Link Selection with Collaborative Low-rank and Sparse Factorization for Community Detection in Multiplex Networks

Dengdi Sun [1], Xinxin Wang [1], Zhuanlian Ding [2] [+] and Bin Luo [1]

[1] Anhui Provincial Key Lab of Multimodal Cognitive Computing, School of Computer Science and Technology, Anhui University, 230601, Hefei China

[2] School of Internet, Anhui University, 230039, Hefei China

Abstract. With the advent of multiple types of proximities between nodes, multiplex networks have emerged widely in the real world and been attracting increasing attention recently. The existing researches on community detection in multiplex networks usually assume that all layers come from a same latent consistent topology structure, and learn compatible and complementary information from different layers together, so as to dig out a shared community structure. However, this assumption is not satisfied in many real-world networks due to the existence of noisy/irrelevant links. To address this problem, in this paper we propose a multiplex network structure optimization algorithm based on collaborative low-rank and sparse factorization, which promotes the collaboration of different layers and lets them decompose as the robust consistent representations. In addition, an effective iterative algorithm is designed to optimize and solve the model. The experimental results in multiple ground-truth datasets show that this method can significantly improve the community detection performance of multiplex complex networks.

Keywords: Multiplex networks, Low-rank sparse representation, Link Selection, Community detection.

1. Introduction

Complex networks are both irregular and not completely random networks, and widely existed in real life, such as social networks, World Wide Web, protein interaction networks, flight network and so on. One of the most prominent properties of such networks is the community structure, which closely related to the attribute, function, organization and dynamic of a network. Although there is no precise definition of community structure at present, Newman [1] et al. put forward a widely accepted definition that the community should satisfy the characteristics of dense node connections within the community and relatively sparse node connections between the communities. Dividing the complex network into several local and close community structures is helpful for us to analyse the functional characteristics of the network, the topology of the network, and to understand the relationship between potential network nodes.

As researches on network science have become more and more mature, the research direction has started to transfer from simple graph theory to more complicated real system. In the same time, the focus has also gradually transferred from single isolated network to coupling network or multi-layer with mutual influence on each other, or multiplex networks [2]. The characteristics of multiplex networks are high dimension, complex structure and great difference between layers. However, existing algorithms cannot effectively detect its community structure, due to the ubiquitous noisy, meaningless or even inaccurate links within layer and the inconsistency between layers in real networks. In recent years, with the deepening of community detection research, community detection algorithms based on network structure optimization are emerging. By selecting the correct and effective links, structure optimization can make the communities clear and get better detection results. Unfortunately, the existing network structure optimization model is only suitable for single-layer network, so it is necessary to extend it to multiplex networks and deeply mine the complementary information between layers, so as to achieve better multi-layer fusion and collaboration.

Based on the analysis above, in this paper we propose a multi-layer network structure optimization algorithm based on collaborative low rank sparse factorization. Specifically, through the low rank sparse

---

[+] Corresponding author. Tel.: +86 551 65108582.
  E-mail address: dingzhuanlian@163.com

representations within each layer and collaborative constraint between different layers, and the overall factorization error approximation in multiplex networks, we get the link selection model with collaborative low-rank and sparse factorization. In addition, we design an effective iterative algorithm to optimize the model. Finally, experimental results on multiple multiplex networks with ground-truth show that our method is effective.

2. Related work

2.1. Single-Layer Networks

With single layer networks as the object, the most of existing approaches focus on the original topology structure, while the algorithm based on improving network structure is relatively few. The typical algorithms are briefly introduced as follows. GN algorithm [3] is a community detection algorithm that directly changes the network structure, and divides the community according to the proposed edge betweenness; Yang et al. [4] proposed a community detection algorithm based on semi supervised method, which obtains the node's supervision information (whether it belongs to a community or not) according to the corresponding strategy, and then connects the edges belonging to the same community node to delete the edges of different community nodes Edge-boost [5] is a community detection algorithm based on link prediction. According to the similarity index of link prediction, this method adds edges to the network according to the corresponding strategy to get the updated network, and then uses the traditional community detection algorithm to divide these networks.

