International Journal of Modern Physics C
Vol. 28, No. 8 (2017) 1750101 (24 pages)

DOI: 10.1142/S0129183117501017



# Link prediction via layer relevance of multiplex networks

Yabing Yao*, Ruisheng Zhang†, Fan Yang‡, Yongna Yuan§,
Qingshuang Sun¶, Yu Qiu‖ and Rongjing Hu**

*School of Information Science and Engineering*
*Lanzhou University*
*Lanzhou, Gansu 730000, P. R. China*
**yaoyb14@lzu.edu.cn*
†*zhangrs@lzu.edu.cn*
‡*fanyang2014@lzu.edu.cn*
§*yuanyn@lzu.edu.cn*
¶*sunqsh2015@lzu.edu.cn*
‖*qiuy15@lzu.edu.cn*
***hurj@lzu.edu.cn*

Received 25 April 2017
Accepted 12 July 2017
Published 18 August 2017



In complex networks, the existing link prediction methods primarily focus on the internal structural information derived from single-layer networks. However, the role of interlayer information is hardly recognized in multiplex networks, which provide more diverse structural features than single-layer networks. Actually, the structural properties and functions of one layer can affect that of other layers in multiplex networks. In this paper, the effect of interlayer structural properties on the link prediction performance is investigated in multiplex networks. By utilizing the intralayer and interlayer information, we propose a novel "Node Similarity Index" based on "Layer Relevance" (NSILR) of multiplex network for link prediction. The performance of NSILR index is validated on each layer of seven multiplex networks in real-world systems. Experimental results show that the NSILR index can significantly improve the prediction performance compared with the traditional methods, which only consider the intralayer information. Furthermore, the more relevant the layers are, the higher the performance is enhanced.

*Keywords*: Complex networks; layer relevance; link prediction; multiplex networks; node similarity.

PACS Nos.: 89.20.Ff, 89.75.Fb.

## 1. Introduction

In the real world, many systems can be regarded as networks in which nodes stand for individuals and links reflect the interactions between individuals.[1] As one of the

† Corresponding author.

fundamental problems in complex networks, link prediction seeks to estimate the connecting probability of a nonexistent or missing link between a pair of nodes by leveraging the observed information of networks.[2] Recently, the problem of link prediction has attracted a great deal of attention from researchers in different disciplines. On the one hand, the study of link prediction has important theoretical significance. It develops our understanding of the evolution mechanism of networks.[3] On the other hand, link prediction has widespread application value. For instance, based on the link prediction method, the potential interactions among neurons and proteins can be identified in biological networks,[4] the promising friendship can be explored in social networks[5] and the user behaviors can be well recommended in online commercial systems.[6]

Numerous structural similarity-based methods have been put forward for link prediction. Generally, these methods are based on the underlying assumption that a pair of nodes tends to connect if they have similar structural features in networks.[7] Furthermore, the more similar the nodes are, the more likely they will form links in the future. In order to improve the prediction performance, the existing link prediction methods endeavor to mine the internal topological information from single-layer networks as much as possible, such as neighbors of nodes,[8] degree,[9,10] clustering coefficient,[11] paths between two nodes,[12,13] betweenness[14] and community structure,[15,16] and so forth.

Recent studies[17–19] have shown that same individuals always participate in various types of interactive networks (e.g. the communications between people always depend on the systems of e-mail, mobile phone, Facebook or Twitter), which can be described as multiplex networks. In multiplex networks, the same individuals have different types of interactions and each type of interactive network corresponds to one layer between the same set of nodes.[20,21] Several researches have shown that the topological features of a layer are indeed affected by the shape of other layers in multiplex networks.[18,19,22] In other words, the traditional link prediction methods only take into account the topological information derived from the single-layer networks (called the intralayer information) but neglect the additional information originated from other layers in multiplex networks (called the interlayer information). From the macroperspective point of view, if a multiplex network is treated as an integrated network, the interactions between nodes in different layers can be regarded as the overall structural features of these nodes in different aspects. For link prediction, any structural information of networks can improve the prediction performance.[23] Therefore, in order to predict the likelihood of making connections between node pairs in a given layer, it is worth exploring feasible means to enhance the prediction performance with the aid of the interlayer information from other layers in multiplex networks.

In this paper, we propose a novel Node Similarity Index based on Layer Relevance (NSILR) that exploits the intralayer and interlayer information of the multiplex networks for link prediction. Considering two different methods to measure the relevance between layers, i.e. the Global Overlap Rate (GOR) and the Pearson

Correlation Coefficient (PCC), we validate and analyze the prediction performance of NSILR index on seven multiplex networks. The experimental results show that our NSILR index can significantly improve the prediction performance compared with the traditional link prediction methods that are based only on the single-layer networks. Furthermore, the more relevant the layers are, the higher the performance is improved. Meanwhile, our NSILR index can address the cold-start problem of link prediction based on the topological feature of multiplex networks. One can predict the behavior of new or small degree individuals based on their background behaviors in multiplex networks.

This paper is structured as follows. Section 2 introduces the related work of link prediction. In Sec. 3, we present the link prediction problem depending on the multiplex networks and propose our prediction index. In Sec. 4, the experimental datasets and the standard metric for performance evaluation are introduced. The experimental results and analysis are shown in Sec. 5. This work is concluded in Sec. 6.



## 2. Related Work

As a simple and effective framework of link prediction, the structural similarity-based methods can be classified into three categories: local similarity indices, global similarity indices and quasi-local indices.[3] Based on the local structural features of networks, Newman[8] proposed the Common Neighbors (CN) index which measures the number of overlapping neighbors between two nodes as the node similarity. Considering the influence of two endpoints and different roles of common neighbors, the variations of CN index have been proposed, such as Salton,[24] Jaccard,[25] LHN-I,[26] AA[9] and RA.[10] As a representative method of the global similarity indices, Katz index considers all connected paths between two nodes via a weighted form.[12] In most cases, local similarity indices have lower computational complexity and lower prediction accuracy as a result of requiring limited structural features. Global similarity indices have higher prediction accuracy, certainly they also associate with the higher computational complexity. In order to improve prediction accuracy and reduce computational complexity, Lü *et al.*[13] proposed the Local Path index (LP) which is a typical method of quasi-local indices.

Recently, several new link prediction methods have been developed based on different insights. From the microscale perspective[27] that only takes into account the properties of nodes (e.g. degree, clustering), the hybrid-based methods have been studied. Liao *et al.*[28] presented a novel correlation-based index and combined it with the AA index to achieve better performance. Zeng[29] presented a hybrid index through combining the CN index with the preferential attachment index (PA).[30] Li *et al.*[31] exploited the coupling degrees and the clustering coefficient of common neighbor nodes for link prediction. Meanwhile, from the mesoscale perspective[27] that analyzes the substructures of networks, the community-based methods are developed. Yan and Gregory[16] combined the community detection methods with the

existing link prediction indices for predicting missing links. Ding *et al.*[15] studied the relevance between communities and presented a link prediction model that fully utilized the information of intra-community and inter-community.

All the methods mentioned above explore the intralayer information from the single-layer network as much as possible to improve the prediction performance. However, few works employ the interlayer information from other layers in multiplex networks for link prediction. Pujari *et al.*[32] built a three-layer multiplex scientific collaboration network (corresponding to the relationships of co-authorship, co-venue and co-citing, respectively). Based on three multiplex attributes across all layers, the interaction of co-author layer was predicted via the decision tree algorithm of supervised machine learning. Hristova *et al.*[33] constructed a two-layer multilayer social network (corresponding to the relationships of Twitter and Foursquare) and defined several structural features on each layer. Based on the supervised random forest classifier of machine learning, the relationship of one layer was predicted in helping with the feature information gained from another layer. To utilize interlayer information of multiplex networks for link prediction, all the above two methods exploit supervised machine learning approaches and evaluate on a multiplex network for the specific context. Different from the above two methods, Sharma *et al.*[34,35] proposed a structural similarity-based method in multiplex network that considered the different weights of other layers for a predicted target layer by identifying the link correspondence between layers. However, the intralayer information of the predicted layer was not fully exploited by this method and the prediction performance was validated only on one specific multiplex network.

In this paper, our NSILR index comprehensively exploits the intralayer information from single-layer networks and the interlayer information from other layers in multiplex networks based on two layer relevance algorithms, i.e. GOR and PCC. The method of Refs. 34 and 35 can be considered as a special case of our index when the GOR algorithm is adopted and the intralayer similarities of node pairs are neglected for the NSILR index, (i.e. $\varphi = 1$ in Eq. (2)). The experimental results on seven multiplex networks show that the prediction performance can be further improved when the intralayer similarities of node pairs are simultaneously considered. Moreover, the PCC algorithm has advantage over GOR in performance improvement.

