results of Community Detection

In Table II, we summarize the community detection results of the proposed CDMG and other competing methods on the real-world networks in terms of NMI and Omega Index. Compared with other algorithms, CDMG can achieve the highest NMI score and Omega Index in most cases, which means CDMG can provide a more robust community structure with higher accuracy. Specifically, in Amazon, ComE outperforms other algorithms, achieving the highest NMI and Omega Index. In LiveJournal, YouTube and DBLP, our method performs the best, especially, the Omega Index of CDMG is much larger than other algorithms' in these networks. One possible explanation for the superior performance of CDMG in these experimental networks is that because both GCN and Markov Stability deeply depend on network structure, CDMG

TABLE II

NMI SCORE AND OMEGA INDEX OF DIFFERENT METHODS ON SEVERAL EXPERIMENTAL NETWORKS

Network	LiveJournal		Amazon		YouTube		DBLP	
Metric	NMI	Ω-Index	NMI	Ω-Index	NMI	Ω-Index	NMI	Ω-Index
CPM	0.2196	0.0607	0.0995	0.0937	0.0000	0.0152	0.0439	0.0144
NSED	0.0293	0.0408	0.1104	0.0496	0.052	0.0543	0.0004	0.0000
SymNMF	0.0442	0.0556	0.1699	0.1621	0.1084	0.0963	0.0201	0.0458
MNMF	0.0219	0.0571	0.0000	0.0768	0.0872	0.1475	0.0030	0.0178
LINE	0.1324	0.0250	0.1815	0.0465	0.0473	0.0198	0.0171	0.0041
Node2Vec	0.2102	0.0636	0.2236	0.1793	0.0659	0.0487	0.0242	0.0223
NOCD	0.2007	0.1157	0.2323	0.1945	0.1773	0.2145	0.0631	0.0742
ComE	0.2542	0.0674	**0.2643**	**0.2447**	0.0973	0.0551	0.0419	0.0429
CDMG	**0.3015**	**0.5503**	0.1656	0.1909	**0.2453**	**0.4846**	**0.1141**	**0.2099**

TABLE III

NMI SCORE AND OMEGA INDEX OF CDMG ON SEVERAL EXPERIMENTAL NETWORKS WHILE MARKOV TIME T INCREASING.

Network	LiveJournal		Amazon		YouTube		DBLP	
Metric	NMI	Ω-Index	NMI	Ω-Index	NMI	Ω-Index	NMI	Ω-Index
1	0.3015	0.5503	0.1656	0.1909	0.2453	0.4846	0.1141	0.2099
2	0.2625	0.4236	0.1822	0.2041	0.2439	**0.5201**	0.1040	0.2665
3	0.2915	0.4956	0.2212	0.2135	0.2469	0.5127	0.1205	0.2779
4	0.3423	0.5656	0.2061	0.2041	**0.2604**	0.5080	0.1259	0.3661
5	0.3540	0.6445	0.2338	0.2415	0.2467	0.4813	0.1474	0.4353
10	0.4014	0.7429	0.2329	0.2591	0.2273	0.4883	0.1982	0.5273
20	**0.4670**	**0.8295**	0.2937	0.2803	0.1852	0.3683	0.2779	0.6060
50	0.4531	0.7720	0.3276	0.3242	0.1575	0.3109	0.3664	0.6320
100	0.4430	0.7898	0.3679	0.4081	0.0761	0.1028	**0.3991**	**0.6585**
200	0.4308	0.7851	0.4012	0.3971	0.0242	0.0091	0.3943	0.2692
500	0.4257	0.8060	**0.4212**	0.3983	0.0226	0.0132	0.1328	0.2692
1000	0.4177	0.8187	0.3894	0.3983	0.0214	0.0155	0.0872	0.0994
2000	0.2800	0.5485	0.4044	**0.4185**	0.0233	0.0140	0.0748	0.0788

is sensitive to the modification of network structure causing that CDMG can capture community structure well.

