Expert Systems With Applications 177 (2021) 114973

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Effective link prediction in multiplex networks: A TOPSIS method

Shenshen Bai, Yakun Zhang, Longjie Li *, Na Shan, Xiaoyun Chen

School of Information Science and Engineering, Lanzhou University, Lanzhou 730000, China

ARTICLE INFO

Keywords:
Link prediction
Multiplex network
TOPSIS method
Interlayer relevance

ABSTRACT

This paper investigates the link prediction in multiplex networks. Multiplex networks that represent multiple types of interaction between the same group of individuals are a special case of complex networks. Each type of interaction is modeled as a layer in a multiplex network. Usually, the topological structures between different layers of a multiplex network have a certain extent of correlation. As a result, the accuracy of link prediction in multiplex networks can be enhanced by combining the information of different layers. In this paper, link prediction in multiplex networks is regarded as a multiple-attribute decision-making problem, in which the potential links in the target layer are considered as alternatives, layers are viewed as attributes, and the similarity score of a potential link in each layer is an attribute value. In implementation, the TOPSIS method is employed to rank alternatives, and interlayer relevance is used to weight the attributes. The experimental results show that the proposed method is not sensitive to the parameter and the interlayer relevance measure, and achieves superior prediction accuracy.

1. Introduction

In real-world, a considerable amount of complex systems can be represented as complex networks, where nodes denote individuals or entities, and links (or edges) depict the interactions or relationships between nodes (Albert & Barabási, 2002; Shen & Barabási, 2014; Boccaletti, Latora, Moreno, Chavez, & Hwang, 2006). Examples include social networks, biological networks, communication networks, and transportation networks. Accordingly, the research of complex networks has attracted growing interests from diverse disciplines (Wang, Xu, Wu, & Zhou, 2015; Pandey, Bhanodia, Khamparia, & Pandey, 2019; Divakaran & Mohan, 2020; Kumar, Singh, Singh, & Biswas, 2020), such as computer science, physics and economics. Due to its wide range of applications in both theory and reality (Lü Zhou, 2011; Martínez, Berzal, & Cubero, 2016), link prediction has become one of the research hotspots in the study of complex networks, and consequently a multitude of link prediction methods have been proposed with various motivations (Lü Zhou, 2011; Wang et al., 2015).

The task of link prediction is to find missing links that are not observed at present and forecast new links that will appear in the near future in a network (Lü Zhou, 2011; Liben-Nowell & Kleinberg, 2007). Therefore, many applications, such as friend suggestions in social networks (Ma, Zhou, & Zhang, 2016; Li et al., 2020), trade relations finding in international trade (Feng, Li, Qi, Guan, & Wen, 2017; Liu, Dong, Ding, Wang, & Zhang, 2020), product recommendations in e-commerce systems (Huang, Li, & Chen, 2005; Lü et al., 2012), identification of protein–protein interactions in biological networks (Cannistraci, Alanis-Lobato, & Ravasi, 2013; Guimerà & Sales-Pardo, 2009), can benefit from link prediction. Furthermore, link prediction is employed to explore network evolution mechanisms (Zhang, Lü, Wang, & Zhou, 2013; Zhang, Xu, Zhu, & Zhou, 2015) and reconstruct networks like protein interaction networks (Guimerà & Sales-Pardo, 2009; Lei & Ruan, 2013).

To date, various approaches have been designed to address the link prediction problem. Among those approaches, the so-called *similarity-based methods* are very popular for their simplicity, interpretability and decent accuracy (Li, Qian, Cheng, Ma, & Chen, 2015; Ayoub, Lotfi, El Marraki, & Hammouch, 2020; Bai, Li, Cheng, Xu, & Chen, 2018). The hypothesis of similarity-based methods is that node pairs with higher similarity scores are more likely to form links (Lü Zhou, 2011; Li et al., 2019). So, the central mission of similarity-based methods is to estimate the similarities between unconnected nodes by using the known information including node attributes and network structures. However, due to privacy policy in real scenario, node attributes are usually inaccessible and unreliable (Lü Zhou, 2011). On the other hand, estimating similarities of nodes based on network structures can produce better

* Corresponding author.
E-mail addresses: baishsh16@lzu.edu.cn (S. Bai), zhangyk2015@lzu.edu.cn (Y. Zhang), ljli@lzu.edu.cn (L. Li), dann18@lzu.edu.cn (N. Shan), chenxy@lzu.edu.cn (X. Chen).

https://doi.org/10.1016/j.eswa.2021.114973
Received 27 May 2020; Received in revised form 23 June 2020; Accepted 27 March 2021
Available online 2 April 2021

correlation with human judgments than based on node attributes (Maguitman, Menczer, Erdinc, Roinestad, & Vespignani, 2006). As a result, most of the similarity-based link prediction methods in the literature only use observed network structures, such as common neighbors (Newman, 2001; Adamic & Adar, 2003; Zhou, Lü, & Zhang, 2009), paths (Katz, 1953; Lü, Jin, & Zhou, 2009; Zhou et al., 2009; Ayoub et al., 2020), and triangles (Cannistraci et al., 2013; Wu, Lin, Wang, & Gregory, 2016; Bai et al., 2018). Besides, some studies have proposed to fuse several structural features to achieve stable performance (Li, Bai, Leng, Wang, & Chen, 2018; Ma, Bao, & Zhang, 2017; He, Liu, Hu, & Wang, 2015; Ahmad, Akhtar, Noor, & Shahnaz, 2020).

However, the majority of the existing methods focus on the link prediction problem in monolayer (or single-layer) networks, in which both nodes and links belong to one type. Some recent studies (Szell, Lambiotte, & Thurner, 2010; Cardillo et al., 2013; Nicosia, Bianconi, Latora, & Barthelemy, 2013) have shown that different types of interaction usually exist among the same individuals. For example, a group of cities can be connected by various kinds of vehicle, such as coach, train and airplane. This kind of network that is composed of different types of interaction between the same set of nodes is defined as *multi-layer network* or *multiplex network* (Boccaletti et al., 2014; Kivela et al., 2014). In a multiplex network, each type of interaction is regared as a separate layer. Moreover, several researches have reported that the topological characteristics of different layers are indeed interrelated in a multiplex network (Nicosia et al., 2013; Szell et al., 2010; Lee, Min, & Goh, 2015; Bródka, Chmiel, Magnani, & Ragozini, 2018). Therefore, the prediction accuracy in multiplex network can be improved by taking the structure information of all layers into account. This assumption has been proved in the literature (Yao et al., 2017; Najari, Salehi, Ranjbar, & Jalili, 2019; Hristova, Noulas, Brown, Musolesi, & Mascolo, 2016). For this reason, traditional link prediction methods that only use structure information derived from monolayer networks are not appropriate for multiplex networks.

In this paper, to fuse the structure information of separate layers, we consider multiplex network link prediction as a multiple-attribute decision-making (MADM) problem. The MADM problem concentrates on how to choose the preferred solutions from a set of alternatives with respect to multiple attributes (Rao & Series, 2007; Ma, Zhan, Ali, & Mehmood, 2018; Behzadian, Otaghsara, Yazdani, & Ignatius, 2012), which is employed in a wide range of applications like material selection (Chatterjee & Chakraborty, 2012; Zhang, Peng, Tian, Wang, & Xie, 2017; Mousavi-Nasab Sotoudeh-Anvari, 2017), supplier selection (Liu, Deng, & Chan, 2018; Fei, Deng, & Hu, 2019), pattern recognition (Sun, Guan, Yi, & Zhou, 2018), influence maximization (Zareie, Sheikhahmadi, & Khamforoosh, 2018; Yang, Xu, & Chen, 2018). Li et al. (2018) used MADM to solve the link prediction in monolayer networks; they regarded potential missing links as alternatives, and different similarity indexes as attributes. In this study, the unconnected node pairs in target layer are viewed as alternatives, separate layers are considered as multiple attributes, the similarity score of a node pair in one layer is the corresponding attribute value. Specifically, the technique for order preference by similarity to ideal solution (TOPSIS) (Behzadian et al., 2012) is employed to rank node pairs. In TOPSIS, a weight for each attribute is a necessary. To this end, the interlayer relevance is used to weight all attributes. To the best of our knowledge, this is the first work that regards link prediction in multiplex networks as an MADM problem. To evaluate the prediction performance of the proposed link prediction method, extensive experiments are conducted on a group of multiplex networks. The experimental results indicate that the proposed method is not sensitive to the parameter and the interlayer relevance measure, and outperforms the state-of-the-art methods for multiplex networks.

The reminder of the paper is structured as follows: First, we briefly summarize the related work in Section 2, and then describe the proposed method in Section 3. Next, the datasets and evaluation metrics are given in Section 4 and experimental results are elaborated in Section 5. Finally, Section 6 concludes this work.

2. Related work

The study on the problem of link prediction has attracted growing attention from researchers due to its numerous applications including friend suggestions in social networks, protein interaction detection in biological networks, network evolution simulation, and network reconstruction. So far, a myriad of methods have been proposed to solve this problem with various motivations. Those methods are mainly grouped into similarity-based methods and supervised learning-based methods. Some recent surveys (Kumar et al., 2020; Haghani & Keyvanpour, 2019; Pandey et al., 2019; Divakaran & Mohan, 2020) about link prediction are available for interested readers.

Similarity-based methods, which estimate the connection likelihood between two nodes by computing their similarity score, are the mainstream as aforementioned. The PA (Preferential Attachment) index (Barabási & Albert, 1999) is the simplest one, which defines the similarity score between two nodes as the product of their degrees. The CN (Common Neighbors) index (Newman, 2001) calculates the similarity between two nodes by counting the number of their shared neighbors. Different with CN, both AA (Adamic-Adar) index (Adamic & Adar, 2003) and RA (Resource Allocation) index (Zhou et al., 2009) assume the contribution of a shared neighbor to the similarity is inversely proportional to its degree. However, CN, AA and RA do not take the degrees of endpoints into account. Some indexes, such as Jaccard index (Liben-Nowell & Kleinberg, 2007), Hub Promoted index (Ravasz, Somera, Mongru, Oltvai, & Barabási, 2002) and Hub Depressed index (Zhou et al., 2009), harbor the idea that the degrees of endpoints have negative contribution to their similarity, and then they divide the number of common neighbors by different combinations of the degrees of endpoints. The LNB (Local Naïve Bayes) model (Liu, Zhang, Lü, & Zhou, 2011), MI (Mutual Information) index (Tan, Xia, & Zhu, 2014), CCLP (Clustering Coefficient for Link Prediction) index (Wu et al., 2016), CAR index (Cannistraci et al., 2013), TRA index (Bai et al., 2018), etc. gauge the similarity score between two nodes by considering the diverse clustering capabilities of common neighbors. The Katz index (Katz, 1953) counts all paths between two nodes with longer paths have smaller weights. The computation of Katz index is time-consuming since it uses all paths connecting two nodes. To reduce the running time of Katz index, the LP (Local Path) index (Lü et al., 2009; Zhou et al., 2009) proposes to use only paths with length 2 and 3, which strikes a balance between CN and Katz. Likewise, the FriendLink (Papadimitriou, Symeonidis, & Manolopoulos, 2012) uses the paths of varying length connecting two nodes to measure their similarity. Besides, some methods like SP (Significant Paths) index (Zhu, Tian, Cai, Huang, & Zhou, 2014), RP (Relative-path-based) method (Li, Huang, Liu, Huang, & Chen, 2020), and SHOPI (Significance of Higher-Order Path Index) (Kumar, Mishra, Singh, Singh, & Biswas, 2020), take into account the significance of paths while using paths of various lengths.

On the other hand, supervised learning-based methods regard link prediction as a classification problem (Wang et al., 2015; Martínez et al., 2016), and hence employ supervised learning algorithms or deep learning algorithms, such as SVM, Decision trees and deep neural networks, to predict the existence of links between nodes. To this end, features of a node pair are extracted from the network (Feyessa, Bikdash, & Lebby, 2011; Pecli, Cavalcanti, & Goldschmidt, 2018). To unfold new links for Twitter, Ahmed, Elkorany, and Bahgat (2016) proposed a supervised learning framework, in which features of node pairs contain connectivity-based, community-based, interaction-based and trust-based features, and classifiers include classical ones like *K*NN and ensemble ones like AdaBoost. Fire et al. (2013) extracted a set of topological features including vertex features, link features and path features to feed the classification models trained by one Decision Trees algorithm, *i.e.*, J48, and two ensemble learning, namely Bagging and RandomForest. Chiu and Zhan (2018) presented a deep learning-based

approach, where traditional similarity metrics, such as CN, AA, Jaccard, are selected to represent features of node pairs, and a deep learning network is trained using weak estimators. Li, Tu, and Chai (2020) proposed the ensemble-model-based link prediction algorithm, in which four kinds of similarity indexes representing different characteristics of a network are considered and several machine-learning algorithms are employed.

