

# **A Multi-objective Evolutionary Algorithm Based on Multi-layer Network Reduction for Community Detection**

Xin Qi<sup>1</sup>, Langzhou He<sup>1</sup>, Jiaxin Wang<sup>4</sup>, Zhanwei Du<sup>3</sup>, Zheng Luo<sup>1</sup>, and Xianghua Li<sup>1,2( $\boxtimes$ )</sup>

 $1$  College of Computer and Information Science, Southwest University, Chongqing 400715, China

li xianghua@163.com

 $2^2$  School of Artificial Intelligence, Optics and Electronics (iOPEN), Northwestern Polytechnical University, Xi'an 710072, China

<sup>3</sup> School of Public Health, The University of Hong Kong, Hong Kong, China zwdu@hku.hk

<sup>4</sup> State Grid Information and Telecommunication Group Co. Ltd., Beijing, China wangjiaxin1@sgitg.sgcc.com.cn

**Abstract.** Community detection is an important method to reveal the characteristics of complex systems, which usually requires the system to meet the conditions of close connections within communities but sparse connections between communities. In view of this, community detection has been proven to be an NP-hard problem. Multi-objective evolutionary algorithm (MOEA) is an indispensable aspect of multi-layer network community detection. However, most MOEA-based multi-layer network detection algorithms only take the acquired prior information as the network preprocessing method and ignore its full utilization in optimization, resulting in the accuracy of network partition cannot be guaranteed. To this end, this paper proposes a multi-objective community detection algorithm based on multi-layer network reduction (MOEA-MR). Specifically, we use the non-negative matrix factorization method to generate the consistent prior information layer of multi-layer network. Based on this, a network reduction strategy based on node degree is constructed to recursively reduce the size of the prior information network. In addition, in the evolution process, we consider using the multi-layer network similarity to correct the mis-divided nodes in the local reduction community. Compared with other advanced community detection algorithms, the experimental results on the real-world and synthetic multi-layer networks proved the superiority of MOEA-MR.

**Keywords:** Multi-layer network reduction  $\cdot$  Community detection  $\cdot$  Multi-objective evolution  $\cdot$  Consensus prior information  $\cdot$  Dice similarity

# **1 Introduction**

A variety of complex systems in the real world can be modeled as complex networks [\[1](#page-11-0),[2\]](#page-11-1). The traditional single-layer network can no longer meet the requirement of the existing research which focuses on diversified entities, so the multi-layer network with richer attributes has gradually become a research hotspot [\[3](#page-11-2)]. Community detection is essential to the understanding of the information and function of networks. Its overall goal is to divide a network into multiple clusters (communities). Many existing community detection algorithms focus on the topology of network, and require the sparse connections of intracluster but dense connections of inter-cluster. The current single-layer network clustering methods are relatively mature. For example, Girvan et al. propose the GN algorithm, which applies the concept of modularity function to network clustering for the first time [\[4\]](#page-11-3). Louvain algorithm obtains the network clustering results by optimizing the modularity repeatedly [\[5\]](#page-11-4). However, due to the complex characteristics of multi-layer networks, the single-layer network clustering methods applied on multi-layer networks have suboptimal performance.