## 3. Methods

### 3.1. *Problem definition and motivation*

Generally, a simple undirected unweighted network (i.e. a single-layer network) can be denoted by $G = (V, E)$, where $V$ and $E$ are the set of nodes and links, respectively. A multiplex network with $M$ layers and $N$ nodes can be denoted by $\mathbf{G} = (G^1, G^2, \ldots, G^M)$, where $G^\alpha = (V^\alpha, E^\alpha)$ represents the network of layer $\alpha$.[20,21,36] Meanwhile, each layer of the multiplex network $\mathbf{G}$ has the same set of nodes,

i.e. $|V^1| = |V^2| = \cdots = |V^M| = N$. The adjacency matrix of each layer $G^\alpha$ can be denoted by $A^{[\alpha]} = \{a_{ij}^{[\alpha]}\}$, where

$$a_{ij}^{[\alpha]} = \begin{cases} 1 & \text{if } (V_i^\alpha, V_j^\alpha) \in E^\alpha, \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$

for $1 \leq \alpha \leq M$ and $1 \leq i, j \leq N^\alpha$ ($N^\alpha = |V^\alpha|$).[37] In fact, not all nodes interact with other nodes at all layers in multiplex networks and there are some isolated nodes in some of the layers. A node $i$ is active at layer $\alpha$ if it has at least one connection at layer $\alpha$ and is inactive otherwise.[38]

For link prediction, given a node pair $(x, y)$, its similarity score in layer $\alpha$ is denoted by $\text{sim}_{xy}^\alpha$. All potential node pairs in layer $\alpha$ are ranked in descending order by their similarity scores and the top ranked pairs have more chance to connect. Hence, the key of link prediction is how to calculate the similarity of one node pair based on the known topology of networks.

The recent studies have shown that there are dependencies between layers in multiplex networks.[39,40] The existence of links in one layer is relevant to the existence of corresponding links in other layers.[36] Motivated by this, we take into account the relationship between layers to improve prediction performance for a particular layer in multiplex networks. For a layer $\alpha$ to be predicted ($G^\alpha \in \mathbf{G}$, called the predicted layer) and any other layer $\beta$ ($G^\beta \in \mathbf{G}, \beta \neq \alpha$), the higher the relevance between layers $\alpha$ and $\beta$, the more interlayer information of layer $\beta$ can be adopted to predict missing links in layer $\alpha$. In other words, for a node pair in the predicted layer $\alpha$, its final similarity score is not only contributed by the intralayer structure information from the layer $\alpha$ but also depends on the interlayer structure information from other layers $\beta$. Moreover, the more the layer $\beta$ is relevant to layer $\alpha$, the greater the contribution of layer $\beta$ is.

### 3.2. *Node similarity index based on layer relevance of multiplex networks*

In order to employ the interlayer information of other layers in a multiplex network, for a node pair $(x, y)$ in the predicted layer $\alpha$, we first calculate its similarity within each layer based on the traditional existing methods (denoted by $\text{sim}_{xy}^\alpha$). These methods include CN,[8] RA,[10] AA,[9] LP[13] and Katz[12] to compare the different prediction effect of different methods. The detailed calculation of these five baseline measures are given in Appendix A. Then, for the node pair $(x, y)$ in the layer $\alpha$, we propose a Node Similarity Index based on Layer Relevance in multiplex network for link prediction. It is defined as

$$S_{xy}^\alpha = (1 - \varphi)\text{sim}_{xy}^\alpha + \varphi \sum_{\beta \neq \alpha}^{M} \mu^{\alpha\beta} \text{sim}_{xy}^\beta, \tag{2}$$

where $\text{sim}_{xy}^\alpha$ and $\text{sim}_{xy}^\beta$ are the similarities depending on the intralayer information from the predicted layer $\alpha$ and the interlayer information from other layers $\beta$,

respectively. $\mu^{\alpha\beta}$ is the relevance between layers $\alpha$ and $\beta$, which represents the weight of interlayer information from any other layer $\beta$ for predicting missing links in layer $\alpha$. The computing method of $\mu^{\alpha\beta}$ will be given in Sec. 3.3. The tunable parameter of $\varphi$, which lies in the interval [0,1], controls the weight of all the interlayer information from other layers for each predicated layer. $S^{\alpha}_{xy}$ is the final similarity of the node pair $(x, y)$ in layer $\alpha$ that employs the intralayer and interlayer information from multiplex network $\mathbf{G}$. This node similarity method is referred as the NSILR index in this paper.

According to Eq. (2), the first item is the contribution of intralayer information derived from the single predicted layer $\alpha$, the second item is all the contributions of interlayer information derived from other layers. When $\varphi = 0$, it indicates that the NSILR index just utilizes the intralayer information from the single predicted layer $\alpha$. It will degenerate to the traditional baseline similarity measures that consider only the single-layer network. $\varphi = 1$ means that the NSILR index completely exploits the interlayer information from other layers. Obviously, $0 < \varphi < 1$ implies that both the intralayer and interlayer information derived from layer $\alpha$ and other layers are exploited.

### 3.3. *Relevance between layers in multiplex networks*

In order to compare the prediction performance, we exploit two algorithms for quantifying the relevance between layers in multiplex networks, i.e. $\mu$ in Eq. (2).

(1) Global Overlap Rate (GOR) between layers. Given two layers $\alpha$ and $\beta$ in a multiplex network, an overlap link means that a same node pair simultaneously connects in layers $\alpha$ and $\beta$. The global overlap between layers $\alpha$ and $\beta$ is denoted by $O^{\alpha\beta}$, which is the total number of overlap links observed in layers $\alpha$ and $\beta$.[20,36] The global overlap rate[39] between layers $\alpha$ and $\beta$ can be defined as

$$\mu^{\alpha\beta}_{\mathrm{GOR}} = \frac{2O^{\alpha\beta}}{L^{\alpha} + L^{\beta}} = \frac{2\sum_{i<j} a^{\alpha}_{ij} a^{\beta}_{ij}}{\sum_{i<j} a^{\alpha}_{ij} + \sum_{i<j} a^{\beta}_{ij}}, \tag{3}$$

where $L^{\alpha}$ is the total number of observed links in the layer $\alpha$ and $a^{\alpha}_{ij} = 0$ or 1 depends on whether there exists a link between nodes $i$ and $j$ in the layer $\alpha$. $\mu^{\alpha\beta}_{\mathrm{GOR}}$ ranges in [0, 1]. 0 indicates that there are no overlap links between layers $\alpha$ and $\beta$ corresponding completely irrelevant for them. On the contrary, 1 indicates they are perfectly relevant. Accordingly, we hold that the higher the GOR between two layers is, the more relevant or similar they are.

(2) Pearson Correlation Coefficient (PCC) between layers. In Ref. 28, the feature of one node $i$ was characterized as an attribute vector $\mathbf{v}_i$ derived from the adjacency matrix $A$ of network graph, where $\mathbf{v}_i = (a_{i1}, a_{i2}, \ldots, a_{in})$ and $a_{in}$ is equal to 0 or 1. Then, the similarity of the node pair $(i, j)$ was proposed based on the Pearson correlation coefficient for vectors $\mathbf{v}_i$ and $\mathbf{v}_j$. Based on this idea, we view the feature

of the layer $\alpha$ as a long vector $\mathbf{g}_\alpha$ derived from the adjacency matrix $A^{[\alpha]}$, where $\mathbf{g}_\alpha = (a^\alpha_{11}, \ldots, a^\alpha_{1n}, a^\alpha_{21}, \ldots, a^\alpha_{2n}, \ldots, a^\alpha_{n1}, \ldots, a^\alpha_{nn})$. For the sake of simplicity, the alternative representation of $\mathbf{g}_\alpha$ is replaced by $\mathbf{g}_\alpha = (a^\alpha_1, a^\alpha_2, \ldots, a^\alpha_{n^2})$. Therefore, the Pearson correlation coefficient between layers $\alpha$ and $\beta$ is defined as

$$\mu^{\alpha\beta}_{\mathrm{PCC}} = \frac{1}{n^2} \sum_{i=1}^{n^2} \frac{(a^\alpha_i - \overline{\mathbf{g}_\alpha})(a^\beta_i - \overline{\mathbf{g}_\beta})}{\sigma(\mathbf{g}_\alpha)\sigma(\mathbf{g}_\beta)}, \tag{4}$$

where $\overline{\mathbf{g}_\alpha}$ and $\sigma(\mathbf{g}_\alpha)$ represent the mean value and standard deviation of vector $\mathbf{g}_\alpha$, respectively. $\mu^{\alpha\beta}_{\mathrm{PCC}}$ ranges in $[-1, 1]$. $-1$ indicate completely negative correlation for the layer pair, i.e. the links that appear in layer $\alpha$ cannot be present in layer $\beta$. While, 1 indicates perfect positive correlation, i.e. the links that appear in layer $\alpha$ are always present in layer $\beta$. It is slightly different from GOR, we hold that the more PCC tends to 0, the less relevant or similar the two layers are. Conversely, the more PCC tends to 1 (or $-1$), the more positive (or negative) relevant or similar they are.