The aforementioned methods primarily concern monolayer networks. In fact, multiplex networks, which consist of different types of relationship between the same set of nodes (Boccaletti et al., 2014; Kivela et al., 2014), can provide more information for link prediction. In recent years, a number of researchers have explored the link prediction problem in multiplex networks. Sharma and and Singh (2015) proposed a likelihood assignment algorithm to predict links in target layer by considering whether corresponding links appear in other layers. Pujari and Kanawati (2015) proposed a supervised learning-based approach to predict the relationship of co-author in bibliographical multiplex networks. The prediction model is learned by using a set of topological features extracted from all layers (Pujari & Kanawati, 2015). Hajibagheri, Sukthankar, and Lakkaraju (2016) focused on link prediction in dynamic multiplex networks. To this end, their method first weights the target layer in each snapshot by leveraging information about cross-layer link co-occurrences. Subsequently, a collection of similarity indexes are employed to compute the similarity scores and a weighted exponentially decay model (Acar, Dunlavy, & Kolda, 2009) is used to aggregated the similarity scores of different snapshots. Finally, the rank aggregation method of Borda (Sculley, 2007) is adopted to combine the ranked lists of node pairs. Yao et al. (2017) designed a similarity-based method, which aggregates the similarity scores between two nodes in all layers via layer relevance, to predict links in multiplex networks. Mandal, Mirchev, Gramatikov, and Mishkovski (2018) predicted the link formation in a two-layer network composed of Twitter and Foursquare by using features extracted from them. Najari et al. (2019) proposed a link prediction framework, which takes both interlayer similarity and intralayer features into consideration. Samei and Jalili (2019) presented two similarity indexes in line with hyperbolic geometry of network to solve spurious and missing links in multiplex networks.

3. The proposed method

3.1. Problem description

This subsection formalizes the conception of multiplex networks, and describes the problem of link prediction in multiplex networks.

A multiplex network with l layers is represented as $\mathcal{G} = (G^1, G^2, \ldots, G^l)$, where $G^k = (V^k, E^k)$ $(0 < k < l)$ is the kth layer of $\mathcal{G}$; V^k and E^k denote the node set and link set in G^k, respectively. Here, we assume that G^k is undirected and unweighted, and all layers of $\mathcal{G}$ have the same set of nodes, *i.e.*, $V^1 = V^2 = \cdots = V^l$. Suppose the total number of distinct nodes in $\mathcal{G}$ is N, then $N = |V^k|$. Actually, there exist some nodes that do not interact with any others in some layers. In other words, the degrees of those nodes in those layers are zero. If node v_i has zero degree in layer k, it is saied inactive in the layer, otherwise, it is active (Nicosia & Latora, 2015). Let $\mathbf{A}^k = \{a_{ij}^k\}_{N\times N}$ be the adjacency matrix of graph G^k, where $a_{ij}^k = 1$ if there is a link between nodes v_i and v_j in layer k and $a_{ij}^k = 0$ otherwise.

Link prediction in multiplex networks aims to predict links in a specified layer, called *target layer*, based on the structural information of entire networks. Here, the target layer is represented as G^α, and other layers are represented as $G^{\beta_1}, G^{\beta_2}, \ldots, G^{\beta_{l-1}}$. Similar with link prediction in monolayer networks, a similarity-based method computes the similarity scores for node pairs that do not connect in target layer, and assumes that node pairs with higher scores have more chance to connect (Yao et al., 2017).

More explicitly, let U be the universal set of all node pairs in $\mathcal{G}$, and the set of nonexistent links in G^α is $U - E^\alpha$. Given two nodes $v_i, v_j \in V^\alpha$, such that $(v_i, v_j) \in U - E^\alpha$, a link prediction method assigns a similarity score to (v_i, v_j) using the structural information from not only target layer but also other layers. The node pairs that have high similarity score are likely to form links.

3.2. Motivation

A number of researches (Nicosia et al., 2013; Szell et al., 2010; Lee et al., 2015) reported that the topological structures of different layers are interrelated in a multiplex network. Accordingly, the existence of one link in some layer may be pertinent to the existence of corresponding links in other layers (Bianconi, 2013). Therefore, when performing link prediction on the target layer, the prediction performance can be enhanced by adopting the structure information of all layers. To this end, in this paper, link prediction in a multiplex network is regarded as a multiple-attribute decision-making (MADM) problem, where each layer of the network is treated as an attribute. By solving the MADM problem, the information of all layers are taken into account. The task of MADM is to pick out the preferred solutions from a set of alternatives by ranking them with respect to multiple attributes. In the MADM problem in this paper, the potential links (unconnected node pairs) in the target layer are alternatives, and the similarity score of a potential link in each layer is the corresponding attribute value.

3.3. Interlayer relevance method

So far, a number of methods have been proposed to measure the interlayer relevance between layers in a multiplex network (Najari et al., 2019; Bródka et al., 2018). This paper takes the Overlap Rate (Gemmetto & Garlaschelli, 2015), Pearson Correlation Coefficient (Yao et al., 2017), and Asymmetric Average Similarity of the Neighbors (Najari et al., 2019) into account as per their effectiveness and efficiency.

(1) Overlap Rate (OR) between two layers

Given two layers α and β in a multiplex network, and two nodes v_i and v_j in the network, if $(v_i, v_j) \in E^\alpha$ and $(v_i, v_j) \in E^\beta$, the link between v_i and v_j is called an *overlap link* with respect to layers α and β. Suppose the number of overlap links observed in layers α and β is denoted as $O_{\alpha\beta}$, the numbers of links observed in layers α and β are L^α and L^β, respectively, the overlap rate between layers α and β is defined as

$$r_{\alpha\beta}^{OR} = \frac{2O_{\alpha\beta}}{L^\alpha + L^\beta}. \tag{1}$$

$r_{\alpha\beta}^{OR}$ ranges from 0 to 1; 0 indicates that layers α and β are completely irrelevant, whilst 1 means that their links are identical.

(2) Pearson Correlation Coefficient (PCC) between two layers

This measure employs the Pearson Correlation Coefficient to estimate the relevance between two layers. To this end, Yao et al. (2017) represented the adjacency matrix of one layer as a long vector. For layer α, the long vector derived from adjacency matrix $\mathbf{A}^\alpha$ is

$$\mathbf{g}_\alpha = (a_{11}^\alpha, \cdots, a_{1N}^\alpha, a_{21}^\alpha, \cdots, a_{2N}^\alpha, \cdots, a_{N1}^\alpha, \cdots, a_{NN}^\alpha).$$

The interlayer relevance between layers α and β according to Pearson Correlation Coefficient is computed as

$$r_{\alpha\beta}^{PCC} = \frac{1}{N^2}\sum_{i,j=1}^{N}\frac{(a_{ij}^\alpha - \bar{\mathbf{g}}_\alpha)(a_{ij}^\beta - \bar{\mathbf{g}}_\beta)}{\sigma(\mathbf{g}_\alpha)\sigma(\mathbf{g}_\beta)}, \tag{2}$$

where, $\bar{\mathbf{g}}_\alpha$ is the mean of $\mathbf{g}_\alpha$ and $\sigma(\mathbf{g}_\alpha)$ is the standard deviation of $\mathbf{g}_\alpha$.

Table 1
The structural features of the eight multiplex networks. N and M denote the numbers of active nodes and edges in each layer, respectively. C is the clustering coefficient (Watts & Strogatz, 1998), ρ represents the layer density, and $\langle k\rangle$ is the average degree. H and r indicate the degree heterogeneity (Lü Zhou, 2011) and assortative coefficient (Newman, 2003), respectively.

Network	Layer	N	M	C	ρ	$\langle k\rangle$	H	r
Vicker	1	29	240	0.754	0.591	16.552	1.099	−0.161
	2	29	126	0.681	0.310	8.690	1.275	−0.152
	3	29	152	0.713	0.374	10.483	1.250	−0.110
CKM	1	215	449	0.260	0.020	4.177	1.494	−0.137
	2	231	498	0.260	0.019	4.312	1.348	−0.098
	3	228	423	0.211	0.016	3.711	1.237	0.102
Lazega	1	71	717	0.522	0.289	20.197	1.166	0.020
	2	69	399	0.498	0.170	11.565	1.314	0.079
	3	71	726	0.509	0.292	20.451	1.167	−0.079
Aarhus	1	60	139	0.673	0.109	6.433	1.213	0.005
	2	32	124	0.540	0.250	7.750	1.227	0.003
	3	25	21	0.268	0.070	1.680	1.389	0.017
	4	47	88	0.392	0.081	3.745	1.514	−0.010
	5	60	194	0.640	0.110	6.467	1.665	−0.213
Krackhardt	1	21	145	0.765	0.690	13.81	1.064	−0.265
	2	21	79	0.566	0.376	7.524	1.225	−0.224
	3	21	20	0	0.095	1.905	1.916	−0.672
Padgett	1	15	20	0.160	0.190	2.667	1.256	−0.375
	2	11	15	0.433	0.273	2.727	1.247	−0.667
Celegans	1	253	517	0.202	0.016	4.087	2.163	−0.116
	2	260	888	0.186	0.026	6.831	1.788	−0.081
	3	278	1703	0.288	0.044	12.252	1.668	−0.078
TF	1	1564	14090	0.131	0.012	18.018	3.386	−0.098
	2	1508	18471	0.344	0.016	24.498	3.775	−0.041

Evidently, $r_{\alpha\beta}^{PCC}$ is in $[-1,1]$; the more $r_{\alpha\beta}^{PCC}$ is close to 1 (or -1), the more positive (or negative) correlation that layers α and β have.

(3) Asymmetric Average Similarity of the Neighbors (AASN) between two layers

Given two layers α and β in a multiplex network, any node v_i in the network may have three types of links: links in layer α, links in layer β, and links in both layers α and β. Let $k^{\alpha}(i), k^{\beta}(i)$ and $k(i)$ be the numbers of these three types of links, respectively, the relevance of layer α with respect to layer β based on AASN is defined as

$$r_{\alpha\beta}^{AASN} = \frac{\sum_{v_i} k(i)}{\sum_{v_i} k^{\alpha}(i)}. \tag{3}$$

Unlike $r_{\alpha\beta}^{OR}$ and $r_{\alpha\beta}^{PCC}, r_{\alpha\beta}^{AASN}$ is asymmetric, which is used for predicting links in layer α considering the information available in layer β. The range of $r_{\alpha\beta}^{AASN}$ is between 0 and 1; 1 indicates that all links in layer α appear in layer β, while 0 denotes that none of links in layer α appear in layer β.

3.4. TOPSIS method

As described in the previous subsection, this work addresses the link prediction problem in multiplex networks by considering it as an MADM problem. In implementation, the TOPSIS method (Behzadian et al., 2012) is employed among numerous MADM methods because of its satisfactory performance across diverse application areas (Behzadian et al., 2012; Du, Gao, Hu, Mahadevan, & Deng, 2014). In what follows, the procedure of TOPSIS is elaborated.

Let $\mathbf{X}$ be a decision matrix with m alternatives and n attributes, which is

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix}$$

where, x_{ij} is the attribute value of the ith alternative with respect to the jth attribute. According to decision matrix $\mathbf{X}$, TOPSIS first discovers the positive ideal solution and the negative ideal solution. Subsequently, the distance from the positive ideal solution and the distance from the negative ideal solution of each alternative are computed. Then, the relative closeness of each alternative to the ideal solution is calculated

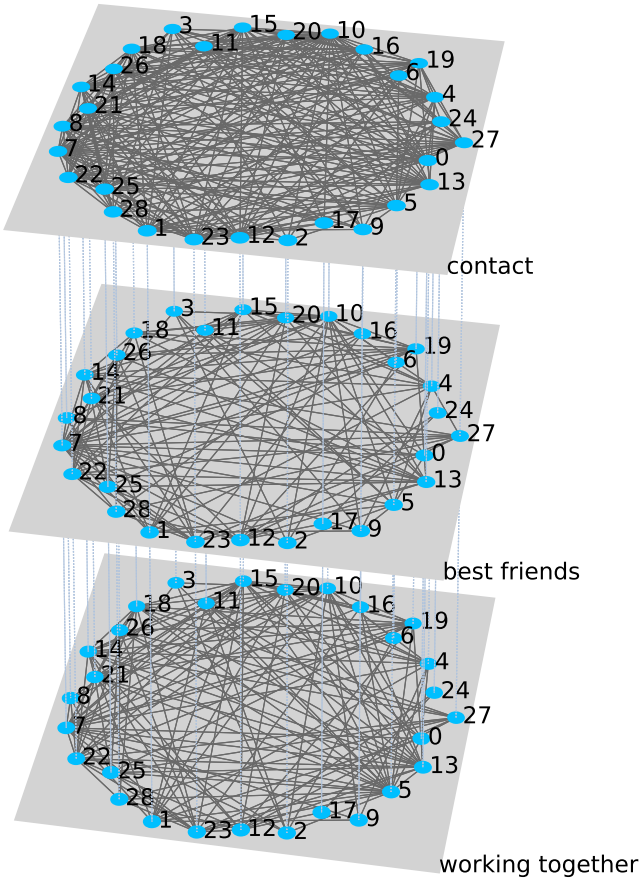

Fig. 1. The three layers of Vicker.

