THE EUROPEAN
PHYSICAL JOURNAL B

Regular Article - Computational Methods

Vital node identification based on cycle structure in a multiplex network

Quan Ye[1] , Guanghui Yan[1,a], Wenwen Chang[1], and Hao Luo[2]

[1] School of Electronics and Information Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China
[2] School of Information Science and Engineering, Gansu University of Traditional Chinese Medicine, Lanzhou 730070, China

Abstract.
Multiplex networks frame the heterogeneous nature of real systems, where the multiple roles of nodes, both functionally and structurally, are well represented. We identify these vital nodes in a multiplex network so that we can control a pandemic outbreak like COVID-19, eliminate damage from a network attack, maintain traffic, and so on. Vital node identification has attracted scientists in various fields for decades. In this paper, we propose a hybrid supra-cycle number and hybrid supra-cycle ratio based on the cycle structure, and present an extensive experimental analysis by comparing our indexes and several different indexes in four real multiplex networks on layer nodes and multiplex nodes. The experimental results show that these proposed indexes have good robustness, synchronization, and transmission dynamics. Finally, we provide an in-depth understanding of multiplex networks and cycle structure, and we sincerely hope more valuable academic achievements are proposed in the future.

1 Introduction

The understanding of complex networks has greatly improved over recent decades, and researchers have made tremendous achievements in both theory and application. However, these remarkable efforts do not mean that we have a sufficient understanding of complex networks, because the framework of a complex network is partly systematic and partly chaotic [1], and becomes increasingly intricate with the acceleration of internet development. Recently, the focus of network science has been shifting from discovering macroscopic statistic regularity [2–4] to decomposing microscopic structural organization [5,6] and further revealing the explicit roles played by such microscopic elements as nodes [7] and edges [8].

Indeed, nodes play different roles not only in structures but also in the function of complex systems, and understanding how nodes affect each other is critical for further research. It is important to identify vital nodes functionally and topologically [9], which helps us to control epidemic outbreaks [10,11] and rumors [12], reduce and eliminate the damage from cyber attacks [13–16], prevent catastrophic outages in power grids [17] and the internet, maintain traffic and telecommunication networks [18], target receptor proteins and key genes for molecular biological drugs [19], and predict success-

ful persons or groups based on collaborative networks [20,21].

However, a perfect method for identifying vital nodes has not yet been found, as it is never a trivial task, and the directions in identification methods and even various criteria have been inconsistent. In the first place, there are two main branches for identifying vital nodes [22], one based on the neighborhood and the other based on the path. In reality, different methods fit different situations, which means that for vital node identification, a universal method is theoretically and practically unavailable. Moreover, the criteria in this task are diverse for the same reason. In some scenarios where we need to protect populations from an epidemic, a susceptible-infected-susceptible (SIS) model, susceptible-infected-recovered (SIR) model [23], or susceptible-infected-recovered-infected) model [24] perform well. And in other scenarios where we have to maintain a network and minimize loss when it is not possible to implement protection for all nodes, node removal is more suitable. Last but not least, sometimes identifying a group of coupled vital nodes is more useful than identifying a series of single nodes; topologically, it means the vital element could be a motif or a community in the identification process. However, identification of vital motifs or vital communities will be explored in future research; this paper will focus on nodes.

As mentioned above, methods for vital node identification can be divided into two branches, methods based on neighborhood and methods based on path, and these

ᵃ e-mail: yanghacademic@163.com (corresponding author)

methods can also be classified according to the topological structure: chain, star, or cycle [25]. Researchers have pieced together an increasingly clear picture of the function and interaction of disparate structures in a variety of dynamical processes, including the roles of different motifs in biological networks, the transmission of information and behaviors, and the self-sustaining process of star structures in an epidemic. Aside from chain and star structures, cycles are a ubiquitous part of networks and play an unalterable role in storage, synchronizability, and controllability of network function. A cycle is a measurement of similitude between a target network and tree network, which means the model network cannot reproduce cycles as well as the real network [26]. The methods proposed in this paper, a supra-cycle number and supra-cycle ratio, are based on a cycle structure, where the former index describes the importance of the node in a global scale, while the latter index describes the same in a local scale. A cycle ratio is designed like a clustering coefficient and edge multiplicity [7]; the three methods focus on the local proportion of a certain structure in a network.

Methods of identifying vital nodes were originally designed for single-layer networks, but in reality, individuals have multiple relationships, and it is biased to describe a complex system as a single-layer network. A multiplex network is a pragmatic framework in which multiple relationships of individuals are represented in different layers (same relations are represented in a single layer). There are same nodes in each layer, and inter-layer connection only exists between the same nodes represented in a different layer. The framework of a multiplex network is shown in Fig. 1. A multiplex network is a special multilayer network. On the one hand, it retains some characteristics of a single-layer network, in that every layer in a multiplex network can be considered a single-layer network because there are same nodes in each layer, and inter-layer connections are special; on the other hand, inter-layer connections do exist, which is the hallmark of a multilayer network.

Considering the special nature of a multiplex network, the symbolically basic element we focus on can be different. Zhao et al. [27,28] divide the nodes of a multiplex network into two categories: a multiplex node (MN, a collection of the same node in different layers) and a layer node (LN, the replica of a MN in each layer). Researchers have made much progress based on the MN. Halu et al. [29] define Multiplex PageRank by drawing on biased random walks. Solé-Ribalta et al. [30] redefine the betweenness centrality and proposed an algorithm to compute it in an efficient way. Solá et al. [31] extend the concept of eigenvector centrality to a multiplex network. But by contrast, the LN has received less attention. Zhao et al. [32] extend several indexes or algorithms from single-layer networks to multiplex networks and find explosive immunization is always best for the identification of vital LNs by performing them on different kinds of multiplex model networks.

In this paper, we focus on cycle number and cycle ratio [33], which were newly proposed indexes for vital node identification, and extend them to multi-

plex networks. We compare the evaluation results to other indexes through robustness, synchronization, and immunization, and we make a further conclusion from two viewpoints (MN and LN), which reveals the applicability of our method.

This paper is organized as follow. Initially, we mathematically define the problem we are focused upon and introduce several indexes and our indexes in the Methods section. Then we present the main idea of our algorithm and the time complexity in the experiments section, and we analyze the experimental results. Finally, we make a conclusion based on a theoretical perspective and the experimental results in the Conclusion section.

2 Problem definition

A multiplex network is a combination of several single-layer networks which contain the same nodes (or replicas of the same nodes) and a different intra-layer connection. A multiplex network is a framework to perfectly describe the multiple relationships between individuals in real-world complex systems including social networks, transportation networks, biological networks, etc.

Figure 1 gives three illustrations of a multiplex framework. We take Fig. 1a, b as an example. Users are connected on different social apps. Every user is an MN when the user is embedded in the framework as a node. Users interact on TikTok via blue links and on Facebook via green links (as shown in the left panel). When different users are connected with a kind of link in the same layer and there are same the individuals in different layers, the individuals are LNs in each layer (as shown in the right panel). As a result, MNs represent individuals, and LNs represent accounts of social apps, and there is same explanation for Fig. 1c–f. Different cities are connected by aviation via a blue link in Fig. 1c and by railway via a green link in Fig. 1d, where MNs represent cities, and LNs represent stations; different gene fragments are physically connected via a blue link in Fig. 1e and functionally connected via a green link in Fig. 1f, where MNs represent gene fragments, and LNs represent the specific role the gene fragments play.

Per the definitions of MN and LN we show above, the difference between MNs and LNs is whether the basic element is a coupled node or a single node. Mathematically, multiplex network G with N nodes and M layers can be described as follows:

$$D_\mu = \overbrace{\{G_\alpha, G_\beta, \cdots\}}^{M} \tag{1}$$

Research based on MNs help us understand the individuals that have multiple interactions in a complex system, and research based on LNs helps us focus on every single role played by individuals, and benefits us when focusing on a multilayer network with different nodes in each layer. Thus, the task of vital node identification in multiplex networks can be defined as follows.
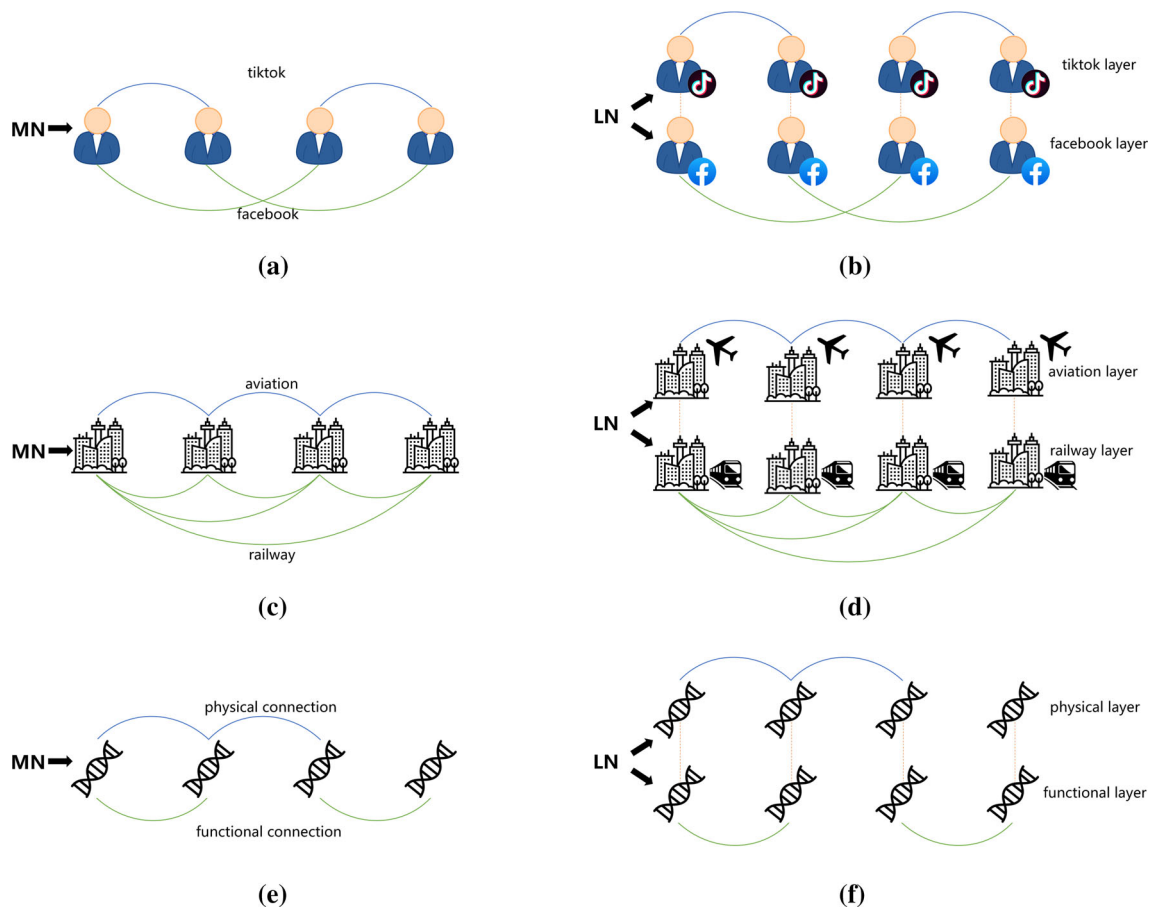
Fig. 1 Illustrations of MNs and LNs in a multiplex network. **a**, **b** Different relationships between four users on TikTok and Facebook. **c**, **d** Aviation and railway relationships between four cities. **e**, **f** Physical and functional relationships between four gene fragments

Problem definition: Propose an index or algorithm to rank MNs/LNs, such that we can find a set of MNs/LNs to better protect or minimize the damage in specific tasks.