Actually, according to Ref. 39, $\mu^{\alpha\beta}_{\mathrm{PCC}}$ can be considered as a comparison of the number of overlap links between layers $\alpha$ and $\beta$ in their observed networks (i.e. $\mu^{\alpha\beta}_{\mathrm{GOR}}$) and their corresponding Random Graph (i.e. $\mu^{\alpha\beta}_{\mathrm{RG}}$). The difference between GOR and PCC can be defined as[39]

$$\mu^{\alpha\beta}_{\mathrm{PCC}} = \frac{L^\alpha + L^\beta}{2\sqrt{L^\alpha L^\beta (1 - \frac{2L^\alpha}{N^2})(1 - \frac{2L^\beta}{N^2})}} (\mu^{\alpha\beta}_{\mathrm{GOR}} - \mu^{\alpha\beta}_{\mathrm{RG}}). \tag{5}$$

The overlap rate of random graphs for layers $\alpha$ and $\beta$ is given by

$$\mu^{\alpha\beta}_{\mathrm{RG}} = \frac{1}{N^2} \frac{4L^\alpha L^\beta}{L^\alpha + L^\beta}. \tag{6}$$

The differences of prediction performance for GOR and PCC are shown in Sec. 5.

## 4. Data and Metric

### 4.1. *Datasets*

In order to validate the prediction accuracy of the NSILR index, we perform experiments on seven multiplex networks in real-world systems.

(1) Vicker[41]: It is a multiplex social network in a school in Victoria, Australia. It contains 3 layers and 29 nodes. Nodes represent students in the class. Each layer corresponds to the relations of contact, best friends and working together, respectively.
(2) CKM[42]: It is a multiplex network between physicians when they adopt a new drug. It is composed of 3 layers and 246 nodes. In this network, nodes stand for physicians, layers indicate three different interactions between physicians, corresponding to ask for advice, discussing and friendship.

(3) Lazega[43,44]: It is a multiplex network of a corporate law partnership between partners and associates. It includes three kinds of relations, such as co-work, friendship and advice, corresponding to 3 layers, 71 nodes that represent partners and 2223 directed links that indicate the connections between them. Herein, we treat this network as an undirected multiplex network in our experiment.
(4) CSAar[45]: It is a multiplex network among employees in the Department of Computer Science at Aarhus University. This network consists of 5 layers and 61 employees. The five layers represent five kinds of online and offline relationships between employees, corresponding to the interactions of lunch, Facebook, co-authorship, leisure and work, respectively.
(5) Celegans[46,47]: It is a multiplex neuronal network between Caenorhabditis elegans. It includes 3 layers and 279 nodes. Each layer corresponds to different synaptic junctions: electric, chemical monadic and chemical polyadic.



Table 1. The basic topological features of the experimental networks. The values in brackets are the total number of nodes in multiplex networks.

| Networks | $i^\alpha$ | $N^\alpha$ | $E^\alpha$ | $\langle k\rangle$ | $S$ | $r$ | $\langle d\rangle$ | $C$ | $H$ |
|---|---|---|---|---|---|---|---|---|---|
| Vicker (29) | 1 | 29 | 240 | 24.9 | 0.591 | −0.043 | 1.41 | 0.707 | 1.11 |
| | 2 | 29 | 126 | 12.48 | 0.31 | −0.083 | 1.81 | 0.532 | 1.26 |
| | 3 | 29 | 152 | 13.66 | 0.374 | 0.026 | 1.76 | 0.599 | 1.26 |
| CKM (246) | 1 | 215 | 449 | 4.47 | 0.02 | −0.113 | 3.15 | 0.159 | 1.51 |
| | 2 | 231 | 498 | 4.89 | 0.019 | −0.073 | 3.37 | 0.188 | 1.36 |
| | 3 | 228 | 423 | 4.44 | 0.016 | 0.153 | 3.94 | 0.207 | 1.26 |
| Lazega (71) | 1 | 71 | 556 | 17.24 | 0.224 | 0.002 | 1.91 | 0.381 | 1.33 |
| | 2 | 71 | 725 | 24.08 | 0.292 | 0.051 | 1.75 | 0.44 | 1.2 |
| | 3 | 70 | 378 | 21.6 | 0.156 | −0.176 | 2.1 | 0.307 | 1.27 |
| CSAar (61) | 1 | 60 | 193 | 6.43 | 0.109 | 0.005 | 3.19 | 0.569 | 1.21 |
| | 2 | 32 | 124 | 7.75 | 0.25 | 0.003 | 1.96 | 0.481 | 1.23 |
| | 3 | 25 | 21 | 1.68 | 0.07 | 0.017 | 1.5 | 0.429 | 1.39 |
| | 4 | 47 | 88 | 3.74 | 0.081 | −0.01 | 3.12 | 0.343 | 1.51 |
| | 5 | 60 | 194 | 6.47 | 0.11 | −0.213 | 2.39 | 0.339 | 1.67 |
| Celegans (279) | 1 | 253 | 514 | 8.15 | 0.016 | −0.118 | 4.52 | 0.128 | 2.15 |
| | 2 | 260 | 888 | 12.61 | 0.026 | −0.092 | 3.37 | 0.137 | 1.76 |
| | 3 | 278 | 1703 | 22.97 | 0.044 | −0.057 | 2.71 | 0.194 | 1.71 |
| Terrorist (78) | 1 | 74 | 200 | 5.41 | 0.074 | −0.176 | 2.7 | 0.266 | 2.14 |
| | 2 | 14 | 15 | 2.14 | 0.165 | 0.25 | 1.43 | 0.72 | 1.24 |
| | 3 | 51 | 79 | 3.1 | 0.062 | −0.665 | 3.01 | 0.098 | 2.94 |
| | 4 | 70 | 259 | 7.4 | 0.107 | 0.222 | 2.83 | 0.531 | 1.79 |
| Yeast (4458) | 1 | 4422 | 59705 | 30.41 | 0.006 | −0.162 | 2.84 | 0.029 | 2.84 |
| | 2 | 4432 | 108015 | 55.99 | 0.011 | −0.084 | 2.58 | 0.055 | 2.54 |
| | 3 | 4458 | 4395844 | 1972.11 | 0.442 | −0.043 | 1.56 | 0.52 | 1.05 |
| | 4 | 4458 | 3886844 | 1743.76 | 0.391 | 0.163 | 1.61 | 0.42 | 1.06 |

*Notes*: $i^\alpha$ represents the sequence number of the layer $\alpha$ in a multiplex network. $N^\alpha$ and $E^\alpha$ are the size of active nodes and link number in the layer $\alpha$, respectively. $\langle k\rangle$ is the average degree. $S$ is the layer density ($S = \frac{2E}{N(N-1)}$). $r$ is the assortative coefficient. $\langle d\rangle$ and $C$ are the average shortest path length and clustering coefficient, respectively. $H$ is the degree heterogeneity ($H = \frac{\langle k^2\rangle}{\langle k\rangle^2}$).

(6) Terrorist[37]: It is a multiplex network among Indonesian terrorists containing 4 layers and 78 nodes in total. Nodes represent terrorists and links correspond to the interactions between terrorists. Four layers represent the communication, financial, operation and trust interactions between terrorists, respectively.
(7) Yeast[47,48]: It is a multiplex genetic interaction network of the Saccharomyces Cervisiae that is a species of yeast. This network consists of 4 layers and 4458 nodes. Nodes represent genes, links stand for the connections between them. Each layer corresponds to the positive, negative interactions, as well as the positive and negative correlations.

The basic topological features of these networks are shown in Table 1.

