



OPEN

Motif-based community detection in heterogeneous multilayer networks

Yafang Liu, Aiwèn Li, An Zeng, Jianlin Zhou[✉], Ying Fan[✉] & Zengru Di

Multilayer networks composed of intralayer edges and interlayer edges are an important type of complex networks. Considering the heterogeneity of nodes and edges, it is necessary to design more reasonable and diverse community detection methods for multilayer networks. Existing research on community detection in multilayer networks mainly focuses on multiplexing networks (where the nodes are homogeneous and the edges are heterogeneous), but few studies have focused on heterogeneous multilayer networks where both nodes and edges represent different semantics. In this paper, we studied community detection on heterogeneous multilayer networks and proposed a motif-based detection algorithm. First, the communities and motifs of multilayer networks are defined, especially the interlayer motifs. Then, the modularity of multilayer networks based on these motifs is designed, and the community structure of the multilayer network is detected by maximizing the modularity of multilayer networks. Finally, we verify the effectiveness of the detection algorithm on synthetic networks. In the experiments on synthetic networks, comparing with the classical community detection algorithms (without considering interlayer heterogeneity), the motif-based modularity community detection algorithm can obtain better results under different evaluation indexes, and we found that there exists a certain relationship between motifs and communities. In addition, the proposed algorithm is applied in the empirical network, which shows its practicability in the real world. This study provides a solution for the investigation of heterogeneous information in multilayer networks.

The research of multilayer networks is a current frontier and a hot issue in the field of complex networks, which considers multiple types of nodes and edges relationships (including intralayer and interlayer edges), reflecting the heterogeneity of nodes and edges in networks. In general, a network in which the nodes are homogeneous and the edges are heterogeneous is called a multiplexing network^{1–4}, for example, in social networks, there are two different social relationships between the same users, friendship and work, which can be abstracted into different layers in networks^{5,6}. At present, research on multiplexing networks has covered many aspects, such as robustness^{7,8}, dynamics^{9,10}, community structure^{3,11,12}, disease transmission^{13,14}, etc. Further, a network in which both nodes and edges are heterogeneous is called the heterogeneous multilayer network^{15–17}. For example, in financial systems, the heterogeneity of nodes is reflected in the two different individuals of stocks and users, and the heterogeneity of edges is reflected in the stock relationship, the user relationship, and the relationship between the user and the stock they hold can be constituted as a two-layer network¹⁸. In ecological networks, the heterogeneity of nodes is reflected in the two different individuals of plants and animals, and the heterogeneity of edges is reflected in the relationship between plants, the relationship between animals, and the relationship between plants and animals^{19–21}. Due to the complexity caused by the heterogeneity of nodes and edges, most of the existing studies only focus on the robustness^{22,23} and cascade failure²⁴ of the network, and there are few studies focusing on community detection in heterogeneous multilayer networks.

Community structure is the main macroscopic feature of complex networks²⁵. Most of the current research on the community structure of multilayer networks is based on multiplexing networks, in which the community refers to the structure consisting of homogeneous nodes that are all more tightly connected in different layers¹¹. Community detection on multilayer networks is usually implemented based on three methods: the algorithms based on modularity optimization^{26–28}, network layer aggregation-based algorithms^{29,30}, and dynamics-based algorithms^{31,32}, among which modularity is the most widely used^{4,33}. In 2010, Mucha et al.²⁶ summarized the previous studies to obtain a modularity function for multilayer networks. In 2018, Pamfil et al.²⁷ obtained various

School of Systems Science, Beijing Normal University, Beijing 100875, People's Republic of China. ✉ email: jianlinzhou@bnu.edu.cn; yfan@bnu.edu.cn

types of multilayer networks and demonstrated its effectiveness in synthetic and empirical networks. In 2019, Zhang et al.²⁸ proposed a multilayer edge mixture model, and identifies different communities. Although the existing research on community detection of multiplexing networks has been relatively mature, it does not take into account the heterogeneity of nodes.

In the heterogeneous multilayer network, the community refers to the structure consisting of heterogeneous nodes that are all more tightly connected in different layers. Compared with the community of multiplexing networks, the community of heterogeneous multilayer networks can reflect the relationship between different types of individuals in different environments. Therefore, it is very necessary to study the community detection of the heterogeneous multilayer network. At present, there are some researches about it. Lin et al. proposed MetaGraph decomposition framework to extract communities from networks containing various social backgrounds and interaction relations³⁴. Liu et al. proposed the use of composite modularity for community detection of heterogeneous multi-relational networks, and realized community detection of multi-relational networks³⁵. Pramanik et al. proposed the multilayer modularity index to detect communities consisting of only one type or multiple types of nodes (and edges)³⁶. However, the existing research mainly explores the low-order structure by using the information of nodes and edges in the network alone, without considering the high-order structure (e.g. motif), which contains more meso-scale information about the network^{37–39}.

The motif is a form of a higher-order structure, which refers to a network subgraph with a higher probability of occurrence than in a random network. The existing research shows that the use of motifs for community detection can get a better effect. Reference³⁹ proposes a community detection method using motif weighted tags based on single-layer networks. Reference³⁷ proposed a generalized framework for community detection based on high-order structure, and obtained good results in single layer network. Reference³⁸ extended the motif to multiplexing networks. It integrated the method of network layer aggregation, used multilayer topology information to construct a single-layer network, and the results of community detection are satisfactory. However, the above studies are based on single-layer networks and multiplexing networks, which do not consider the heterogeneous multilayer networks.

In this paper, we consider the research of community detection algorithm based on motif in heterogeneous multilayer networks. First, we redefine motifs of multilayer networks to break through the problem of heterogeneity in the study of multilayer networks and provide a solution for the study of heterogeneity. Then, we modified the modularity of multilayer networks by interlayer motifs and intralayer motifs, and proposed an Motif-Based Community Detection in Heterogeneous Multilayer Networks (CDMMHN) Finally, we performed an experimental evaluation on synthetic networks. To investigate the effectiveness of the proposed algorithm, we conducted an extensive experimental evaluation on networks containing different densities of community structures, and the results show that our algorithm is effective.

The information for each section of this article is as follows. In Section 1, we summarize some related research on multilayer networks. In Section 2, we describe the problem and define the community structure of the heterogeneous multilayer network. In Section 3, we define the motifs of multilayer networks and propose a motif-based modularity that is suitable for heterogeneous multilayer networks, and propose a community detection algorithm by motif-based modularity. In Section 4, we use evaluation indicators to measure the algorithmic results of the community detection for the more general multilayer networks which are synthetic, and the relationship between the motif and community structure in multilayer networks is analyzed, and community detection was performed in empirical networks. In Section 5, we summarize our work and provide an outlook on future research directions.

Related work

Current research for multilayer networks involves a variety of forms of multilayer networks, especially multi-relational networks. One of these classes refers to networks in which different layers of networks represent different interactions between the same individuals (nodes in different layers may be missing and increasing, but the vast majority of nodes in different layers of networks are consistent). Take two-layer social network as an example, the nodes represent users, and the connected edges refer to the network built by two kinds of relationships (e.g., interaction, following, etc.) of these users. There is a one-to-one correspondence between the nodes in different layers of the network, but there are no substantial edges, all such networks are called multiplexed networks.

At present, many researches on community detection are based on multiplex networks, in 2019, Alimadadi et al.⁴⁰ proposed a semi-supervised joint symmetric non-negative matrix decomposition using topological information of the network as well as prior information algorithm for community detection in multilayer networks; in 2022, Venturini et al.⁴¹ investigated multilayer networks with the same set of nodes but without interlayer contiguous edges and proposed a filter-based multi-objective optimization approach for community detection by maximizing the modularity of different layers; in 2022, Ortiz-Bouza et al.⁴² proposed a multiple orthogonal nonnegative matrix TriFactorization method to achieve the detection of cross-layer communities in multilayer networks as well as unique communities on a single layer; in 2023, Roozbahani et al.⁴³ designed a multi-relational directed network based on a semi-supervised approach for overlapping community detection; in 2023, Cai et al.¹² proposed a graph convolution fusion model based on intralayer and interlayer information to achieve community detection for multiplexing networks. These studies are based on the topology information of the network, and some studies are based on the motif. In 2018, Pizzuti et al.⁴⁴ proposed a motif-based community detection method based on multi-objective optimization, the main idea of which is based on the number of motifs; in 2023, Li et al.⁴⁵ proposed a community detection algorithm for multiplexing networks based on motif awareness, which reduces the loss of information during network aggregation and improves the quality of community detection.