3 Methods

Many methods of identifying vital nodes are designed for single-layer networks, and researchers put effort into reframing them for multiplex networks. In this section, we introduce several methods and compare them with our methods through percolation, synchronization, and immunization.

3.1 Methods based on a star structure

3.1.1 Degree centrality

Degree centrality is the easiest way to achieve identification, and it has been well performed in all sorts of networks. When it comes to a multiplex network with M layers and N nodes in each layer, the degree cen-

trality of an LN v_α in layer α, denoted as DCL(i^α), is shown as follows:

$$\mathrm{DC}(i^\alpha) = \sum_{i^\alpha, j^\alpha \in V^\alpha} a_{i^\alpha j^\alpha} \tag{2}$$

where $a_{i^\alpha j^\alpha}$ is the element of the adjacency matrix for layer α in row i column j, and $a_{i^\alpha j^\alpha}$ equals 1 if v_i^α and v_j^α are connected and 0 otherwise; V^α is the set of nodes in layer α.

Similarly, the degree centrality of an MN v_i, denoted as DCM(i), is shown as follows:

$$\mathrm{DC}(i^\alpha) = \sum_{m \in L}^{M} \sum_{i^m, j^m \in V^m} a_{i^m j^m} \tag{3}$$

where L is the set of layers. Mathematically, the degree centrality of an MN couples the degree centrality of LNs.

3.1.2 K-core decomposition

K-core decomposition is a process to obtain important nodes, which decomposes graph step by step. According to degree centrality, nodes with higher degree are more important than the others, which means these vital nodes are in the inner layer. The steps are as follows: in the first stage of removal, nodes with degree one are removed at first, then if other nodes with degree one appear because of the removal, the network is repeatedly processed until there is no node with degree one. Thus, all the removed nodes are considered to have the same importance in this method, and these removed nodes form the 1-core. The next stage of removal is conducted for the nodes with degree two. Repeat the removal stage by stage until the network is completely decomposed.

The k-core decomposition is applied in each single layer of the multiplex network [34], which means for LNs, the coreness (k-core value) of nodes is still a number, but for MNs, the k-core value of nodes is a label consisting of several number; for example, a node with k-core value (a_α, b_β) means the coreness is a in layer α, and the coreness is b in layer β). For the sake of obtaining a united standard, we sum up the coreness of the LN as the coreness of the MN.

3.1.3 ClusterRank

Unlike the degree and k-core decomposition, for the ClusterRank [35], researchers not only consider the number of nearest neighbors but also take into account the interaction among them. ClusterRank was originally designed for a directed network, but it can also be applied to an undirected network. In a multiplex network with M layers and N nodes in each layer, the ClusterRank value of LN v_i^α in layer α can be redefined as follows:

$$\mathrm{DR}(i^\alpha) = f(c_{i\alpha}) \sum_{j^\alpha \in \Gamma_{i_\alpha}} (k_{\mathrm{out}} + 1) \qquad (4)$$

where $f(c_{i\alpha})$ is a function of the clustering coefficient $c_{i\alpha}$ of v_i^α, k_out is the out degree of v_i^α, and Γ_{i_α} is the set of nearest neighbors of v_i^α.

Similarly, the ClusterRank value of the MN v_i can be redefined as follows:

$$\mathrm{DR}(i) = \sum_{m \in L}^{M} f(c_{i^m}) \sum_{j^m \in \Gamma_{i_m}} (k_{\mathrm{out}} + 1) \qquad (5)$$

where L is the set of layers in the multiplex network.

3.1.4 Semi-local centrality

Semi-local centrality [36] makes a good balance between the neighbor and path. These vital methods of node identification based on neighbors have the advantage of simplicity, but the topological structure is never taken into consideration. These methods based on the path effectively process the topological structure, but they are not suitable for large-scale networks because of great complexity. Essentially, semi-local centrality takes into account the direct neighbors of nodes, including first-order and second-order neighbors, and the first-order and second-order neighbors of these direct neighbors. In semi-local centrality, researchers partly introduce the characteristics of the path and keep the simplicity at the same time. In a multiplex network with M layers and N nodes in each layer, the semi-local centrality of LN v_i^α in layer α can be redefined as follows:

$$\mathrm{SCL}(i^\alpha) = \sum_{j^\alpha \in \Gamma(i^\alpha)} \sum_{k^\alpha \in \Gamma(j^\alpha)} N(k^\alpha) \qquad (6)$$

where $\Gamma(i^\alpha)$ is the set of first-order neighbors of LN v_i^α, $\Gamma(j^\alpha)$ is the set of first-order neighbors of LN v_j^α, and $N(k^\alpha)$ is the amount of neighbors which v_i^α can reach in a 2-hop.

When it comes to an MN, we take the inter-layer structure into consideration, and the semi-local centrality of the MN is redefined as follows:

$$\mathrm{SCL}(i^\alpha) = \sum_{m,n \in L}^{M} \sum_{j^m \in \Gamma(i^m)} \sum_{k^\alpha \in \Gamma(j^m) \cup \Gamma(i^m)} N(k) \quad (7)$$

Topologically, this definition couples the intra-layer and the inter-layer neighbors which LNs can reach in a 4-hop.

3.2 Methods based on a chain structure

3.2.1 Betweenness centrality

Betweenness centrality in a single-layer network is defined as the proportion of shortest paths that pass through a given node. Nodes with high betweenness centrality usually act as hub, and thus a bridge which connects subgraphs is constructed when several such nodes are connected (a bridge may also consist of one hub node), and the network is probably cut into pieces when such nodes are removed. In a multiplex network with M layers and N nodes in each layer, because paths can be inter-layer or intra-layer, the betweenness centrality of LN v_i^α is redefined as follows:

$$\mathrm{BC}(i^\alpha) = \sum_{s^m \in V^m, t^n \in V^n, i^\alpha \in V^\alpha} \frac{\sigma(s^m, t^n \mid i^\alpha)}{\sigma(s^m, t^n)} \quad (8)$$

where V^m, V^n, V^α is the set of nodes in layer m, n, α, $\sigma(s^m, t^n)$ is the number of shortest paths from LN v_s^m in layer m to LN v_t^n in layer n, and $\sigma(s^m, t^n \| i^\alpha)$ is the number of shortest paths that pass through the given LN v_i^α in $\sigma(s^m, t^n)$. Specifically, if $v_i^\alpha \in \{v_s^m, v_t^n\}$, $\sigma(s^m, t^n \| i^\alpha) = 0$.

Similarly, the betweenness centrality of MN v_i is redefined as follows:

$$\text{BC}(i) = \sum_{m,n,r \in L} \sum_{s^m \in V^m, t^n \in V^n, i^r \in V^r} \frac{\sigma(s^m, t^n \mid i^r)}{\sigma(s^m, t^n)} \tag{9}$$

where L is the set of layers in the multiplex network.

3.2.2 Closeness centrality

Closeness centrality eliminates the disturbance by summarizing the distance from the target node to all other nodes while evaluating the importance of nodes. When it comes to a multiplex network with M layers and N nodes in each layer, the closeness centrality of LN v_i^α is redefined as follows:

$$\text{CC}(i^\alpha) = \frac{N-1}{\sum_{i^\alpha \in V^\alpha, j^m \in V^m} d(j^m, i^\alpha)} \tag{10}$$

where v^m, v^α is the set of nodes in layer m, α, and $d(j^m, i^\alpha)$ is the distance from LN j^m in layer m to LN i^α in layer α.

Similarly, the closeness centrality of MN v_i is redefined as follows:

$$\text{CC}(i) = \sum_{n,m \in L}^{M} \frac{N-1}{\sum_{i^n \in V^n, j^m \in V^m} d(j^m, i^n)} \tag{11}$$

where L is the set of layers in the multiplex network.

3.3 Methods based on a cycle structure

3.3.1 Supra-cycle number

A cycle is a path which contains three or more nodes; the nodes are connected end to end, and the size of cycle is the number of nodes it contains. Once a cycle does not contain a smaller cycle inside, it is the basic cycle, and the smallest basic cycle is the smallest one in the basic cycle of a given node. For a node v_i, the length of the smallest basic cycle is the girth of v_i when the cycle contains the node. As shown in Fig. 2, we can obtain the incidence matrix by enumerating all the smallest basic cycles in the network, and the cycle number matrix from the incidence matrix (algebraically, the cycle number matrix equals the incidence matrix multiplied by the inverse of the incidence matrix, and topologically, the cycle number matrix is the distribution of the smallest basic cycle where the element in row i, column j equals the number of smallest basic cycle containing node v_i and node v_j).

When the network shifts from a single-layer network to a multiplex network, the cycle is not only intra-layer but also inter-layer; thus, we set a weight for the inter-layer cycle because network dynamics and behaviors change when inter-layer interactions happen. Also, the

situation changes when the basic elements shift from LNs to MNs. We set the quotient between the maximum in the spectrum of inter-layer networks and the maximum in the spectrum of intra-layer networks as the weight of inter-layer edges [37], and this process is shown in Fig. 3.