2.2. Multiplex Networks

Although community detection algorithms for single-layer networks have achieved some results, these algorithms cannot be directly used for community detection in multi-layer networks. At present, the most commonly used method is based on single-layer network algorithm, and further extended to multi-layer network. It is mainly divided into the following categories: Based on modularization, the concept of modularization is extended from single-layer network to multi-layer network, such as (PMM) [6] algorithm and Didier, The main disadvantage of this method is that it can't distinguish the information rich network layer or complementary layer, and lose the global information; based on spectral clustering method, the representative algorithms are (SC-ML) [7] and (SC-GED) [8], the problem of this kind of method is that it can't fully deal with the network layer which is lack of connection or weak connection, or the layer which has broken part; The method based on information diffusion uses the concept of network diffusion to integrate the network layer. One of the methods is similar network fusion (SNF) proposed by Wang et al. [9]. However, for sparse networks, this method has low information dissemination efficiency, which may lead to poor clustering performance.

3. Methodology

3.1. Problem Formulation

A multiplex network is a set of K layer networks, $G_k(V, E_k)$, for $k = 1, \ldots, K$. The number of nodes in each layer is the same, $n = |V|$. For each individual layer of the adjacency matrix $\mathbf{A}^{(i)} \in \mathbb{R}^{n \times n}$, for $i = 1, \ldots, N$, if there is an edge connection between node v_i and node v_j, $A_{ij} = 1$, otherwise $A_{ij} = 0$.

3.2. Proposed Framework

In real networks, the widespread disturbances and noises obscure seriously the real community structures. Therefore, we desire to seek an efficient recovery of adjacency matrix to discover underlying community structures, reserve smooth structural low rank effective links(must-links), and get rid of the sparse noisy links(cannot-links). It is worth noting that, in the ideal recovery matrix, links appears more within communities, which makes the recovery matrix is approximately block diagonal after sorting node properly [10]. Thus, the column vectors of recovery matrix corresponding the nodes in the same community are also linearly correlated with each other, forming a low-rank recovery matrix L. With the low-rank constrain, the matrix recovery problem can be directly written in the following form:

$$\min_{L} \text{rank}(L) \text{ s.t. } L_{ij} = A_{ij}, \forall i,j \in \Omega \tag{1}$$

where $A \in \mathbb{R}^{n \times n}$ and $L \in \mathbb{R}^{n \times n}$ designate the given adjacency matrix, the desired recovery matrix respectively. $\Omega \in \{0,1\}^{m \times n}$ is a binary matrix denoting the must-link support, which if e_{ij} is a cannot-link $\Omega_{ij} = 0$; when e_{ij} is a must-link $\Omega_{ij} = 1$. And we define $\Omega^{\perp} + \Omega = \mathcal{P}_{0/1}(A)$, where $\mathcal{P}_{0/1}(A)$ represent the binaryzation of the matrix A:

$$\mathcal{P}_{0/1}(A)_{ij} = \begin{cases} 0, & \text{if } A_{ij} = 0 \\ 1, & \text{if } A_{ij} \neq 1 \end{cases} \tag{2}$$

For unweighted networks, $A = \mathcal{P}_{0/1}(A)$. Furthermore, considering the noise interference, for the must-links, where $\Omega_{ij} = 1$, we assume that $A_{ij} = L_{ij} + \varepsilon_{ij}$, where ε_{ij} denotes i.i.d. Gaussian noise. Thus, L_{ij} should be the best fitting to A_{ij} the least squares sense when $\Omega_{ij} = 0$, while it is meaningless to approximate the zero elements in matrix A. Based on the above assumptions, Eq.1 can be rewritten as,

$$\min_{L, \Omega_{i,j} \in \{0,1\}} \frac{1}{2} \|\Omega \circ (A - L)\|_F^2 + \beta \|\Omega^{\perp}\|_0, \text{ s.t. rank}(L) \leq \delta \tag{3}$$