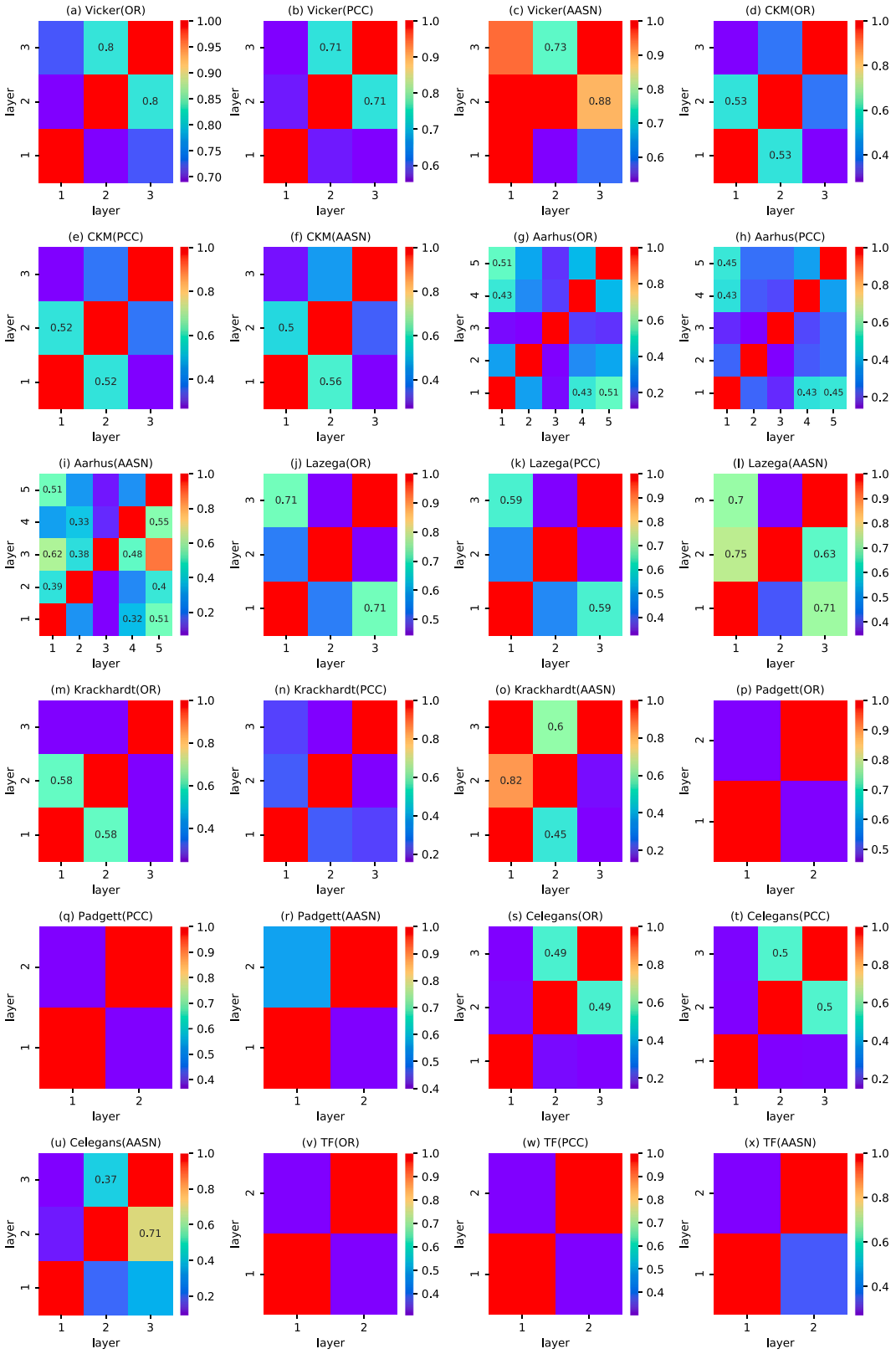

Fig. 2. Interlayer relevance between layers of the eight multiplex networks.

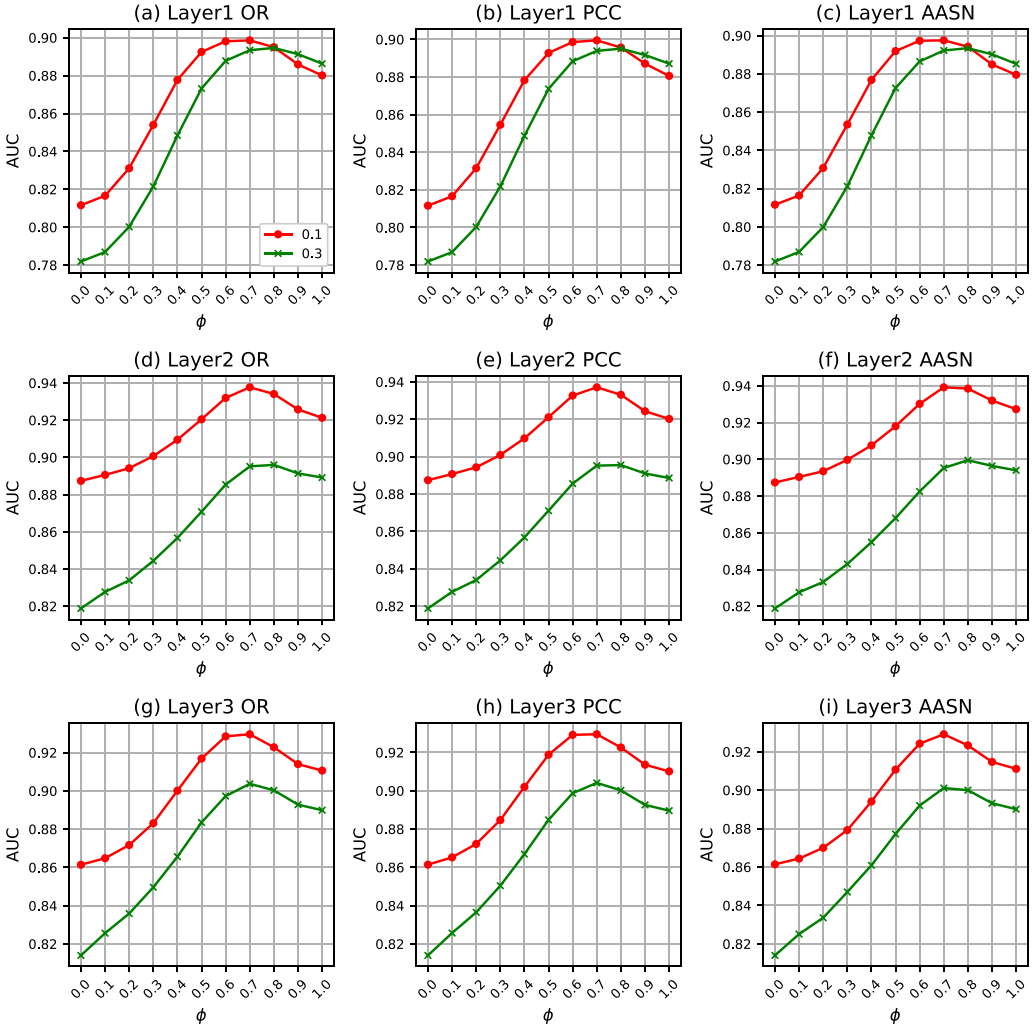

Fig. 3. AUC of the proposed method on Vicker with different values of ϕ.

by using the two distances. Superior alternatives are close to the positive ideal solution and simultaneously far from the negative ideal solution. Finally, all alternatives are ranked depending on their relative closeness. The specific steps of TOPSIS method are described at follows.

Step 1: Normalize the decision matrix.

This step normalizes the decision matrix using the vector-normalized technique. Suppose $\mathbf{Y} = \{y_{ij}\}_{m\times n}$ is the normalized matrix, where

$$y_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^{m} x_{ij}^2}}. \tag{4}$$

Step 2: Weight the decision matrix.

In general, the importance of different attributes is diverse. Therefore, TOPSIS further weights the decision matrix by considering the weight of each attribute. Let $\mathbf{Z} = \{z_{ij}\}_{m\times n}$ be the weighted matrix, such that

$$z_{ij} = w_j \times y_{ij}, \tag{5}$$

here, w_j $(j = 1, 2, \ldots, n)$ is the weight of jth attribute.

Step 3: Determine the positive ideal solution and the negative ideal solution.

In TOPSIS, the positive ideal solution, denoted as $\mathbf{S}^+$, is assumed to be the best solution available, of which each attribute value takes the optimal value among all alternatives. On the contrary, the negative ideal solution, marked as $\mathbf{S}^-$, is supposed to be the worst solution available, of which each attribute value takes the worst value among all alternatives. Their mathematical expressions read as

$$\mathbf{S}^+ = \{s_1^+, s_2^+, \ldots, s_n^+\}, \tag{6}$$

$$\mathbf{S}^- = \{s_1^-, s_2^-, \ldots, s_n^-\}. \tag{7}$$

If the jth attribute is a benefit one,

$$s_j^+ = \max_i\{z_{ij}\}, \quad s_j^- = \min_i\{z_{ij}\},$$

otherwise,

$$s_j^+ = \min_i\{z_{ij}\}, \quad s_j^- = \max_i\{z_{ij}\}.$$

Step 4: Compute the distance of each alternative from the positive ideal and the negative ideal solutions.

TOPSIS takes the attitude that a favorable alternative should be close to the positive ideal and far away from the negative ideal solutions. Therefore, this step calculates the distances of each alternative from both solutions. Let d_i^+ and d_i^- represent the distances of the ith

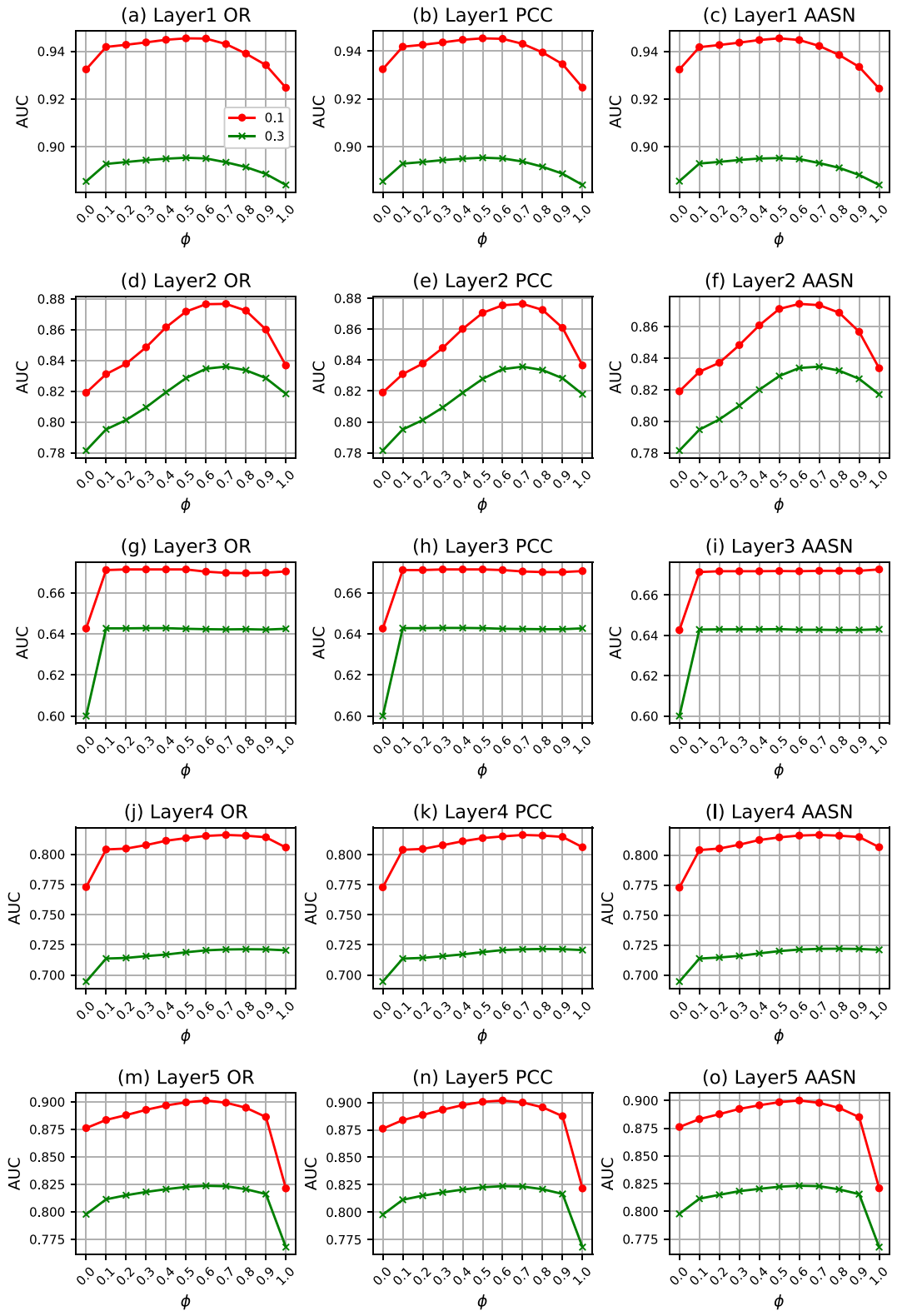

Fig. 4. AUC of the proposed method on Aarhus with different values of ϕ.