The other category refers to networks in which other interrelationships are added to the underlying relationships. In the case of two-layer networks, for example, a two-layer network with interlayer edges is constructed by introducing a third type of interaction between individuals (e.g., friendship) into the two-layer network as an interlayer edge of the network. In 2020, Contisciani et al.⁴⁶ proposed a principled probabilistic approach for community detection in multilayer networks by fusing the attributes of nodes and network structure information; in 2022, Al-sharoa et al.³ proposed a joint non-negative matrix decomposition method for the community detection of multilayer networks by dividing the multilayer network design into a combination of multiplexing and dichotomous networks.

However, in the existing studies, there are few studies focusing on community detection in heterogeneous multilayer networks, and most of the existing studies are directed to the robustness^{22,23} and cascade failure²⁴ of the network, and there are almost no studies on the community division of the network. We consider the design of a method for community detection for multilayer networks that are heterogeneous in terms of both nodes and edges.

Community structure of heterogeneous multilayer networks

In multiplexing networks, because the nodes in different layers represent the same individuals, community detection needs to integrate the information of all layers and obtain a unified community detection result of the multiplexing network. However, for the heterogeneous multilayer network, because the nodes at different layers are not consistent, it is necessary to combine the information of intralayer edges and interlayer edges to make the nodes at different layers get a unified division. As shown in Fig. 1a, taking two-layer networks as an example, layer L_1 and layer L_2 are two networks constructed based on the intralayer edges, a bipartite network can be constructed based on the interlayer edges. For this two-layer network, $G = \{G_{in}, G_{out}\}$, where G_{in} represents the set of intralayer networks and G_{out} represents the set of interlayer networks. $G_{in} = \{G_{in}^{L_1}, G_{in}^{L_2}\}$, $G_{out} = \{G_{out}^{L_1 L_2}\}$, where L_1 and L_2 represent different layers of the multilayer network. Because the nodes of each layer in the network represent a type of individuals, N_{L_1} (N_{L_2}) is used to represent the nodes of layer L_1 (L_2), E_{L_1} (E_{L_2}) is used to represent the intralayer edges of layer L_1 (L_2), and $E_{L_1 L_2}$ is used to represent the interlayer edges between layer L_1 and layer L_2 .

Next, the community structure of heterogeneous multilayer networks is described in detail. In Fig. 1a, different colors represent different communities. To community C_1 , it includes three parts: $C_1^{L_1}$, $C_1^{L_2}$, and the joined edges between the nodes (1, 2, 6, 8, 9, 10 and 17) of the two parts. In the layer of L_1 , the nodes (1, 2, 6) in $C_1^{L_1}$ are closely connected, and these nodes are more sparsely connected to nodes in $C_2^{L_1}$ and $C_3^{L_1}$ in L_1 . In the layer of L_2 , the nodes (8, 9, 10 and 17) in $C_1^{L_2}$ are closely connected, and these nodes are more sparsely connected to nodes in $C_2^{L_2}$ and $C_3^{L_2}$ in L_2 . For interlayer edges, $C_1^{L_1}$ and $C_1^{L_2}$ belong to the same community, the edges between nodes in $C_1^{L_1}$ of layer L_1 and nodes in $C_1^{L_2}$ of layer L_2 are close, the edges between the nodes in $C_1^{L_1}$ and the nodes in $C_2^{L_2}$ or $C_3^{L_2}$ are sparser, the edges between the nodes in $C_1^{L_1}$ and the nodes in $C_2^{L_1}$ or $C_3^{L_1}$ are sparser. The same is true for the other communities (C_2 and C_3).

Most of the existing studies on community detection methods for multilayer networks are based on the edge information in the network, and rarely consider the local structure (e.g., motif) that can be constructed by the edges in the network. As a special kind of multilayer network, there have been some studies proposing community detection for multiplexing networks based on motifs^{38,47}, and better results have been achieved. In heterogeneous multilayer networks, due to the heterogeneity of both its nodes and edges, simply using edges for community detection of the network can not make good use of the heterogeneous characteristics of the network. Motif, as a

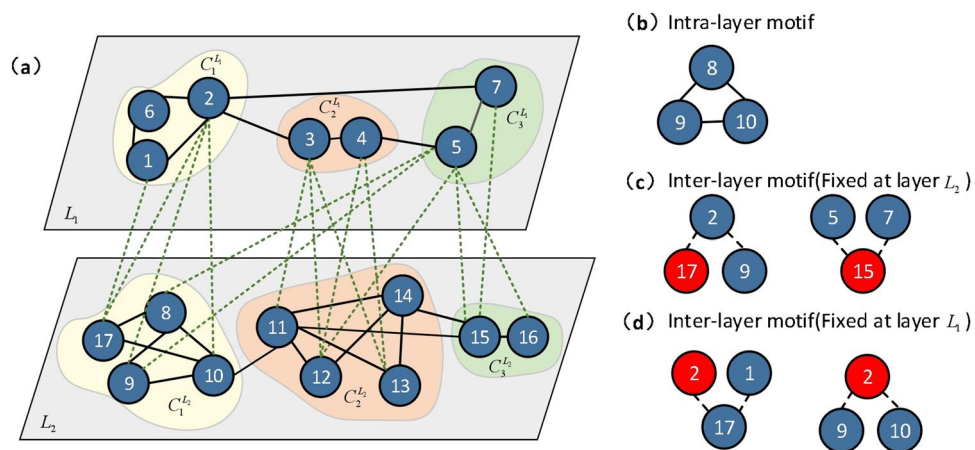


Figure 1. Community structure of heterogeneous multilayer networks. (a) represents the community structure of the multilayer network. Different colors represent different communities, which are divided into three communities (C_1 , C_2 and C_3). (b) shows the structure of the intralayer motif for the network. (c) shows the structure of the interlayer motifs for the network based on layer L_2 . (d) shows the structure of the interlayer motif for the network based on layer L_1 .

local network structure formed by the combination of nodes and edges, is able to interpret the heterogeneity of the network in a better way, and utilize the information of the network in a more comprehensive way. Therefore, in the next study of community detection for heterogeneous networks, we consider to do it based on motifs.

The proposed CDMMHN method

We consider designing an algorithm to break through the heterogeneity of multilayer networks and achieve the detection of community structures in multilayer networks where both nodes and edges are heterogeneous. Because motifs can well describe the heterogeneity of edges in more general multilayer networks, this section modifies a function of motif-based modularity for multilayer networks, and proposes a community detection algorithm suitable for heterogeneous multilayer networks.

Traditional modularity function of multilayer networks

The traditional modularity function for multilayer networks targets multiplexing networks, referring to multilayer networks with consistent nodes and different edge semantics., the modularity²⁷ is calculated as:

$$Q = \underbrace{\sum_{t=1}^T \sum_{i,j \in N_t} \left(A_{ij}^t - \gamma_t \frac{d_i^t d_j^t}{2m_t} \right) \delta(g_i^t, g_j^t)}_{\text{intralayer modularity}} + \omega \underbrace{\sum_{t=2}^T \sum_{i \in N_t} \delta(g_{\pi_i^{t-1}}^{t-1}, g_i^t)}_{\text{interlayer modularity}}, \quad (1)$$

where, A_{ij}^t represents the adjacency matrix of the network, t represents the layer of the network, d_i^t represents the degree of node i in the network on layer t , m_t represents the number of edges in the current layer (layer t) of the network, $\delta(g_i^t, g_j^t)$ determines whether node i in layer t and node j in layer t belong to the same community. If node i and node j are in the same community, $\delta(g_i^t, g_j^t) = 1$, and if node i and node j are in the different community, $\delta(g_i^t, g_j^t) = 0$. π_i^t refers to the parent of node i , which is the node that has a relationship with node i at another layer, and in a multiplexing network, it refers to node i at another layer. $\delta(g_{\pi_i^{t-1}}^{t-1}, g_i^t)$ determines whether node i in layer t and its parent node (node i in layer $t - 1$) belong to the same community. If they are in the same community, $\delta(g_{\pi_i^{t-1}}^{t-1}, g_i^t) = 1$, and if they are not in the same community, $\delta(g_{\pi_i^{t-1}}^{t-1}, g_i^t) = 0$.

Motif-based modularity function of multilayer networks

Because of the heterogeneity of edges in multilayer networks, the motif is considered to study the network structure in multilayer networks. For a heterogeneous multilayer network, the edges of the same layer are homogenous, so the motif structure in the layer is consistent with the single-layer network. However, due to the heterogeneity of the two nodes that make up the connecting edges between layers, the traditional method compresses the two-layer network and puts all edges on the single-layer network, which cannot reflect the heterogeneity of nodes in different layers and the heterogeneity of edges in the network. Therefore, we redefine the motif structure of multilayer networks with heterogeneous nodes and edges.