where δ is a constant to be predefined. Intrinsically, δ constrains the complexity of the recovery networks. β is a penalized factor, $\|\cdot\|_0$ denotes the ℓ_0-norm, which counts the number of nonzero entries. $\|\cdot\|_F$ is the Frobenius norm, and the operator "o" denotes element-wise multiplication of two matrices. In addition, in order to preserve the community structure among must-links, we construct an undirected edge graph weighted with link similarity matrix S, including all the links from original network as the vertices in the edge graph. The similarity between link e_{ip} and link e_{jp} can be calculated by the Jaccard index [11] With the link similarity matrix constructed $S \in \mathbb{R}^{m \times m}$, it is essential to penalise the neighbor links with distant similarity, while enhancing the ones with close similarity. Therefore, we enforce the link similarity into the network topology with structural sparsity as follow:

$$\sum_{(e_{ij}, e_{pq}) \in \varepsilon} S_{(e_{ij}, e_{pq})} |\Omega_{ij} - \Omega_{pq}| \tag{4}$$

Finally, above process is structural optimization on the basis of the single-layer network. In the multi-layer network, two strategies are adopted in this paper to ensure that must-links can be extracted in different network layers in a collaborative way. In the single-layer network, the recovery matrix of the network is required to be low-rank. In the multi-layer network, we intend to cascade the background matrix L^k of each layer to form the joint background matrix L and requires L to maintain a low rank on the whole. In order to maintain multi-layer collaborative computing, we obtain a group of multi-layer consistent edge sets, including the common edges of the must-links edge sets of each layer to a certain extent, while computing the must-links edge sets of each layer. Assume that the indicator matrix of each layer is Ω^k. Since the number of nodes in each layer is the same, set the multi-layer consistent indicator matrix Ω of the same size as all Ω^k calibrate the multi-layer consistent must-links, and restrict the indicator matrix Ω^k of each layer to be as close as possible to the multi-layer consistent indicator matrix Ω.

Combining the above two collaborative constraints and the global decomposition error approximation of multi-layer network, a multi-layer network model based on structured low rank sparse decomposition is formed:

$$\min_{L^k, \Omega^k, \Omega} \sum_{k=1}^{K} \frac{1}{2} \|\Omega_{\perp}^k \circ (A^k - L^k)\|_F^2 + \alpha \|\Omega^k\|_0 + \beta \sum_{(e_{ij}, e_{pq}) \in \varepsilon^k} S_{il,jl}^k |\Omega_{ij}^k - \Omega_{pq}^k| + \eta \|\Omega^k - \Omega\|_F^2 \tag{5}$$

where, $L = [L^1, \cdots, L^K]$ is the joint background matrix, $\|\Omega^k - \Omega\|_F^2$ is the multi-layer consistent confidence constraint.

4. Model Optimization

Since Eq.5 is a NP problem, it is very difficult to optimize the indicator matrix Ω, by using the kernel norm to relax the rank operator on Ω, the kernel norm has been proved to be an effective convex substitution of the rank operator in [12]. Therefore, Eq.5 can be rewritten as follows:

$$\min_{L^k, \Omega_{i,j}^k \in \{0,1\}} \sum_{k=1}^{K} \frac{1}{2} \|\Omega^k \circ (A^k - L^k)\|_F^2 + \beta \|\Omega_{\perp}^k\|_1 + \alpha \|L^k\|_*$$
$$+ \gamma \sum_{(e_{ij}, e_{pq}) \in \varepsilon^k} S_{(e_{ij}, e_{pq})}^k |\Omega_{ij}^k - \Omega_{pq}^k| + \eta \|\Omega^k - \Omega\|_F^2 \tag{6}$$

where β, α, γ and η are the tuning parameters. Therefore, the Eq.6 is decomposed into energy minimization L^k and Ω^k by using the alternating iterative algorithm, and the following three sub-problems are solved to separate the optimal solutions of L^k, Ω^k and Ω.