Therefore, multi-layer network community detection algorithms based on different strategies were proposed in the past few years. Since the process of community detection is an NP-hard problem and Pareto optimal framework can provide a set of optimal compromise solutions based on optimization goals, the optimization algorithm has received great attention. According to the number of optimization goals, these algorithms can be divided into two types: single-objective optimization and multi-objective optimization. For single-objective optimization methods, only one objective function is selected for optimization in the iterative process. For example, GACD algorithm obtains good clustering results by optimizing modularity [\[6](#page-11-5)]. Pizzuti et al. implement network clustering by optimizing the fitness function [\[7\]](#page-11-6). For multi-objective methods that consider multiple objective functions simultaneously, the multi-objective evolution algorithm (MOEA) is strongly competitive in optimization calculations. For example, MOGA-Net algorithm introduces two functions (i.e., the community score and community fitness) for optimization calculation [\[8\]](#page-11-7). Moreover, Shi et al. use the concept of inter-objective correlation to develop a multi-objective optimization framework [\[9\]](#page-11-8). Furthermore, MOGA-@Net algorithm takes the structure of community and the similarity of nodes into consideration in order to obtain a high-quality solution [\[10](#page-11-9)]. Compared with single-objective optimization algorithms, multiobjective methods perform better because they focus on the links between the inter-community and intra-community simultaneously. However, the individual length of network code is proportional to the number of network nodes, which means that the search space of MOEA increases exponentially. What's more, they usually only take the topology information of network into full consideration without paying attention to the prior information.

To address the above problems, this paper proposes a multi-objective evolutionary algorithm based on multi-layer network reduction (MOEA-MR), which is used for community clustering in multi-layer networks. In the proposed MOEA-MR, a reduction strategy is used for network processing, which is based on the network consensus prior information layer, and the repairing strategy in the optimization process further improves the quality of network division. Specifically, the main contributions of this paper can be summarized as follows:

- The symmetric non-negative matrix factorization method is suggested for better applying the prior information of multi-layer networks. With this strategy, we obtain the non-negative low-dimensional representations of each network layer, and fuse them into a consensus prior information representation to enable clustering interpretation.
- A network reduction strategy based on the node degree is proposed for complex network clustering to reduce the computational complexity effectively. Specifically, local communities can be transformed into nodes of reduced network during the optimization, which is gainful for improving the scalability of large-scale multi-layer network.
- A network repairing strategy is proposed to correct the misidentified nodes after multi-layer network reduction, which helps improve the accuracy of clustering result.

The rest of the paper is organized as follows. Section [2](#page-2-0) introduces the related work. The proposed MOEA-MR method is described in detail in Sect. [3.](#page-3-0) After that, Sect. [4](#page-8-0) compares our proposed algorithm with several advanced network clustering methods. Finally, Sect. [5](#page-10-0) gives a summary of this paper.

## <span id="page-2-0"></span>**2 Related Work**

The main goal of community detection is to divide a network into different subgraphs, among which the internal connections are maximized while the external links are minimized. In complex networks, these subgraphs are also called communities. Unlike single-layer networks, each layer of multi-layer network represents a kind of relationship. This means that we need to consider the information of different layers when clustering the network, and find a final partition that is most suitable for each layer of multi-layer network. An example of multi-layer network clustering is provided in Fig. [1.](#page-2-1)



<span id="page-2-1"></span>**Fig. 1.** Illustration of multi-layer networks with 8 nodes given in three different layers (i.e.,  $G_1$ ,  $G_2$ , and  $G_3$ ). Two different color communities (i.e.,  $C_1$ ,  $C_2$ ) can be identified in all layers.

For better multi-layer network clustering, several standard functions are proposed to measure the link density of inter-community and intra-community. Here we formulate the multi-layer network clustering into a two-objective clustering optimization problem, utilizing ratio cut (RC) and kernel k-means (KKM) [\[11\]](#page-11-10) as the objective function. Given a multi-layer network  $G = (V, E_m)$  $(m = 1, ..., M)$ , where M represents the total number layers of multi-layer network and  $E_m$  denotes the edge set of the  $m^{th}$  layer. Moreover,  $V = \{v_i\}_{i=1}^N$ <br>indicates a set of nodes shared by all layers in G. The adjacency matrix of mulindicates a set of nodes shared by all layers in G. The adjacency matrix of multilayer network is expressed as:  $A^{(m)} \in R^{N*N}(m = 1 - M)$ . The definition of tilayer network is expressed as:  $A^{(m)} \in R_+^{N*N}(m = 1, ..., M)$ . The definition of bi-objective minimization problem is shown in Eq. (1) bi-objective minimization problem is shown in Eq. [\(1\)](#page-3-1).