Triangles are a more classical higher-order structure that is often used for community detection research^{45,48,49}. In 2016 Benson et al.³⁷ proposed a variety of triangular subgraphs with orientations (only two structures, closed triangles and open triangles, are applicable to undirected networks) and experimentally demonstrated that triangular subgraphs are very important for social networks. In the study, we consider the three-node motif structure.

Motifs refers to subgraphs that have a much higher probability of occurring in a real network than in a random network, and for the three-node subgraph, the number of closed triangles occurring in the network is much higher than in a random network. Therefore, for the intra-layer motifs, we chose the structure shown in Fig. 1b. For the interlayer modal of the heterogeneous multilayer network, since the composition of closed triangles requires intra-layer edges, the interlayer motif chooses an open triangle structure, which means that both edges composing the motif are heterogeneous edges, which is able to show the heterogeneity of the edges. If the structure of Fig. 1b is chosen as the interlayer modifiers, the heterogeneity of the interlayer edges cannot be represented. The structure of the intralayer motif is shown in Fig. 1b, and the interlayer motif of the network comes in two forms, as shown in Fig. 1c,d. Because the nodes of the multilayer network are heterogeneous, nodes of different layers refer to different individuals, so the layer where the fixed node of the motif is selected has an impact on the structure of the motif. The network layer to which the fixed node belongs is different, and the motif structure is inconsistent. We take the red nodes in the figure as fixed nodes, in Fig. 1c structure 17-2-9 (or in Fig. 1d structure 2-17-1) is defined as a regular triplet. For the structure 17-2-9 in Fig. 1c, the fixed-node 17 and node 9 belong to layer L_2 and they have the same properties. The fixed-node 17 and node 2 belong to different layers and have different properties. Therefore, the connection structure that includes a fixed node, a node that is homogeneous to the fixed node, and a node that is heterogeneous to the fixed node is called the regular triplet (the structure 2-17-1 in Fig. 1d also applies). Besides, the structure consisting of nodes 5-15-7 in Fig. 1c (or structure 9-2-10 in Fig. 1d) is defined as an inverted triplet, in which the connection structure that includes a fixed node, two nodes that are heterogeneous to the fixed node.

In addition, for different fixed nodes, the number of nodes that can form regular triples and have heterogeneity with them is different. For example, in Fig. 1a, node 12 as a fixed node can form regular triplets with node 3 and node 4. Among them, there are two regular triples with node 3 (12-3-11, 12-3-13), and one regular triplet with node 4 (12-4-13). However, for any fixed node, the other two nodes which can form an inverted triple with the fixed node belong to different layers. For any node that has a different property from the fixed node, the number of inverted triples that can be formed with the fixed node is the same. For example, in Fig. 1a, node 12 as a fixed node can form inverted triples with node 3, node 4 and node 5. For any of them, the number of inverted triples that can be formed with a fixed node is 3 (3-12-4, 3-12-5, 4-12-5). Compared with the regular triple, the

function of nodes is not discriminative, so the inverted triple is not considered in the experiment. This paper studies the community structure of multilayer networks based on the above motif structure.

In more general multilayer networks, different from the multiplexing network, the network individuals in different layers are different, there are multiple nodes in neighboring layers connected to the same node. Therefore, considering the motif can reflect the heterogeneity of the network, a calculation method of motif-based modularity for more general multilayer networks is designed. The modularity obtained based on the motif can be calculated as:

$$Q_{motif} = \frac{1}{2\mu_1} \left[\underbrace{\sum_{t=1}^T \sum_{i,j \in N_t} \left(W_{ij}^t - \frac{w_i^t w_j^t}{2w_t} \right) \delta(g_i^t, g_j^t)}_{intra\text{-layer modularity}} \right] + \frac{1}{2\mu_2} \left[\underbrace{\sum_{t=1}^T \sum_{i \in N_t} w_{i\pi_i^t} \delta(g_{\pi_i^t}^{t+1}, g_i^t)}_{inter\text{-layer modularity}} \right], \tag{2}$$

where $\mu_1 = \sum_{t=1}^T \sum_{i,j \in N_t} W_{ij}^t$, $\mu_2 = \sum_{t=1}^T \sum_{i \in N_t} w_{i\pi_i^t}$, W_{ij}^t represents the adjacency matrix based on the number of the intralayer motif of t layer network, t represents the layer of the network, w_t is the motif number of intralayer that can be formed in layer t , w_i^t represents the motif number of intralayer that node i can form in layer t . $\delta(g_i^t, g_j^t)$ determines whether node i in layer t and node j in layer t belong to the same community. If so, $\delta(g_i^t, g_j^t) = 1$, and if not, $\delta(g_i^t, g_j^t) = 0$. π_i^t refers to the set of nodes in other layers that can form the motif with nodes i in t layer. $\delta(g_{\pi_i^t}^{t+1}, g_i^t)$ determines whether node i in layer t and its parent node (the set of nodes in layer $t + 1$ that can form the interlayer motif with node i) belong to the same community. $w_{i\pi_i^t}$ refers to the number of the interlayer motif that can be formed by node i and its parent node. For example, as shown in Fig. 1a, for node 9 in layer L_2 , nodes 2 and 5 can form regular triplets with it. For node 5, only one regular triplet can be formed. For node 2, there are two regular triplets can be formed. Set node 9 as fixed node i , for the interlayer edge(2-9), $w_{i\pi_i^t} = 2$, for the interlayer edge(5-9), $w_{i\pi_i^t} = 1$.

Community detection algorithm by motif-based modularity for heterogeneous multilayer networks

Next, based on the modularity function of the motifs, we design a community detection algorithm for heterogeneous multilayer networks. For a multilayer network with t layers, the network structure is represented by a

matrix $G = \begin{bmatrix} A_{L_1} & C_{L_1 L_2} & \cdots & C_{L_1 L_t} \\ C_{L_1 L_2}^T & A_{L_2} & \cdots & C_{L_2 L_t} \\ \vdots & \vdots & \ddots & \vdots \\ C_{L_1 L_t}^T & C_{L_2 L_t}^T & \cdots & A_{L_t} \end{bmatrix}$. In the process of research, we use the two-layer network for research,

and the matrix of the two-layer network can be expressed as $G = \begin{bmatrix} A_{L_1} & C_{L_1 L_2} \\ C_{L_1 L_2}^T & B_{L_2} \end{bmatrix}$, where A_{L_1} and B_{L_2} represent the information matrix of the intralayer structure of the network in the different layer. The number of nodes in the network at layer L_1 is n_1 , the number of nodes in the network at layer L_2 is n_2 . And there is no self-loop between nodes. The matrix $C_{L_1 L_2}$ is used to represent interlayer edges of a multilayer network. Take $C_{L_1 L_2}$ for example,

$$C_{L_1 L_2} = \begin{bmatrix} c_{a_1 b_1} & c_{a_1 b_2} & \cdots & c_{a_1 b_{n_2}} \\ c_{a_2 b_1} & c_{a_2 b_2} & \cdots & c_{a_2 b_{n_2}} \\ \vdots & \vdots & \ddots & \vdots \\ c_{a_{n_1} b_1} & c_{a_{n_1} b_2} & \cdots & c_{a_{n_1} b_{n_2}} \end{bmatrix}$$