3.3.2 Supra-cycle ratio

The cycle ratio of node v_i in a single-layer network with n nodes is defined as follows based on the cycle number:

$$\text{CR}(i) = \begin{cases} 0, & c_{ij} = 0 \\ \sum_{j=1}^{n} \frac{c_{ij}}{c_{jj}}, & \text{otherwise} \end{cases} \tag{12}$$

where c_{ij} is the element of the cycle number matrix in row i column j. The process of calculation is shown in Fig. 2.

When the ratio index is based on the supra-cycle number rather than the cycle number, the index becomes a supra-cycle ratio, and the process is shown in Fig. 3.

When the cycle ratio is applied in a multiplex network, the results of the supra-cycle ratio change with supra-cycle number, but the process from cycle number to cycle ratio, which is shown in Fig. 3, does not change (Table 1).

If we treat the inter-layer edge the same as the intra-layer edge, the process of cycle number and cycle ratio calculation in the single-layer network can be performed without changing the supra-cycle number. The unweighted supra-cycle number and supra-cycle ratio are not performed well enough because the inter-layer edge usually plays different roles in the multiplex network, and the difference between the inter-layer edge and intra-layer edge does not influence the index. We reset the importance of the inter-layer cycle that is adaptive for a multiplex network structure (Table 2).

We designed an algorithm to calculate the supra-cycle number and supra-cycle ratio. Firstly, we cut off all nodes with coreness one because a coreness greater than one is a necessary condition for a cycle structure. Secondly, we enumerate all the smallest basic cycles. Each search for the smallest basic cycle can be regarded as the shortest path problem, and the complexity of the shortest path problem is $o(N + E)$ in a network with N nodes and M edges. Finally, we get the supra-cycle number and supra-cycle ratio. The multiplex network has E edges, M layers, and N nodes in each layer. The complexity of our algorithm is $o(N(N + E))$, and in a real network, the complexity is $o(N'(N' + E'))$, where $N' \ll N$, $E' \ll E$. A real network is usually sparse, and an amount of nodes are removed from the original network initially; thus, the network we actually process is N' nodes and E' edges, which is much smaller than the original network).

Cycle Number Matrix

$$CR(1) = \frac{1}{1} + \frac{1}{1} + \frac{1}{2} + \frac{1}{1} + 0 + 0 + 0 + \frac{1}{3} + 0 + 0 = \frac{23}{6}$$
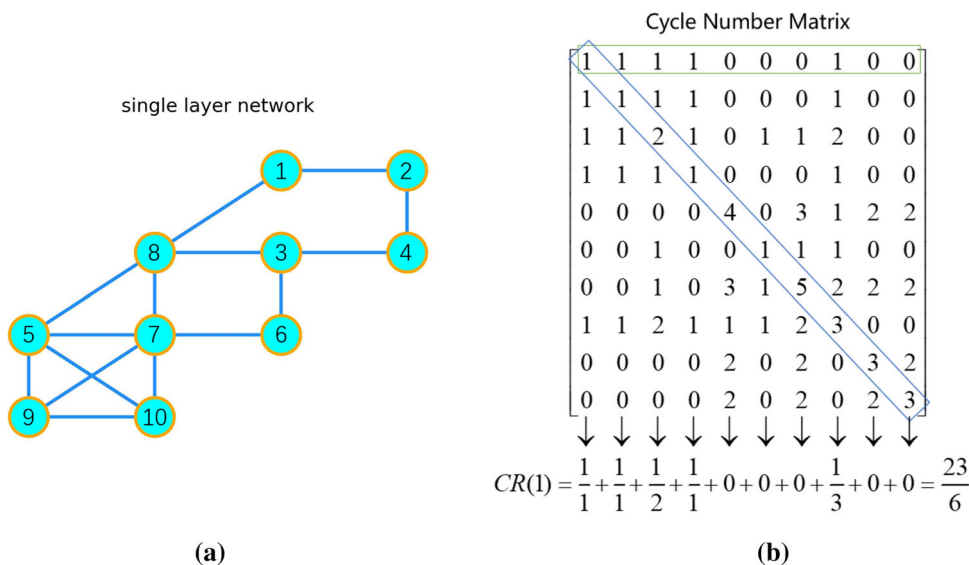
(a)   **(b)**

Fig. 2 Cycle number and cycle ratio in a single-layer network. **a** An example of a single-layer network with ten nodes. **b** Cycle number matrix of the network example, and how the cycle ratio is obtained from the cycle number matrix

Table 1 Enumerating the smallest basic cycles, the cycle number (CN) and the cycle ratio (CR) of a network example are calculated

Node	Associated smallest basic cycle 2	CN	CR
1	{1, 2, 3, 4, 8}	1	23/6
2	{1, 2, 3, 4, 8}	1	23/6
3	{1, 2, 3, 4, 8}, {3, 6, 7, 8}	2	88/15
4	{1, 2, 3, 4, 8}	1	23/6
5	{5, 7, 10}, {5, 7, 9}, {5, 7, 8}, {5, 9, 10}	4	49/15
6	{3, 6, 7, 8}	1	61/30
7	{3, 6, 7, 8}, {5, 7, 10}, {5, 7, 9}, {7, 9, 10}, {5, 7, 8}	5	21/4
8	{1, 2, 3, 4, 8}, {3, 6, 7, 8}, {5, 7, 8}	3	133/20
9	{5, 7, 9}, {7, 9, 10}, {5, 9, 10}	3	77/30
10	{5, 7, 10}, {7, 9, 10}, {5, 9, 10}	3	77/30

3.3.3 Minimum flux criterion

The weight setting of an inter-layer edge is usually put into use on layers pair by pair, and it is not only for precision of weights but also for complexity, because the complexity of spectrum calculation increases exponentially when a network grows. But an inter-layer cycle structure is not limited to two layers; when an inter-layer cycle stretches across more than two layers, we introduce the minimum flux criterion to evaluate the significance.

The minimum flux criterion is stated as follows: the importance of the inter-layer cycle structure equals the minimum weight of the inter-layer edges it contains. One of the characteristics of a path is its flux depends on the minimum flux along the whole path, which works like a bottleneck, and the cycle structure can be regarded as a specific path structure with directional or bidirectional infinite continuity.

3.3.4 Graph mapping simplification

This algorithm will not perform well when a multiplex network explosively grows. Specifically, we can map all layers into a single layer, which is shown as Fig. 4. This mapping process will cause distortion (in Fig. 4, inter-layer cycle {3, 5, 3′, 5′} disappear after mapping), but in comparison with the simplification in complexity, the distortion is acceptable.

3.3.5 Hybrid indexes based on Bayesian inference

Although we introduce an inter-layer cycle structure to identify vital nodes in the framework of a multiplex network, and make adjustment to the nonequivalent importance between inter-layer edges and intra-layer edges, the reliability of the adjustment is not quantified. For a convincing and precise theoretical method, we introduce Bayesian inference to make a choice between the unweighted indexes and the weighted indexes.
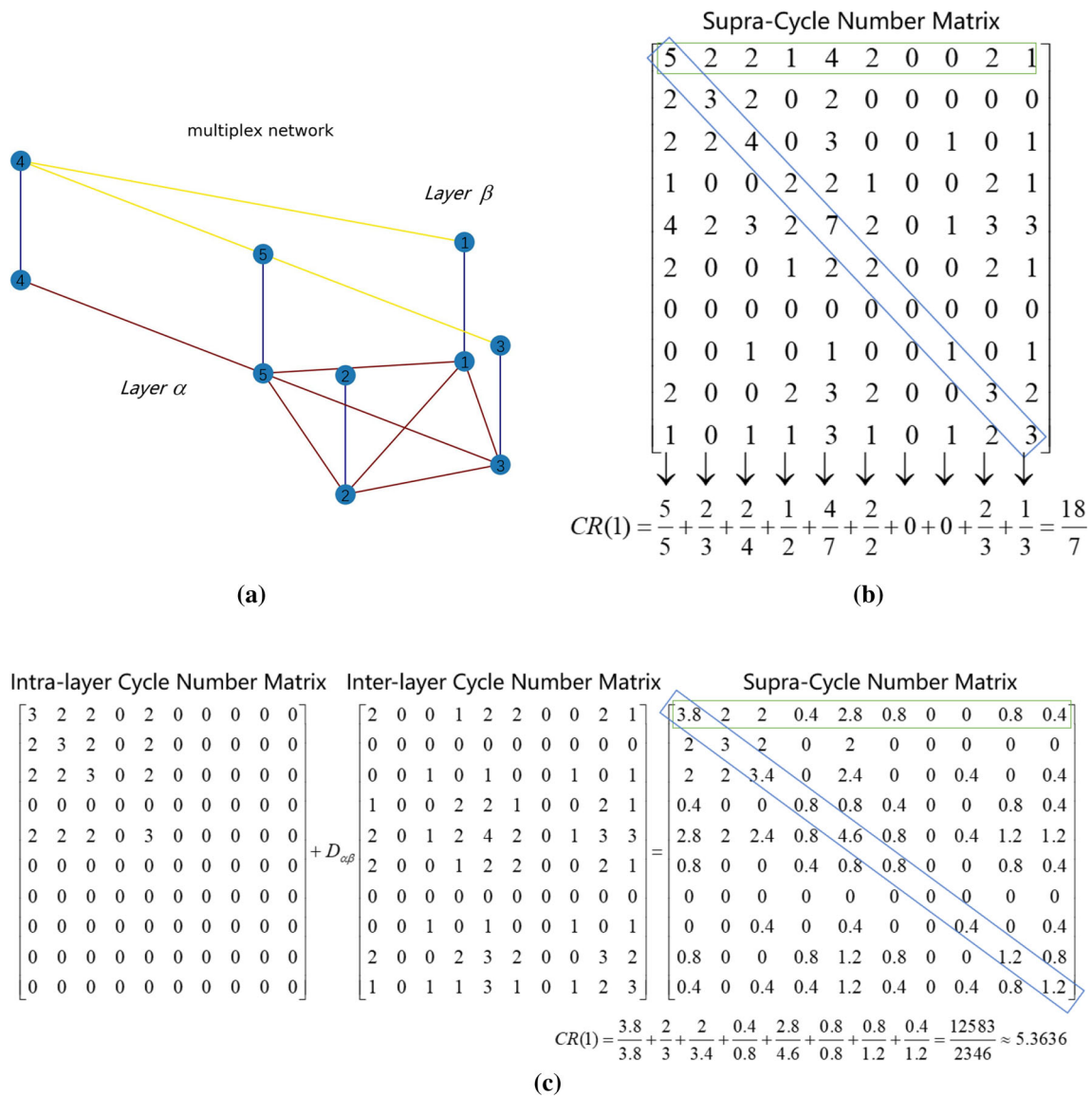
Supra-Cycle Number Matrix

$$
\begin{bmatrix}
5 & 2 & 2 & 1 & 4 & 2 & 0 & 0 & 2 & 1 \\
2 & 3 & 2 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\
2 & 2 & 4 & 0 & 3 & 0 & 0 & 1 & 0 & 1 \\
1 & 0 & 0 & 2 & 2 & 1 & 0 & 0 & 2 & 1 \\
4 & 2 & 3 & 2 & 7 & 2 & 0 & 1 & 3 & 3 \\
2 & 0 & 0 & 1 & 2 & 2 & 0 & 0 & 2 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \\
2 & 0 & 0 & 2 & 3 & 2 & 0 & 0 & 3 & 2 \\
1 & 0 & 1 & 1 & 3 & 1 & 0 & 1 & 2 & 3
\end{bmatrix}
$$

$$
CR(1) = \frac{5}{5} + \frac{2}{3} + \frac{2}{4} + \frac{1}{2} + \frac{4}{7} + \frac{2}{2} + 0 + 0 + \frac{2}{3} + \frac{1}{3} = \frac{18}{7}
$$