<span id="page-3-1"></span>
$$
\min \begin{cases} KKM = 2(N-c) - \sum_{i=1}^{c} \frac{L(V_i, V_i)}{|V_i|} \\ RC = \sum_{i=1}^{c} \frac{L(V_i, \overline{V}_i)}{|V_i|} \end{cases}
$$
(1)

where c is the number of multi-layer network division.  $L(V_i, V_i)$  and  $L(V_i, V_i)$ <br>represent the internal and external connection density in the same community where c is the number of multi-layer network division.  $L(v_i, v_i)$  and  $L(v_i, v_i)$  represent the internal and external connection density in the same community i, respectively.  $KKM$  means the density of intra-community, and RC denotes the density of inter-community. So far,  $KKM$  and RC have broad applications in MOEA algorithm and achieve a good performance. Therefore, we also adopt the same functions as optimization objectives.

# <span id="page-3-0"></span>**3 Proposed Method**

In this section, we present a detailed introduction of the proposed MOEA-MR. First, owing to the fact that the construction of multi-layer network consensus information and the network reduction method are essential to MOEA-MR, we describe them in Sect. [3.1](#page-3-2) and [3.2](#page-4-0) in detail respectively. After that, we explain the community repairing strategy based on the similarity information in Sect. [3.3,](#page-5-0) which is used to modify and simplify the dis-divided nodes in the local network. Section [3.4](#page-6-0) demonstrates the genetic operators. Finally, the comprehensive framework of MOEA-MR is illustrated in Sect. [3.5.](#page-7-0)

#### <span id="page-3-2"></span>**3.1 Consensus Information Layer Construction**

The difference between multi-layer network and single-layer network is that the former has more complex features, and the noise produced by different layers can lead to a poor clustering result. Existing research has proved that the accuracy of clustering results can be significantly improved after integrating the network prior information [\[12\]](#page-11-11). However, most semi-supervised optimized clustering methods only focus on single-layer networks. Therefore, we propose a concept of multi-layer network optimization clustering based on the prior information.

Up to now, the symmetric non-negative matrix factorization method has been successfully applied to network clustering problem [\[13](#page-11-12)], so our work also takes the idea into consideration to obtain the prior information of multi-layer networks. Firstly, each network layer is denoised by calculating its non-negative lowdimensional representation. After that, we fuse the low-dimensional representations of all layers into a common consensus information layer, which is applied to subsequent optimization process to improve the effectiveness of MOEA-MR algorithm. In particular, given a network represented by an N-layered adjacency matrix  $(A_1, A_2, ..., A_n)$ , Eq. [\(2\)](#page-4-1) is utilized to get the consensus information layer of the multi-layer network.

<span id="page-4-1"></span>
$$
A_{\text{cons}} = \frac{\sum_{i=1}^{n} A^{(i)} + \gamma H^{(i)} H^{(i)T}}{n(1+\gamma)}
$$
(2)

where  $\gamma$  represents a constant. For each network layer *i*, we have the following constraints:  $A^{(i)} \approx H H^T H > 0$  and  $H^T H = I$ . We interpret  $A_{\text{max}}$  as the conconstraints:  $A^{(i)} \approx HH^T, H > 0$  and  $H^TH = I$ . We interpret  $A_{\text{cons}}$  as the con-<br>sensus information matrix of extracting potential communities shared by the sensus information matrix of extracting potential communities shared by the multilayer network and apply it to the clustering coding.

#### <span id="page-4-0"></span>**3.2 Network Reduction Strategy**

To reduce the computational complexity of algorithm, a network reduction strategy is proposed to reduce the network scale during the process of optimization. Moreover, when the network topology is used for multi-objective network clustering based on genetic operations, some nodes are always in an indivisible group. Therefore, we can regard these indivisible nodes as a whole in the calculation, which will remain unchanged even in the later optimization.