4.1. L^k- subproblem

Given an current estimate of the must-link support Ω^k, the minimization in Eq.6 over L turns out to be the matrix completion problem:

$$\min_{L^k} \frac{1}{2}\|\Omega^k \circ (A^k - L^k)\|_F^2 + \alpha\|L^k\|_* \tag{7}$$

There is a clear drawback in Eq.7 above, that is the nuclear norm minimization requires to execute expensive singular value decompositions on the whole data. To mitigate the computational pressure, we decompose the recovery matrix L^k into two sub-matrices as $L^k = U^k V^k$, where $U^k \in \mathbb{R}^{n \times r}$ and $V^k \in \mathbb{R}^{r \times n}$ based on Mazumder's study [13]. If $rank\,(L^k) \le min(m,n)$, then the minimum solution of (7) is attained as:

$$\min_{L^k}\|P^k - U^k V^k\|_F^2 + \alpha(\|U^k\|_F^2 + \|V^k\|_F^2) \tag{8}$$

where $P^k = \Omega^k \circ A + \Omega_{\perp}^k \circ L^k$, which can be minimized along one coordinate direction at each iteration. Fix the other variables and update U^k and V^k respectively, so this problem has a closed-form solution:

$$U^k = P^k V^{k^\top}(V^k V^{k^\top} + \alpha I)^{-1} \tag{9}$$

$$V^k = (U^{k^\top} U^k + \alpha I)^{-1} U^{k^\top} P^k \tag{10}$$

where I is the identity matrix. Finally, the recovery matrix L^k can be combined via $L^k = U^k V^k$.

4.2. Ω^k- subproblem

Given a current estimate of the low-rank recovery matrix L^k, Eq.6 can be transferred into following optimization function:

$$\min_{\Omega^k} \frac{1}{2}\|\Omega^k \circ (A^k - L^k)\|_F^2 + \beta\|\Omega_{\perp}^k\|_1 + \gamma \sum_{(e_{ij},e_{pq}) \in \varepsilon^k} S_{(e_{ij},e_{pq})}^k |\Omega_{ij}^k - \Omega_{pq}^k| \tag{11}$$

Leveraging the complementary, the $\Omega_{\perp}^k$ can be transformed in to $\Omega_{\perp}^k = \mathcal{P}_{0/1}(A) - \Omega^k$. Noticing that $S_{ij}^k \in \{0,1\}$, so Eq.13 can be expanded as follows:

$$\frac{1}{2}\sum_{ij}(A_{ij}^k - L_{ij}^k)^2 \Omega_{ij}^k + \beta \sum_{ij}(\mathcal{P}_{0/1}(A^k)_{ij} - \Omega_{ij}^k) + \gamma \sum_{(e_{ij},e_{pq}) \in \varepsilon^k} S_{(e_{ij},e_{pq})}^k |\Omega_{ij}^k - \Omega_{pq}^k|$$

$$= \sum_{ij}(\frac{1}{2}(A_{ij}^k - L_{ij}^k)^2 - \beta)\,\Omega_{ij}^k + \gamma \sum_{(e_{ij},e_{pq}) \in \varepsilon^k} S_{(e_{ij},e_{pq})}^k |\Omega_{ij}^k - \Omega_{pq}^k| + C \tag{12}$$

Where L^k is fixed, and $\beta \sum_{ij}(\mathcal{P}_{0/1}(A^k)_{ij})$ is a constant, independent of Ω^k. The above function is similar to the first-order Markov Random Field model, so we can easily solve for the must-link support Ω^k by graph cutting theory.

4.3. Ω- subproblem

Obtain the low-rank matrix L^k and current mask matrix Ω^k under a certain layer of multi-layer network, and update Ω by $\min_{\Omega} \sum_{k=1}^{K}\|\Omega - \Omega^k\|_F^2$ .Finally we have a closed solution:

$$\Omega = \frac{1}{K}\sum_{k=1}^{K}\Omega^k \tag{13}$$

Above all, the complete pseudocode of our sparse and structured low-rank representation for multi-layer network is described in Algorithm 1. According to Algorithm 2, we make use of our model to obtain the optimized multi-network which we finally need. The framework of the whole model is shown in Fig 1.

Algorithm 1: Sparse and Structured Low-rank Representation for Multi-layer network

Input: Multilayer network G_k, k = 1, ..., K, parameter γ and δ.

Initialize: $\alpha, \beta, S^k \leftarrow 0, L^k \leftarrow A^k, U_O, V_0, \tau = 1e - 4$, maxIter $= 30$.