, where $c_{a_1 b_1}$ represents whether the node a_1 at layer L_1 and the node b_1 at layer L_2 have edge. If there is, $c_{a_1 b_1} = 1$, otherwise $c_{a_1 b_1} = 0$. On the basis of the edge matrix, the intralayer and interlayer motif in the network are calculated respectively, so as to obtain the network motif matrix W (the intralayer and interlayer motif with three

nodes). $W = \begin{bmatrix} W_{motif}^{L_1} & W_{motif}^{L_1 L_2} \\ W_{motif}^{L_2 L_1} & W_{motif}^{L_2} \end{bmatrix}$, where $W_{motif}^{L_1}$ and $W_{motif}^{L_2}$ both represent the intralayer motif matrix. Take $W_{motif}^{L_1}$ for example,

$$W_{motif}^{L_1} = \begin{bmatrix} 0 & w_{a_1 a_2} & \cdots & w_{a_{n_1} a_{n_1}} \\ w_{a_2 a_1} & 0 & \cdots & w_{a_2 a_{n_1}} \\ \vdots & \vdots & \ddots & \vdots \\ w_{a_{n_1} a_1} & w_{a_{n_1} a_2} & \cdots & 0 \end{bmatrix}$$

represents the intralayer motif matrix obtained from the network at layer L_1 , where, $w_{a_1 a_2}$ represents the number of intralayer motif obtained from the node a_1 and the node a_2 at layer L_1 . For the interlayer network, take $W_{motif}^{L_1 L_2}$ for example,

$$W_{motif}^{L_1 L_2} = \begin{bmatrix} w_{a_1 b_1}^{L_1} & w_{a_1 b_2}^{L_1} & \dots & w_{a_1 b_{n_2}}^{L_1} \\ w_{a_2 b_1}^{L_1} & w_{a_2 b_2}^{L_1} & \dots & w_{a_2 b_{n_2}}^{L_1} \\ \vdots & \vdots & \ddots & \vdots \\ w_{a_{n_1} b_1}^{L_1} & w_{a_{n_1} b_2}^{L_1} & \dots & w_{a_{n_1} b_{n_2}}^{L_1} \end{bmatrix}$$

refers to the number matrix of interlayer motif obtained from L_1 -based network, where $w_{a_1 b_1}^{L_1}$ refers to node a_1 in the network of layer L_1 as the fixed-node, the number of interlayer motifs that can be formed by an edge of the node a_1 in the network at layer L_1 and the node b_1 in the network at layer L_2 . $W_{motif}^{L_2 L_1}$ refers to the number matrix of interlayer motif obtained from L_2 -based network.

We use motif-based modularity to detect communities in heterogeneous multilayer networks, the implementation process of the algorithm is shown in algorithm 1. Firstly, the motif matrix of the network W is obtained according to the structure matrix G . In the design, in order to improve the efficiency of the algorithm, we use the community detection algorithm for a single-layer network to divide the community of the network with fewer nodes (default $n_1 < n_2$), the community ids obtained is denoted as g^{L_1} , g^{L_2} can be obtained according to the relationship between the number of motifs among layers, but considering that there are few edges between layers, it is not possible to obtain the belonging community of all nodes in the L_2 layer. So we consider using the community detection algorithm for a single-layer network to get the initial community ids of nodes in the network of layer L_2 , integrating all community ids in the initial community $g^t = (g_1^t, g^{L_2})$. Using the original community to calculate modularity, and then transforming the community ids of nodes, updated community id and modularity (Step6 in algorithm 1). In the process of Step7 in algorithm 1, we need to reach the final result of the conditions if the modularity reaches its maximum value or the number of communities in different layers is the same. The final output Q_{new} and g_{new} meet the conditions. In the process, it should be noted that the community of nodes at different layers should be unified to ensure the correctness of the community results and the correlation between nodes at different layers.

Input: $G, N = \{n_1, n_2\}$.

Output: Community structure g_{end} for the multilayer network.

- 1 The motif number matrix W was constructed by counting the number of intralayer and interlayer motif.
 - 2 The initial community ids g^{L_1} of the network of layer L_1 with a small number of nodes is obtained.
 - 3 According to the interlayer motif, the initial community ids g^{L_2} of nodes of layer L_2 is obtained.
 - 4 The community ids g^t of each node obtained above is used as the initial community for community detection.
 - 5 Use the community ids g^t obtained above for community division, according to Eq.2.
 - 6 Adjust the community of the node, repeat Step5, and iteratively compare the modularity Q_{new} and the community ids g_{new} obtained.
 - 7 Get the best detection result g_{end} by loop iteration.
-

Algorithm 1. Community Detection by Motif-Based Modularity for Heterogeneous Two-layer Networks.

In the processing, there are two aspects need to be noticed. First, in the iteration, the calculation of the modularity of heterogeneous multilayer networks contains two parts, intralayer modularity and interlayer modularity. The community detection of nodes in different layers can be realized by calculating the intralayer modularity, the calculation of interlayer modularity is to unify the community of nodes of different layers, to realize the integration of multilayer information. Take a two-layer network as an example, the intralayer modularity includes the modularity calculation of the layer L_1 and the modularity calculation of the layer L_2 . For the two-layer network, interlayer modularity includes modularity calculation that takes layer L_1 as the base network, changes the community number of nodes in layer L_2 , and takes layer L_2 as the base network, change the community number of nodes in layer L_1 . Second, because there are few edges between layers in multilayer networks, the community ids of two nodes in different layers can be obtained in the following three ways: 1) When two nodes can form the interlayer motif, these two nodes are likely to the same community. 2) When two nodes cannot form the interlayer motif, but there is an edge between them, these two nodes are likely to the same community. 3) When two nodes cannot form the interlayer motif and they are not connected, the community id of the node should be obtained according to the community ids of other nodes in the layer to which the node belongs that can form the motif with it.

Experiment

In this section, the benchmark model is constructed for more general multilayer networks with community and we verify the algorithm on the multilayer networks. Besides, the evaluation index NMI and R_Inter are used to measure the accuracy of the community detection results.

Synthetic networks

Based on the relationship between the community structure and the degree of density between nodes, this paper constructs a synthetic heterogeneous two-layer network model with a community structure. The edge density

inside and outside the community within the layer as well as the effect of the specificity of fewer inter-edges on the multilayer network structure are also considered. The generation algorithm for the artificial two-layer network is shown in algorithm 2, and the network generation model is controlled by the following parameters: $c, n_1, n_2, z, z_1, p_{in}, p_{layer}$, where c is the number of communities in the two-layer network, n_1 and n_2 indicates the number of nodes at different layers in the two-layer network, z represents the average degree of nodes in the whole network, z_1 represents the average interlayer degree of nodes in the network, p_{in} represents the probability that two nodes belong to the same community, $1 - p_{in}$ represents the probability that two nodes are in different communities, and p_{layer} represents the probability that two nodes belong to the same layer.

Input: $c, n_1, n_2, z, z_1, p_{in}, p_{layer}$
Output: Network edge connection matrix G

- 1 Build an all-zero matrix G of size $((n_1 + n_2) * (n_1 + n_2))$ and loop through each element of the upper triangle of the matrix to operate on behavior i , column j .
- 2 **if** Node i and node j belong to the same community **then**
- 3 Random generation of a probability p
- 4 **if** $p < (p_{in} * z / (n_1 + n_2 / c - 1))$ **then**
- 5 **if** node i and node j not belong to the same layer **then**
- 6 Random generation of a probability p
- 7 **if** the number of nodes on the network to which node i belongs is n_1 **then**
- 8 **if** $p < ((1 - p_{layer}) * z_1 / (n_2 / c))$ **then**
- 9 Add Edge (i, j) , $G[i, j] = 1$
- 10 **end**
- 11 **end**
- 12 **else** the number of nodes on the network to which node i belongs is n_2
- 13 **if** $p < ((1 - p_{layer}) * z_1 / n_2 - (n_2 / c))$ **then**
- 14 Add Edge (i, j) , $G[i, j] = 1$
- 15 **end**
- 16 **end**
- 17 **end**
- 18 **end**
- 19 **end**
- 20 **else**
- 21 Random generation of a probability p
- 22 **if** $p < ((1 - p_{in}) * z / ((n_1 + n_2) - (n_1 + n_2) / c))$ **then**
- 23 **if** node i and node j not belong to the same layer **then**
- 24 Random generation of a probability p **if** $p < ((1 - p_{layer}) * z_1 / (n_1 / c))$ **then**
- 25 Add Edge (i, j) , $G[i, j] = 1$
- 26 **end**
- 27 **end**
- 28 **end**
- 29 **end**

Algorithm 2. Building a benchmark model of heterogeneous multilayer networks with community structure.

During the experiment, we generated random benchmark networks with the number of communities of 3, 4, 5 and 6, the parameter settings are as follows: $c = 3, n_1 = 60, n_2 = 90, p_{layer} = 0.8; c = 4, n_1 = 80, n_2 = 100, p_{layer} = 0.8; c = 5, n_1 = 75, n_2 = 105, p_{layer} = 0.8$ and $c = 6, n_1 = 90, n_2 = 120, p_{layer} = 0.8$. The community to which each node belongs is fixed, and the layer to which each node belongs is determined. Once the parameter c, n_1, n_2, p_{layer} is determined, the structure of the two-layer network only depends on the p_{in}, z, z_1 .