The description of network reduction process proposed in this paper is given as follows. Considering that the network degree is an important feature of network adjacency matrix, our reduction strategy takes the node degree as an important division criterion. At first, the node with the largest degree is selected as an initial node after calculating degrees of all nodes in a network. Then we divide the initial node and its neighbors into a tentative community, and check all the nodes in it. Note that the node in the tentative community will be removed if the external connections number is more than half degree of itself. Finally, the tentative community after inspection can be guaranteed as a strong community, and all components in the tentative community are reduced to one node. Checking all remaining nodes of network until they have been completely represented as the reduced community.



<span id="page-4-2"></span>**Fig. 2.** Illustration of reduction strategies. (a)  $v<sub>5</sub>$  is selected due to its largest node degree. (b) The tentative community determined by  $v_5$  contains  $\{v_1,v_2,v_3,v_4,v_5,v_6\}$ . (c) Delete  $v_6$  from the subgraph. (d) The subgraph  $C'$  consists  $\{v_1, v_2, v_3, v_4, v_5\}$  is compressed into one node based on the reduction strategy.

Figure [2](#page-4-2) gives an example about representing the main idea of prereduction process, which considers a network with 8 nodes. First, the tentative community  $\{v_1,v_2,v_3,v_4,v_5,v_6\}$  depicted in Fig. [2\(](#page-4-2)b) is obtained based on  $v_5$ . Then,  $v_6$  is eliminated because the links between this node and tentative subgraph  $\{v_1,v_2,v_3,v_4,v_5,v_6\}$  are less than half of its degree, as shown in Fig. [2\(](#page-4-2)c). Finally, by merging  $\{v_1,v_2,v_3,v_4,v_5\}$  into one node, Fig. [2\(](#page-4-2)d) shows the reduced network  $\{C', v_6, v_7, v_8\}.$ 

#### <span id="page-5-0"></span>**3.3 Network Repairing Method**

There are some reduced networks being found incorrect in the iterative evolution. To put it simply, the reduced network may find some nodes do not belong to itself, which is prone to lead to the poor performance of MOEA-MR. In order to settle this problem, we propose a reduced network repairing strategy based on the similarity prior information to correct the mis-divided nodes.

The similarity index of node is one of the most powerful indicators to evaluate the connectivity strength with node pairs. In order to preserve the structural characteristics of multi-layer networks, Dice index [\[14\]](#page-11-13) based on the connections of all layers is determined to calculate the similarity prior information of multilayer network, which can be depicted in Eq. [\(3\)](#page-5-1).

<span id="page-5-1"></span>
$$
D_s = \frac{2 \times Neighbours_{com}(v_i, v_j)}{Deg(v_i) + Deg(v_j)}
$$
\n(3)

where  $Neighbours_{\text{com}}(v_i, v_j)$  indicates the links of common neighbors of nodes  $v_i$  and  $v_j$ , and  $Deg(v_i)$  means the degree of node  $v_i$ . It can be seen that the equation mentioned above considers the neighbors information of all layers when calculating the neighbor links of nodes.



<span id="page-5-2"></span>**Fig. 3.** Illustration of network repairing method. (a) The Dice similarity of  $v_1$  and two local reduced network (e.g.,  $C'$ ,  $C''$ ). (b) The repairing results corresponding to the local network  $C'$  and  $C''$  with threshold  $\sigma = 0.2$ .