### 4.2. *Evaluation metric*

Given a layer $G^\alpha = (V^\alpha, E^\alpha)$, the universal set of possible links in layer $\alpha$ is denoted by $U^\alpha = N(N-1)/2$, where $N = |V^\alpha|$ is the total number of nodes in $G^\alpha$. Each nonexistent link $l_{xy} \in U^\alpha \backslash E^\alpha$ $(x, y \in V^\alpha)$ is assigned a similarity score by one link prediction method. To evaluate the prediction accuracy of the prediction indices, the observed links in layer $\alpha$ $(E^\alpha)$ are randomly divided into the training set $E_t^\alpha$ and the probe set $E_p^\alpha$. $E_t^\alpha$ represents the known information and $E_p^\alpha$ is used to test the prediction performance. It is obvious that $E_t^\alpha \cup E_p^\alpha = E^\alpha$ and $E_t^\alpha \cap E_p^\alpha = \emptyset$. In our experiments, we randomly chose a certain proportion (denoted by $r$) of the observed links as the training set and the residual proportion $(1-r)$ of the observed links as the probe set in a layer.

In our experiments, we adopt a standard metric called *area under the receiver operating characteristic curve* (AUC)[49] to evaluate the accuracy of prediction indices. AUC can be considered as a probability where the scores of node pairs randomly chosen from the probe set (i.e. $E_p^\alpha$) are higher than those randomly chosen from the nonexistent link sets (i.e. $U^\alpha - E^\alpha$). If we perform $n$ times of independent comparisons, there are $n_1$ times that the probe links have higher scores and $n_2$ times they are equivalent. Then, AUC can be defined as

$$\text{AUC} = \frac{n_1 + 0.5 \times n_2}{n}. \tag{7}$$

It is roughly equivalent to 0.5 if all scores are generated from an independent and identical distribution. Therefore, the extent to which AUC exceeds 0.5 represents how much the prediction method performs better than pure chance.

## 5. Experiments and Analysis

### 5.1. *Layer relevance of the observed multiplex networks*

In this section, the layer relevance between layers is explored in different multiplex networks. We calculate the relevance between layers by GOR and PCC in each multiplex network, respectively, and present the results in Fig. 1. As shown in the

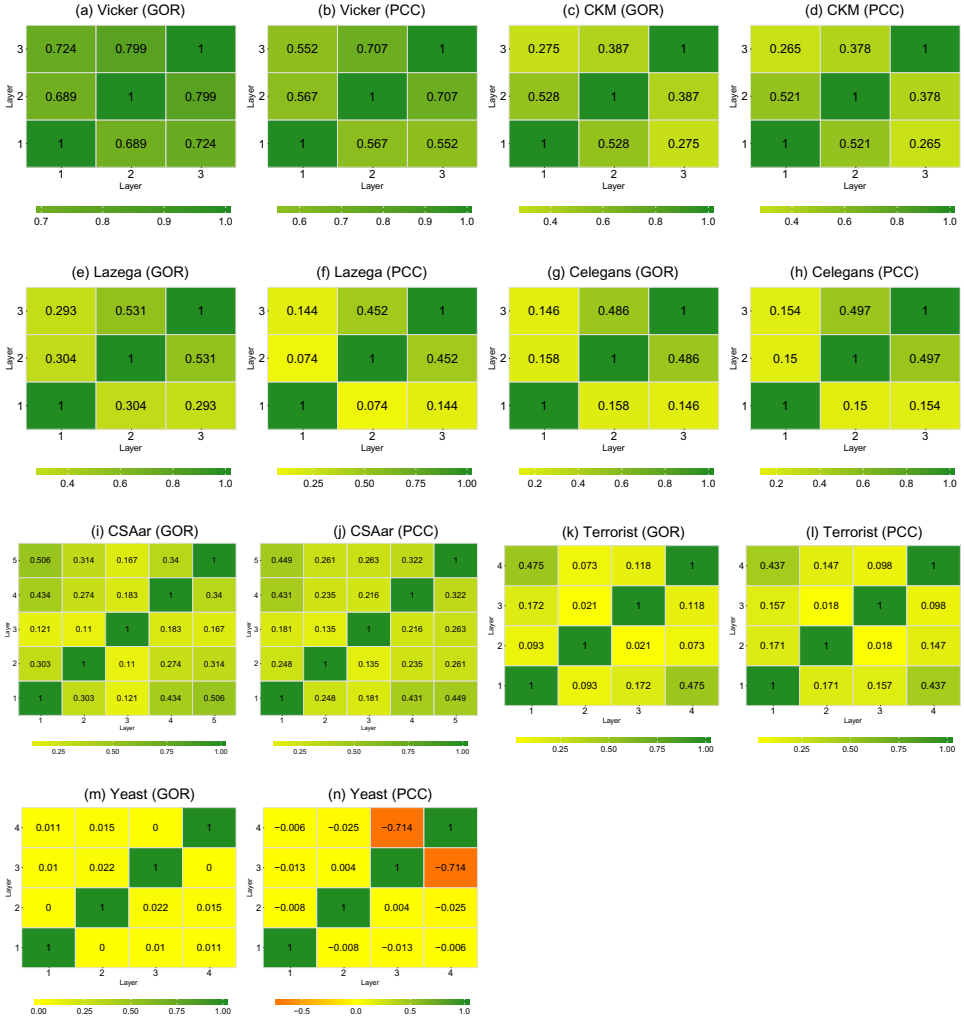


Fig. 1. (Color online) Layer relevance of seven multiplex networks. The heat maps reflect the relevance between layers in different networks. Each subfigure corresponds to the result of a network, which is measured by GOR or PCC.

figure, it has the highest layer relevance for each layer pair in the Vicker network. The relevance values of layer pairs approximately range from 0.68 to 0.8 with GOR as shown in Fig. 1(a) (0.55 to 0.7 with PCC in Fig. 1(b)), except for the layers' self-relevance. This result means that each layer pair is quite similar in this network. Conversely, the Yeast network has the lowest layer relevance for most layer pairs (see Figs. 1(m) and 1(n)). It indicates that most layers are less similar in this network.

Although the layer relevance calculated by GOR is consistent with that by PCC in most cases, yet there are differences in some multiplex networks. We notice that different relevance measurements of layer pairs even draw the opposite conclusion

for the same layer pairs in a multiplex network. For example, in the Vicker network (see Figs. 1(a) and 1(b)), the methods of GOR and PCC all agree $\mu^{32} > \mu^{31}$ for layer 3 (i.e. $\mu^{32}_{\mathrm{GOR}} = 0.799 > \mu^{31}_{\mathrm{GOR}} = 0.724$ and $\mu^{32}_{\mathrm{PCC}} = 0.707 > \mu^{31}_{\mathrm{PCC}} = 0.552$). However, for layer 1, the GOR method considers $\mu^{13}_{\mathrm{GOR}} > \mu^{12}_{\mathrm{GOR}}$ (i.e. $\mu^{13}_{\mathrm{GOR}} = 0.724$, $\mu^{12}_{\mathrm{GOR}} = 0.689$), the PCC method considers $\mu^{13}_{\mathrm{PCC}} < \mu^{12}_{\mathrm{PCC}}$ (i.e. $\mu^{13}_{\mathrm{PCC}} = 0.552$, $\mu^{12}_{\mathrm{PCC}} = 0.567$). In particular for the Yeast network (see Figs. 1(m) and 1(n)), there is little relevance for each layer pair with GOR because its layer relevance is nearly close or equal to 0. However, for layers 3 and 4 with PCC, there is high negative relevance (i.e. $\mu^{34}_{\mathrm{PCC}}$ is equal to $-0.714$). Thus, for the NSILR index, different methods that characterize the layer relevance may have distinctive prediction accuracy. This influence on prediction performance will be analyzed in the following sections.

### 5.2. *The influence of tunable parameter on prediction performance*

Since the NSILR index depends on the tunable parameter $\varphi$, here, we investigate its prediction performance with variation of $\varphi$. In our experiments, we adopt different fractions $r$ of all observed links in a layer as the training set. The layer relevance $\mu$ is calculated by GOR and PCC, respectively. Note that, given a predicted layer $\alpha$ and any other layer $\beta$ in a multiplex network, in order to calculate the weight that layer $\beta$ contributes to layer $\alpha$, we make use of the information of the training set of layer $\alpha$ and all the observed information of layer $\beta$ to obtain $\mu^{\alpha\beta}$. For the NSILR index, the node similarity on each layer (i.e. $\mathrm{sim}_{xy}$ in Eq. (2)) is calculated based on five baseline measures, i.e. CN, AA, RA, LP and Katz. Here, we mainly show the results of the Vicker (Fig. 2) and Yeast (Fig. 3) multiplex networks based on the CN measure and the results of other baseline measures are similar with CN. The prediction performances on other five multiplex networks are shown in Supplementary Material Note 1.