The average degree in the network includes two parts: intra-community degree (z_{in}) and inter-community degree (z_{out}), where, $z_{in} = p_{in} * ((n_1 + n_2) / c - 1) / 2, z_{out} = (1 - p_{in}) * ((n_1 + n_2) / c * 2) / 2$, the average degree z of the network is the sum of z_{in} and z_{out} . The degree of interlayer includes two parts: the degree of interlayer within the community and the degree between the communities of interlayer. Among them, the degree of intra-layer within the community is $z_{layer}^{in} = p_{layer} * ((n_1 + n_2) / c) / 2$, and the degree of interlayer within the same community is $z_{in} - z_{layer}^{in}$; the degree of interlayer between the communities is $z_{out} * (1 - p_{layer})$; the average degree of the interlayer network (z_1) is the sum of the degree of interlayer within the community and the degree between the communities of interlayer. During the experiment, we also considered the different numbers of communities in different layers, and the parameters were set as $c_1 = 3, c_2 = 4, n_1 = 60, n_2 = 80, p_{layer} = 0.8$, where c_1 and c_2 represent the number of communities in different layers.

The advantage of the constructed network in this way is that we can know the community id of each node in the network, which can be compared with the partition result obtained by the algorithm, so as to

determine the effectiveness of the algorithm and the correctness of the partition result. According to the model of network construction, Fig. 2a is one of the network structure diagrams constructed when the parameters were $c_1 = 3, c_2 = 3, n_1 = 60, n_2 = 90, p_{layer} = 0.8, p_{in} = 0.9$, Fig. 2b is one of the network structure diagrams constructed when the parameters were $c_1 = 3, c_2 = 4, n_1 = 60, n_2 = 80, p_{layer} = 0.8, p_{in} = 0.9$.

To access the code and supplementary data used in this research, please contact the corresponding author.

Algorithm performance

To evaluate the accuracy of the community detection results on the synthetic networks, *NMI* (Normalized Mutual Information) was used to measure the results. This metric is one of the commonly used measures of partition similarity⁵⁰, and it can measure the accuracy of community detection as a whole. *NMI* is calculated as follows:

$$NMI = \frac{2I(X; Y)}{H(X) + H(Y)}, \quad (3)$$

where, X refers to the real category. Y refers to the category obtained by clustering algorithm, and $H(\cdot)$ refers to the cross entropy, which is calculated as follows:

$$H(X) = - \sum_{i=1}^{|X|} P(i) \log P(i), \quad (4)$$

$I(X; Y)$ is the mutual information, $I(X; Y) = H(X) - H(X|Y)$. *NMI* is obtained by normalizing mutual information, so the range of *NMI* is $[0, 1]$. The more similar the two vectors are, the closer the similarity is to 1. If *NMI* = 1, the result of community detection obtained by the algorithm is exactly the same as the real community IDs.

In addition, to measure the accuracy of community correspondence between individuals of different layers, we designed an evaluation index *R_Inter*. By comparing the community ids of individuals in different layers, an interlayer correspondence matrix C is obtained to determine whether two nodes belong to the same community. If two nodes belong to the same community, $C_{ij} = 1$, otherwise $C_{ij} = 0$. *R_Inter* is calculated as follows:

$$R_Inter = 1 - \frac{\sum_{i=1}^{n_1} \sum_{j=1}^{n_2} |C_{real}^{ij} - C_{alg}^{ij}|}{n_1 * n_2}, \quad (5)$$

where C_{real}^{ij} represents the real interlayer correspondence matrix of the network, C_{alg}^{ij} represents the correspondence matrix obtained by the detection algorithm, n_1 represents the number of nodes at Layer L_1 , n_2 represents the number of nodes at Layer L_2 , and *R_Inter* represents the ratio of correct detection in the interlayer relationship. $|C_{real}^{ij} - C_{alg}^{ij}|$ is used to determine whether node i belonging to layer L_1 and node j belonging to layer L_2 belong to the same community is correctly identified. When $|C_{real}^{ij} - C_{alg}^{ij}| = 1$, it indicated detecting errors. Therefore, the value range of *R_Inter* is $[0, 1]$. When the interlayer community correspondence is completely correct, *R_Inter* = 1, when the interlayer community correspondence is completely incorrect, *R_Inter* = 0.

To verify the effectiveness of the community detection algorithm in multilayer networks based on motifs, we conducted experiments on synthetic networks. At the same time, to confirm the role of heterogeneity in community detection, the two-layer network was compressed into a single-layer network, in other words, all nodes (nodes in different layers) and edges (intralayer edges and interlayer edges) were input into a network. The intralayer motif in multilayer networks is regarded as the motif of a single-layer network, which is then divided based on the motif of a single-layer network. The community detection algorithm that considers heterogeneity is compared with the algorithm that does not consider heterogeneity. To avoid the specificity of the results, we

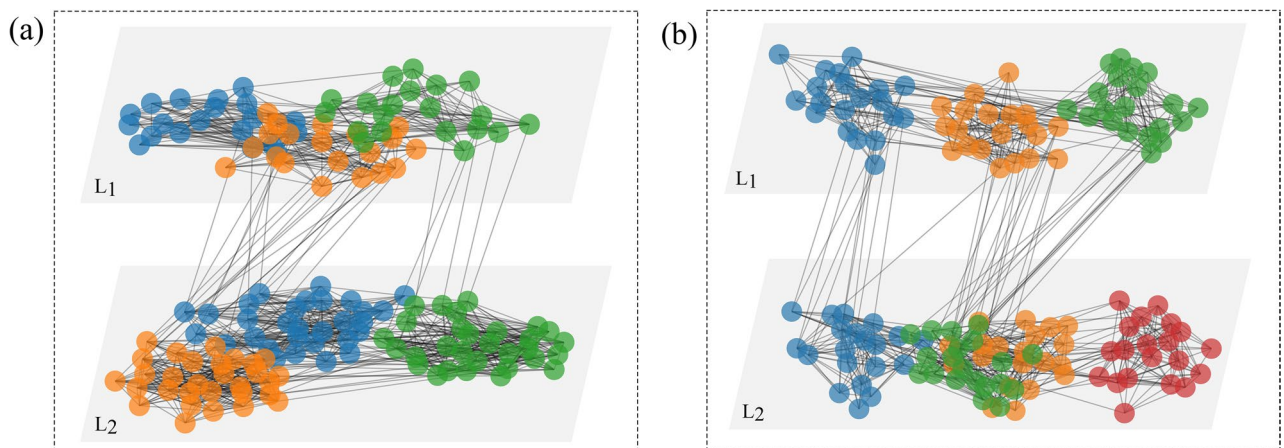


Figure 2. Structure diagram of a two-layer network with a known community structure. (a) shows a two-layer network with the same numbers of communities at different layers, and (b) shows a two-layer network with different numbers of communities at different layers.

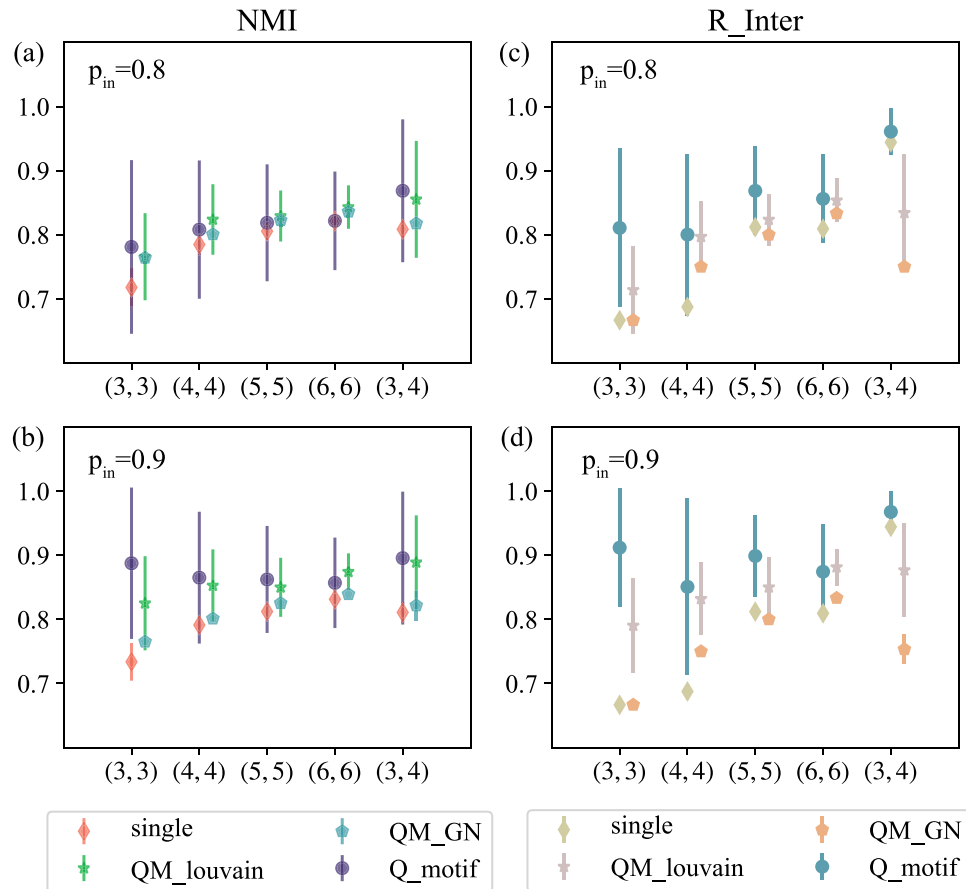


Figure 3. Performance of different detection algorithms in identifying community members. (a) and (c) depict the NMI and R_Inter values obtained based on the following parameters: $p_{in} = 0.8$ and $p_{layer} = 0.8$. (b) and (d) present the NMI and R_Inter values obtained based on the following parameters: $p_{in} = 0.9$ and $p_{layer} = 0.8$. The first number in the abscissa is c_1 (the number of communities at layer L_1), and the second number is c_2 (the number of communities at layer L_2).