After obtaining the similarity prior information of multi-layer network, the steps of the reduced network repairing strategy in this paper perform as follows. Specifically, for each community in the reduced network, checking whether the nodes it contains are selected in the same community with their most similar

neighbor. If not, the nodes verified are mis-divided and need to be moved to other appropriate community. What's more, we give a threshold  $\sigma$  to judge the similarity of the node pairs, which means only the first  $\sigma\%$  of node similarities are taken into consideration. In that case, we need to give up those smaller similarity value. Figure [3](#page-5-2) shows the repairing strategy based on the reduced network with 8 nodes, and threshold  $\sigma$  is equal to 0.2. After repairing, the final reduced network changes to  $\{v_1,v_6,v_7,v_8\}$  and  $\{v_2,v_3,v_4,v_5\}$ . By correcting the mis-divided nodes in reduction network, the quality of the final community division is greatly improved.

#### <span id="page-6-0"></span>**3.4 Encoding Method and Genetic Operators**

This paper adopts the locus-based encoding method. In this coding method, each chromosome gene corresponds to a network node, and the gene value represents a neighbor of the node. In other words, there is a connection between nodes  $v_i$  and  $v_j$  if the *i*<sup>th</sup> gene value is *j*, and they are also in the same clustering. Figure [4\(](#page-6-1)a) shows the locus-based encoding method with 7 network nodes, and *Pop*, represhows the locus-based encoding method with 7 network nodes, and  $Pop_1$  represents the corresponding gene value. Different from the label-based coding, the method mentioned above does not need to define the number of communities beforehand, which greatly reduces the search space but also conducives to the genetic operator operation in the evolution.



<span id="page-6-1"></span>**Fig. 4.** Illustration of genetic operators. (a) The complex network used for the crossover operation and its two locus-based encoding individuals. (b) The value of offspring individual New is determined by Rand. If Rand  $= 1$ , the offspring chooses the value of  $Pop_2$ , otherwise  $Pop_1$  will be taken into consideration. (c) The mutation operation adopts the V ector selected randomly in range of  $[0,1]$ , in which the gene value of node is mutated to the index of its adjacent node if the value of corresponding vector is less than 0.1. (d) The newly mutated individual  $New'$  and its structure.

Genetic operators are important for exploration and exploitation of MOEA. In this paper, we use the uniform crossover and neighborhood-based mutation to improve the population diversity. The former is adopted because of its randomness. Specifically, we select parents to crossover and generate the offspring by randomly yielding a set of binary values, which own the same length with population. If the binary value is 0, the first parent is selected, otherwise, the second is selected. At the same time, the local information is taken into neighborbased mutation, which selects the neighbors of parent-population for mutation according to the predefined mutation probability. The illustration of crossover and mutation operators is shown in Fig. [4.](#page-6-1)

#### <span id="page-7-0"></span>**3.5 General Framework of MOEA-MR**

The proposed MOEA-MR utilizes a decomposition-based genetic algorithm framework (MOEA/D) [\[15](#page-11-14)], in which the multi-objective optimization problem is disassembled for a set of single-objective sub-problems according to the Tchebycheff definition in Eq. [\(4\)](#page-7-1).



<span id="page-7-2"></span>**Fig. 5.** The flow chart of the proposed MOEA-MR algorithm with four main steps.

<span id="page-7-1"></span>
$$
\min g^{te}(x \mid w, z^*) = \max_{i=1}^{2} \{w_i \cdot (|F_i(x) - z_i^*)|\} \tag{4}
$$

where  $w = (w_1, w_2)$  with the constraints  $||w||_2 = 1, w_1, w_2 > 0$ .  $z^* = (z_1^*, z_2^*)$  is<br>a reference point and  $z^*$  is the minimal value of the *i*<sup>th</sup> objective function a reference point and  $z_i^*$  is the minimal value of the  $i^{th}$  objective function.<br>Figure 5 shows the overall flow of our proposed MOEA-MR algorithm

Figure [5](#page-7-2) shows the overall flow of our proposed MOEA-MR algorithm with the following four main steps. At the beginning, the consensus information layer of a multilayer network is generated by exploiting the non-negative matrix factorization technology. Next, compressing the consensus information layer into a reduction network with the reduction strategy. At the third step, the population is initialized based on the locus-based encoding method. Meanwhile, the reference point  $z^*$  assigned by using the minimal values of  $KKM$  and RC is prepared for clustering. Finally, for each individual population, applying the uniform crossover and neighborhood-based mutation stagey to generate offspring chromosome. If the Tchebycheff value of offspring chromosome performs better

than any existing chromosomes, the latter will be replaced. In particular, the repairing strategy suggested above is applied to correct misidentified nodes in the reduced network until reaching the maximal iteration of MOEA-MR.