As shown in the figures, although the NSILR index shows different trends in different multiplex networks when $\varphi$ changes, in most cases, its performance can be improved and achieves a peak value (corresponding $\varphi$ denoted as $\varphi^*$), regardless of GOR or PCC. This result suggests that the interlayer information from other layers can further improve the prediction performance compared with that based only on the intralayer information from single-layer network (i.e. $\varphi = 0$). Additionally, we find that the PCC has more advantages than GOR in performance improvement. As shown in Fig. 3, for all layers in Yeast, their prediction performances are not improved by GOR. Conversely, the prediction performance of each layer is improved to some extent by PCC, especially for layers 3 and 4. This is because the method of PCC considers the null models (i.e. random graphs) of the observed layer networks and can identify the negative relevance between layers (see Figs. 1(m) and 1(n), for details). Therefore, we can conclude that the negative layer relevance can improve performance as well.

We also find other phenomena and describe them through three phases: (1) $\varphi$ from 0 to 0.01; (2) $\varphi$ from 0.01 to $\varphi^*$; (3) $\varphi$ from $\varphi^*$ to 1. Here, we pay special attention to the prediction performance of NSILR when a slight contribution of all

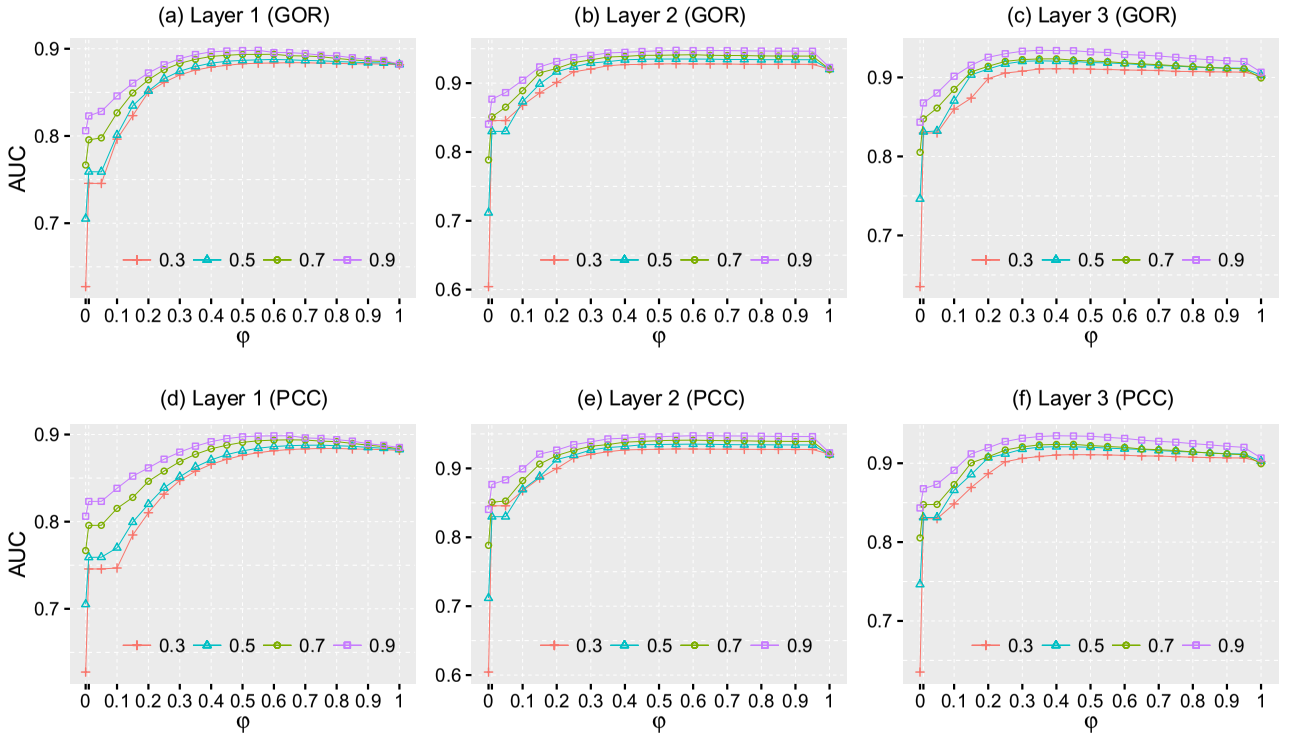


Fig. 2. (Color online) Prediction performance of NSILR index on each layer of the Vicker network based on the CN baseline measure. Each subfigure corresponds to the performance with GOR or PCC in a layer over 20 independent experiments. $\varphi$ controls the weight of interlayer information from other layers ($\varphi = 0.01$ which is very close to 0 in the horizontal axis is labeled). Each curve represents the performance of AUC corresponding to different fractions $r$ 30%, 50%, 70%, 90%, respectively.

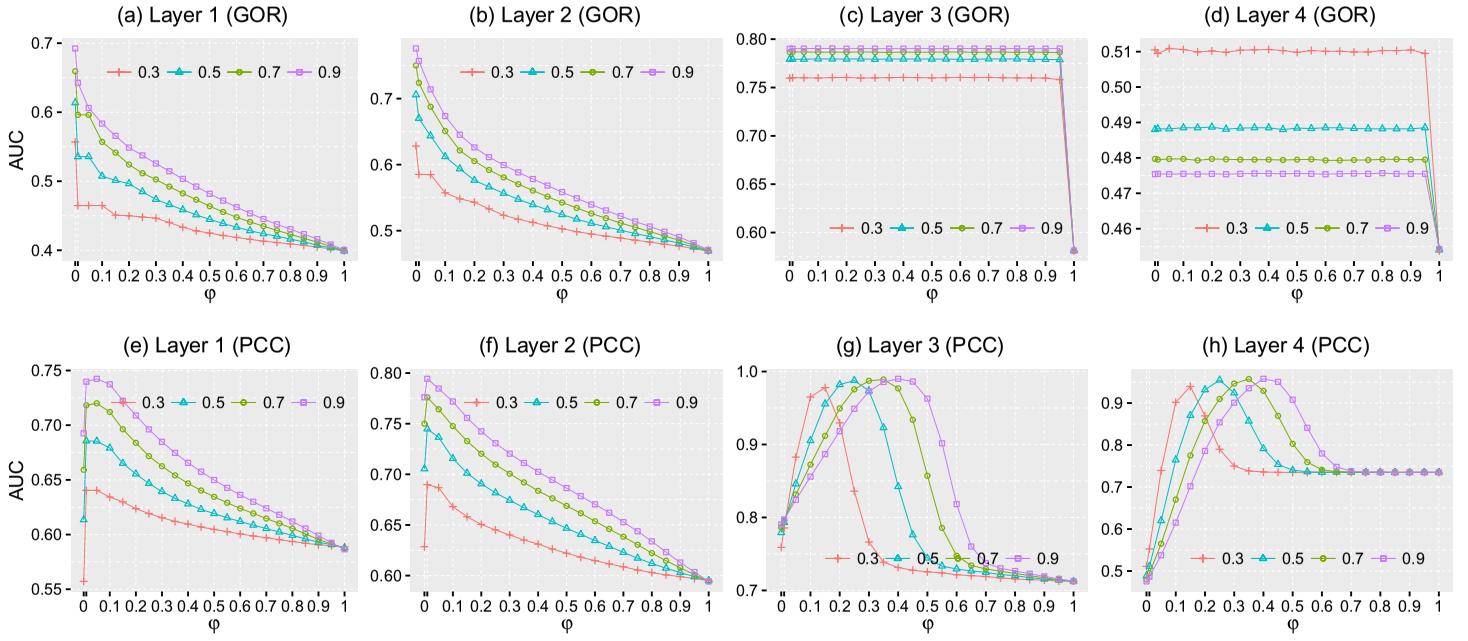


Fig. 3. (Color online) Prediction performance of NSILR index on each layer of the Yeast network based on the CN baseline measure. Each subfigure corresponds to the performance with GOR or PCC in a layer over 20 independent experiments. $\varphi$ controls the weight of interlayer information from other layers ($\varphi = 0.01$ which is very close to 0 in the horizontal axis is labeled). Each curve represents the performance of AUC corresponding to different fractions $r$ 30%, 50%, 70%, 90%, respectively.

other layers is considered, i.e. $\varphi = 0.01$. The reason is that, for LP and Katz indices, their relative optimal performance improvement corresponds to a slight contribution of the additional internal structural information from single layers (i.e. $\epsilon = 0.01$ for LP in Eq. (A.5), $\beta = 0.01$ for Katz in Eq. (A.4)).