randomly generated 100 synthetic networks for the same parameters, and the results are shown in Fig. 3. The figure shows the mean and standard deviation of the accuracy of community detection for multiple experiments. Q_motif denotes the result obtained based on algorithm 1, and *single* denotes the result of partitioning the two-layer network by compressing it into a single-layer network without considering heterogeneity. QM_GN ³⁶ and $QM_louvain$ ³⁶ denote the results obtained by other algorithms. Figure 3a,b provide a comparison of the overall network partition obtained with different probabilities of edges within the community, and Fig. 3c,d compare the community correspondence between nodes of different layers of the network under different edge connectivity probabilities within the community. As seen from the figure, the more obvious the community structure is, the more accurate the detected community structure is, and the nodes of different layers can be effectively unified when considering heterogeneity.

In Fig. 3, we evaluated the accuracy of the community detection results for members of a heterogeneous multilayer network, but did not account for the accuracy of the number of communities detected by the network as a whole. Therefore, in Fig. 4, the number of communities obtained by community detection algorithms with the real number of communities are compared. The graph shows the number of communities for different probabilities of edges within a community and for different numbers of communities. In the figure, *num_real* represents the number of real communities in the network, Q_motif represents the number of communities detected based on Algorithm 1, *num_single* represents the number of communities obtained when the network is compressed to a single layer, and QM_GN and $QM_louvain$ represent the number of communities obtained by other algorithms. As seen from the figure, with the increase in the probability of edges within a community, that is, the more obvious the community structure is, the closer the number of detected communities is to the real number of communities, and the number of communities obtained based on the algorithm is closer to the real number of communities than that obtained without considering the heterogeneity.

Relationship between the intralayer motif and community

In order to prove the validity of community detection using intra-motifs, we identify whether each fixed node in the layer belongs to the same community as the nodes that can form the maximum number of intralayer motifs

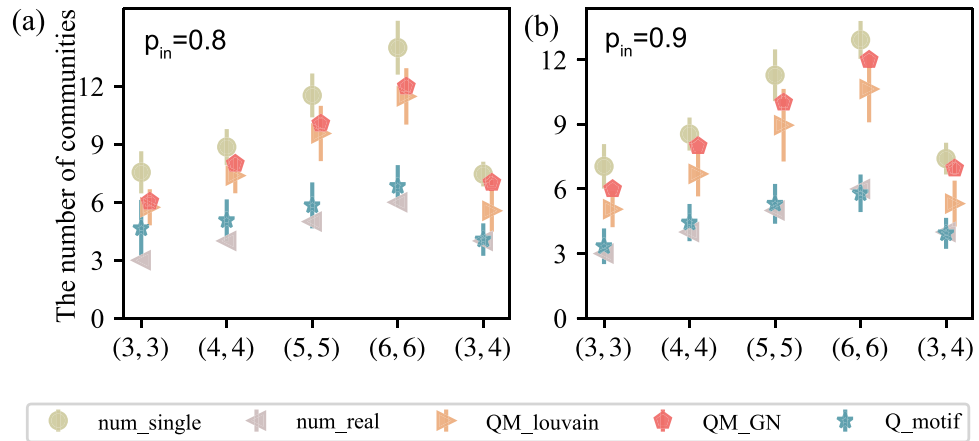


Figure 4. Performance of the detection algorithm in terms of identifying the number of communities. The purple triangle represents the real number of communities in the network, the yellow circle represents the number of communities detected in the single-layer network, and the blue star represents the number of communities obtained based on Algorithm 1. The red and orange colors indicate the results obtained by other algorithms.

with it and calculate the probability that the fixed node belongs to the same community as these nodes that can form the intralayer motif with it. We randomly selected one of the networks constructed based on the following parameters: $p_{in} = 0.9$, $c_1 = 3$, $c_2 = 3$, $n_1 = 60$, $n_2 = 90$, and $p_{layer} = 0.8$. The results are shown in Fig. 5a,b. The brown line shows the proportion of fixed nodes that belong to the same community as other nodes in the same network that can form more than one intralayer motif, and the red line indicates the proportion of fixed nodes that belong to the same community as the other nodes in the same network that can form the largest number intralayer motifs. As seen from the figure, in this network, the fixed node must be in the same community as the nodes that can form the intralayer motif with it, and the node that has the largest number of intra-motifs must be in the same community.

In addition, we calculated the number of motifs that can be formed by nodes that belong to the same community as the fixed node and the number that are not in the same community. Taking Fig. 1a as an example, for

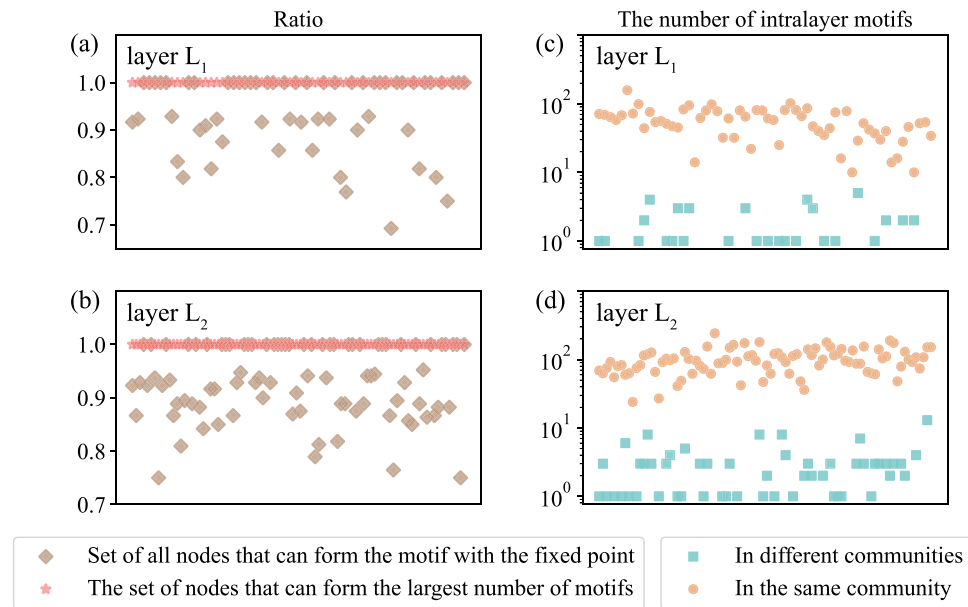


Figure 5. Statistical analysis of intralayer motifs and experimental results of an intralayer motif for a benchmark network with three communities. The horizontal coordinate represents each node in the corresponding network. (a) and (b) show the ratio in which motifs can be formed in layers L_1 and L_2 , respectively, and (c) and (d) show the number of motifs that can be formed in the same community in the same layer L_1 and L_2 , respectively, compared with the number of motifs that can be formed between the communities in the same layer.

layer L_2 , the nodes that can compose the motif structure with the fixed node 11 are: 12, 14, 13, and 15. Node 11 and nodes 12, 13, and 14 are in the same community, and node 15 is not in the same community. The number of the intralayer motifs that nodes 12, 13, 14, and 15 can form is 1, 1, 3, and 1, respectively. Thus, the number of motifs that belong to the same community is 5, and the number of motifs that do not belong to the same community is 1. The statistical results are shown in Fig. 5c,d. The abscissa represents nodes, and the ordinate represents the number of motifs. The orange line is the number of motifs that can be formed within a community, and the blue line is the number of motifs that can be formed between communities. The number of motifs belonging to the same community is far greater than the number of motifs formed between nodes that are not in the same community.