### <span id="page-8-0"></span>**4 Experiment**

#### **4.1 Experimental Design**

**Datasets and Comparison Algorithms.** We test all algorithms on synthetic datasets and real-world datasets, where the former is produced based on the multilayer LFR benchmark (mLFR) [\[3\]](#page-11-2). Compared with other forms, the mLFR benchmark can control the community structure by adjusting the node degrees and the mixing parameters  $(\mu)$  in multilayer networks. The value of  $\mu$  ranges in  $(0, 1)$ , and the larger the  $\mu$  value is, the more complex the community structure is. In addition to synthetic networks (e.g., Sdata1, Sdata2 and Sdata3), the real-world networks we consider include five different types and sizes (i.e., SND  $[16]$ , MPD  $[16]$  $[16]$ , WBN  $[16]$ , CoRA  $[16]$ , CiteSeer  $[16]$ ). The basic description of datasets is listed in Table [1.](#page-8-1) What's more, the performance of MOEA-MR is compared with other classic community detection algorithms, namely, MOEA-MultiNet [\[17\]](#page-11-16) (based on MOEA), GMC [\[19\]](#page-11-17) ( based on multi-view clustering), S2-jNMF [\[3](#page-11-2)], COMCLUS [\[18](#page-11-18)] and CSNMF [\[13\]](#page-11-12) (based on matrix factorization).



<span id="page-8-1"></span>

**Evaluation Metrics and Parameter Settings.** To evaluate the quality of network clustering, this paper adopts two widely-used indicators, namely Normalized Mutual Information (NMI) [\[13](#page-11-12)] and Adjusted Rand Index (ARI) [\[13\]](#page-11-12). Both synthetic network and real-world networks can evaluate the similarity between the ground truth and division detected by algorithms. NMI and ARI use the value range from [0, 1]. If  $NMI(A, B) = 0$ , the division A and B are completely different; if  $NMI(A, B) = 1$ , the division A is the same with B, so ARI does. In the proposed MOEA-MR, the number of iterations is fixed as 200, which is the same as chromosomes, and  $\sigma$  value is set to 30%. It is worth remarking that all algorithms attain the mean value based on 10 dependent runs.

Algorithms			$S2-jNMF$	<b>CSNMF</b>	ComClus	GMC
NMI	0.727	0.437	0.582	0.681	0.555	0.597
ARI	0.772	0.201	0.452	0.493	0.481	0.428
NMI	0.588	0.494	0.516	0.504	0.421	0.451
ARI	0.499	0.385	0.396	0.397	0.365	0.248
NMI	0.457	0.309	0.347	0.436	0.343	0.164
ARI	0.237	0.142	0.191	0.225	0.167	0.048
NMI	0.723	0.317	0.719	0.514	0.471	0.519
ARI	0.768	0.243	0.626	0.491	0.447	0.526
NMI	0.301	$\ast$	0.149	0.237	0.182	0.042
ARI	0.265	$\ast$	0.147	0.207	0.119	0.012
NMI	1	$\ast$	0.924	0.962	0.943	0.954
ARI	1	$\ast$	0.858	0.894	0.722	0.915
NMI	1	$\ast$	0.827	0.980	0.965	0.903
ARI	1	$\ast$	0.716	0.991	0.748	0.999
NMI	1	$\ast$	0.452	0.946	0.894	0.999
ARI	0.954	$\ast$	0.074	0.874	0.814	0.963
			$MOEA-MR$	MOEA-MultiNet		

<span id="page-9-0"></span>**Table 2.** Comparisons of several algorithms on real-world and synthetic networks. Note that ∗ means that the algorithm failed to run on this network.