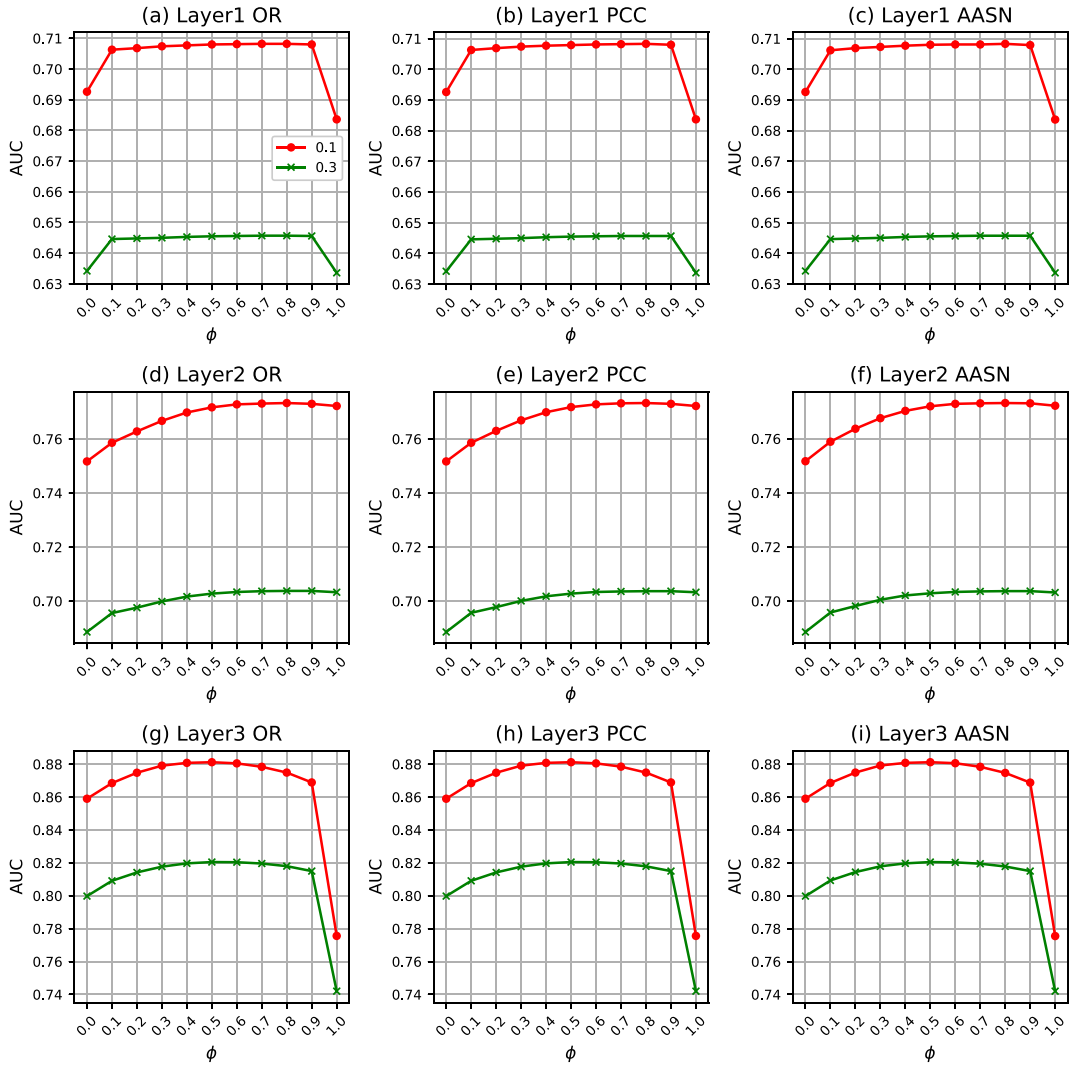

Fig. 5. AUC of the proposed method on Celegans with different values of ϕ.

alternative from the positive ideal solution and the negative ideal solution, respectively, which are computed as

$$d_i^+ = \sqrt{\sum_{j=1}^{n} (v_{ij} - s_j^+)^2}, \tag{8}$$

$$d_i^- = \sqrt{\sum_{j=1}^{n} (v_{ij} - s_j^-)^2}. \tag{9}$$

Step 5: Obtain the relative closeness of each alternative to the ideal solution.

After getting both d_i^+ and d_i^-, the relative closeness of the ith alternative to the ideal solution is defined as

$$C_i = \frac{d_i^-}{d_i^- + d_i^+}. \tag{10}$$

Higher C_i indicates that the alternative is more superior.

3.5. Algorithm of the proposed method

.

Algorithm 1. Algorithm of the proposed method.

Input: Multiplex network $\mathscr{G}$with $l$layers; target layer α; basic similarity index $\mathscr{S}$; interlayer relevance measure $\mathscr{R}$; free parameter ϕ.
Output: Ranked potential links.
1: Obtain decision matrix **X** according to $\mathscr{S}$;
2: Normalize **X** using Eq. (4), **Y**← normalized **X**;
3: **for** k←1 **to** $l-1$ **do**
4: $r_{\alpha\beta_k} \leftarrow \mathscr{R}(G^\alpha, G^{\beta_k})$;
5: **end for**
6: $w_\alpha \leftarrow 1-\phi$;
7: **for** k←1 **to** $l-1$ **do**
8: $w_{\beta_k} \leftarrow \phi \cdot \frac{r_{\alpha\beta_k}}{\sum_{\beta_k} r_{\alpha\beta_k}}$;
9: **end for**
10: Weight **Y** using Eq. (5), **Z**← weighted **Y**;
11: Determine $\mathbf{S}^+$ and $\mathbf{S}^-$ based on Eqs. (6) and (7), respectively;
12: **for each** potential link in G^α **do**
13: Compute d_i^+, d_i^- and C_i according to Eqs. ()()()(8)–(10), respectively;
14: **end for**
15: Rank all potential links according to C_i in descending order;

Algorithm 1 outlines the proposed method. Besides the target layer, the basic similarity index $\mathscr{S}$, the interlayer relevance measure $\mathscr{R}$ and the free parameter ϕ are required. $\mathscr{S}$ is used to compute the similarity of two nodes in separate layer; $\mathscr{R}$ is employed to measure the relevance between target layer and others; ϕ is adopted to adjust the weights of

target layer and others. Some key steps of the proposed method are elaborated in the following.

The first step of our method is to compute the decision matrix. As aforementioned, each latent link in target layer α is regarded as an alternative. Then, the similarity scores of all latent links in all layers are calculated using the basic similarity index $\mathscr{S}$. Here, $\mathscr{S}$ can be any similarity index designed for monolayer networks, such as CN (Newman, 2001), AA (Adamic & Adar, 2003) and RA (Zhou et al., 2009). Owning to the promising performance of RA (Zhou et al., 2009), our method picks it as the basic similarity index in default. Suppose there are m latent links in G^{α}, the similarity scores of the ith latent link in G^{α} and G^{β_k} are denoted as sim_i^{α} and $sim_i^{\beta_k}$, respectively. The decision matrix $\mathbf{X}$ is represented as

$$\mathbf{X} = \begin{bmatrix} sim_1^{\alpha} & sim_1^{\beta_1} & \cdots & sim_1^{\beta_{l-1}} \\ sim_2^{\alpha} & sim_2^{\beta_1} & \cdots & sim_2^{\beta_{l-1}} \\ \vdots & \vdots & \ddots & \vdots \\ sim_m^{\alpha} & sim_m^{\beta_1} & \cdots & sim_m^{\beta_{l-1}} \end{bmatrix}$$

In TOPSIS, each attribute need to be assigned a weight. As a consequence, the proposed method should give a weight to each layer. To do that, we consider the interlayer relevance. Due to the correlation between layers, the structural information of other layers has a certain extent of influence to the formation of links in target layer. If G^{β_k} has high correlation with G^{α}, the information of G^{β_k} has a great influence to G^{α}. So the weights of other layers are proportional to their relevance degrees to target layer. In addition, we introduce a free parameter ϕ to further adjust the importance of target layer and others. Let w_{α} and w_{β_k} respectively denote the weights of G^{α} and G^{β_k}, which are formally defined as

$$w_{\alpha} = 1 - \phi, \tag{11}$$

$$w_{\beta_k} = \phi \cdot \frac{r_{\alpha\beta_k}}{\sum_{\beta_k} r_{\alpha\beta_k}}, \tag{12}$$

where, $r_{\alpha\beta_k}$ represents the interlayer relevance between G^{α} and G^{β_k}, which is computed by using the interlayer relevance measure $\mathscr{R}$. In this work, $\mathscr{R}$ can be any one of the measures introduced in Section 3.3. The free parameter ϕ is in $[0, 1]$; $\phi = 0$ means only the information of target layer is used, on the other hand, $\phi = 1$ indicates only the information of other layers is considered.

Line 11 in Algorithm 1 is to determine the positive ideal solution $\mathbf{S}^{+}$ and the negative ideal solution $\mathbf{S}^{-}$. Since similarity scores are the higher the better, all attributes in our method are benefit attributes. Therefore, $\mathbf{S}^{+}$ consists of the maximum similarity score of each layer, whereas $\mathbf{S}^{-}$ is composed of the minimum similarity score of each layer.

Then, the distances of each potential link to both $\mathbf{S}^{+}$ and $\mathbf{S}^{-}$ can be computed according to Eqs. (8) and (9). After that, the relative closeness of a potential link to the ideal solution is calculated based on Eq. (10). Finally, all potential links are ranked according to the relative closeness in descending order. The potential links with high closeness are assumed to be missing ones.

4. Datasets and evaluation metrics

4.1. Datasets

In this paper, we use eight multiplex networks to experimentally analyze the performance of the proposed method. In what follows, the descriptions of these networks are briefly introduced.

(1) Vicker (Vickers & Chan, 1981): This multiplex network consists of 29 students and their social relationships in a school in Victoria, Australia. Three kinds of relationship are contact, best friends, and working together, respectively.
(2) CKM (Coleman, Katz, & Menzel, 1957): This is a network about the connection of physicians when they adopt a new drug. This multiplex network has three layers, which denote the relations of seeking advice, discussing cases, and friendship, respectively.
(3) Lazega (Emmanuel, 2001; Snijders, Pattison, Robins, & Handcock, 2006): This is a social network between partners and associates of a corporate law partnership. It contains three layers that correspond to co-working, friendship, and advice.
(4) Aarhus (Magnani, Micenková, & Rossi, 2013): This is a multiplex network between the employees of the Department of Computer Science at Aarhus. This network is composed of five kinds of online and offline relationships, *i.e.*, Facebook, leisure, work, co-authorship, and lunch.
(5) Krackhardt (Krackhardt, 1987): This multiplex network consists of three kinds of relationship, *i.e.*, advice, friendship, and "reports to", between managers of a high-tech company.
(6) Padgett (Padgett & Ansell, 1993): This network describes the marriage alliances and business relationship of Florentine families in the Renaissance.
(7) Celegans (Chen, Hall, & Chklovskii, 2006; De Domenico, Porter, & Arenas, 2015): This is a multiplex network of Caenorhabditis elegans connectome, in which three layers correspond to different synaptic junctions: electric, chemical monadic, and polyadic.
(8) TF (Jalili, Orouskhani, Asgari, Alipourfard, & Perc, 2017): This is a two layers social network consisting of the common users between Twitter and Foursquare.

The basic structural features of these multiplex networks are outlined in Table 1.[1] Table 1 apparently shows that the structural characteristics

Table 2
AUC of 8 networks with three interlayer relevance measures.

Network	ϕ	OR	PCC	AASN
Vicker	0	0.8535	0.8535	0.8535
	0.5	0.9100	0.9109	0.9069
	ϕ^*	0.9220	0.9220	0.9219
CKM	0	0.7179	0.7179	0.7179
	0.5	0.7289	0.7289	0.7289
	ϕ^*	0.7292	0.7292	0.7292
Lazega	0	0.8521	0.8521	0.8521
	0.5	0.8939	0.8948	0.8954
	ϕ^*	0.8978	0.8986	0.8991
Aarhus	0	0.8086	0.8086	0.8086
	0.5	0.8403	0.8403	0.8402
	ϕ^*	0.8421	0.8422	0.8416
Krackhardt	0	0.5987	0.5987	0.5987
	0.5	0.6094	0.6091	0.5994
	ϕ^*	0.6325	0.6338	0.6215
Padgett	0	0.5998	0.5998	0.5998
	0.5	0.6459	0.6459	0.6459
	ϕ^*	0.6534	0.6534	0.6534
Celegans	0	0.7678	0.7678	0.7678
	0.5	0.7870	0.7870	0.7871
	ϕ^*	0.7876	0.7876	0.7876
TF	0	0.8861	0.8861	0.8861
	0.5	0.8918	0.8918	0.8918
	ϕ^*	0.8923	0.8923	0.8923

hline

[1] The TF network was provided by the authors of Jalili et al. (2017) as a supplemental material. The others are download from https://comunelab.fbk.eu/data.php.