In the first phase, the NSILR index can well solve the network sparsity problem. The sparser the predicted layer is, the more significantly the prediction performance is improved. Take the Vicker network as an example (see Fig. 2), when $\varphi$ changes from 0 to 0.01, the AUC with PCC on average improves by 2.59%, 4.46%, 8.57% and 18.47% corresponding to $r$ of 90%, 70%, 50% and 30% for all three layers, respectively. The reason is that the smaller $r$ always indicates less known information in the predicted layer and consequently most disconnected node pairs of this layer are assigned to 0 when the single layer is only considered, i.e. $\varphi = 0$. Therefore, it is difficult to distinguish the scores between the nonexistent links of the training set and the observed links of the probe set. However, when the interlayer information of other layers is simultaneously considered, i.e. $\varphi = 0.01$, their distinction becomes increasingly clear, which results in the improvement of prediction performance.

In the second phase, the prediction performance is further enhanced with increment of the weight of other layers in most cases. The reason is that the disconnected node pairs which are more likely to connect in the predicted layer are assigned to high similarity scores due to the interlayer information from other layers. Thus, these node pairs have top ranks so that the potential missing links are easily identified. In addition, the performance improvement is more significant when $r$ tends to 1 in this phase. It is because the higher $r$ corresponds to the more known information in the predicted layer and this information can enhance the relevance between layers as a result, which makes the highly relevant layers have heavier weights in the prediction procedure. However, no improvement in the prediction performance is observed for some layers, no matter whatever $r$ is, such as all layers with GOR in the Yeast network (see Figs. 3(a)–3(d)). This is because of the low relevance between them for the GOR method (see Fig. 1(m) for details).

In the last phase, the prediction performance begins to decrease. This is because much noisy information drawn from other layers is taken into the prediction procedure for the predicted layer with increment of $\varphi$. Meanwhile, for some networks such as Vicker (Fig. 2), CKM (Fig. S1 in Supplementary Note 1) and Terrorist (Fig. S4 in Supplementary Note 1), the prediction accuracies corresponding to $\varphi = 1$ outperform those of $\varphi = 0$. It indicates that one can get better prediction accuracy depending only on the interlayer information without considering the similarity scores of the predicted layer. This observation is very useful to solve the cold-start problem, especially for online social networks. Generally, it is difficult to obtain sufficiently available topology information for new or small degree nodes, which is the key difficulty of the cold-start problem.[28] The multiplex networks have natural advantages in topology information because they characterize the relationship of individuals more diverse than the single-layer networks. Therefore, one can predict

the behavior of new or small degree individuals based on their background relationships or behaviors.

Next, we analyze the choice of the optimal parameter $\varphi*$ for each layer of different networks. According to the experimental results of seven multiplex networks, it seems that $\varphi* \approx 0.5$ for all networks except for the Yeast network, of which $\varphi*$ is very different for different layers. For PCC, the optimal values of $\varphi*$ are 0.05, 0.01, 0.4 and 0.4 corresponding to layers 1, 2, 3 and 4 of the Yeast network, respectively. The reason is that the relevance of layer pairs is very different for some predicted layers of the Yeast network, e.g. as to the layer 3, $\mu^{3,1} = -0.013$, $\mu^{3,2} = 0.004$ and $\mu^{3,4} = -0.714$. Whereas, the relevance of all layer pairs is relatively high for the rest of the six networks, e.g. the relevance of all pairs of layers is greater than 0.55 for the Vicker network. Based on the layer relevance shown in Fig. 1 and the results shown in Figs. 2 and 3 (other networks shown in Supplementary Material Note 1), we find that $\varphi*$ can be approximately set to 0.5 for one predicted layer when all other layers have the relatively high relevance to the predicted layer. It indicates that, for a new unknown network, if there is a high relevance for any each pair of layers, we can set 0.5 as the approximately optimal value of $\varphi$ for each layer of this network. However, with regard to the case that there is obvious different layer relevance for layer pairs, we need to further analyze the optimal value $\varphi*$ by separately considering each predicted layer. This is because there may be obvious difference in some features for these layer pairs of networks. The advantage of this way is that we can get the optimal layer weight of the interlayer information for each layer to further improve the prediction performance. Certainly, its disadvantage is that it takes more time to search the optimal value $\varphi*$ for each layer.

### 5.3. *Comparison of the performance based on different baseline measures*

In order to validate the robustness of NSILR index, its prediction performance is analyzed based on different baseline measures in this subsection. Here, we mainly show the results of CN, RA, AA, LP and Katz measures in the Vicker (Fig. 4) and Yeast networks (Fig. 5) under different fractions $r$ of the training set. The results of other five networks are shown in Supplementary Material Note 2. For the sake of contrast, we also show the performance depending only on the intralayer information of the predicted layers (i.e. $\varphi = 0$).

As shown in the figures, the prediction accuracy based on the interlayer information significantly outperforms that depending only on the intralayer information in most networks for all baseline measures. In most cases, when $\varphi = 0$, the global index (i.e. Katz) has better performance than the quasi-local index (i.e. LP), and the local similarity indices (i.e. CN, RA, AA) have the worst performance, particularly for the sparse network (i.e. the small $r$). This result is consistent with the existing study conclusion.[3] However, when we simultaneously consider the interlayer and intralayer information (i.e. $\varphi = \varphi^*$), the differences between different baseline

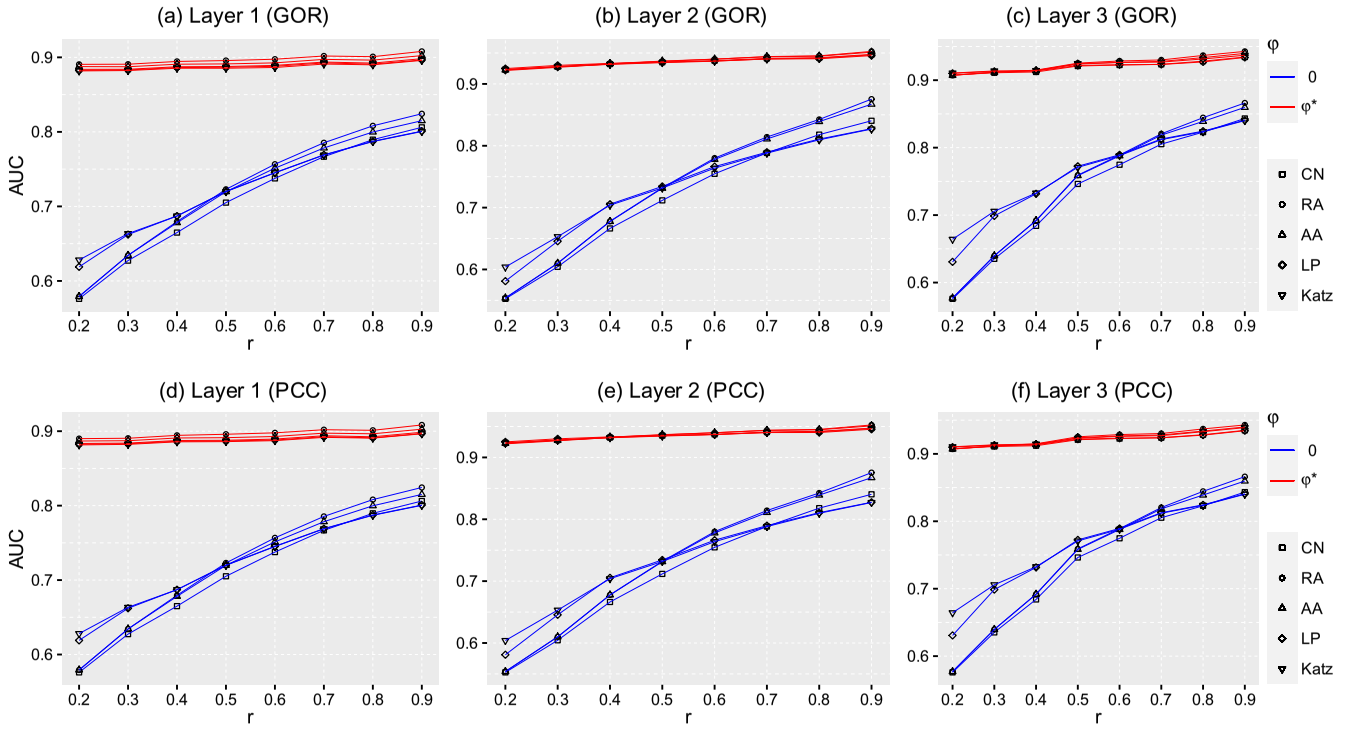


Fig. 4. (Color online) Comparison of the prediction performance among different baseline measures in Vicker network. Each subfigure corresponds to the performance with GOR or PCC in a layer over 20 independent experiments. $r$ is the fraction of training set. The points on blue curves are the performance depending only on the intralayer information (i.e. $\varphi = 0$). The points on red curves are the performance depending on both intralayer and interlayer information when the optimal $\varphi^*$ is employed.