(a)

(b)

Intra-layer Cycle Number Matrix

$$
\begin{bmatrix}
3 & 2 & 2 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\
2 & 3 & 2 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\
2 & 2 & 3 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
2 & 2 & 2 & 0 & 3 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
+ D_{\alpha\beta}
$$

Inter-layer Cycle Number Matrix

$$
\begin{bmatrix}
2 & 0 & 0 & 1 & 2 & 2 & 0 & 0 & 2 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \\
1 & 0 & 0 & 2 & 2 & 1 & 0 & 0 & 2 & 1 \\
2 & 0 & 1 & 2 & 4 & 2 & 0 & 1 & 3 & 3 \\
2 & 0 & 0 & 1 & 2 & 2 & 0 & 0 & 2 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \\
2 & 0 & 0 & 2 & 3 & 2 & 0 & 0 & 3 & 2 \\
1 & 0 & 1 & 1 & 3 & 1 & 0 & 1 & 2 & 3
\end{bmatrix}
=
$$

Supra-Cycle Number Matrix

$$
\begin{bmatrix}
3.8 & 2 & 2 & 0.4 & 2.8 & 0.8 & 0 & 0 & 0.8 & 0.4 \\
2 & 3 & 2 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\
2 & 2 & 3.4 & 0 & 2.4 & 0 & 0 & 0.4 & 0 & 0.4 \\
0.4 & 0 & 0 & 0.8 & 0.8 & 0.4 & 0 & 0 & 0.8 & 0.4 \\
2.8 & 2 & 2.4 & 0.8 & 4.6 & 0.8 & 0 & 0.4 & 1.2 & 1.2 \\
0.8 & 0 & 0 & 0.4 & 0.8 & 0.8 & 0 & 0 & 0.8 & 0.4 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0.4 & 0 & 0.4 & 0 & 0 & 0.4 & 0 & 0.4 \\
0.8 & 0 & 0 & 0.8 & 1.2 & 0.8 & 0 & 0 & 1.2 & 0.8 \\
0.4 & 0 & 0.4 & 0.4 & 1.2 & 0.4 & 0 & 0.4 & 0.8 & 1.2
\end{bmatrix}
$$

$$
CR(1) = \frac{3.8}{3.8} + \frac{2}{3} + \frac{2}{3.4} + \frac{0.4}{0.8} + \frac{2.8}{4.6} + \frac{0.8}{0.8} + \frac{0.8}{1.2} + \frac{0.4}{1.2} = \frac{12583}{2346} \approx 5.3636
$$

(c)

Fig. 3 The supra-cycle number and supra-cycle ratio in a multiplex network. **a** An example of a multiplex network with two layers (layer α and layer β). **b** The "unweighted" supra-cycle number and supra-cycle ratio calculation in the network example. **c** The weighted supra-cycle number and supra-cycle ratio for a multiplex network, where $D_{\alpha\beta}$ is the weight of the inter-layer edge, which is 0.4 in this case

To describe the reliability of an index, we set up a random event set Ω at first, where $\Omega = \{\Theta_1, \Theta_2, \Theta_3\}$, $P(\Theta_1) + P(\Theta_2) + P(\Theta_3) = 1$, event Θ_1 represents the index is highly reliable, event Θ_2 represents the index is unreliable, and event Θ_3 represents the reliability of index is uncertain. Taking the cycle number of node v_i as an example and denoting it as $CN(i)$, the basic probability assignment [38] can be defined as follows:

$$
P_i(\Theta_1) = \frac{|CN(i) - CN_{\min}|}{CN_{\max} - CN_{\min} + \alpha} \tag{13}
$$

$$
P_i(\Theta_2) = \frac{|CN(i) - CN_{\max}|}{CN_{\max} - CN_{\min} + \alpha} \tag{14}
$$

where $CN_{\max} = \max\{CN(1), CN(2), \ldots, CN(n)\}$, $CN_{\min} = \min\{CN(1), CN(2), \ldots, CN(n)\}$, and α is a user-defined parameter, $\alpha \in [0, 1]$. For the sake of simplicity, the α equals 1 in this paper. There are two states Θ_1 and Θ_2 in Θ_3, and the ratio $\beta = \frac{P(\Theta_1\Theta_3)}{\Theta_3}$ is designed to describe preference for the reliability, $\beta \in [0, 1]$. The ratio β is also a user-defined parameter, and it usually equals 0.5. The reliability can be easily rewritten as $P(\Theta_1) + \beta P(\Theta_3)$.

The disadvantage for the reliability mentioned above is obvious; the parameters α and β contain too much subjective decision. Thus, we use Bayesian inference to make a better decision. Firstly, the indexes should be normalized within 0 and 1. Then we set a probability

Table 2 Enumerating the smallest basic cycles, the supra-cycle number and the supra-cycle ratio of the network example are calculated. We denote node v_i in layer α as i, and node v_i in layer β as i'. The wSCN is the weighted supra-cycle number, and the wSCR is the weighted supra-cycle ratio

Node	Associated smallest basic cycle		wSCN	wSCR
1	Inter-layer	$\{1, 5, 1', 4', 5'\}$, $\{1, 4, 5, 1', 4'\}$	3.8	5.3636
	Intra-layer	$\{1, 2, 5\}$, $\{1, 3, 5\}$, $\{1, 2, 3\}$		
2	Inter-layer	null	3	2.5493
	Intra-layer	$\{1, 2, 5\}$, $\{1, 2, 3\}$, $\{2, 3, 5\}\}$		
3	Inter-layer	$\{3, 5, 3', 5'\}$	3.4	4.0481
	Intra-layer	$\{1, 2, 5\}$, $\{1, 3, 5\}$, $\{1, 2, 3\}\}$		
4	Inter-layer	$\{1, 5, 1', 4', 5'\}$, $\{1, 4, 5, 1', 4'\}$	0.8	2.7792
	Intra-layer	null		
5	Inter-layer	$\{1, 5, 1', 4', 5'\}$,$\{1, 4, 5, 1', 4'\}$,$\{3, 5, 3', 5'\}$, $\{4, 5, 4', 5'\}$	4.6	8.1094
	Intra-layer	$\{1, 2, 5\}$, $\{1, 3, 5\}$, $\{2, 3, 5\}\}$		
1'	Inter-layer	$\{1, 5, 1', 4', 5'\}$, $\{1, 4, 5, 1', 4'\}$	0.8	2.8844
	Intra-layer	null		
2'	Inter-layer	null	0	0
	Intra-layer	null		
3'	Inter-layer	$\{3, 5, 3', 5'\}$	0.4	1.5379
	Intra-layer	null		
4'	Inter-layer	$\{1, 5, 1', 4', 5'\}$, $\{1, 4, 5, 1', 4'\}$, $\{4, 5, 4', 5'\}$	1.2	4.1381
	Intra-layer	null		
5'	Inter-layer	$\{1, 5, 1', 4', 5'\}$, $\{3, 5, 3', 5'\}$, $\{4, 5, 4', 5'\}$	1.2	3.1504
	Intra-layer	null		

(a)

(b)
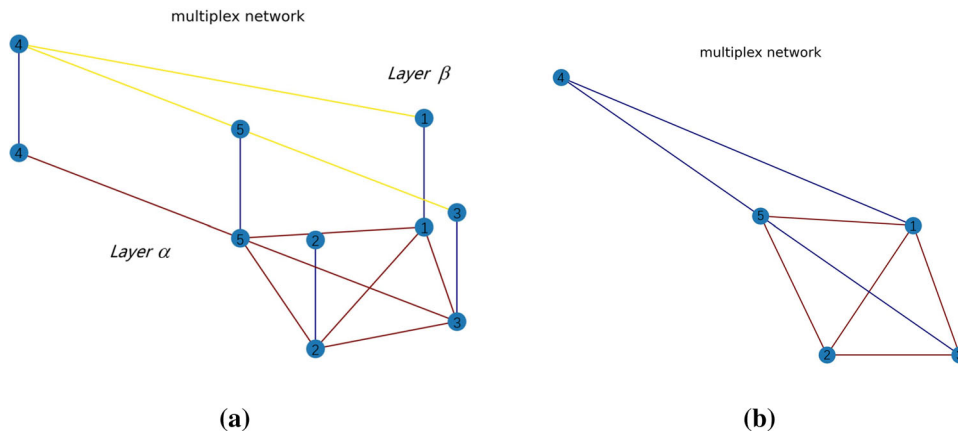
Fig. 4 **a** The original structure of a multiplex network. **b** The structure of a multiplex network after mapping

evaluation function as follows:

$$P(\Theta_1 \mid \neg \Theta'_1) = \frac{P(\Theta_1) \prod_{i=2,3} P(\Theta'_1 \mid \Theta)}{P(\Theta'_1)} \quad (15)$$

where Θ and Θ' represent two different indexes. We choose an index as the reliable index of the node when the probability of the index is higher than another. We denote the index from the unweighted cycle number and weighted cycle number by probability as a hybrid supra-cycle number, and we get a hybrid supra-cycle ratio in the same way. The hybrid supra-cycle number and hybrid supra-cycle ratio of the example graph in Fig. 3 can be recalculated, and the ranks similar to the

weighted supra-cycle number and weighted supra-cycle ratio come out, which is shown in Table 3.