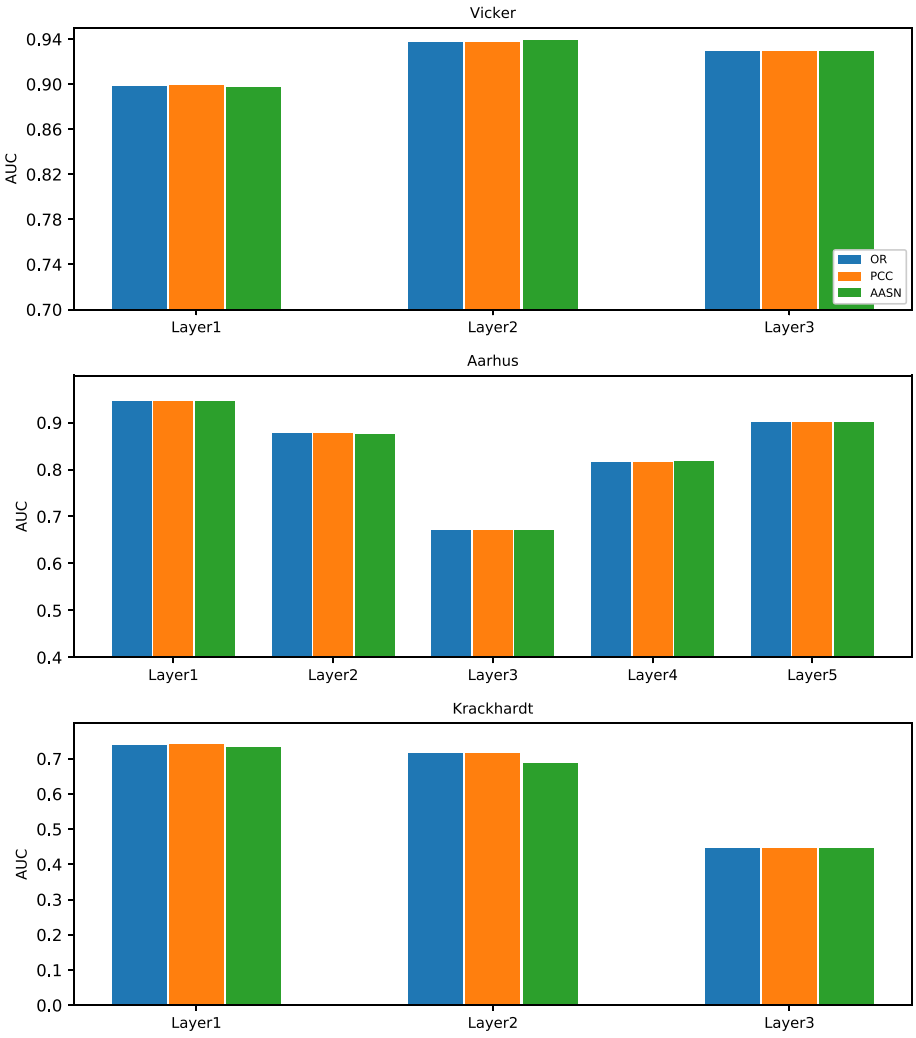

Fig. 6. AUC of each layer in Vicker, Aarhus and Krackhardt with three interlayer relevance measures.

of different networks and even different layers of the same network vary greatly. For instances, Pagett and TF contain two layers, while Aarhus consists of five layers; the clustering coefficient of the first layer of Pagett is only 0.16, whereas that of the second layer is 0.433; all layers of Vicker are dense networks, whilst two layers of TF are sparse ones. As an example, Fig. 1 graphically displays the three layers of Vicker network. We can observe from the figure that the link density of layer 1 (contact) is the highest, while that of layer 2 (best friends) is the lowest. In addition, there are many overlap links between these layers, and thence they should have high interlayer relevance.

Fig. 2 presents the structural correlation between layers of each network, which is separately measured by OR, PCC and AASN. Take Fig. 2(c) as an example to explain the content in the subfigure. This subfigure shows the interlayer relevance, calculated by AASN, between layers of Vicker. The top row indicates that the relevance scores of layer 3 with respect to layer 1, 2 and 3 are 0.93, 0.73 and 1, respectively. The elements in diagonal represent autocorrelation. Generally speaking, the network of Vicker has the highest interlayer relevance. As can be seen from Fig. 1, in Vicker, most of the links in layers 2 and 3 appear in layer 1. As a consequence, the interlayer relevance of Vicker is very high. Specially, the AASN score of layer 2 with respect to layer 1 and that of layer 3 with respect to layer 1 are up to 1 and 0.93, respectively. In other words, all observed links in layer 2 appear in layer 1, and 93% observed links in layer 3 appear in layer 1. Contrary to Vicker, Celegans and TF have low interlayer relevance. Especially, the first layer of Celegans has very few overlap links with the other two layers. In addition, the corresponding results of OR and PCC on most of these networks are consistent, but those of AASN are significantly different from them. This is because that both OR and PCC are symmetry measures, while AASN is asymmetric. Take Fig. 2(c) as an example again, it reports that $r_{21}^{AASN} = 1$ and $r_{12}^{AASN} = 0.53$. Although the links in layer 2 all appear in layer 1, only 53% of the links in layer 1 exist in layer 2. Therefore, the mutual influence of the two layers on link prediction is also different. For the network of Krackhardt, the scores of OR and PCC of layer 3 with respect to other two layers are relatively small ($r_{31}^{OR} = 0.24, r_{32}^{OR} = 0.24, r_{31}^{PCC} = 0.23, r_{32}^{PCC} = 0.16$), but those of AASN are relatively large ($r_{31}^{AASN} = 1$, $r_{32}^{AASN} = 0.6$).

4.2. Evaluation metrics

To evaluate the prediction accuracy of a link prediction method, the observed links in the target layer network G^{α}, *i.e.*, E^{α}, are randomly partitioned into two parts: training set E_t^{α} and testing set E_p^{α}, such that $E_t^{\alpha} \cup E_p^{\alpha} = E^{\alpha}$ and $E_t^{\alpha} \cap E_p^{\alpha} = \emptyset$. Here, E_t^{α} is considered as the known links used for training and E_p^{α} is regarded as the missing links used for testing. Based on this partition, two standard metrics, namely *AUC* (Lü Zhou, 2011) and *Precision* (Herlocker, Konstan, Terveen, & Riedl, 2004), are adopted in this work to quantify the prediction accuracy.

In the situation of link prediction, the value of AUC indicates the probability that a randomly selected missing link (*i.e.*, a link in E_p^{α}) has higher similarity score than a randomly selected nonexistent link (*i.e.*, a

Table 3
AUC of the proposed method with different basic similarity indexes.

Network	Layer	CN	AA	RA	Jaccard	PA
Vicker	1	0.8476	0.8807	**0.8986**	0.8914	0.7424
	2	0.8722	0.9017	0.9377	**0.9419**	0.7453
	3	0.8742	0.9103	**0.9296**	0.9252	0.7805
	Mean	0.8647	0.8976	**0.9220**	0.9195	0.7561
CKM	1	0.7140	0.7274	0.7274	**0.7313**	0.5969
	2	0.7365	0.7554	**0.7559**	0.7550	0.5836
	3	0.6827	0.7043	0.7044	**0.7062**	0.5128
	Mean	0.7111	0.7290	0.7292	**0.7308**	0.5644
Lazega	1	0.8509	0.8800	0.9068	**0.9108**	0.6849
	2	0.8927	0.9054	**0.9155**	0.9145	0.7216
	3	0.8221	0.8453	**0.8711**	0.8674	0.6856
	Mean	0.8552	0.8769	**0.8978**	0.8976	0.6974
Aarhus	1	0.9412	0.9445	**0.9454**	0.9452	0.5924
	2	0.8607	0.8744	**0.8767**	0.8718	0.6797
	3	0.6702	**0.6714**	**0.6714**	0.6702	0.4312
	4	0.8009	0.8134	**0.8160**	0.8115	0.6339
	5	0.8821	0.8971	**0.9012**	0.8699	0.7361
	Mean	0.8310	0.8402	**0.8421**	0.8337	0.6147
Krackhardt	1	0.6871	0.7086	**0.7387**	0.6703	0.7312
	2	0.6708	0.7055	**0.7139**	0.6936	0.6641
	3	**0.5000**	0.4449	0.4449	0.3907	0.4282
	Mean	0.6193	0.6197	**0.6325**	0.5849	0.6078
Padgett	1	**0.5603**	0.5529	0.5544	0.5553	0.5378
	2	0.7263	0.7519	**0.7525**	0.7312	0.5056
	Mean	0.6433	0.6524	**0.6535**	0.6433	0.5217
Celegans	1	0.6932	**0.7084**	0.7082	0.7030	0.6632
	2	0.7630	0.7711	**0.7733**	0.7698	0.7249
	3	0.8424	0.8756	**0.8812**	0.8580	0.7307
	Mean	0.7662	0.7850	**0.7876**	0.7769	0.7063
TF	1	0.8187	0.8312	0.8362	0.8206	**0.8821**
	2	0.9331	0.9430	**0.9484**	0.9329	0.9049
	Mean	0.8759	0.8871	0.8923	0.8768	**0.8935**

link in $U-E^{\alpha}$) (Lü Zhou, 2011). In implementation, if among n times independent comparisons, there are n_1 times that the link in E_p^{α} has higher score than the link in $U-E^{\alpha}$ and n_2 times that they have the same score, the value of AUC is defined as

$$AUC = \frac{n_1 + 0.5n_2}{n}. \tag{13}$$

Precision characterizes the ratio of the number of relevant items predicted to the number of items selected (Herlocker et al., 2004). In link prediction, it is the proportion of missing links correctly predicted in the top-K links predicted. The mathematical expression reads as

$$Precision = \frac{p}{K}, \tag{14}$$

where, p is the number of missing links correctly predicted.

5. Experiments

In this section, a series of experiments will be conducted to investigate the performance of the proposed method. All methods in these experiments are implemented in Python 3.7 with the graph package of NetworkX.[2]

5.1. The influence of other layers

As mentioned before, the information of other layers can enhance the prediction accuracy of link prediction in target layer. In this experiment, we analyze the degree of influence of other layers on link prediction performance by adjusting the value of free parameter ϕ. Figs. 3–5 respectively show the results on Vicker, Aarhus and Celegans in terms of AUC. The circumstances on other networks are similar with those on these three networks. So the results on other networks are not displayed. All results in these figures are the average of 20 independent realizations for each network. In each realization, 10% and 30% of observed links from each layer are randomly selected as testing sets.

The results in Fig. 3(a)–(c) are the changes of AUC on Vicker when ϕ ranges from 0 to 1 with layer 1 is the target layer. In the three subfigures, the corresponding interlayer relevance is gauged by OR, PCC and AASN. Although the relevance measures used in the three subfigures are different, the changing trends of AUC are extremely similar. With the increase of ϕ, the AUC scores gradually increase, and then decrease after reaching the peaks. Similar phenomena can also be observed from Fig. 3 (d)–(f) (layer 2 is the target layer) and Fig. 3(g)–(i) (layer 3 is the target layer). For different target layer, although the corresponding changing trends of AUC with the value of ϕ are not in full agreement, AUC will reach a peak. The corresponding value of ϕ is marked as ϕ^*. When the value of ϕ ranges from ϕ^* to 1, the AUC scores slowly decrease. This is because as the value of ϕ increases, (1) the role of the target layer itself weakens, and (2) some noise information from other layers disturbs the prediction results to some extent. In addition, Fig. 3 manifests that the corresponding scores of AUC when $\phi = 1$ are higher than those when $\phi = 0$. That is to say, the prediction accuracy using only other layer information is superior to that using only target layer information. This finding is very useful for solving the cold-start problem (Yao et al., 2017).

From the results of Aarhus and Celegans respectively presented in Figs. 4 and 5, we can see that the scores of AUC on the first, second, and fifth layers of Aarhus, and the second and third layers of Celegans show similar changing trends with those on Vicker, *i.e.*, the AUC scores gradually increase until reaching peaks, and then decrease. But the prediction accuracy of the third and fourth layers of Aarhus, and the first layer of Celegans changes slightly with the increase of the value of ϕ when $\phi > 0$. This is caused by the intrinsic characteristics of those layers. Table 1 reveals the third and fourth layers of Aarhus, and the first layer of Celegans have very low density. In comparison with other layers, these ones contain very few links. Accordingly, the corresponding testing set is very small. Although the information from other layers can enhance prediction performance, the accuracy tends to stabilize due to small testing set. Additionally, the corresponding AUC scores on the fifth layer of Aarhus, and the third layer of Celegans when $\phi = 1$ are remarkably lower than those when $\phi = 0$, which are significantly different from the phenomenon observed before. The reason is the third layer of Celegans has far more links than the other two layers, so the third layer itself can provide more information. Similar reason is for the fifth layer of Aarhus.

As a consequence, we can conclude from the above outcomes that the information from other layers can indeed improve the accuracy of link prediction compared with using only the information of target layer (when $\phi = 0$).

5.2. The influence of interlayer relevance measure

According to the results in the last subsection, the changing trends of AUC obtained by the proposed method with different interlayer relevance measures are intensely analogous. In this subsection, we further analyze the influence of these measures on the performance of prediction. Table 2 outlines the compared results of three relevance measures under the metric of AUC. In this experiment, testing set contains 10% of observed links, and the values of ϕ are set to be 0, 0.5 and ϕ^*, respectively. $\phi = 0$ indicates that the information of other layers is not used. In this case, results are independent of the interlayer relevance measure.