F. Influence of Markov Time on Community Detection

Table III and Fig. 3 demonstrate the influence of Markov Time t on the performance of CDMG. We can observe that for YouTube, at first the NMI score and Omega Index fluctuate within a certain range while Markov Time t increasing, but when t is above a threshold, the NMI score and Omega Index decay gradually. For LiveJournal, Amazon and DBLP, their NMI score and Omega Index increase gradually as t rises, and their highest NMI score and Omega Index are much larger than other algorithms' respectively. For DBLP when Markov Time t is above 100, both NMI score and Omega Index decrease gradually. For LiveJournal and YouTube their Markov Time thresholds are probably 20 and 4 respectively. To sum up, for CDMG there probably is a Markov Time threshold for a network. When Markov Time t is below the threshold, NMI score and Omega Index of community partition detected by CDMG either keeps relatively steady with a little fluctuation or increases gradually. Oppositely when Markov Time t is above the threshold, NMI score and Omega Index could decay gradually. According to this empirical observation, our method CDMG can provide a much better result when using the Markov Time t around the threshold.

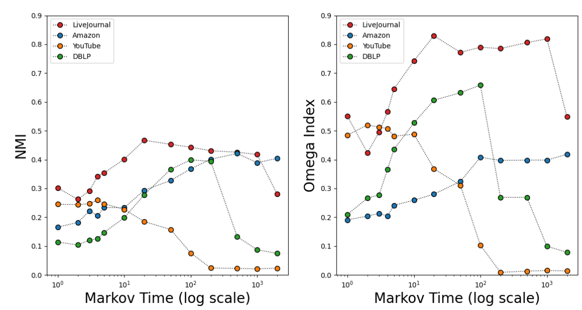

Fig. 3. NMI score and Omega Index of CDMG on the real-world networks including LiveJournal, Amazon, YouTube and DBLP with Markov Time t increasing.

G. Runtime Comparison

We compare the runtime of these community detection algorithms on the experimental networks with different scale including LiveJournal, Amazon, YouTube and DBLP. Fig. 4 illustrates that CPM, NSED, SymNMF and MNMF are faster than other algorithms in these real-world networks. The GNNs-based methods such as CDMG and NOCD have shorter runtime in comparison to Node2Vec and LINE, and CDMG is relatively faster than NOCD when tackling networks with more vertices. As to ComE, its great consumption of time is reasonable for it jointly addresses community detection, community embedding and node embedding. Even though CDMG is not the fastest method, the runtime of CDMG is still acceptable for its performance significantly outperforms other methods.

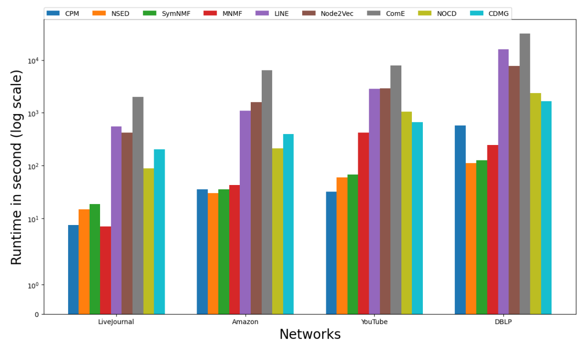

Fig. 4. Runtime of different community detection algorithms on the experimental networks including LiveJournal, Amazon, YouTube and DBLP.

V. Conclusion

In this paper, we propose a graph neural network model CDMG for overlapping community detection from the perspective of optimizing Markov Stability. Specifically, we train a graph neural network to generate the community affiliation weight matrix indicating community structure while maximizing Markov Stability. The experiments confirm that on several real-world networks our method outperforms other baseline methods considering both the accuracy and runtime. And we also explore the influence of Markov Time t on the performance of CDMG and find out that there is a Markov Time threshold for a network. When Markov Time t is around the threshold, CDMG can provide a better result with higher accuracy. The results of CDMG also demonstrate how powerful Graph Network Networks are.

Acknowledgment

This work was supported in part by the National Natural Science Foundation of China under Grants 62072352, U1708262.