For the sake of rigor, we will compare the hybrid supra-cycle number, the hybrid supra-cycle ratio, and some other methods in the Experiments section.

4 Experiments

In this section, we test and compare the indexes mentioned above by their performance on robustness, synchronization, and transmission dynamics. All the experiments for evaluation apply in four real multiplex networks: multiplex social networks of a sample of

Table 3 Recalculation of the indexes, the hybrid supra-cycle number, and the hybrid supra-cycle ratio of a network example. wSCN is the weighted supra-cycle number, wSCR is the weighted supra-cycle ratio, HSCN is the hybrid supra-cycle number, and HSCR is the hybrid supra-cycle ratio

Node	wSCN	wSCR	HSCN	HSCR
1	3.8	5.3636	0.8261	0.6614
2	3	2.5493	0.6522	0.3144
3	3.4	4.0481	0.7391	0.4992
4	0.8	2.7792	0.2857	0.3634
5	4.6	8.1094	1	1
$1'$	0.8	2.8844	0.2857	0.3877
$2'$	0	0	0	0
$3'$	0.4	1.5379	0.1429	0.2101
$4'$	1.2	4.1381	0.4286	0.5471
$5'$	1.2	3.1504	0.4286	0.5532

physicians in the United States [39], a multiplex air transportation network of the European Union [40], a multiplex genetic and protein interaction network of *Caenorhabditis elegans* [41,42], and a multiplex co-authorship network of the arXiv free scientific repository [43]. The essential information of each network is shown in Table 4.

4.1 Robustness analysis

Robustness and stability are always the core issue in network study. In the identification of vital nodes, reinforcing the protection for these vital nodes we select will help us to maintain the network, which can enhance its robustness and stability. Node removal attack [44] is a common method to evaluate the rank of vital nodes, where we assume a network is under attack by removal of nodes, and the removal always takes place on the most important node of the rank in each iteration; and the size of the largest giant connected component (GCC) in the left part of the network represents the performance of our vital node identification method.

The difference between removal of MNs and LNs is obvious: when removal happens to a node, all replicas of the node will be removed from the network if we do research on the MN, and only the node itself will be removed if we do research on the LN. Experiments on

four multiplex networks are shown in Figs. 5, 6, 7, and 8.

In Figs. 5 and 6, we compare indexes, including the unweighted supra-cycle ratio (SCR), unweighted supra-cycle number (SCN), hybrid supra-cycle ratio (HCR), hybrid supra-cycle number (HCN), degree centrality (D), semi-local centrality (S), ClusterRank (Clu), k-core index (Co), betweenness centrality (B), and closeness centrality (Clo).

In Fig. 5, it is obvious that HCN and HCR do not perform as well as SCN and SCR when the network is not huge enough (Fig. 5a), and the situation changes when the network grows (Fig. 5b–d). By the way, degree centrality and betweenness centrality have good performance as well, for which there are two reasons. On one hand, nodes with a high degree work as hubs and nodes with betweenness centrality working as bridges in the network; GCCs are vulnerable on removal of these nodes. On the other hand, node removal attack does not reflect the difference between inter-layer interaction and intra-layer interaction.

In a multiplex network with M layers and N nodes in each layer, robustness [45] of the network is defined as follows:

$$R = \frac{1}{N} \sum_{i}^{N} g(i) \tag{16}$$

where the relative size $g(i)$ is the number of nodes in the largest giant connected component after removing i nodes, and the normalization factor $1/N$ makes sure the robustness of different networks can be compared. Obviously, a smaller R means a quicker collapse, thus better performance. The result is shown in Table 5.

When it comes to MNs, we can get several other indexes by mapping all the layers of the network into a single-layer network. In Fig. 7, SCRM, HCRM, SCNM, and HCNM are SCR, HCR, SCN, and HCN, respectively, in the mapping network.

We select the best performing index in Fig. 7 and compare it with other indexes in Fig. 8. The results reveal that the hybrid indexes based on cycle structure have better performance, and the indexes in the mapping network have advantages both in performance and complexity.

Table 4 The essential information of the test multiplex network

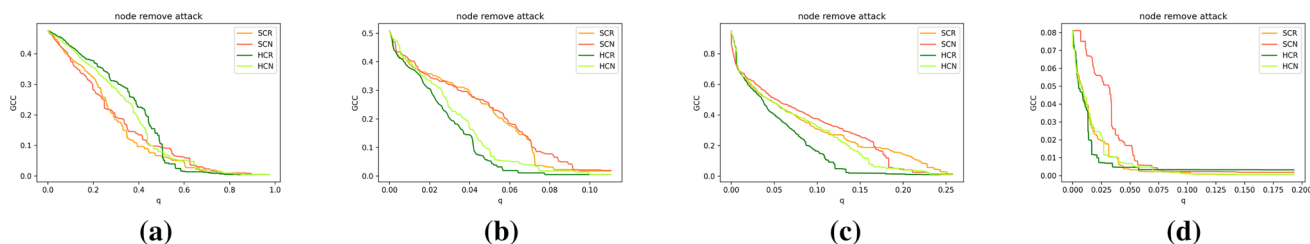| Name	Category	Node	Edge	Layer	Used layer
CKM–physicians–innovation multiplex	Physics	246	1551	3	3
EU air transportation multiplex Network transportation	Transportation	450	3588	37	4
C. elegans multiplex GPI network	Biology	3879	8181	6	4
ArXiv Netscience multiplex coauthorship	Coauthorship	14,489	59,026	13	2

Fig. 5 The GCC size of a multiplex network after a q fraction of LNs is removed. We compare indexes about cycle structure here in these subgraphs. **a** CKM–physicians–innovation multiplex network. **b** EU air transportation multiplex network. **c** *C. elegans* multiplex GPI network. **d** ArXiv Netscience multiplex
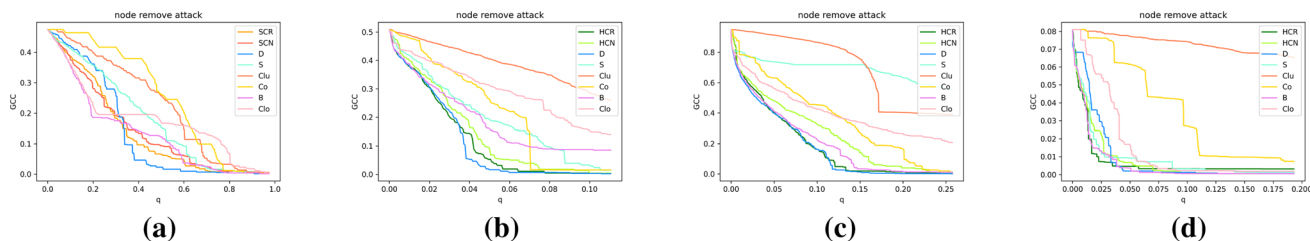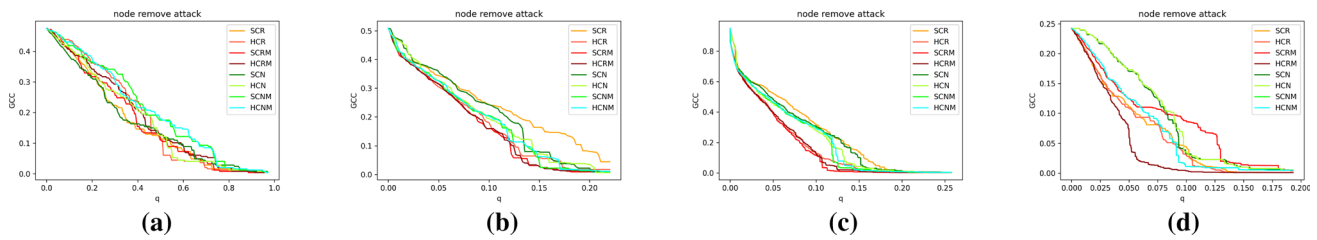

Fig. 6 The GCC size of a multiplex network after a q fraction of LNs is removed. We compare best performing indexes in Fig. 5 with several other indexes. **a** CKM–physicians–innovation multiplex network. **b** EU air transportation multiplex network. **c** *C. elegans* multiplex GPI network. **d** ArXiv Netscience multiplex

There are multiple reasons for the better performance of HCNM and HCRM; the removal of MNs diminishes the influence of the inter-layer edge on GCC, and apparently the weight setting more or less fixes distortion caused by mapping. We quantify the robustness of networks on MN removals, which is shown in Table 6.

4.2 Pinning control

We evaluate the validity of the indexes in synchronization by pinning these selected top nodes of rank [46]. Generally, the interacting dynamics of a network with N nodes is shown as follows:

$$\dot{x}_i = f(x_i) + \sigma \sum_{j=1}^{N} l_{ij}\Gamma(x_{ij}) + U_i(x_1, \dots, x_n) \quad (17)$$

where the $x_i \in \mathbb{R}$ is the state of node i, the coupled strength is denoted as positive constant σ, l_{ij} is the elements of a Laplacian matrix of the network, the inner coupled matrix Γ is positive semi-definite, and U_i is the controller of node i. The goal of pinning control is driving the system from an initial state to a target state in finite time by pinning some nodes. Similar to the robustness test, all the nodes are ranked in descending order by the indexes; then the nodes are successively pinned, which starts from the top rank. The synchronization can be measured by the smallest nonzero eigenvalue of the principal submatrix [46,47], and the higher synchronization is correlated with the higher value. We denote it as $\mu(L_{-Q})$, where Q is the number of pinned nodes, and L_{-Q} is the principal submatrix after deleting Q nodes. Inspired by the evaluation equation of robustness, the analogous evaluation equation for synchronization is proposed by researchers:

Table 5 Robustness of four multiplex networks by the indexes mentioned above on LN removal

	CKM–physicians–innovation	EU air transportation	*C. elegans*	ArXiv netscience
SCR	0.1389	0.1986	0.2937	**0.0074**
SCN	0.1456	0.2079	0.3020	0.0152
HCR	0.1752	**0.1218**	**0.1879**	**0.0061**
HCN	0.1672	0.1466	0.2634	0.0158
D	**0.1381**	**0.1144**	**0.1765**	0.0089
S	0.1825	0.2005	0.6918	0.0090
Clu	0.2326	0.3920	0.7032	0.0735
Co	0.2574	0.2244	0.3722	0.0365
B	**0.1353**	0.1915	0.2081	**0.0078**
Clo	0.1796	0.2966	0.4101	0.0076