Output: L^k, Ω^k, Ω.

1: repeat

2: Updating L^k: optimizing energy function Eq.8 via:
$$U^k \leftarrow P^k V^{k^\top}\left(V^k V^{k^\top} + \alpha I\right)^{-1}$$
$$V^k \leftarrow \left(U^{k^\top} U^k + \alpha I\right)^{-1} U^{k^\top} P^k$$
$$L^k \leftarrow U^k V^k$$

3: if ran k$(L^k) \leq \delta$ then

   $\alpha \leftarrow \eta_1 \alpha$

   go to Step 2

   end if

4: $\beta \leftarrow max(\eta_2 \beta, 4.5\sigma^2)$

5: Updating Ω^k : using graph cuts to optimize energy function Eq.12.

6: Updating Ω : get the Ω by Eq.13.

7: until convergence: if the maximum objective change between two consecutive iterations is less than τ or the maximum number of iterations reaches maxIters, then terminate the loop.

Algorithm 2: Adaptive Link Selection Using Sparse and Structured Low-rank Representation for Multi-layer network

Input: Multilayer network G_k, k = 1, ..., K, parameter γ and δ.

Output: Selected efficient links and noisy links

1: Computing the link similarity matrix S^k using Eq.3.

2: Solve the Eq.6 using Algorithm 1, and obtain an optimal solution (L^k, Ω^k, Ω).

3: Select the links corresponding to$(\Omega^k \circ L^k_{ij}) = 1$ as efficient links.

4: $\Omega^k_\perp = \mathcal{P}_{0/1}(A^k) - \Omega^k$.

5: Select the links corresponding to $(\Omega^k_\perp \circ L^k_{ij}) = 1$ as noisy links.
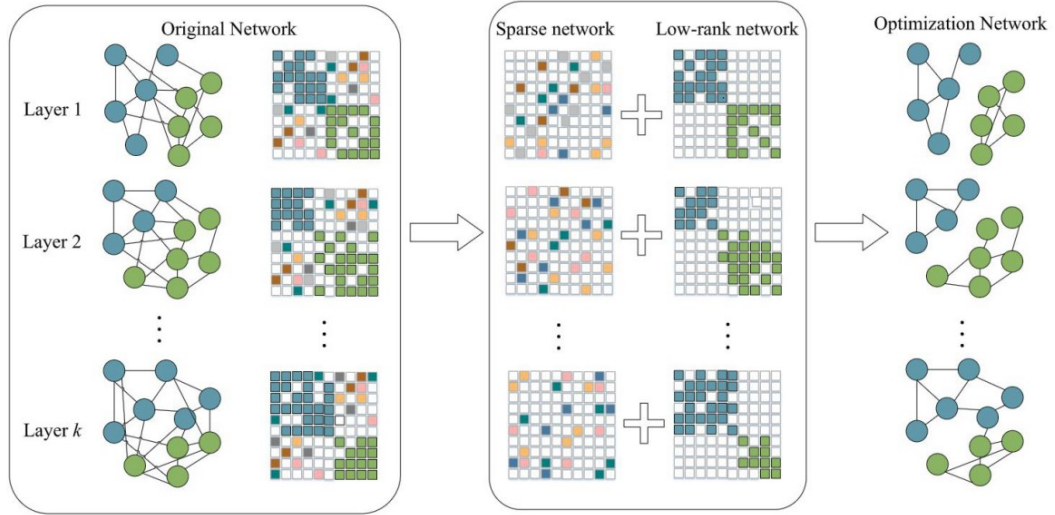

Fig. 1: Overview of the proposed framework

5. Experiments

5.1. Datasets

In order to verify the effect of our method in different multi-layer networks, we selected 7 types of non-overlapping multi-layer networks as datasets, include：Football [14], Olympics [14], WBN [15], WTN (cat) [15], WTN (reg) [16], Cora [17] and CiteSeer [17]. The important statistics and information of multiplex networks are shown in Table 1.

Table 1: Multiplex Networks Used for Our Comparative Study.