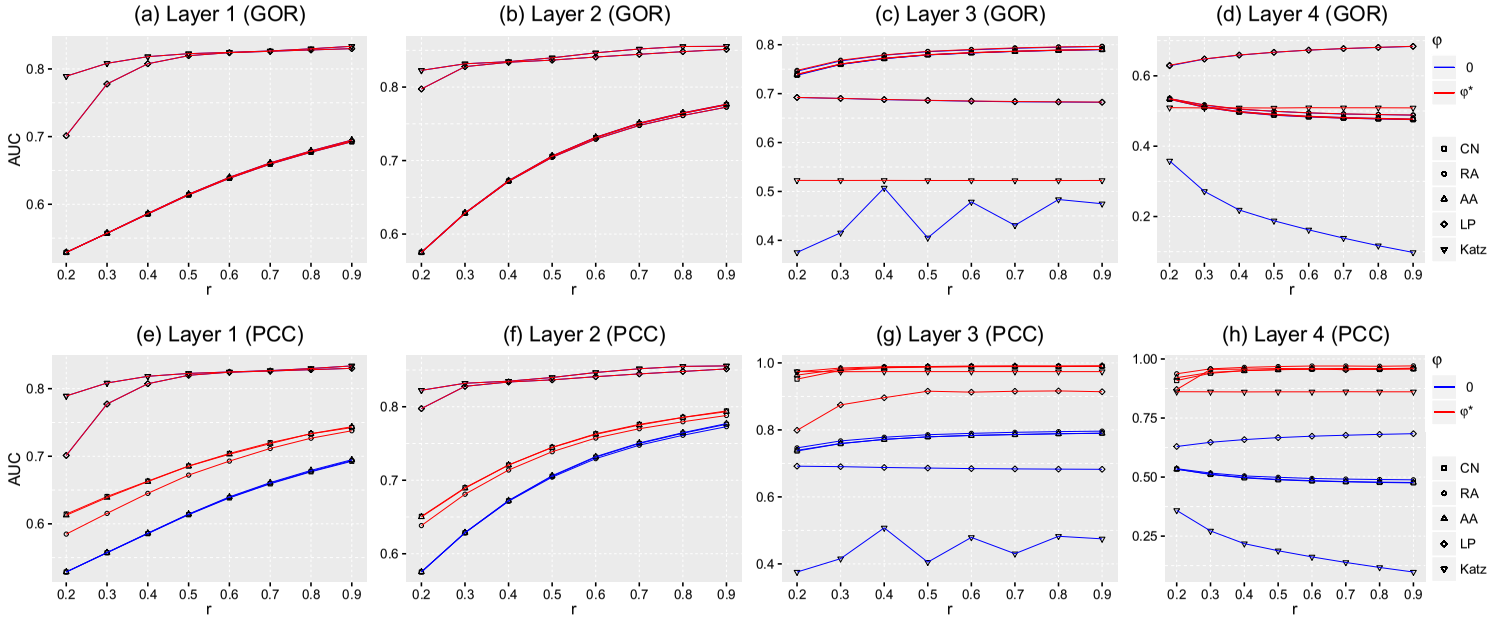


Fig. 5. (Color online) Comparison of the prediction performance among different baseline measures in Yeast network. Each subfigure corresponds to the performance with GOR or PCC in a layer over 20 independent experiments. $r$ is the fraction of training set. The points on blue curves are the performance depending only on the intralayer information (i.e. $\varphi = 0$). The points on red curves are the performance depending on both intralayer and interlayer information when the optimal $\varphi^*$ is employed.

measures are not obvious for one predicted layer in most networks (see Figs. 4 and S8–S10 in Supplementary Material Note 2). In some cases, the performance of CN measure nearly approaches LP and Katz measures and the performance of RA index is even better than that of LP and Katz. It indicates that, when the single layer information is only considered, the global and quasi-local similarity measures can achieve better performance than local similarity measures because they make use of additional internal structural information from layer networks (i.e. $\epsilon A^3$ for LP in Eq. (A.5), $\beta^l \cdot |\text{paths}_{xy}^{\langle l \rangle}|$ for Katz in Eq. (A.4)). When the interlayer and intralayer information is simultaneously considered for the NSILR index, the additional internal information from the predicted layers which is used by the global and quasi-local indices has no significant advantage in the prediction performance. It is suggested that the interlayer information is more valuable than the additional internal information for the NSILR index to some extent. Therefore, the prediction accuracies of NSILR index with local baseline measures are approximately equal to or better than those with LP and Katz measures. Inspired by this result, for the NSILR index, one can adopt local similarity measures that have low time complexity and the interlayer information from other layers to obtain better performance comparing with the global and quasi-local measures that have higher time complexity.

We noticed that in the Yeast networks (see Fig. 5), the GOR performance based on the single layer (i.e. $\varphi = 0$) and that based on other layers (i.e. $\varphi = \varphi^*$) are overlapped for different baseline measures on each predicted layer in some cases. This is because there is low relevance between the predicted layer and other layers when the GOR method is considered, with the result that the optimal value of $\varphi$ equals 0, i.e. $\varphi^* = 0$ (see Fig. 3).

### 5.4. *The impact of layer relevance on prediction performance*

For the NSILR index, the contribution of other layers to the prediction performance is investigated as a whole in the above sections. In this section, we focus attention on the different contribution of each other layer to the prediction performance for a predicted layer in multiplex networks, i.e. the layer relevance $\mu$ of the NSILR index will be studied. In this experiment, the ratio of training set $r$ is kept to a constant value 0.9 and $\varphi$ is fixed at its optimal value on each layer based on CN baseline measure (denoted by $\varphi_{\text{CN}}^*$). To shed light on the different influence of layer relevance $\mu$ on the prediction performance, we explore the changing of prediction performance through removing the contributions of other layers. Given a predicted layer $\alpha$ and a series of numbers $n$ ($n \in [0, 1]$ for GOR and $n \in [-1, 1]$ for PCC), we eliminate the contribution of each other layer $\beta$ to the predicted layer $\alpha$ step by step through setting the layer relevance $\mu^{\alpha\beta}$ to 0 when $\mu^{\alpha\beta} \leq n$. This procedure will be stopped when the prediction performance of the NSILR index does not rely upon the interlayer information from all the other layers. In this case, $\mu^{\alpha\beta}$ is equal to 0 for any other layer $\beta$ and therefore the second item of Eq. (2) has no effect on the prediction