Significant bold values are less robustness value means smaller average GCC, which means the network collapses quicker

Fig. 7 The GCC size of a multiplex network after a q fraction of LNs is removed. We compare indexes about cycle structure here in these subgraphs. **a** CKM–physicians–innovation multiplex network. **b** EU air transportation multiplex network. **c** *C. elegans* multiplex GPI network. **d** ArXiv Netscience multiplex
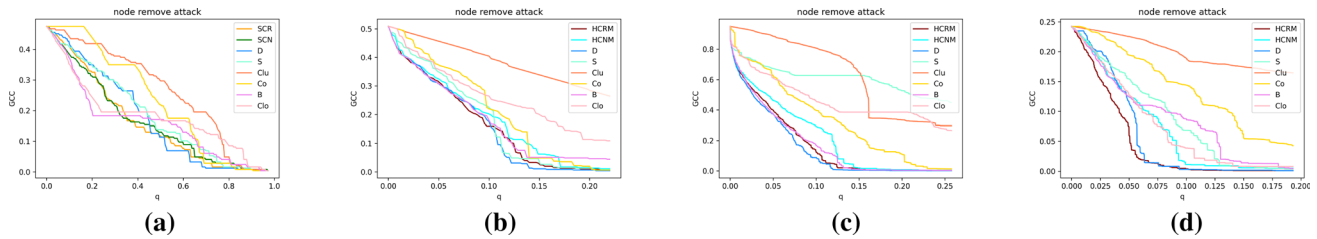

Fig. 8 The GCC size of a multiplex network after a q fraction of LNs is removed. We compare indexes about cycle structure with other indexes in these subgraphs. **a** CKM–physicians–innovation multiplex network. **b** EU air transportation multiplex network. **c** *C. elegans* multiplex GPI network. **d** ArXiv Netscience multiplex

Table 6 Robustness of four multiplex networks by the indexes mentioned above on MN removal

	CKM–physicians–innovation	EU air transportation	*C. elegans*	ArXiv netscience
SCR	**0.1590**	0.2408	0.2590	0.0664
SCN	0.1616	0.2096	0.2435	0.0926
HCR	0.1772	0.1772	0.1783	0.0644
HCN	0.1710	0.1949	0.2243	0.0943
SCRM	0.1649	0.1667	0.1626	0.0893
SCNM	0.2038	0.1838	0.2157	0.0716
HCRM	0.1779	**0.1654**	**0.1717**	**0.0455**
HCNM	0.2017	0.1870	0.2198	0.0716
D	0.1723	**0.1595**	**0.1515**	**0.0569**
S	0.1855	0.1824	0.6098	0.0893
Clu	0.2622	0.3871	0.6353	0.1996
Co	0.2298	0.2194	0.3373	0.1411
B	**0.1604**	0.1823	0.1760	0.0893
Clo	0.1814	0.2668	0.4697	0.0748

Significant bold values are less robustness value means smaller average GCC, which means the network collapses quicker

$$P = \frac{1}{Q_{\max}} + \sum_{Q=1}^{Q_{\max}} \mu(L_{(-Q)}) \qquad (18)$$

where $Q_{\max}$ is the maximum of pinned nodes. In simulations, we set the maximum equal to 10% nodes. The result of pinning LNs by different indexes is presented in Table 6. We highlight the top two indexes in each network, and the result shows that indexes based on a cycle structure have good performance on synchronization but not the best all the time (Tables 7, 8).

4.3 Transmission dynamics

The SIR model is a classical method for network transmission dynamic research, and it can also be used for evaluation of vital nodes identification. Researchers developed the SIR model to simulate the pandemic. Every individual can be in one of three states: susceptible (S), infected (I), recovered (R). Individuals in the S state are regarded as healthy ones and easy to be infected. Individuals in the I state are able to infect healthy ones and can also probably transfer into state R (simulating recovery). Individuals will not be infected or infect others again once transferring into state R. Recently, scholars did abundant research on the SIR model in a disease–information coupled multiplex network [48], which revealed the inhibition effect of information on disease spread. When those vital nodes we select are set as initially infected individuals (seed nodes), the speed and breadth of transmission reflects

Table 7 Synchronization of four multiplex networks by the indexes mentioned above on pinning LNs

	CKM–physicians–innovation	EU air transportation	*C. elegans*	ArXiv netscience
SCR	3.4886E−16	4.3906E−18	1.4376E−18	5.6773E−20
SCN	3.2663E−16	**4.8838E−18**	9.8974E−19	4.5442E−18
HCR	3.6865E−16	3.7540E−18	**2.8520E−18**	**7.9041E−18**
HCN	3.1922E−16	4.2485E−18	1.0656E−18	5.7689E−18
D	3.6354E−16	2.8187E−18	6.1360E−19	4.5442E−18
S	**4.0932E−16**	3.8061E−18	6.7993E−19	2.9700E−18
Clu	3.2718E−16	3.8053E−18	1.1256E−18	1.2458E−18
Co	3.1186E−16	3.9261E−18	1.1242E−18	1.2458E−18
B	3.7463E−16	2.3368E−18	**1.8688E−18**	**1.7515E−17**
Clo	**4.4551E−16**	**4.6898E−18**	5.6352E−19	1.2544E−18

Significant bold values are bigger synchronization value means bigger average smallest nonzero eigen value when more and more nodes are pinned, which means better synchronization

Table 8 Synchronization of four multiplex networks by the indexes mentioned above on pinning MNs

	CKM–physicians–innovation	EU air transportation	*C. elegans*	ArXiv netscience
SCR	1.4657E−16	1.1359E−17	4.0218E−18	1.6203E−17
SCN	1.4453E−16	1.0451E−17	**5.9910E−18**	2.4618E−18
HCR	3.3132E−16	**1.4187E−17**	1.8656E−18	3.1507E−18
HCN	2.4974E−16	1.2156E−17	**4.9718E−18**	6.0328E−18
SCRM	**3.4227E−16**	9.8288E−18	2.2731E−18	6.9271E−18
SCNM	2.1900E−16	1.1158E−17	9.4398E−19	1.6440E−17
HCRM	1.8943E−16	1.1435E−17	9.3987E−19	6.9276E−18
HCNM	2.2931E−16	1.2348E−17	2.2049E−18	**1.6441E−17**
D	1.9295E−16	1.3035E−17	1.1866E−18	1.4862E−17
S	**4.4781E−16**	9.2381E−18	4.5992E−18	2.9334E−18
Clu	2.8172E−16	1.2269E−17	1.0051E−18	1.1762E−17
Co	2.9362E−16	1.8097E−17	2.0366E−18	1.8794E−17
B	1.4632E−16	1.2092E−17	2.0366E−18	**2.1846E−17**
Clo	1.9109E−16	**1.4626E−17**	1.9267E−18	6.1986E−18

Significant bold values are bigger synchronization value means bigger average smallest nonzero eigen value when more and more nodes are pinned, which means better synchronization

the importance of seed nodes, which is deemed as the performance of the vital node identification method.

In this paper, epidemic spread is not limited in every single layer; nodes in the I state can infect others in different layers. Unlike the single-layer network, when transmission happens in the same layer, nodes in the I state infect others with a probability β_1, and with probability β_2 when transmission happens in different layers. Nodes in the I state recover at recovery rate γ. For LNs, inter-layer transmission is easy to define: transmission between a node and its replica is inter-layer. But for MNs, the node is infected in every layer once any one of its replicas is infected, and in this circumstance, inter-layer transmission should be defined in another way: a transmission process is inter-layer if an MN gets infected in a layer and infects others in different layers. Figure 9 provides a diagram of inter-layer transmission for LNs and MNs.

The dynamic equations of the SIR model in a multiplex network are shown as follows:

$$\frac{\mathrm{d}(s)}{\mathrm{d}t} = -\beta_1 s(t)i(t) - \beta_2 s(t)i(t) \qquad (19)$$

$$\frac{\mathrm{d}i(s)}{\mathrm{d}t} = \beta_1 s(t)i(t) + \beta_2 s(t)i(t) - \gamma i(t) \qquad (20)$$

$$\frac{\mathrm{d}y(t)}{\mathrm{d}t} = \gamma i(t) \qquad (21)$$

where t is the time (iteration steps), and S, $i(t)$, and $\gamma(t)$ are the number of nodes in states S, I, and R. When the intra-layer infection rate β_1 equals 0.01, the inter-layer infection rate β_2 equals 0.005, and the recovery rate γ equals 0.05. We select the top 1% nodes as seed nodes and simulate the whole process of the transmission 50 times, then compare the average of variable proportion of nodes in state I under different vital node identification methods, which is shown in Figs. 10, 11, 12, and 13.