Networks	Nodes	Communities	Layers
Football	248	20	3
Olympics	464	28	3
WBN	279	10	5
WTN (cat)	214	3	7
WTN (reg)	214	5	7
Cora	2708	7	2
CiteSeer	3312	6	2

5.2. Evaluation Indicator

In this paper, we mainly use three kinds of community detection indexes, including Normalized Mutual Information (NMI) [18], Accuracy(ACC) [19] and Multislice Modularity($Q_{\text{multislice}}$) [20]. The first two measures provide a quantitative way to compare the computed clusters with respect to the ground truth classes. To trade off the quality of the clustering against the number of clusters we use NMI. Accuracy (ACC) is used to measure the percentage of correct labels obtained. Given a data set containing n vertexes, for each sample network, let l_i be the cluster label we obtain by applying different algorithms and r_i be the label provided by the data set. When the real community result in the dataset is unknown, we use the $Q_{\text{multislice}}$ as the evaluation index of the method.

5.3. Parameter Setting

There are four important parameters in our model. In the experiment, we first estimate roughly the value of δ in Eq.12, and reduces α from a larger value. After each iteration, if rank $(L^k) \leq \delta$, we reduce α by a factor $\eta_1 < 1$ and repeat until rank $(L^k) \geq \delta$. In practice, we initialize empirically α to be the second largest singular value of A^k, and $\eta_1 = 1/\sqrt{2}$. Meanwhile, the parameter β in controls the sparsity of the outlier support. From Eq.14, we can see that Ω_{ij}^k is more likely to be 1 if $(A^k_{ij} - L^k_{ij})^2/2 > \beta$. Thus, the choice of β should depend or the noise level in network. Typically, we set $\beta = 4.5\sigma^2$ in algorithm 1, where σ^2 is estimated online by the variance of $A^k_{ij} - L^k_{ij}$. In addition, for the other two parameters δ and γ, we will tune them and show their influence in our experiments in following sections.

5.4. Experiment Results

● Community Detection with Known Ground-truth

In order to highlight the optimization of multi-layer network topology by our method, we conduct clustering experiments on the original network with ground-truth datasets and the network optimized by our method. Here we use four common multi-layer network clustering methods, including: SNF [9], SC-GED [10], PMM [6] and SC-ML [7]. The experimental results are shown in Table 2-5. The experiment shows that our method has significant effect on multi-layer network structure promote.

Table 2: Community Detection Results of SNF Algorithms in Datasets with Ground-truth

Metrics	Methods	Football	Olympics	WBN	WTN (cat)	WTN(reg)	Cora	CiteSeer
NMI	original	0.7412	0.8193	0.2265	0.3178	0.2943	0.0176	0.0489
	ours	**0.7688**	**0.8126**	**0.2419**	**0.3814**	**0.4296**	**0.1348**	**0.1502**
AC	original	0.7108	**0.7619**	**0.7632**	0.7920	0.5211	0.2229	0.2914
	ours	**0.8042**	0.7527	0.7479	**0.8216**	**0.5593**	**0.3198**	**0.4269**

Table 3: Community Detection Results of SC-GED Algorithms in Datasets with Ground-truth

Metrics	Methods	Football	Olympics	WBN	WTN (cat)	WTN(reg)	Cora	CiteSeer
NMI	original	0.6593	0.5936	0.3649	0.0341	0.0432	0.0037	0.0022
	ours	**0.8123**	**0.7243**	**0.4312**	**0.2286**	**0.1297**	**0.2213**	**0.0149**
AC	original	0.6471	0.5214	0.2176	0.5116	0.5603	0.2136	0.4391
	ours	**0.7732**	**0.6552**	**0.4859**	**0.5693**	**0.6319**	**0.3972**	**0.4828**

Table 4: Community Detection Results of PMM Algorithms in Datasets with Ground-truth

Metrics	Methods	Football	Olympics	WBN	WTN (cat)	WTN(reg)	Cora	CiteSeer
NMI	original	0.2473	0.1892	0.2086	0.0640	0.0321	0.0098	0.0067
	ours	**0.4418**	**0.3274**	**0.3714**	**0.1423**	**0.1239**	**0.0871**	**0.0594**
AC	original	0.1856	0.1307	0.3147	0.4421	0.2744	0.1647	0.2047
	ours	**0.2430**	**0.2396**	**0.3892**	**0.4919**	**0.4473**	**0.2864**	**0.3428**