Relationship between the interlayer motif and community

In order to prove the validity of community detection using inter-motifs, we calculated the formation of regular triplets. We used layer L_1 and layer L_2 as the base layer, assumed that each node in the layer was a fixed node, and then calculated the probability that the fixed node belongs to the same community as the nodes that can form the motif with it, and the probability that the nodes that have interlayer edges with the fixed node belonging to the same community. As shown in Fig. 1a, with layer L_2 as the base layer and node 9 as the fixed node, there are two nodes (2 and 5) belonging to layer L_1 and connected to fixed node 9, one of which (node 2) is in the same community as node 9. Therefore, among the nodes with which node 9 can form an interlayer edge, the probability that node 9 belongs to the same community is 0.5. Nodes 2 and 5 form the motif with node 9. Nodes 2 and 9 can form two interlayer motifs, and node 5 can form one interlayer motif. Therefore, node 9 tends to be in the same community as node 2.

In a heterogeneous multilayer network, there are few edges between layers. Therefore, we analyze the statistics of the edges between layers in the whole network, and the results are shown in Table 1. The results in the table are based on one of the networks with $c_1 = c_2 = 3$, $n_1 = 60$, $n_2 = 90$, $p_{layer} = 0.8$, and $p_{in} = 0.9$. It can be seen that fixed nodes and nodes that can form interlayer modules with them are essentially in the same cluster, and most of the nodes with which they have interlayer edges belong to the same community. Table 2 shows the mean values of the metric *ratio* in Table 1, which obtained based on multiple networks, and it can be seen that this phenomenon exists in all networks.

Applications in empirical networks

Finally, the algorithm is applied to the empirical network. For an example of scientist-research topic two-layer networks, we looked at the network surrounding HQL, an academician with the Chinese Academy of Sciences (CAS). He is mainly engaged in condensed matter theory and related computational physics research, and his main research interests include the following aspects: strongly correlated systems, quantum entanglement and quantum phase transitions, and numerical methods for many-body systems. Based on the journal data (1893–2010) provided by the American Physical Society, we collected the (1984) papers published by Haiqing Lin's collaborators as well as the PACS codes of each paper, where each PACS code corresponds to a specific research topic in physics.

Based on the above data, we constructed a collaborative network of scientists with Haiqing Lin's collaborators as nodes at the scientist level. We also constructed a co-occurrence network between the PACS codes of the papers published by these scientists at the research topic level. In the scientist collaboration network we constructed, there are 844 nodes and 146817 edges, the nodes in the network represent Haiqing Lin's collaborators, and the

Network	Interlayer motif			Interlayer edges		
	Motif ¹	Same community ²	ratio	Edge ³	Same community ⁴	Ratio ⁵
L_1	3	3	1.0	22	20	0.9
L_2	6	6	1.0	22	20	0.9

Table 1. Statistical analysis of interlayer edges and motif based on a network. ¹ It refers to the number of objects that can be made up. ² The number of nodes that make up the motif that belong to the same community as the consent community. ³ It refers to the number of edges that exist between layers. ⁴ It refers to the number of nodes in the same community of two nodes with edges between layers. ⁵ ratio=same community/edge.

Network	p_{in}	(3,3) ¹	(4,4) ²	(5,5) ³	(6,6) ⁴	(3,4) ⁵
L_1	0.9	0.957/0.961	0.966/0.967	0.975/0.976	0.975/0.984	0.960/0.972
L_1	0.8	0.893/0.895	0.887/0.912	0.923/0.945	0.912/0.961	0.898/0.931
L_2	0.9	0.958/1.000	0.968/0.995	0.976/1.000	0.981/1.000	0.964/1.000
L_2	0.8	0.890/0.995	0.905/0.959	0.938/0.986	0.954/1.000	0.916/0.996

Table 2. Statistical analysis of interlayer edges and motif of different structural networks. ¹ $c_1 = c_2 = 3$. ² $c_1 = c_2 = 4$. ³ $c_1 = c_2 = 5$. ⁴ $c_1 = c_2 = 6$. ⁵ $c_1 = 3, c_2 = 4$.

edges represent the two scientists who have coauthored at least one *APS* journal paper. In the research topic co-occurrence network, which contains 78 nodes and 241 edges, the nodes in the network represent the *PACS* codes included in the *APS* journal papers published by the above scientists, and the edges represent that the two *PACS* codes have co-occurred at least once in the *APS* journal papers published by these scientists. At the same time, based on the information that a scientist has published a paper in a particular *PACS* code, we built an edge between the scientist layer and the research topic layer. Finally, we obtained a scientist-research topic two-layer network with heterogeneous nodes and edges. Based on algorithm 1, we performed community detection on this heterogeneous scientist collaboration network, and the whole network was divided into 4 groups, which allowed us to determine in what research areas scientists mainly collaborate with each other.

If community detection for single-layer networks is performed on the scientist cooperation network and the research topic co-occurrence network, we can determine which scientists cooperate more closely with each other and which research fields are more closely connected, but we cannot integrate the scientist groups and research fields. To fuse information about scientists and research fields on top of the above two single-layer networks, it is also necessary to add information about which papers scientists collaborated on and in which papers research fields co-occur, i.e., to construct a bipartite network between scientists and papers, and a bipartite network between research fields and papers. Our approach is able to directly relate scientists and research fields together to achieve a more rational division of scientists and their research topics.

Conclusion and future work

In this paper, the community detection algorithm of more general multilayer networks in which both nodes and edges contain heterogeneity are proposed. Among, the motif structure are used to calculate the modularity of multilayer networks, which can distinguish the intralayer and interlayer edges of networks well. We used two metrics (*NMI*, *R_Index*) to measure the accuracy of community detection in terms of the overall distribution of communities and the integration of communities between different layers. The results of proposed algorithm in synthetic networks showed the more accurate community structure (the number of communities and the community to which members belong), comparing with other community detection algorithms.

We applied algorithm 1 on an empirical network, which enables the problem of detecting communities uniformly between different types of individuals. Finally, this paper analyzed the relationship between the intralayer motifs (interlayer motifs) and communities and further explains the feasibility of the algorithm.

This study provides an understanding of the community structure of more general multilayer networks, and our algorithm mainly addresses the problem of community detection in undirected multilayer networks. In the following research, we hope to conduct more in-depth research in more multi-level empirical networks. Moreover, there are many unequal relationships in real networks, such as directed networks. In the future, we will conduct an in-depth study of community detection in multilayer-directed networks.

Data availability

The APS data can be downloaded at <https://journals.aps.org/datasets>.

Received: 10 January 2024; Accepted: 8 April 2024

Published online: 16 April 2024

References

1. Wang, Z., Guo, Q., Sun, S. & Xia, C. The impact of awareness diffusion on sir-like epidemics in multiplex networks. *Appl. Math. Comput.* **349**, 134–147 (2019).
2. Xia, C. *et al.* A new coupled disease-awareness spreading model with mass media on multiplex networks. *Inf. Sci.* **471**, 185–200 (2019).
3. Al-Sharoa, E. M. & Aviyente, S. Community detection in fully-connected multi-layer networks through joint nonnegative matrix factorization. *IEEE Access* **10**, 43022–43043 (2022).
4. Bazzi, M. *et al.* Community detection in temporal multilayer networks, with an application to correlation networks. *Multiscale Model. Simulat.* **14**, 1–41 (2016).
5. Battiston, F., Nicosia, V. & Latora, V. Structural measures for multiplex networks. *Phys. Rev. E* **89**, 032804 (2014).
6. Granell, C., Gómez, S. & Arenas, A. Dynamical interplay between awareness and epidemic spreading in multiplex networks. *Phys. Rev. Lett.* **111**, 128701 (2013).
7. Shekhtman, L. M. & Havlin, S. Percolation of hierarchical networks and networks of networks. *Phys. Rev. E* **98**, 052305 (2018).
8. Turalska, M., Burghardt, K., Rohden, M., Swami, A. & D'Souza, R. M. Cascading failures in scale-free interdependent networks. *Phys. Rev. E* **99**, 032308 (2019).
9. Cencetti, G. & Battiston, F. Diffusive behavior of multiplex networks. *New J. Phys.* **21**, 035006 (2019).
10. Yang, Y., Tu, L., Guo, T. & Chen, J. Spectral properties of supra-laplacian for partially interdependent networks. *Appl. Math. Comput.* **365**, 124740 (2020).
11. Zhai, X. *et al.* Null model and community structure in multiplex networks. *Sci. Rep.* **8**, 3245 (2018).
12. Xiang, C. & Bang, W. A graph convolutional fusion model for community detection in multiplex networks. *Data Min. Knowl. Disc.* **37**, 1518–1547 (2023).
13. Faryad, *et al.* Generalized epidemic mean-field model for spreading processes over multilayer complex networks. *IEEE/ACM Trans. Netw. (TON)* **21**, 1609–1620 (2013).
14. Darabi Sahneh, F. & Scoglio, C. Competitive epidemic spreading over arbitrary multilayer networks. *Phys. Rev. E* **89**, 062817 (2014).
15. Poccock, M. J. O., Evans, D. M. & Memmott, J. The robustness and restoration of a network of ecological networks. *Science* **335**, 973–977 (2012).
16. Li, G. Y., Cheng, B. S., Zhang, P. & Li, D. Q. Review of the interdependent networks. *Dianzi Keji Daxue Xuebao/J. Univ. Electronic ence Technol. China* **42**, 23–28 (2013).
17. Pio-Lopez, L., Valdeolivas, A., Tichit, L., Remy, E. & Baudot, A. Multiplex-heterogeneous network embedding for drug repositioning. *Sci. Rep.* **11**, 8794 (2021).
18. Boccaletti, S. *et al.* The structure and dynamics of multilayer networks. *Phys. Rep.* **544**, 1–122 (2014). The structure and dynamics of multilayer networks.