When an LN is the basic element in our research, we make comparison between SCN and HCN, and SCR and HCR in multiplex networks we used above. The results show that HCN and HCR have better performance most of time, and the advantages of HCN and HCR are not clear when the network is way too small. The transmission of LNs shows some similar characteristics to transmission in a single-layer network, because
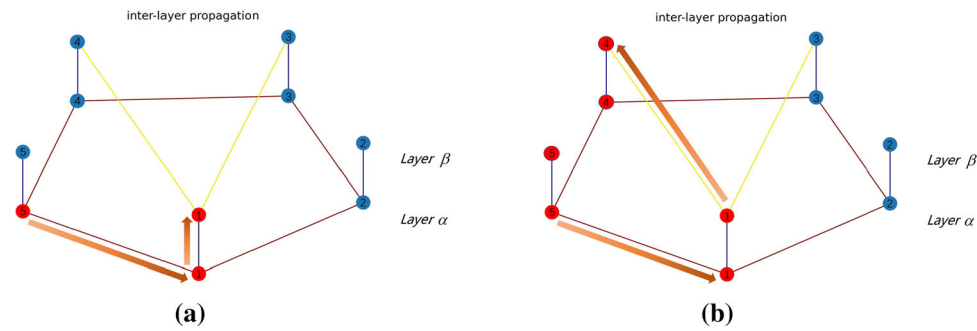
Fig. 9 Inter-layer transmission in a multiplex network consisting of two layers and five nodes in each layer. The nodes in the S state are blue, the nodes in the I state are red, and the orange arrow is the transmission process. **a** Inter-layer transmission for LNs: node v_1^α infects v_1^β. **b** Inter-layer transmission for MNs: node v_1 gets infected in layer α and infects node v_4 in layer β
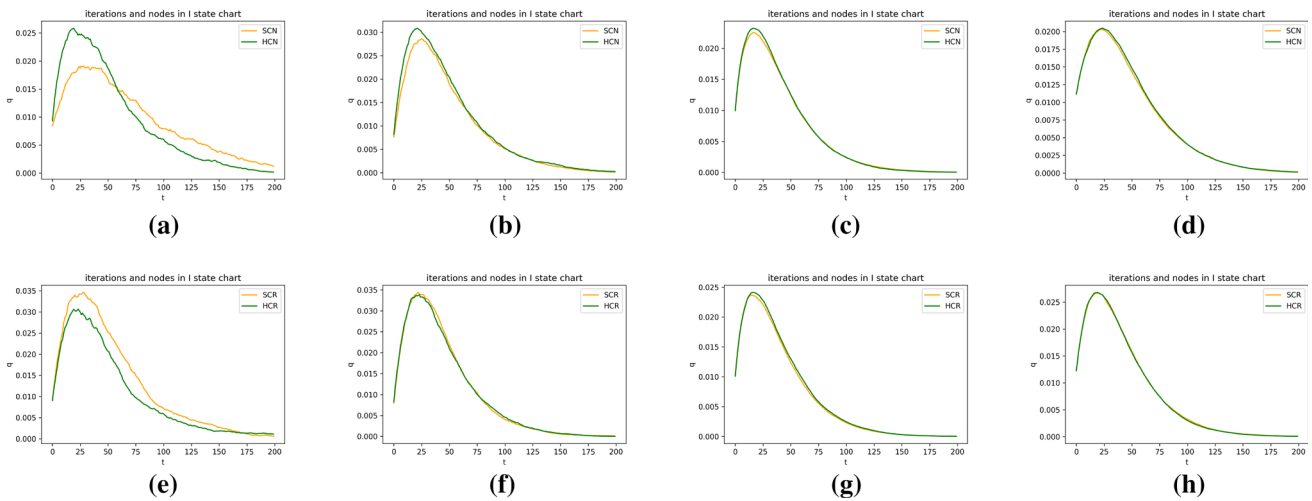
Fig. 10 When the LN is the subject of study, comparison between SCN and HCN is in subgraph (**a**–**d**), and comparison between SCR and HCR is in subgraph (**e**–**h**). Subgraphs (**a**) and (**e**) are simulations in the CKM–physicians–innovation multiplex network. Subgraphs (**b**) and (**f**) are simulations in the EU air transportation multiplex network. Subgraphs (**c**) and (**g**) are simulations in the *C. elegans* multiplex GPI network. Subgraphs (**d**) and (**h**) are simulations in the ArXiv Netscience multiplex
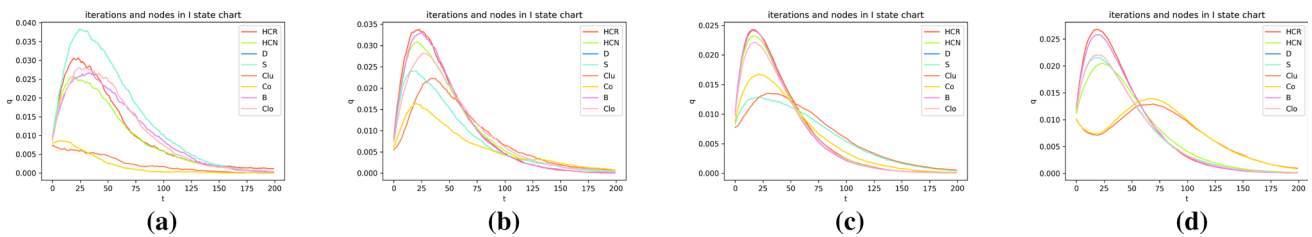
Fig. 11 Comparison among several sets of LNs from different indexes. **a** CKM–physicians–innovation multiplex network. **b** EU air transportation multiplex network. **c** *C. elegans* multiplex GPI network. **d** ArXiv Netscience multiplex

the transmission of LNs can be regarded as transmission in a single-layer network with two infection rates. Because of the similitude, HCN and HCR do not show overwhelming advantages, and if we take simplicity into consideration, indexes based on an unweighted cycle structure show less but similar practicability in the transmission task.

We make comparison among HCN, HCR, and some other indexes in four multiplex networks. Although HCN and HCR do not have more superb performance than SCN and SCR all the time, situations with other indexes are much better, where HCR is the best for both speed and breadth of transmission in three cases [subgraph (b), subgraph (c), and subgraph (d)] and the second best once [subgraph (a)]; HCN just has less
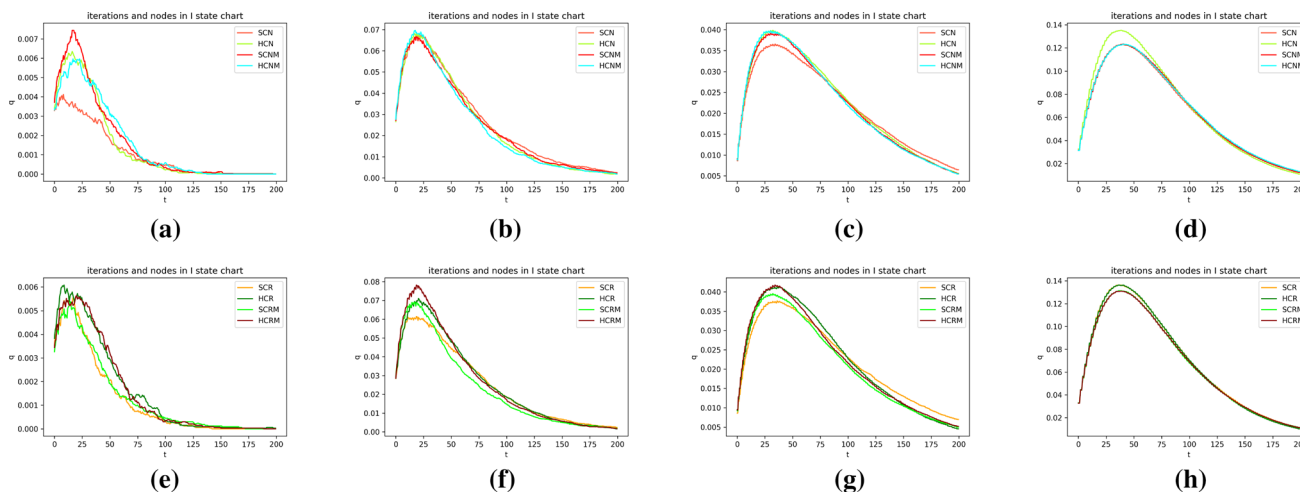
Fig. 12 When MN is the subject of study, comparison between SCN, HCN, SCNM, and HCNM is in subgraph (**a**–**d**), and comparison between SCR, HCR, SCRM, and HCRM is in subgraph (**e**–**h**). Subgraphs (**a**) and (**e**) are simulations in a CKM–physicians–innovation multiplex network. Subgraphs (**b**) and (**f**) are simulations in an EU air transportation multiplex network. Subgraphs (**c**) and (**g**) are simulations in a *C. elegans* multiplex GPI network. Subgraphs (**d**) and (**h**) are simulations in an ArXiv Netscience multiplex
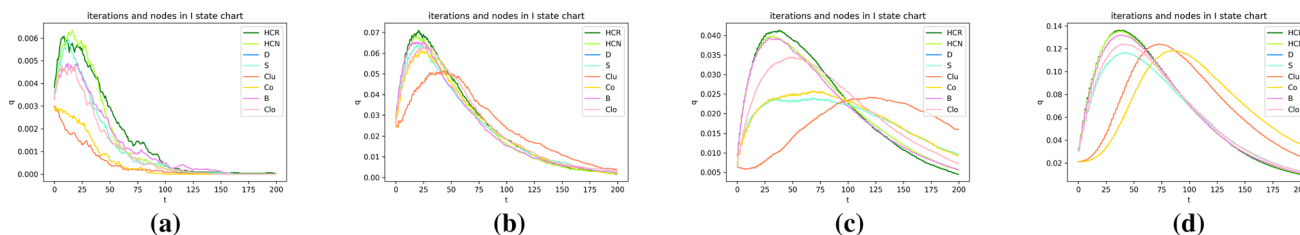
Fig. 13 Comparison among several sets of MNs from different indexes. **a** CKM–physicians–innovation multiplex network. **b** EU air transportation multiplex network. **c** *C. elegans* multiplex GPI network. **d** ArXiv Netscience multiplex

performance than HCR and betweenness centrality in two cases [subgraph (b) and subgraph (d)]. There is an exception that betweenness centrality is also one of the best methods for the transmission task, but betweenness centrality has tremendous disadvantages for complexity, which is a disaster when a network grows. And comparing with betweenness centrality, our algorithm of HCN and HCR has remarkable success for complexity in a huge sparse network.

When we work on MNs, the performance of HCN and HCR gets better. In Fig. 12, we make a comparison with CNs and CRs. The data prove that HCN and HCR have advantages for both speed and breadth. But the interesting thing is that HCN and HCR not only perform well; HCNM and HCRM also have advantages for both speed and breadth, and sometimes they are even better than HCN and HCR (Fig. 12b, f). This phenomenon suggests that HCNM and HCRM have good performance in transmission tasks as well, and when the algorithm is applied to a very huge multiplex network, especially a very dense multiplex network, mapping the network into a single layer can achieve good compromise between complexity and performance.

The reason why the transmission process performs well based on HCNM and HCRM is very clear. Although

an inter-layer transmission is redefined on MNs, it is not real layer-to-layer transmission because this inter-layer transmission requires continuity in the process or step. Real inter-layer transmission is a topological process, but the principle of transmission for an MN is once a node gets infected, all replicas get infected; thus, it is impossible to redefine inter-layer transmission topologically. With the limit of the principle, the effect of inter-layer edges is ignored for a multiplex network.

The disadvantage of transmission for an MN causes similar performance between HCN and HCNM, and HCR and HCRM, but it does not mean the inter-layer cycle structure and the weight setting are worthless. The limit of the transmission principle for the MN does not change the fact that the transmission process still works on inter-layer cycle structures, and these inter-layer cycle structures are not as important as the intra-layer cycle structures, which makes sure that the hybrid indexes based on cycle structure are an effective way to deal with transmission tasks.