Table 5: Community Detection Results of SC-ML Algorithms in Datasets with Ground-truth

Metrics	Methods	Football	Olympics	WBN	WTN (cat)	WTN(reg)	Cora	CiteSeer
NMI	original	0.8403	0.9241	0.3477	0.1016	0.2649	0.1137	0.0741
	ours	**0.8367**	**0.9310**	**0.4143**	**0.1849**	**0.3941**	**0.1439**	**0.1082**
AC	original	0.8037	**0.8929**	0.4190	0.4473	0.3826	**0.3063**	0.2768
	ours	**0.8341**	0.8847	**0.5397**	**0.4634**	**0.5147**	0.2954	**0.3690**

- Community Detection with Unknown Ground-truth

To quantify or evaluate the level of community partition of unlabelled multi-layer networks, the Multislice Modularity $Q_{\text{multislice}}$ is a viable indicator. In this section, we will compare the modularity of the original network, our method in 7 real network datasets without ground-truth, as shown in Table 6.

Table 6: $Q_{\text{multislice}}$ of Different Network Recovery Methods in Datasets without Ground-truth.

	Football	Olympics	WBN	WTN (cat)	WTN (reg)	Cora	CiteSeer
original	0.7435	0.6811	0.5324	0.7206	0.7127	0.4275	0.3921
ours	**0.7629**	**0.7353**	**0.6152**	**0.7461**	**0.7583**	**0.4526**	**0.4178**

Here, for network without ground-truth, we identify the different recovered networks using unified the method proposed by Mucha [20]. From the results of Table 6, we can see that our matrix recovery method has greatest improved the modularity of the original network, which shows that our method has obvious advantageous on the community partition task with unknown labels.
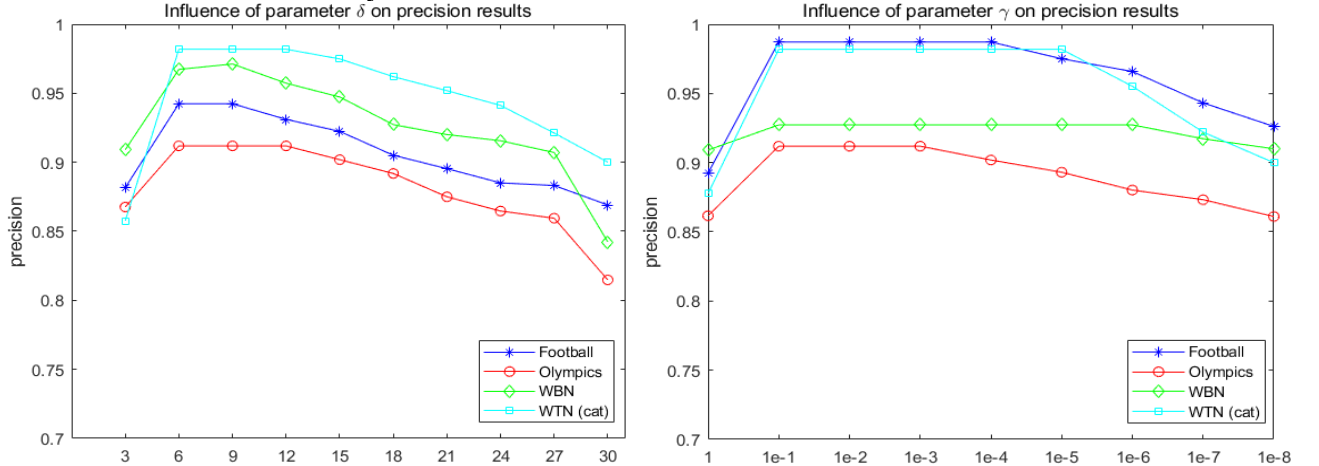
5.5. Parameter Sensitivity

Fig. 2: The influence of parameters to precision of our model.