#### **4.2 Experimental Result**

In this section, the experiment results of all algorithms mentioned on realworld networks and synthetic networks are shown in Table [2.](#page-9-0) The experimental results demonstrate that the averaged NMI and ARI of MOEA-MR on Sdata1 and Sdata2 are 1, which means the real partition can be detected. Compared with other algorithms, the accuracy of MOEA-MR on the real-world network increased by nearly double on average. In addition, the performance of MOEA-MultiNet that uses the genetic algorithm framework is not good. This is because the traditional genetic algorithm is prone to fall into the local optimum and the search speed is very slow. Different from the framework mentioned above, the proposed MOEA-MR treats the local community as a node for calculating, which greatly reduces the search space in the large-scale networks, and the subsequent repairing strategy has further improved the accuracy of network division.

It is also shown that GMC has a better performance than MOEA-MR on Sdata3. That's probably due to the node degrees of mLFR network obey the power-law distribution, which makes it difficult to generate the dense local communities when the structure is fuzzy. Given that, we can conclude that for the most datasets, the proposed MOEA-MR is superior to other algorithms in terms of NMI and ARI, which also means that MOEA-MR is more competitive than other algorithms in detection performance.

#### **4.3 Parameter Analysis**

Figure [6](#page-10-1) provides the experimental results of 10 independent runs based on 4 algorithms with varying similarity thresholds and mixing parameters. For synthetic networks, the larger the mixing parameter is, the fuzzier the network structure is. Figure [6](#page-10-1) presents that MOEA-MultiNet algorithm is the most unstable, whereas the proposed MOEA-MR algorithm is completely opposite. In particular, when  $\mu > 0.4$ , the stability of MOEA-MR begins to decrease. What's more, compared with other options, the similarity threshold  $\sigma$  should also be relatively moderate like the  $\mu$  value. It is obvious that  $\sigma = 0.3$  has the best experimental effect on the algorithm. In short, the results mentioned above prove the influence of parameter settings on the experimental performance and MOEA-MR is superior to other methods on the synthetic dataset.



<span id="page-10-1"></span>**Fig. 6.** The experimental result with different  $\mu$  and  $\sigma$  values on synthetic networks implemented in 4 algorithms.

### <span id="page-10-0"></span>**5 Conclusion**

This paper proposed the MOEA-MR for community clustering in multi-layer complex networks. In MOEA-MR, the non-negative matrix factorization is used to generate the consensus prior information layer of multi-layer network. On this basis, a network reduction method based on the node degree is suggested to compress the size of common network in advance. After that, MOEA-MR adopts an algorithm framework based on MOEA/D, and formulates a repairing strategy based on the network similarity to correct the mis-divided nodes in the local network. The extensive results show the effectiveness of MOEA-MR on the network clustering problem based on the bi-objective optimization. In the future, it is desirable to consider extending the network reduction to cluster detection based on more targets owing to the multi-layer network structure is more complicated in reality.

**Acknowledgements.** This work was supported by the National Key R&D Program of China (2019YFB2102300), National Natural Science Foundation of China (61976181, 11931015), Natural Science Basic Research Plan in Shaanxi Province of China (2022JM-325) and Fundamental Research Funds for the Central Universities (D5000210738).