19. Smith-Aguilar, S., Aureli, F., Busia, L., Schaffner, C. & Ramos-Fernández, G. Using multiplex networks to capture the multidimensional nature of social structure. *Primates* **60**, 277–295 (2019).
20. Gilarranz, L. J., Sabatino, M., Aizen, M. A. & Bascompte, J. Hot spots of mutualistic networks. *J. Anim. Ecol.* **84**, 407–413 (2015).
21. Pilosof, S., Porter, M. A., Pascual, M. & Kéfi, S. The multilayer nature of ecological networks. *Nat. Ecol. Evol.* (2017).
22. Zheng, K., Liu, Y., Gong, J. & Wang, W. Robustness of circularly interdependent networks. *Chaos, Solitons & Fractals* **157**, 111934 (2022).
23. Chen, C.-Y., Zhao, Y., Qin, H., Meng, X. & Gao, J. Robustness of interdependent scale-free networks based on link addition strategies. *Physica A* **604**, 127851 (2022).
24. Lin, I.-C., Yağan, O. & Joe-Wong, C. Dynamic coupling strategy for interdependent network systems against cascading failures. *IEEE Trans. Network Sci. Eng.* **10**, 2265–2282 (2023).
25. Jusup, M. *et al.* Social physics. *Phys. Rep.-Rev. Sec. Phys. Lett.* **948**, 1–148 (2022).
26. Mucha, P. J., Richardson, T., Macon, K., Porter, M. A. & Onnela, J.-P. Community structure in time-dependent, multiscale, and multiplex networks. *Science* **328**, 876–878 (2010).
27. Pamfil, A. R., Howison, S. D., Lambiotte, R. & Porter, M. A. Relating modularity maximization and stochastic block models in multilayer networks. *SIAM J. Math. Data Sci.* **1**, 667–698 (2019).
28. Zhang, P. S. & Huangyu, Jian. Community detection using multilayer edge mixture model. *Knowl. Inf. Syst.* **60**, 757–779 (2019).
29. Tang, L., Wang, X. & Liu, H. Community detection via heterogeneous interaction analysis. *Data Min. Knowl. Disc.* **25**, 1–33 (2012).
30. Ghasemian, A., Zhang, P., Clauset, A., Moore, C. & Peel, L. Detectability thresholds and optimal algorithms for community structure in dynamic networks. *Phys. Rev. X* **6**, 031005 (2016).
31. Kuncheva, Z. & Montana, G. Community detection in multiplex networks using locally adaptive random walks. In *2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 1308–1315 (2015).
32. De Domenico, M., Lancichinetti, A., Arenas, A. & Rosvall, M. Identifying modular flows on multilayer networks reveals highly overlapping organization in interconnected systems. *Phys. Rev. X* **5**, 011027 (2015).
33. Amelio, A., Mangioni, G. & Tagarelli, A. Modularity in multilayer networks using redundancy-based resolution and projection-based inter-layer coupling. *IEEE Trans. Network Sci. Eng.* **7**, 1198–1214 (2020).
34. Lin, Y. *et al.* Metafac: community discovery via relational hypergraph factorization. In *Knowledge Discovery and Data Mining*, 527–535 (2009).
35. Liu, X., Liu, W., Murata, T. & Wakita, K. A framework for community detection in heterogeneous multi-relational networks. *Adv. Complex Syst.* **17**, 1450018 (2014).
36. Pramanik, S., Tackx, R., Navelkar, A., Guillaume, J.-L. & Mitra, B. Discovering community structure in multilayer networks. In *2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, 611–620 (2017).
37. Benson, A. R., Gleich, D. F. & Leskovec, J. Higher-order organization of complex networks. *Science* **353**, 163–166 (2016).
38. Huang, L., Wang, C.-D. & Chao, H.-Y. Hm-modularity: A harmonic motif modularity approach for multi-layer network community detection. *IEEE Trans. Knowl. Data Eng.* **33**, 2520–2533 (2021).
39. Huang, L., Wang, C.-D. & Yu, P. S. Higher order connection enhanced community detection in adversarial multiview networks. *IEEE Trans. Cybern.* **53**, 3060–3074 (2023).
40. Alimadadi, A., Ehsan Bagheri. Community detection in facebook activity networks and presenting a new multilayer label propagation algorithm for community detection. *International Journal of Modern Physics, B. Condensed Matter Physics, Statistical Physics, Applied Physics* **33** (2019).
41. Venturini, S., Cristofari, A., Rinaldi, F. & Tudisco, F. A variance-aware multiobjective Louvain-like method for community detection in multiplex networks. *J. Complex Netw.* **10**, cnac048 (2022).
42. Ortiz-Bouza, M. & Aviyente, S. Community detection in multiplex networks based on orthogonal nonnegative matrix tri-factorization (2023).
43. Roozbahani, Z., Rezaeenour, J. & Katanforoush, A. Community detection in multi-relational directional networks. *J. Comput. Sci.* **67**, 101962 (2023).
44. Pizzuti, C. & Socievole, A. Motif-based community detection in multiplex networks. In Cherifi, C., Cherifi, H., Karsai, M. & Musolesi, M. (eds.) *Complex Networks & Their Applications VI*, 190–201 (Springer International Publishing, Cham, 2018).
45. Li, C. *et al.* Multiplex network community detection algorithm based on motif awareness. *Knowl.-Based Syst.* **260**, 110136 (2023).
46. Xiang, C. & Bang, W. Community detection with node attributes in multilayer networks. *Sci. Rep.* **10**, 15736 (2020).
47. Li, P.-Z., Huang, L., Wang, C.-D., Lai, J.-H. & Huang, D. Community detection by motif-aware label propagation. *ACM Trans. Knowl. Discov. Data* **14** (2020).
48. Prat-Pérez, A., Dominguez-Sal, D., Brunat, J.-M. & Larriba-Pey, J.-L. Put three and three together: Triangle-driven community detection. *ACM Trans. Knowl. Discov. Data* **10**, <https://doi.org/10.1145/2775108> (2016).
49. Wu, X., Zhang, H., Quan, Y., Miao, Q. & Sun, P. G. Graph embedding based on motif-aware feature propagation for community detection. *Physica A* **630**, 129205. <https://doi.org/10.1016/j.physa.2023.129205> (2023).
50. Fred, A. L. N. & Jain, A. K. Robust data clustering. In *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, 128–133 (2003).

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant No. 72371031 and 71731002) and the MOE (Ministry of Education in China) Liberal Arts and Social Sciences Foundation (Grant No. 20YJC870015), the Fundamental Research Funds for the Central Universities (Grant No. 2021NTST33).

Author contributions

Yafang Liu: Conceptualization, Methodology, Software, Formal analysis, Writing – original draft. **Aiwen Li:** Writing – original draft, Writing – review & editing. **An Zeng:** Conceptualization, Supervision. **Jianlin Zhou:** Supervision, Conceptualization, Methodology, Formal analysis, Writing–review & editing. **Ying Fan:** Supervision, Conceptualization, Methodology, Formal analysis, Writing–review & editing. **Zengru Di:** Conceptualization, Supervision.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to J.Z. or Y.F.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024