References

[1] Bacik, Karol A., et al. "Flow-based network analysis of the Caenorhabditis elegans connectome." PLoS computational biology 12.8 (2016): e1005055.
[2] Zhang, Muhan, and Yixin Chen. "Link prediction based on graph neural networks." Advances in Neural Information Processing Systems 31 (2018): 5165-5175.Atwood, James, and Don Towsley. "
[3] Cai, Lei, et al. "Line graph neural networks for link prediction." IEEE Transactions on Pattern Analysis and Machine Intelligence (2021).
[4] Wu, Jun, Jingrui He, and Jiejun Xu. "Net: Degree-specific graph neural networks for node and graph classification." Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2019.
[5] Tsitsulin, Anton, et al. "Graph clustering with graph neural networks." arXiv preprint arXiv:2006.16904 (2020).
[6] Shchur, Oleksandr, and Stephan Günnemann. "Overlapping community detection with graph neural networks." arXiv preprint arXiv:1909.12201 (2019).
[7] Lambiotte, Renaud, Jean-Charles Delvenne, and Mau-ricio Barahona. "Random walks, Markov processes and the multiscale modular organization of complex networks." IEEE Transactions on Network Science and Engineering 1.2 (2014): 76-90.
[8] Delvenne, J-C., Sophia N. Yaliraki, and Mauricio Barahona. "Stability of graph communities across time scales." Proceedings of the national academy of sciences 107.29 (2010): 12755-12760.
[9] Delvenne, Jean-Charles, et al. "The stability of a graph partition: A dynamics-based framework for community detection." Dynamics On and Of Complex Networks, Volume 2. Birkhäuser, New York, NY, 2013. 221-242.
[10] Zachary, Wayne W. "An information flow model for conflict and fission in small groups." Journal of anthropological research 33.4 (1977): 452-473.
[11] Newman, Mark EJ. "Modularity and community structure in networks." Proceedings of the national academy of sciences 103.23 (2006): 8577-8582.
[12] Kuang, Da, Chris Ding, and Haesun Park. "Symmetric nonnegative matrix factorization for graph clustering." Proceedings of the 2012 SIAM international conference on data mining. Society for Industrial and Applied Mathematics, 2012.
[13] Sun, Bing-Jie, et al. "A non-negative symmetric encoder-decoder approach for community detection." Proceedings of the 2017 ACM on Conference on Information and Knowledge Management. 2017.
[14] Li, Ye, et al. "Community detection in attributed graphs: An embedding approach." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 32. No. 1. 2018.
[15] Wang, Fei, et al. "Community discovery using nonnegative matrix factorization." Data Mining and Knowledge Discovery 22.3 (2011): 493-521.
[16] Jia, Yuting, et al. "CommunityGAN: Community detection with generative adversarial nets." The World Wide Web Conference. 2019.
[17] Lobov, Ivan, and Sergey Ivanov. "Unsupervised community detection with modularity-based attention model." arXiv preprint arXiv:1905.10350 (2019).
[18] Hu, Fang, et al. "Community detection in complex networks using Node2vec with spectral clustering." Physica A: Statistical Mechanics and its Applications 545 (2020): 123633.
[19] Perozzi, Bryan, Rami Al-Rfou, and Steven Skiena. "Deepwalk: Online learning of social representations." Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining. 2014.
[20] Grover, Aditya, and Jure Leskovec. "node2vec: Scalable feature learning for networks." Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining. 2016.
[21] Tang, Jian, et al. "Line: Large-scale information network embedding." The World Wide Web Conference. 2015.
[22] Wang, Xiao, et al. "Community preserving network embedding." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 31. No. 1. 2017.
[23] Vaswani, Ashish, et al. "Attention is all you need." arXiv preprint arXiv:1706.03762 (2017).
[24] Cavallari, Sandro, et al. "Learning community embedding with community detection and node embedding on graphs." Proceedings of the 2017 ACM on Conference on Information and Knowledge Management. 2017.
[25] Beguerisse-Díaz, Mariano, Borislav Vangelov, and Mauricio Barahona. "Finding role communities in directed networks using role-based similarity, markov stability and the relaxed minimum spanning tree." 2013 IEEE Global Conference on Signal and Information Processing. IEEE, 2013.
[26] Kipf, Thomas N., and Max Welling. "Semi-supervised classification with graph convolutional networks." arXiv preprint arXiv:1609.02907 (2016).
[27] Aaron F McDaid, Derek Greene, and Neil Hurley. 2011. Normalized mutual information to evaluate overlapping community finding algorithms. arXiv:1110.2515 (2011).
[28] Gregory, Steve. "Fuzzy overlapping communities in networks." Journal of Statistical Mechanics: Theory and Experiment 2011.02 (2011): P02017.
[29] Palla, Gergely, et al. "Uncovering the overlapping community structure of complex networks in nature and society." nature 435.7043 (2005): 814-818.