5 Conclusion

In this paper, we introduce the cycle number and cycle ratio at first and develop them into a multiplex network. We integrate the nonequivalent importance between inter-layer edges and intra-layer edges and the reliability of the nonequivalence by Bayesian inference, then propose two indexes, the hybrid supra-cycle number and hybrid supra-cycle ratio, to identify two different types of vital nodes (LNs and MNs), and evaluate their performance with several different indexes in three aspects: robustness, synchronization, and transmission dynamics. In most cases, the two indexes have good performance, and we draw the following conclusions based on these experimental results: (i) Comparing with hybrid supra-cycle number, the hybrid supra-cycle ratio has better performance on a local scale; (ii) Our index is more suitable for study on MNs than LNs; and (iii) Mapping a multilayer network into a single-layer network is a time-saving and space-saving preprocessing step, especially for research on MNs in very huge multiplex networks.

We end this paper by presenting several open issues. Firstly, we work a lot on the LN because there is more than one kind of multilayer network, and we try to develop the cycle structure into a general multilayer network which may contain different nodes in different layers, and the MN will not exist; but the hybrid supra-cycle number and hybrid supra-cycle ratio on the LN do not work as well as on the MN, which means more reasonable, more effective, and more practical methods are needed. Moreover, although we preprocess the multiplex network by removing nodes in which the k-core number is equal or less than 1, the complexity of our algorithm on a huge but sparse network greatly decreases, which causes uncertainty in vital node identification. Besides, the order of cycle is not involved in the framework. Similar to the high-order cluster coefficient, we try to frame and analyze the order of cycle on network function and dynamics, and quantify these effects and interactive behaviors. Last but not least, our work focuses on topology, and the properties of nodes are not, but should be, involved for a better understanding of complex networks. We hope these missing parts are fulfilled in the future.

Author contributions

The research was planned by QY, GY, and HL. The simulations and data analysis were carried out by QY. The manuscript was written by QY and WC.

Data Availability The manuscript has associated data in a data repository. [Authors' comment: The network data that support the findings of this study are available through the corresponding references [39–43].]

References

1. Watts D.J. The "new" science of networks. Annu. Rev. Sociol., 243–270 (2004)
2. A.L. Barabási, R. Albert, Emergence of scaling in random networks. Science **286**(5439), 509–512 (1999)
3. M.E. Newman, Assortative mixing in networks. Phys. Rev. Lett. **89**(20), 208701 (2002)
4. D.J. Watts, S.H. Strogatz, Collective dynamics of 'small-world' networks. Nature **393**(6684), 440–442 (1998)
5. U. Alon, Network motifs: theory and experimental approaches. Nat. Rev. Genet. **8**(6), 450–461 (2007)
6. M.E. Newman, M. Girvan, Finding and evaluating community structure in networks. Phys. Rev. E **69**(2), 026113 (2004)
7. S. Pei, H.A. Makse, Spreading dynamics in complex networks. J. Stat. Mech: Theory Exp. **2013**(12), 12002 (2013)
8. P. Csermely, *Weak Links: Stabilizers of Complex Systems from Proteins to Social Networks* (Springer, Berlin, 2006), p.37
9. M. De Domenico, A. Solé-Ribalta, E. Omodei, S. Gómez, A. Arenas, Ranking in interconnected multilayer networks reveals versatile nodes. Nat. Commun. **6**(1), 1–6 (2015)
10. R. Cohen, S. Havlin, D. Ben-Avraham, Efficient immunization strategies for computer networks and populations. Phys. Rev. Lett. **91**(24), 247901 (2003)
11. R. Pastor-Satorras, A. Vespignani, Immunization of complex networks. Phys. Rev. E **65**(3), 036104 (2002)
12. W. Chen, L.V. Lakshmanan, C. Castillo, Information and influence propagation in social networks. Synthes. Lect. Data Manag. **5**(4), 1–177 (2013)
13. A.E. Motter, Y.C. Lai, Cascade-based attacks on complex networks. Phys. Rev. E **66**(6), 065102 (2002)
14. A.E. Motter, Cascade control and defense in complex networks. Phys. Rev. Lett. **93**(9), 098701 (2004)
15. R. Albert, H. Jeong, A.-L. Barabási, Error and attack tolerance of complex networks. Nature **406**(6794), 378–382 (2000)

16. R. Cohen, K. Erez, D. Ben-Avraham, S. Havlin, Breakdown of the internet under intentional attack. Phys. Rev. Lett. **86**(16), 3682 (2001)

17. R. Albert, I. Albert, G.L. Nakarado, Structural vulnerability of the north American power grid. Phys. Rev. E **69**(2), 025103 (2004)

18. M.G. Resende, P.M. Pardalos, *Handbook of Optimization in Telecommunications* (Springer, Berlin, 2008)

19. P. Csermely, T. Korcsmáros, H.J. Kiss, G. London, R. Nussinov, Structure and dynamics of molecular networks: a novel paradigm of drug discovery: a comprehensive review. Pharmacol. Therapeut. **138**(3), 333–408 (2013)

20. F. Radicchi, S. Fortunato, B. Markines, A. Vespignani, Diffusion of scientific credits and the ranking of scientists. Phys. Rev. E **80**(5), 056103 (2009)

21. Y.B. Zhou, L. Lü, M. Li, Quantifying the influence of scientists and their publications: distinguishing between prestige and popularity. New J. Phys. **14**(3), 033033 (2012)

22. L. Lü, D. Chen, X.L. Ren, Q.M. Zhang, Y.C. Zhang, T. Zhou, Vital nodes identification in complex networks. Phys. Rep. **650**, 1–63 (2016)

23. M. Martcheva, *An Introduction to Mathematical Epidemiology* (Springer, Berlin, 2015), p.61

24. P. Georgescu, H. Zhang, A lyapunov functional for a siri model with nonlinear incidence of infection and relapse. Appl. Math. Comput. **219**(16), 8496–8507 (2013)

25. Fan, T., Lü, L., Shi, D.: Towards the cycle structures in complex network: a new perspective (2019). arXiv:1903.01397

26. W. Zhang, W. Li, W. Deng, The characteristics of cycle-nodes-ratio and its application to network classification. Commun. Nonlinear Sci. Numer. Simul. **99**, 105804 (2021)

27. D. Zhao, L. Wang, S. Xu, G. Liu, X. Han, S. Li, Vital layer nodes of multiplex networks for immunization and attack. Chaos Solit. Fract. **105**, 169–175 (2017)

28. D. Zhao, L. Wang, Y. Zhi, J. Zhang, Z. Wang, The robustness of multiplex networks under layer node-based attack. Sci. Rep. **6**(1), 1–9 (2016)

29. A. Halu, R.J. Mondragón, P. Panzarasa, G. Bianconi, Multiplex pagerank. PLoS ONE **8**(10), 78293 (2013)

30. Solé-Ribalta, A., De Domenico, M., Gómez, S., Arenas, A.: Centrality rankings in multiplex networks. In: Proceedings of the 2014 ACM Conference on Web Science, pp. 149–155 (2014)

31. L. Solá, M. Romance, R. Criado, J. Flores, A. García del Amo, S. Boccaletti, Eigenvector centrality of nodes in multiplex networks. Chaos Interdiscip. J. Nonlinear Sci. **23**(3), 033131 (2013)

32. D. Zhao, L. Li, S. Li, Y. Huo, Y. Yang, Identifying influential spreaders in interconnected networks. Phys. Scr. **89**(1), 015203 (2013)

33. T. Fan, L. Lü, D. Shi, T. Zhou, Characterizing cycle structure in complex networks. Commun. Phys. **4**(1), 1–9 (2021)

34. N. Azimi-Tafreshi, J. Gómez-Gardenes, S. Dorogovtsev, k- core percolation on multiplex networks. Phys. Rev. E **90**(3), 032816 (2014)

35. D.B. Chen, H. Gao, L. Lü, T. Zhou, Identifying influential nodes in large-scale directed networks: the role of clustering. PLoS ONE **8**(10), 77455 (2013)

36. D. Chen, L. Lü, M.-S. Shang, Y.-C. Zhang, T. Zhou, Identifying influential nodes in complex networks. Phys. A **391**(4), 1777–1787 (2012)

37. A. Sole-Ribalta, M. De Domenico, N.E. Kouvaris, A. Diaz-Guilera, S. Gomez, A. Arenas, Spectral properties of the Laplacian of multiplex networks. Phys. Rev. E **88**(3), 032807 (2013)

38. H. Mo, C. Gao, Y. Deng, Evidential method to identify influential nodes in complex networks. J. Syst. Eng. Electron. **26**(2), 381–387 (2015)

39. J. Coleman, E. Katz, H. Menzel, The diffusion of an innovation among physicians. Sociometry **20**(4), 253–270 (1957)

40. A. Cardillo, J. Gómez-Gardenes, M. Zanin, M. Romance, D. Papo, F.D. Pozo, S. Boccaletti, Emergence of network features from multiplexity. Sci. Rep. **3**(1), 1–6 (2013)

41. M. De Domenico, V. Nicosia, A. Arenas, V. Latora, Structural reducibility of multilayer networks. Nat. Commun. **6**(1), 1–9 (2015)

42. C. Stark, B.-J. Breitkreutz, T. Reguly, L. Boucher, A. Breitkreutz, M. Tyers, Biogrid: a general repository for interaction datasets. Nucleic Acids Res. **34**(suppl-1), 535–539 (2006)

43. M. De Domenico, A. Lancichinetti, A. Arenas, M. Rosvall, Identifying modular flows on multilayer networks reveals highly overlapping organization in interconnected systems. Phys. Rev. X **5**(1), 011027 (2015)

44. S. Osat, A. Faqeeh, F. Radicchi, Optimal percolation on multiplex networks. Nat. Commun. **8**(1), 1–7 (2017)

45. C.M. Schneider, A.A. Moreira, J.S. Andrade Jr., S. Havlin, H.J. Herrmann, Mitigation of malicious attacks on networks. Proc. Natl. Acad. Sci. **108**(10), 3838–3841 (2011)

46. X.F. Wang, G. Chen, Pinning control of scale-free dynamical networks. Phys. A **310**(3–4), 521–531 (2002)

47. H. Liu, X. Xu, J.-A. Lu, G. Chen, Z. Zeng, Optimizing pinning control of complex dynamical networks based on spectral properties of grounded Laplacian matrices. IEEE Trans. Syst. Man Cybernet. Syst. **51**(2), 786–796 (2018)

48. Q. Zeng, Y. Liu, M. Tang, J. Gong, Identifying superspreaders in information-epidemic coevolving dynamics on multiplex networks. Knowl.-Based Syst. **229**, 107365 (2021)