It can be seen from Table 2 that the prediction performance on all networks has some extent of increase after the use of the information of other layers. In addition, on most networks, the AUCs when $\phi = 0.5$ approximate to those when $\phi = \phi^*$. This discovery is conducive to our method; the value of parameter ϕ can be directly set to be 0.5 when it is

[2] http://networkx.github.io.

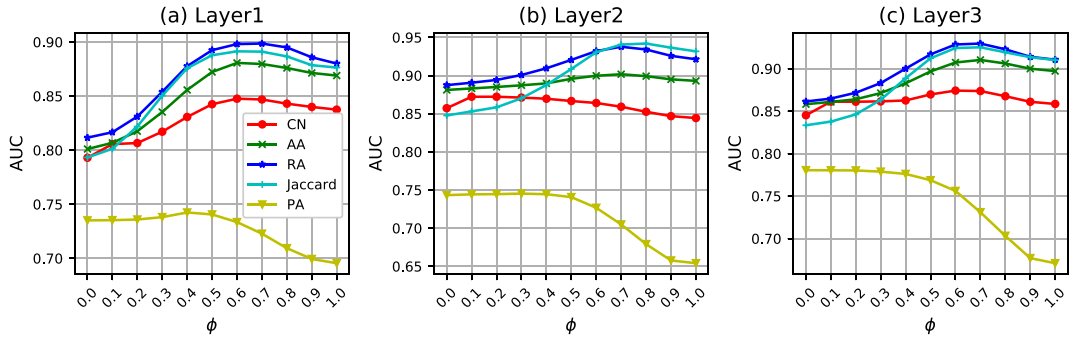

Fig. 7. AUC on Vicker with different basic similarity indexes.

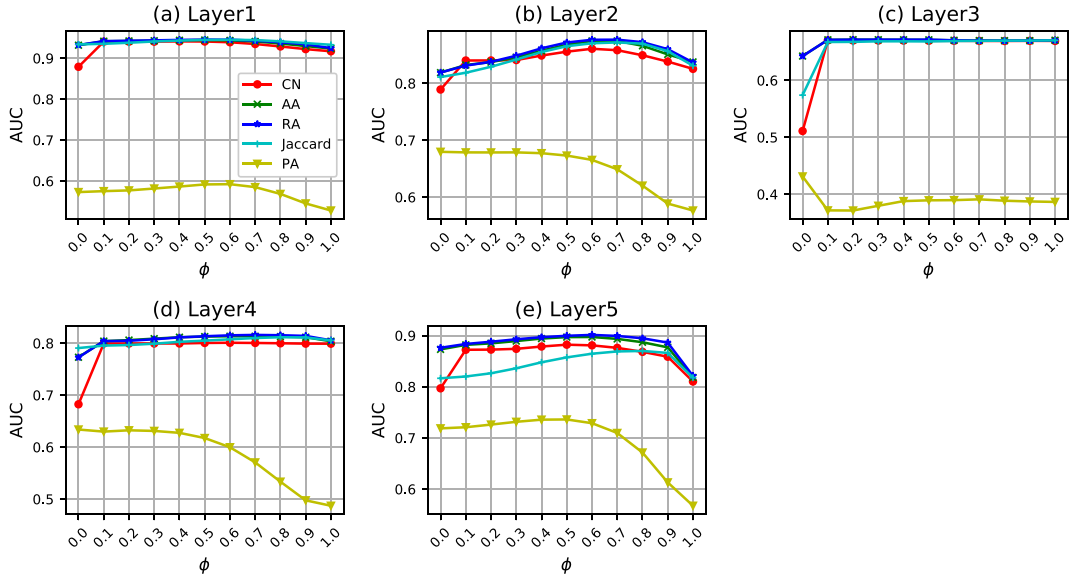

Fig. 8. AUC on Aarhus with different basic similarity indexes.

hard to determine.

Furthermore, Table 2 evidentially shows that the corresponding AUCs on each network attained with different measures when $\phi = 0.5$ or $\phi = \phi^*$ have no significant differences. Specifically, the corresponding AUCs on CKM, Padgett and TF agree on all relevance measures. Because TOPSIS ranks alternatives according to the relative closeness to the ideal solution, slight changes in weights have little effect on the predicted results if the preference relationship among attributes remains unchanged. Results in Table 2 are the average of all layers of each network. Fig. 6 further presents the results on each layer in Vicker, Aarhus and Krackhardt. Likewise, the differences in the results base on different relevance measures are very slight. As per these results, we conclude that the proposed method is robust to interlayer relevance measure.

5.3. The influence of basic similarity index

In our method, to compute the similarity score between two nodes in each layer, the index of RA is adopted in default. This subsection explores the performance of the proposed method with diverse basic similarity indexes. In this experiment, the indexes of CN (Newman, 2001), AA (Adamic & Adar, 2003), RA (Zhou et al., 2009), Jaccard (Liben-Nowell & Kleinberg, 2007) and PA (Barabási & Albert, 1999) are compared, and the results are listed in Table 3. From the previous subsection, the proposed method is not sensitive to interlayer relevance measure, thus this experiment uses only OR to estimate interlayer relevance. In Table 3, the AUC values of each index are the optimal results. In a nutshell, the results based on RA are the best, whereas those based on PA are the worst. In the viewpoint of PA, nodes with large degrees are likely to generate links. Nevertheless, in a multiplex network, nodes with large degrees in target layer may have small degrees in other layers. In consequence, the accuracy based on PA is poor on most networks. However, both of the two layers of TF have high average degree and degree heterogeneity (see Table 1). That is in line with the mechanism of PA, and thence causes the superior performance of PA on TF. Besides, the AUC values obtained by all basic similarity indexes on the third layer of Krackhardt are very low, even worse than those attained by random selection. It can be observed from Table 1 that the third layer of Krackhardt is extremely sparse. Actually, this layer has no any triadic closures. The known information is too little to link prediction.

Moreover, Figs. 7–9 show the changes of AUC based on different similarity indexes with the changes of ϕ on Vicker, Aarhus and TF, respectively. In these figures, the changing trends of CN, AA, RA and Jaccard are roughly the same, because all four indexes are common neighbors-based approaches. Nonetheless, the changing trends of PA are obviously different from those of others, due to its different mechanism.

5.4. Comparison with other multiplex link prediction methods

Finally, we compare the accuracy of the proposed method with three multiplex link prediction methods, namely NSILR (Yao et al., 2017), LPIS (Najari et al., 2019) and LAA (Sharma & and Singh, 2015). For the sake of fairness, the interlayer relevance measure is AASN, basic similarity index is RA for the proposed method, NSILR and LPIS in this

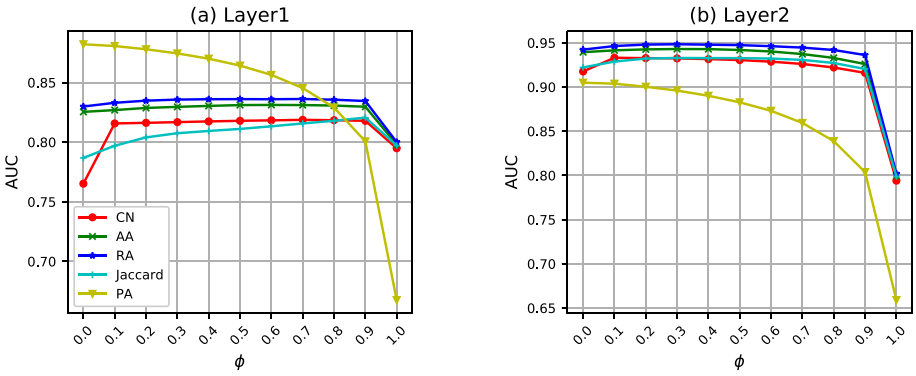

Fig. 9. AUC on TF with different basic similarity indexes.

Table 4
AUC of four multiplex link prediction methods.

Network	Layer	Proposed	NSILR	LPIS	LAA
Vicker	1	**0.8975**	0.8482	0.8119	0.7896
	2	**0.9392**	0.8852	0.8803	0.9177
	3	**0.9291**	0.8656	0.8617	0.8873
	Mean	**0.9219**	0.8663	0.8513	0.8649
CKM	1	0.7274	**0.8327**	0.7054	0.7830
	2	0.7559	**0.8618**	0.7632	0.8302
	3	0.7044	**0.8366**	0.6957	0.7431
	Mean	0.7292	**0.8437**	0.7214	0.7854
Lazega	1	**0.9074**	0.8411	0.8328	0.8401
	2	**0.9154**	0.8891	0.8846	0.7874
	3	**0.8746**	0.8166	0.8163	0.7890
	Mean	**0.8991**	0.8489	0.8446	0.8055
Aarhus	1	0.9455	**0.9564**	0.9482	0.8182
	2	**0.8744**	0.8687	0.8540	0.7685
	3	0.6713	0.9438	0.6691	**0.9622**
	4	0.8167	0.8942	0.8021	**0.8954**
	5	0.9000	**0.9174**	0.8924	0.7912
	Mean	0.8416	**0.9161**	0.8332	0.8471
Krackhardt	1	0.7329	0.7141	**0.7447**	0.6975
	2	0.6868	0.6781	**0.7092**	0.6370
	3	0.4449	0.4612	0.6528	**0.7095**
	Mean	0.6215	0.6178	**0.7022**	0.6813
Padgett	1	0.5544	0.5487	0.4634	**0.6338**
	2	0.7525	**0.9000**	0.8794	0.6250
	Mean	0.6535	**0.7244**	0.6714	0.6294
Celegans	1	0.7083	**0.8532**	0.7030	0.6661
	2	0.7732	**0.8578**	0.7611	0.8503
	3	**0.8812**	0.8690	0.8541	0.7061
	Mean	0.7876	**0.8600**	0.7727	0.7408
TF	1	0.8362	**0.8808**	0.8391	0.6747
	2	**0.9484**	0.9478	0.9426	0.6325
	Mean	0.8923	**0.9143**	0.8909	0.6536

experiment. Tables 4 and 5 give the experimental results under the metrics of AUC and Precision, respectively. These results are the average of 20 independent realizations with 10% of links from each layer are randomly selected as testing set. To compute Precision, K is set to be the size of the testing set.

Table 4 demonstrates that the proposed method significantly outnumbers baselines on Vicker and Lazega, and achieves the second AUC results on Celegans and TF. On other networks, the proposed method ranks at the third place. According to the interlayer relevance in Fig. 2, we find that the performance of the proposed method has positive correlation with the interlayer relevance. In the future, we hope to further improve the performance of the proposed method by adjusting the weight setting strategy. The mean average AUC scores of these four methods shown in Fig. 10(a) present that NSILR achieves the best AUC scores, the proposed method attains the second, while LAA obtains the worst on the whole. According to the main idea of LAA, it is only applicable to multiplex networks with high interlayer correlation. From Fig. 2, the interlayer correlation of most networks are not high. Therefore, LAA cannot achieve satisfactory results.

From the results of Precision listed in Table 5, we observe that the proposed method ranks first on six networks, second on one network, and third on one network. This indicates that the proposed method remarkably outperforms the baselines under the metric of Precision. In addition, for the layers with low clustering coefficient like layer 3 of Krackhardt, layer 1 of Padgett and layer 1 of TF, the corresponding Precision values are very small. Specially, the Precision values obtained by all methods are 0 on layer 3 of Krackhardt because it is extremely sparse. Similar with monolayer networks, these layers are prediction unfriendly (Gao, Musial, Cooper, & Tsoka, 2015). The mean average Precision values of these four methods in Fig. 10(b) manifest that the proposed method is clearly superior to others in terms of Precision.

6. Conclusion

This paper explored the problem of link prediction in multiplex

Table 5
Precision of four multiplex link prediction methods.