# **References**

- <span id="page-11-0"></span>1. Gao, C., Su, Z., Liu, J., Kurths, J.: Even central users do not always drive information diffusion. Commun. ACM **62**(2), 61–67 (2019)
- <span id="page-11-1"></span>2. Gao, C., Fan, Y., Jiang, S., Deng, Y., Liu, J., Li, X.: Dynamic robustness analysis of a two-layer rail transit network model. IEEE Trans. Intel. Trans. Sys. (2021). <https://doi.org/10.1109/TITS.2021.3058185>
- <span id="page-11-2"></span>3. Ma, X., Dong, D.: Community detection in multi-layer networks using joint nonnegative matrix factorization. IEEE Trans. Knowl. Data Eng. **31**(2), 273–286 (2019)
- <span id="page-11-3"></span>4. Newman, M.E., Girvan, M.: Finding and evaluating community structure in networks. Phys. Rev. E **69**(2), 026113 (2004)
- <span id="page-11-4"></span>5. Blondel, V., Guillaume, J., Lambiotte, R., Lefebvre, E.: Fast unfolding of communities in large networks. J. Stat. Mech. **8**, P10008 (2008)
- <span id="page-11-5"></span>6. Shi, C., Yan, Z., Wang, Y.: A genetic algorithm for detecting communities in largescale complex networks. Adv. Complex Syst. **13**(1), 3–17 (2010)
- <span id="page-11-6"></span>7. Pizzuti, C.: GA-Net: a genetic algorithm for community detection in social networks. In: The Proceedings of 10th International Conference on PPSN, pp. 1081– 1090 (2008)
- <span id="page-11-7"></span>8. Pizzuti, C.: A multi-objective genetic algorithm for community detection in networks. In: The 2009 IEEE International Conference on Tools Artificial Intelligence, pp. 379–386 (2009)
- <span id="page-11-8"></span>9. Shi, C., Yan, Z., Cai, Y.: Multi-objective community detection in complex networks. Appl. Soft Comput. **12**(2), 850–859 (2012)
- <span id="page-11-9"></span>10. Pizzuti, C.: Multiobjective optimization and local merge for clustering attributed graphs. IEEE Trans. Cyber. **50**(12), 4997–5009 (2020)
- <span id="page-11-10"></span>11. Li, X., Qi, X., Liu, X.: A discrete moth-flame optimization with an  $l_2$ -norm constraint for network clustering. IEEE Trans. Net. Sci. Eng. **9**(3), 1776–1788 (2022)
- <span id="page-11-11"></span>12. Yang, L., Cao, X.: A unified semi-supervised community detection framework using latent space graph regularization. IEEE Trans. Cybern. **45**(11), 2585–2598 (2015)
- <span id="page-11-12"></span>13. Gligorijevic, V., Zafeiriou, S.: Non-negative matrix factorizations for multiplex network analysis. IEEE Trans. Pattern Anal. Mach. Intell. **41**(4), 928–940 (2019)
- <span id="page-11-13"></span>14. Xie, Y., Gong, M., Wang, S., Yu, B.: Community discovery in networks with deep sparse filtering. Pattern Recogn. **81**, 50–59 (2018)
- <span id="page-11-14"></span>15. Zhang, Q., Li, H.: MOEA/D: a multi-objective evolutionary algorithm based on decomposition. IEEE Trans. Evol. Comput. **11**(6), 712–731 (2007)
- <span id="page-11-15"></span>16. Br´odka, P.: A method for group extraction and analysis in multilayer social networks. CoRR abs/1612.02377 (2016)
- <span id="page-11-16"></span>17. Liu, W., Wang, S.: An improved multiobjective evolutionary approach for community detection in multilayer networks. In: The 2017 IEEE Congress on Evolutionary Computation, Donostia, pp. 443–449 (2017)
- <span id="page-11-18"></span>18. Ni, J., Cheng, W., Fan, W., Zhang, X.: ComClus: a self-grouping framework for multi-network clustering. IEEE Trans. Knowl. Data Eng. **30**(3), 435–448 (2018)
- <span id="page-11-17"></span>19. Wang, H., Yang, Y., Liu, B.: GMC: graph-based multi-view clustering. IEEE Trans. Knowl. Data Eng. **32**(6), 1116–1129 (2020)