As mentioned above, we have only two parameters to choose, δ and γ. To simplify, here we calculate the proportion of the same elements between the recovery matrix and the original adjacency matrix as the evaluation precision. From Fig.2, for δ value, after repeated tests, we find that the change of δ value will not affect the results on 4 different networks, as long as it is within a reasonable range. Our results show that the change of δ value does not affect precision, δ is desired rank of the estimated low-rank component. In addition, the parameter γ is a constant force controlling smoothing normalization in Eq.14. Experiments show that the value of γ are also robust to the results.

6. Conclusion

In this paper, we studied selecting links for complex networks with multiple layers. We proposed an effective framework to let different layers collaborate with each other and capture the consistent low-rank and sparse representation across different layers. In addition, we proposed an effective optimization scheme, which does not require any prior labelled data. We evaluated the performance of our proposed approach on many real-world databases. Experimental results on both the multiplex with ground-truth and without ground-truth demonstrated the effectiveness of our proposed framework. In the future, we plan to apply our framework to more applications, such as large scale multiplex and heterogeneous networks.

7. Acknowledgements

8. References

[1]   Newman, M. Networks: an introduction. Oxford university press, 2010.

[2]   Kivel, M, et al. Multilayer networks[J]. Journal of Complex Networks, 2018, 2(3):203-271.

[3]   M. Girvan, et al. Community structure in social and biological networks, PNAS 2020, 99 (12): 7821-7826.

[4]   L. Yang, et al. Active link selection for efficient semi-supervised community detection, Scientific reports 2015.

[5]   Burgess M, et al. Link-Prediction Enhanced Consensus Clustering for Complex Networks[J]. PLoS ONE, 2015.

[6]   L. Tang, et al. Uncovering groups via heterogeneous interaction analysis, ICDM 2009, pp. 503–512.

[7]   X. Dong, et al. Clustering on multi-layer graphs via subspace analysis on Grassmann manifolds, IEEE Transactions on signal processing 2013, 62 (4) 905 918.

[8]   X. Dong, et al. Clustering with multi-layer graphs: A spectral perspective, IEEE Transactions on Signal Processing 2012, 60 (11) 5820–5831.

[9]   B. Wang, et al. Similarity network fusion for aggregating data types on a genomic scale, Nature methods 2014,11.

[10] C. Lu, et al. Subspace clustering by block diagonal representation, IEEE Transactions on Pattern Analysis and Machine Intelligence 2019, 41 (2) 487–501.

[11] Garcia J O, et al. Applications of Community Detection Techniques to Brain Graphs: Algorithmic Considerations and Implications for Neural Function[J]. Proceedings of the IEEE, 2018.

[12] Sidiropoulos N D, et al. Tensor Decomposition for Signal Processing and Machine Learning[J]. IEEE Transactions on Signal Processing, 2017, PP (13):3551-3582.

[13] R. Mazumder, et al. Spectral regularization algorithms for learning large incomplete matrices, Journal of machine learning research, 2010, 11: 2287–2322.

[14] D. Greene, P. Cunningham, Producing a unified graph representation from multiple social network views, in: Proceedings of the 5th annual ACM web science conference, ACM, 2013, pp. 118–121.

[15] White J G, et al. The Structure of the Nervous System of the Nematode Caenorhabditis elegans[J]. Philosophical Transactions of the Royal Society of London, 1986, 314(1165):1-340.

[16] M. De Domenico, et al. Structural reducibility of multilayer networks, Nature communications 2015,6: 6864.

[17] P. Sen, et al. Collective classification in network data, AI magazine 2008, 29 (3): 93–93.

[18] Lancichinetti A, et al. Detecting the overlapping and hierarchical community structure of complex networks[J]. New Journal of Physics, 2009, 11(3).

[19] H. Liu, et al. Constrained nonnegative matrix factorization for image representation, IEEE Transactions on Pattern Analysis and Machine Intelligence 34 (7) (2011) 1299–1311.

[20] Mucha P J, et al. Community structure in time-dependent, multiscale, and multiplex networks[J]. Science, 2010, 328(5980): 876-878