Network	Layer	Proposed	NSILR	LPIS	LAA
Vicker	1	**0.6521**	0.4854	0.3854	0.6000
	2	**0.4833**	0.3458	0.3583	0.4000
	3	**0.5633**	0.3667	0.3800	0.4667
	Mean	**0.5662**	0.3993	0.3746	0.4889
CKM	1	**0.1545**	0.0761	0.0648	0.1341
	2	0.2296	0.0990	0.0827	**0.2378**
	3	**0.1190**	0.0774	0.0774	0.0762
	Mean	**0.1677**	0.0842	0.0750	0.1494
Lazega	1	**0.4669**	0.3063	0.3190	0.3444
	2	0.3359	**0.3436**	0.3372	0.1269
	3	**0.3688**	0.2681	0.2674	0.2326
	Mean	**0.3905**	0.3060	0.3079	0.2346
Aarhus	1	**0.4447**	0.4368	0.4368	0.2526
	2	0.3667	**0.3875**	0.3042	0.3458
	3	0.2500	0.3250	0.2000	**0.3500**
	4	**0.2750**	0.1750	0.1688	0.1750
	5	0.2763	**0.3053**	0.2684	0.2105
	Mean	0.3225	**0.3259**	0.2756	0.2668
Krackhardt	1	**0.4250**	0.4179	0.4179	0.2607
	2	0.1714	**0.2429**	0.2357	0.0714
	3	0	0	0	0
	Mean	0.1988	**0.2203**	0.2179	0.1107
Padgett	1	0.0500	0	0	**0.0750**
	2	**0.3000**	0.1000	0.1000	0
	Mean	**0.1750**	0.0500	0.0500	0.0375
Celegans	1	**0.1029**	0.0784	0.0647	0.0275
	2	0.1011	0.0506	0.046	**0.1364**
	3	**0.2621**	0.0918	0.0918	0.2250
	Mean	**0.1554**	0.0736	0.0675	0.1296
TF	1	0.0258	0.0248	0.0280	**0.0412**
	2	**0.2399**	0.2301	0.2283	0.0528
	Mean	**0.1329**	0.1275	0.1282	0.0470

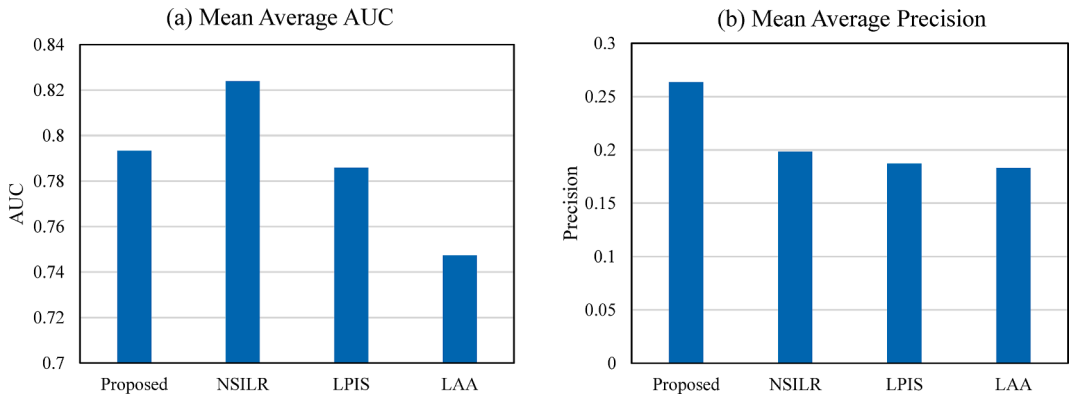

Fig. 10. The mean average AUC and Precision of four methods on eight networks.

networks. To resolve this problem, we regarded it as a multiple-attribute decision-making problem, by which the structural information from all layers were integrated. Specifically, the decision matrix in the proposed method consists of similarity scores of all potential links computed in each layer. To weight each attribute, the interlayer relevance between layers are employed to gauge the weight for each layer with the help of a free parameter.

To evaluate the performance of the proposed method, extensive experiments were implemented on eight multiplex networks under the metric of AUC and Precision. Experimental results show that (1) the information from other layers is really conducive to the accuracy of link prediction; (2) the proposed method is robust to the parameter and interlayer relevance measure; and (3) the proposed method achieves promising prediction results, especially in terms of Precision.

In this paper, the correlation between layers were measured in terms of three interlayer relevance methods that mainly consider the overlap links. Since one basic idea in the proposed method is that high correlated layers have great contribution to the formation of links in target layer, the proposed method assigns large weights to high correlated layers. Correspondingly, the influence of the uncorrelated and anti-correlated layers are weakened. However, some anti-correlation information could also facilitate the link prediction in target layer if the information is reasonably used. Take two layers that have no overlap links as an example, a link is more likely to be predicted in one layer if the corresponding link is absent in another layer. Nevertheless, the effect of anti-correlation information is out of the scope of the proposed method. But, in the future work, it is an interesting topic.

CRediT authorship contribution statement

Shenshen Bai: Methodology, Software, Writing - original draft. **Yakun Zhang:** Methodology, Software, Resources. **Longjie Li:** Conceptualization, Formal analysis, Writing - original draft, Writing - review & editing, Funding acquisition. **Na Shan:** Resources, Formal analysis. **Xiaoyun Chen:** Investigation, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No. 61602225) and the Fundamental Research Funds for the Central Universities (No. lzujbky-2019-90).

References

Acar, E., Dunlavy, D. M., & Kolda, T. G. (2009). Link prediction on evolving data using matrix and tensor factorizations. In *2009 IEEE international conference on data mining workshops* (pp. 262–269). Miami, FL, USA: IEEE. https://doi.org/10.1109/ICDMW.2009.54.

Adamic, L. A., & Adar, E. (2003). Friends and neighbors on the Web. *Social Networks, 25*, 211–230. https://doi.org/10.1016/S0378-8733(03)00009-1

Ahmad, I., Akhtar, M. U., Noor, S., & Shahnaz, A. (2020). Missing link prediction using common neighbor and centrality based parameterized algorithm. *Scientific Reports, 10*, 364. https://doi.org/10.1038/s41598-019-57304-y

Ahmed, C., Elkorany, A., & Bahgat, R. (2016). A supervised learning approach to link prediction in Twitter. *Social Network Analysis and Mining, 6*, 1–11. https://doi.org/10.1007/s13278-016-0333-1

Albert, R., & Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics, 74*, 47–97. https://doi.org/10.1103/RevModPhys.74.47

Ayoub, J., Lotfi, D., El Marraki, M., & Hammouch, A. (2020). Accurate link prediction method based on path length between a pair of unlinked nodes and their degree. *Social Network Analysis and Mining, 10*, 9. https://doi.org/10.1007/s13278-019-0618-2

Bai, S., Li, L., Cheng, J., Xu, S., & Chen, X. (2018). Predicting missing links based on a new triangle structure. *Complexity, 2018*, 1–11. https://doi.org/10.1155/2018/7312603

Barabási, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *Science, 286*, 509–512. https://doi.org/10.1126/science.286.5439.509

Behzadian, M., Otaghsara, S. K., Yazdani, M., & Ignatius, J. (2012). A state-of-the-art survey of TOPSIS applications. *Expert Systems with Applications, 39*, 13051–13069. https://doi.org/10.1016/j.eswa.2012.05.056

Bianconi, G. (2013). Statistical mechanics of multiplex networks: Entropy and overlap. *Physical Review E, 87*, Article 062806. https://doi.org/10.1103/PhysRevE.87.062806

Boccaletti, S., Bianconi, G., Criado, R., del Genio, C., Gómez-Gardeñes, J., Romance, M., Sendiña-Nadal, I., Wang, Z., & Zanin, M. (2014). The structure and dynamics of multilayer networks. *Physics Reports, 544*, 1–122. https://doi.org/10.1016/j.physrep.2014.07.001

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., & Hwang, D.-U. (2006). Complex networks: Structure and dynamics. *Physics Reports, 424*, 175–308. https://doi.org/10.1016/j.physrep.2005.10.009

Bródka, P., Chmiel, A., Magnani, M., & Ragozini, G. (2018). Quantifying layer similarity in multiplex networks: A systematic study. *Royal Society Open Science, 5*, Article 171747. https://doi.org/10.1098/rsos.171747

Cannistraci, C. V., Alanis-Lobato, G. G., & Ravasi, T. (2013). From link-prediction in brain connectomes and protein interactomes to the local-community-paradigm in complex networks. *Scientific Reports, 3*, 1613. https://doi.org/10.1038/srep01613

Cardillo, A., Gómez-Gardeñes, J., Zanin, M., Romance, M., Papo, D., del Pozo, F., & Boccaletti, S. (2013). Emergence of network features from multiplexity. *Scientific Reports, 3*, 1344. https://doi.org/10.1038/srep01344

Chatterjee, P., & Chakraborty, S. (2012). Material selection using preferential ranking methods. *Materials and Design, 35*, 384–393. https://doi.org/10.1016/j.matdes.2011.09.027

Chen, B. L., Hall, D. H., & Chklovskii, D. B. (2006). Wiring optimization can relate neuronal structure and function. *Proceedings of the National Academy of Sciences, 103*, 4723–4728. https://doi.org/10.1073/pnas.0506806103

Chiu, C., & Zhan, J. (2018). Deep learning for link prediction in dynamic networks using weak estimators. *IEEE Access, 6*, 35937–35945. https://doi.org/10.1109/ACCESS.2018.2845876

Coleman, J., Katz, E., & Menzel, H. (1957). The diffusion of an innovation among physicians. *Sociometry, 20*, 253. https://doi.org/10.2307/2785979

De Domenico, M., Porter, M. A. & Arenas, A. (2015). MuxViz: A tool for multilayer analysis and visualization of networks. Journal of Complex Networks, 3, 159–176. doi: 10.1093/comnet/cnu038.

Divakaran, A., & Mohan, A. (2020). Temporal link prediction: A survey. *New Generation Computing, 38*, 213–258. https://doi.org/10.1007/s00354-019-00065-z

Du, Y., Gao, C., Hu, Y., Mahadevan, S., & Deng, Y. (2014). A new method of identifying influential nodes in complex networks based on TOPSIS. *Physica A, 399*, 57–69. https://doi.org/10.1016/j.physa.2013.12.031

Emmanuel, L. (2001). *The collegial phenomenon: The social mechanisms of cooperation among peers in a corporate law partnership*. Oxford University Press.

Fei, L., Deng, Y., & Hu, Y. (2019). DS-VIKOR: A new multi-criteria decision-making method for supplier selection. *International Journal of Fuzzy Systems, 21*, 157–175. https://doi.org/10.1007/s40815-018-0543-y

Feng, S., Li, H., Qi, Y., Guan, Q., & Wen, S. (2017). Who will build new trade relations? Finding potential relations in international liquefied natural gas trade. *Energy, 141*, 1226–1238. https://doi.org/10.1016/j.energy.2017.09.030

Feyessa, T., Bikdash, M. & Lebby, G. (2011). Node-pair feature extraction for link prediction. In 2011 IEEE third int'l conference on privacy, security, risk and trust and 2011 IEEE third int'l conference on social computing (pp. 1421–1424). IEEE. doi: 10.1109/PASSAT/SocialCom.2011.244.

Fire, M., Tenenboim-Chekina, L., Puzis, R., Lesser, O., Rokach, L., & Elovici, Y. (2013). Computationally efficient link prediction in a variety of social networks. *ACM Transactions on Intelligent Systems and Technology, 5*. https://doi.org/10.1145/2542182.2542192

Gao, F., Musial, K., Cooper, C., & Tsoka, S. (2015). Link prediction methods and their accuracy for different social networks and network metrics. *Scientific Programming, 2015*, 1–13. https://doi.org/10.1155/2015/172879

Gemmetto, V., & Garlaschelli, D. (2015). Multiplexity versus correlation: The role of local constraints in real multiplexes. *Scientific Reports, 5*, 9120. https://doi.org/10.1038/srep09120

Guimerà, R., & Sales-Pardo, M. (2009). Missing and spurious interactions and the reconstruction of complex networks. *Proceedings of the National Academy of Sciences, 106*, 22073–22078. https://doi.org/10.1073/pnas.0908366106

Haghani, S., & Keyvanpour, M. R. (2019). A systemic analysis of link prediction in social network. *Artificial Intelligence Review, 52*, 1961–1995. https://doi.org/10.1007/s10462-017-9590-2

Hajibagheri, A., Sukthankar, G., & Lakkaraju, K. (2016). A holistic approach for link prediction in multiplex networks. In *Social informatics* (pp. 55–70). Springer International Publishing. https://doi.org/10.1007/978-3-319-47874-6_5.

He, Y.-L., Liu, J. N., Hu, Y.-X., & Wang, X.-Z. (2015). OWA operator based link prediction ensemble for social network. *Expert Systems with Applications, 42*, 21–50. https://doi.org/10.1016/j.eswa.2014.07.018

Herlocker, J. L., Konstan, J. A., Terveen, L. G., & Riedl, J. T. (2004). Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems (TOIS), 22*, 5–53. https://doi.org/10.1145/963770.963772

Hristova, D., Noulas, A., Brown, C., Musolesi, M., & Mascolo, C. (2016). A multilayer approach to multiplexity and link prediction in online geo-social networks. *EPJ Data Science, 5*, 24. https://doi.org/10.1140/epjds/s13688-016-0087-z

Huang, Z., Li, X., & Chen, H. (2005). Link prediction approach to collaborative filtering. In *Proceedings of the ACM/IEEE joint conference on digital libraries JCDL '05* (pp. 141–142). New York, NY, USA: ACM. https://doi.org/10.1145/1065385.1065415.

Jalili, M., Orouskhani, Y., Asgari, M., Alipourfard, N., & Perc, M. (2017). Link prediction in multiplex online social networks. *Royal Society Open Science, 4*, Article 160863. https://doi.org/10.1098/rsos.160863

Katz, L. (1953). A new status index derived from sociometric analysis. *Psychometrika, 18*, 39–43. https://doi.org/10.1007/BF02289026

Kivela, M., Arenas, A., Barthelemy, M., Gleeson, J. P., Moreno, Y., & Porter, M. A. (2014). Multilayer networks. *Journal of Complex Networks, 2*, 203–271. https://doi.org/10.1093/comnet/cnu016

Krackhardt, D. (1987). Cognitive social structures. *Social Networks, 9*, 109–134. https://doi.org/10.1016/0378-8733(87)90009-8

Kumar, A., Mishra, S., Singh, S. S., Singh, K., & Biswas, B. (2020). Link prediction in complex networks based on Significance of Higher-Order Path Index (SHOPI). *Physica A, 545*, Article 123790. https://doi.org/10.1016/j.physa.2019.123790

Kumar, A., Singh, S. S., Singh, K., & Biswas, B. (2020). Link prediction techniques, applications, and performance: A survey. *Physica A*, 124289. https://doi.org/10.1016/j.physa.2020.124289

Lee, K.-M., Min, B., & Goh, K.-I. (2015). Towards real-world complexity: An introduction to multiplex networks. *The European Physical Journal B, 88*, 48. https://doi.org/10.1140/epjb/e2015-50742-1

Lei, C., & Ruan, J. (2013). A novel link prediction algorithm for reconstructing protein–protein interaction networks by topological similarity. *Bioinformatics, 29*, 355–364. https://doi.org/10.1093/bioinformatics/bts688

Li, K., Tu, L., & Chai, L. (2020). Ensemble-model-based link prediction of complex networks. *Computer Networks, 166*, Article 106978. https://doi.org/10.1016/j.comnet.2019.106978

Li, L., Bai, S., Leng, M., Wang, L., & Chen, X. (2018). Finding missing links in complex networks: A multiple-attribute decision-making method. *Complexity, 2018*, 1–16. https://doi.org/10.1155/2018/3579758

Li, L., Fang, S., Bai, S., Xu, S., Cheng, J., & Chen, X. (2019). Effective link prediction based on community relationship strength. *IEEE Access, 7*, 43233–43248. https://doi.org/10.1109/ACCESS.2019.2908208

Li, L., Qian, L., Cheng, J., Ma, M., & Chen, X. (2015). Accurate similarity index based on the contributions of paths and end nodes for link prediction. *Journal of Information Science, 41*, 167–177. https://doi.org/10.1177/0165551514560121

Li, S., Huang, J., Liu, J., Huang, T., & Chen, H. (2020). Relative-path-based algorithm for link prediction on complex networks using a basic similarity factor. *Chaos, 30*. https://doi.org/10.1063/1.5094448

Li, S., Song, X., Lu, H., Zeng, L., Shi, M., & Liu, F. (2020). Friend recommendation for cross marketing in online brand community based on intelligent attention allocation link prediction algorithm. *Expert Systems with Applications, 139*, Article 112839. https://doi.org/10.1016/j.eswa.2019.112839

Liben-Nowell, D., & Kleinberg, J. (2007). The link-prediction problem for social networks. *Journal of the American Society for Information Science and Technology, 58*, 1019–1031. https://doi.org/10.1002/asi.20591

Liu, S., Dong, Z., Ding, C., Wang, T., & Zhang, Y. (2020). Do you need cobalt ore? Estimating potential trade relations through link prediction. *Resources Policy, 66*, Article 101632. https://doi.org/10.1016/j.resourpol.2020.101632

Liu, T., Deng, Y., & Chan, F. (2018). Evidential supplier selection based on DEMATEL and game theory. *International Journal of Fuzzy Systems, 20*, 1321–1333. https://doi.org/10.1007/s40815-017-0400-4

Liu, Z., Zhang, Q.-M., Lü, L., & Zhou, T. (2011). Link prediction in complex networks: A local naïve Bayes model. *EPL (Europhysics Letters), 96*, 48007. https://doi.org/10.1209/0295-5075/96/48007

Lü, L., Jin, C. H., & Zhou, T. (2009). Similarity index based on local paths for link prediction of complex networks. *Physical Review E, 80*, Article 046122. https://doi.org/10.1103/PhysRevE.80.046122

Lü, L., Medo, M., Yeung, C. H., Zhang, Y.-C., Zhang, Z.-K., & Zhou, T. (2012). Recommender systems. *Physics Reports, 519*, 1–49. https://doi.org/10.1016/j.physrep.2012.02.006

Lü, L., & Zhou, T. (2011). Link prediction in complex networks: A survey. *Physica A, 390*, 1150–1170. https://doi.org/10.1016/j.physa.2010.11.027

Ma, C., Bao, Z.-K., & Zhang, H.-F. (2017). Improving link prediction in complex networks by adaptively exploiting multiple structural features of networks. *Physics Letters A, 381*, 3369–3376. https://doi.org/10.1016/j.physleta.2017.08.047

Ma, C., Zhou, T., & Zhang, H. F. (2016). Playing the role of weak clique property in link prediction: A friend recommendation model. *Scientific Reports, 6*, 1–12. https://doi.org/10.1038/srep30098

Ma, X., Zhan, J., Ali, M. I., & Mehmood, N. (2018). A survey of decision making methods based on two classes of hybrid soft set models. *Artificial Intelligence Review, 49*, 511–529. https://doi.org/10.1007/s10462-016-9534-2

Magnani, M., Micenková, B. & Rossi, L. (2013). Combinatorial analysis of multiple networks. CoRR, abs/1303.4986.

Maguitman, A. G., Menczer, F., Erdinc, F., Roinestad, H., & Vespignani, A. (2006). Algorithmic computation and approximation of semantic similarity. *World Wide Web, 9*, 431–456. https://doi.org/10.1007/s11280-006-8562-2

Mandal, H., Mirchev, M., Gramatikov, S., & Mishkovski, I. (2018). Multilayer Link Prediction in Online Social Networks. In 2018 26th Telecommunications Forum (TELFOR) (pp. 1–4). IEEE. doi: 10.1109/TELFOR.2018.8612122.

Martínez, V., Berzal, F., & Cubero, J.-C. (2016). A survey of link prediction in complex networks. *ACM Computing Surveys, 49*, 1–33. https://doi.org/10.1145/3012704

Mousavi-Nasab, S. H., & Sotoudeh-Anvari, A. (2017). A comprehensive MCDM-based approach using TOPSIS, COPRAS and DEA as an auxiliary tool for material selection problems. *Materials & Design, 121*, 237–253. https://doi.org/10.1016/j.matdes.2017.02.041

Najari, S., Salehi, M., Ranjbar, V., & Jalili, M. (2019). Link prediction in multiplex networks based on interlayer similarity. *Physica A, 536*, Article 120978. https://doi.org/10.1016/j.physa.2019.04.214

Newman, M. E. (2001). Clustering and preferential attachment in growing networks. *Physical Review E, 64*, 4. https://doi.org/10.1103/PhysRevE.64.025102

Newman, M. E. (2003). Mixing patterns in networks. *Physical Review E, 67*, 13. https://doi.org/10.1103/PhysRevE.67.026126

Nicosia, V., Bianconi, G., Latora, V., & Barthelemy, M. (2013). Growing multiplex networks. *Physical Review Letters, 111*, Article 058701. https://doi.org/10.1103/PhysRevLett.111.058701

Nicosia, V., & Latora, V. (2015). Measuring and modeling correlations in multiplex networks. *Physical Review E, 92*, Article 032805. https://doi.org/10.1103/PhysRevE.92.032805

Padgett, J. F., & Ansell, C. K. (1993). Robust action and the rise of the medici, 1400–1434. *American Journal of Sociology, 98*, 1259–1319. https://doi.org/10.1086/230190

Pandey, B., Bhanodia, P. K., Khamparia, A., & Pandey, D. K. (2019). A comprehensive survey of edge prediction in social networks: techniques, parameters and challenges. *Expert Systems with Applications, 124*, 164–181. https://doi.org/10.1016/J.ESWA.2019.01.040

Papadimitriou, A., Symeonidis, P., & Manolopoulos, Y. (2012). Fast and accurate link prediction in social networking systems. *Journal of Systems and Software, 85*, 2119–2132. https://doi.org/10.1016/j.jss.2012.04.019

Pecli, A., Cavalcanti, M. C., & Goldschmidt, R. (2018). Automatic feature selection for supervised learning in link prediction applications: A comparative study. *Knowledge and Information Systems, 56*, 85–121. https://doi.org/10.1007/s10115-017-1121-6

Pujari, M., & Kanawati, R. (2015). Link prediction in multiplex networks. *Networks and Heterogeneous Media, 10*, 17–35. https://doi.org/10.3934/nhm.2015.10.17

Rao, R. V. & Series, S. (2007). Introduction to multiple attribute decision-making (MADM) methods. In Decision making in the manufacturing environment (pp. 27–41). London: Springer, London. doi: 10.1007/978-1-84628-819-7_3.

Ravasz, E., Somera, A. L., Mongru, D. A., Oltvai, Z. N., & Barabási, A. L. (2002). Hierarchical organization of modularity in metabolic networks. *Science, 297*, 1551–1555. https://doi.org/10.1126/science.1073374

Samei, Z., & Jalili, M. (2019). Application of hyperbolic geometry in link prediction of multiplex networks. *Scientific Reports, 9*, 12604. https://doi.org/10.1038/s41598-019-49001-7

Sculley, D. (2007). Rank aggregation for similar items. In *Proceedings of the 2007 SIAM international conference on data mining* (pp. 587–592). Philadelphia, PA: Society for Industrial and Applied Mathematics. https://doi.org/10.1137/1.9781611972771.66.

Sharma, S. & Singh, A. (2015). An efficient method for link prediction in complex multiplex networks. In 2015 11th International conference on signal-image technology & internet-based systems (SITIS) (pp. 453–459). IEEE. doi: 10.1109/SITIS.2015.93.

Shen, H.-W., & Barabási, A.-L. (2014). Collective credit allocation in science. *Proceedings of the National Academy of Sciences, 111*, 12325–12330. https://doi.org/10.1073/pnas.1401992111

Snijders, T. A. B., Pattison, P. E., Robins, G. L., & Handcock, M. S. (2006). New specifications for exponential random graph models. *Sociological Methodology, 36*, 99–153. https://doi.org/10.1111/j.1467-9531.2006.00176.x

Sun, G., Guan, X., Yi, X., & Zhou, Z. (2018). Grey relational analysis between hesitant fuzzy sets with applications to pattern recognition. *Expert Systems with Applications, 92*, 521–532. https://doi.org/10.1016/j.eswa.2017.09.048

Szell, M., Lambiotte, R., & Thurner, S. (2010). Multirelational organization of large-scale social networks in an online world. *Proceedings of the National Academy of Sciences, 107*, 13636–13641. https://doi.org/10.1073/pnas.1004008107

Tan, F., Xia, Y., & Zhu, B. (2014). Link prediction in complex networks: A mutual information perspective. *PLoS One, 9*, Article e107056. https://doi.org/10.1371/journal.pone.0107056

Vickers, M., & Chan, S. (1981). *Representing classroom social structure*. Melbourne: Victoria Institute of Secondary Education.

Wang, P., Xu, B., Wu, Y., & Zhou, X. (2015). Link prediction in social networks: the state-of-the-art. *Science China Information Sciences, 58*, 1–38. https://doi.org/10.1007/s11432-014-5237-y

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of small-world networks. *Nature, 393*, 440–442. https://doi.org/10.1038/30918

Wu, Z., Lin, Y., Wang, J., & Gregory, S. (2016). Link prediction with node clustering coefficient. *Physica A, 452*, 1–8. https://doi.org/10.1016/j.physa.2016.01.038

Yang, P., Xu, G., & Chen, H. (2018). Multi-attribute ranking method for identifying key nodes in complex networks based on GRA. *International Journal of Modern Physics B, 32*, 1850363. https://doi.org/10.1142/S0217979218503630

Yao, Y., Zhang, R., Yang, F., Yuan, Y., Sun, Q., Qiu, Y., & Hu, R. (2017). Link prediction via layer relevance of multiplex networks. *International Journal of Modern Physics C, 28*, 1750101. https://doi.org/10.1142/S0129183117501017

Zareie, A., Sheikhahmadi, A., & Khamforoosh, K. (2018). Influence maximization in social networks based on TOPSIS. *Expert Systems with Applications, 108*, 96–107. https://doi.org/10.1016/j.eswa.2018.05.001

Zhang, H., Peng, Y., Tian, G., Wang, D., & Xie, P. (2017). Green material selection for sustainability: A hybrid MCDM approach. *Plos One, 12*, Article e0177578. https://doi.org/10.1371/journal.pone.0177578

Zhang, Q.-M., Lü, L., Wang, W.-Q., & Zhou, T. (2013). Potential theory for directed networks. *PLoS One, 8*, Article e55437. https://doi.org/10.1371/journal.pone.0055437

Zhang, Q.-M., Xu, X.-K., Zhu, Y.-X., & Zhou, T. (2015). Measuring multiple evolution mechanisms of complex networks. *Scientific Reports, 5*, 10350. https://doi.org/10.1038/srep10350

Zhou, T., Lü, L., & Zhang, Y.-C. (2009). Predicting missing links via local information. *The European Physical Journal B, 71*, 623–630. https://doi.org/10.1140/epjb/e2009-00335-8

Zhu, X., Tian, H., Cai, S., Huang, J., & Zhou, T. (2014). Predicting missing links via significant paths. *EPL (Europhysics Letters), 106*, 18008. https://doi.org/10.1209/0295-5075/106/18008