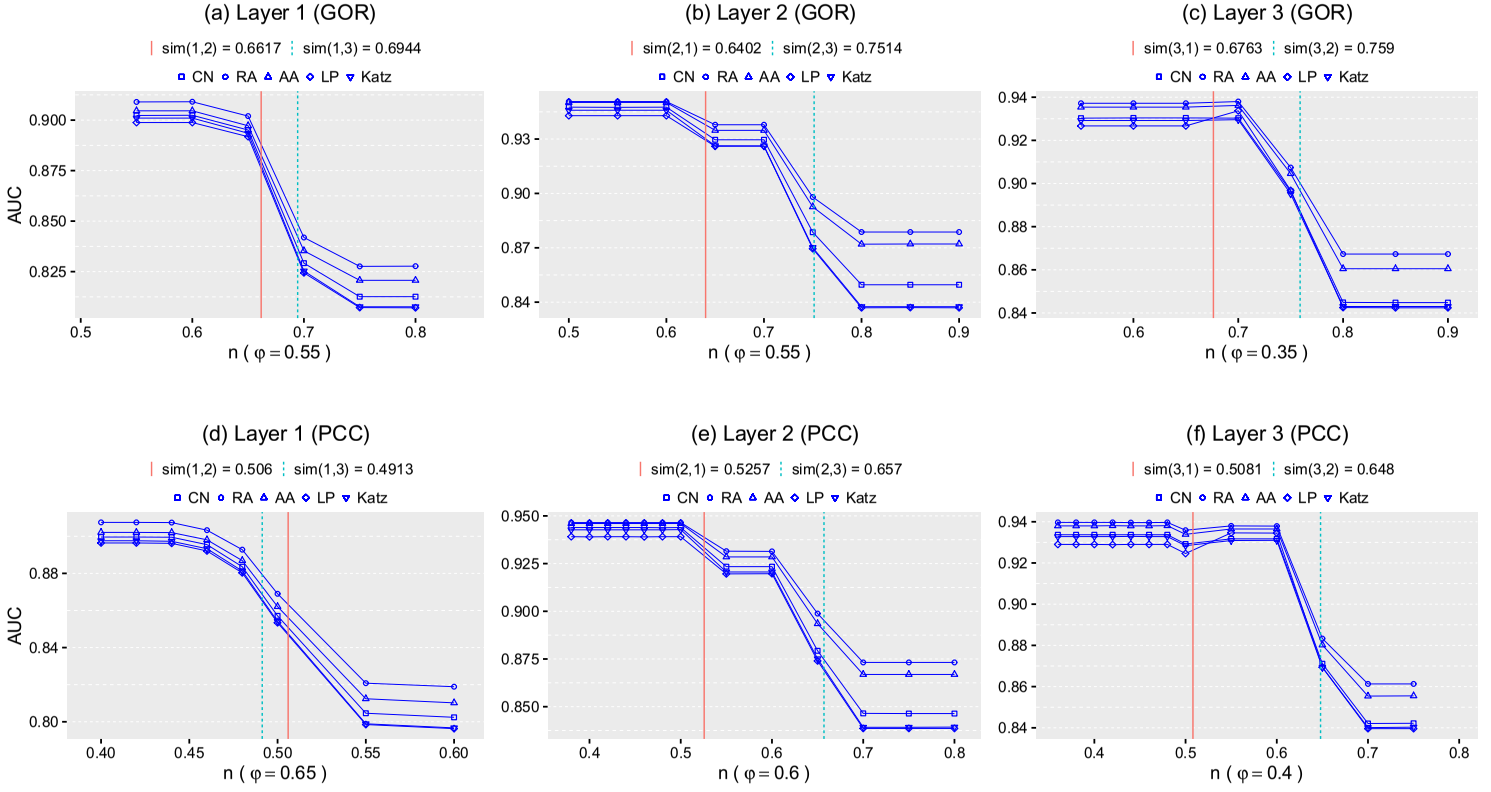


Fig. 6. (Color online) Comparison of the prediction performance of each layer in the Vicker network. Each subfigure corresponds to the performance with GOR or PCC in a layer over 20 independent experiments. The blue curves are the prediction performance based on different baseline measures (i.e. CN, RA, AA, LP and Katz). For a predicted layer, $\varphi$ is equal to $\varphi^*_{\mathrm{CN}}$ and the training set ratio $r$ is 0.9. Each vertical line presents the relevance between the predicted layer and other layer on average.

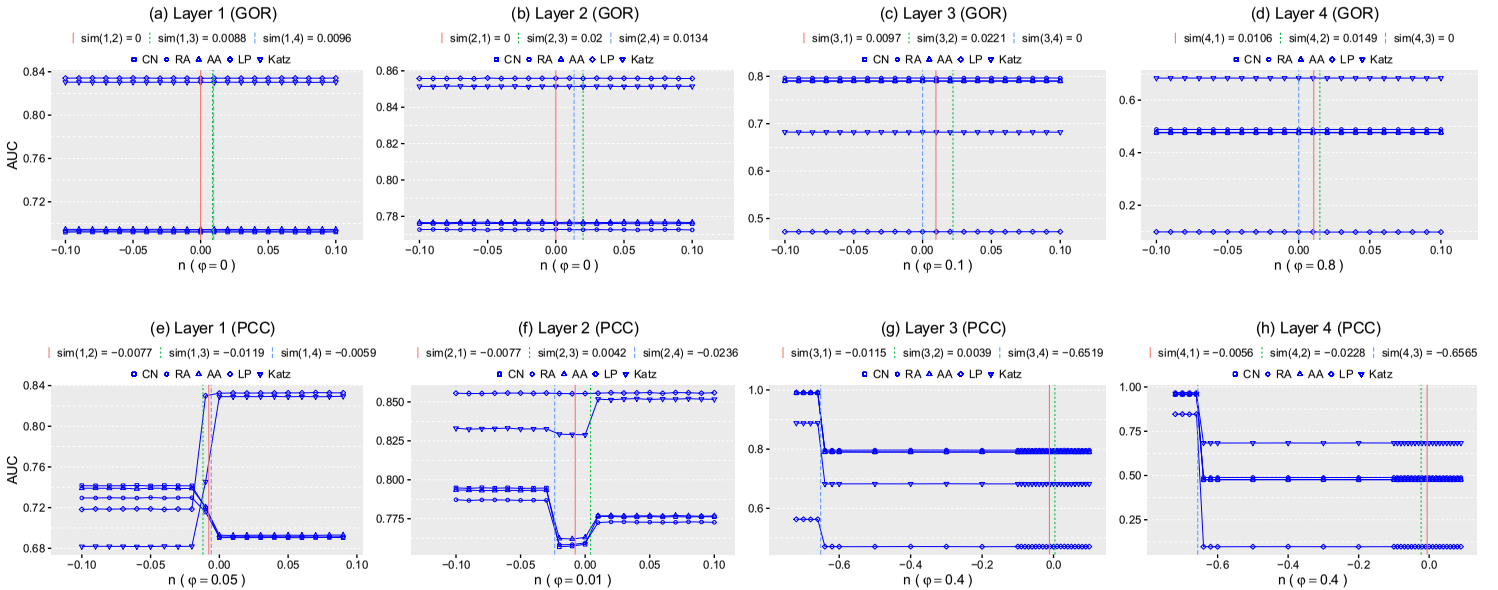


Fig. 7. (Color online) Comparison of the prediction performance of each layer in the Yeast network. Each subfigure corresponds to the performance with GOR or PCC in a layer over 20 independent experiments. The blue curves are the prediction performance based on different baseline measures (i.e. CN, RA, AA, LP and Katz). For a predicted layer, $\varphi$ is equal to $\varphi^*_{\mathrm{CN}}$ and the training set ratio $r$ is 0.9. Each vertical line presents the relevance between the predicted layer and other layer on average.

results, i.e. the Eq. (2) only has the first item. Then, the change in prediction performance is explored during this procedure finally. The experimental results on the Vicker and Yeast multiplex networks are plotted in Figs. 6 and 7, respectively (the results of other five networks are shown in Supplementary Material Note 3). Here, we also plot the performance of other baseline measures corresponding $\varphi = \varphi^*_{\mathrm{CN}}$ in these networks.

As shown in the figures, the prediction performance always decreases in most networks when the high relevance layers are removed for the predicted layer. Moreover, the higher the layer relevance is, the more the decrement of prediction performance is. Take layer 2 of Vicker network as an example (see Fig. 6(b)), after removing the contribution of layer 1 ($\mathrm{sim}(2,1) = 0.6402$, which is the value of layer relevance between layers 2 and 1 on average), the performance of NSILR index that is based on the GOR layer relevance decreases from 0.9507 to 0.9379 for the RA baseline measure. Then, its performance decreases from 0.9379 to 0.8787 with the same baseline measure when the contribution of layer 3 is removed ($\mathrm{sim}(2,3) = 0.7514$). In some cases, the prediction performance is even improved when the low relevance layers are removed (see Figs. 7(e) and 7(f)). Therefore, we can conclude that the layer relevance plays a significant role for the NSILR index and the high layer relevance always corresponds to the high prediction performance.

In most cases, the prediction performance with GOR is consistent with that with PCC for NSILR index. However, PCC is more effective than GOR in recognizing the layer pairs that have low relevance, particularly for identifying the negative relevance between layers. Take the Yeast multiplex network as an example (see Fig. 7), when we remove the contributions of all other layers, the performance of NSILR index with GOR never decreases for all predicted layers. Conversely, their performance with PCC significantly decreases when the highly negative relevance layer is removed for the predicted layer, e.g. layers 3 and 4 of this network. It is because the method of PCC can recognize the negative relevance between layers, but the GOR method regards that there is no relevance between these layers (see Figs. 1(m) and 1(n) for details). Therefore, we can conclude that PCC has more advantage than GOR for NSILR index.

## 6. Conclusion

In this paper, considering the relevance between layers in multiplex networks, a link prediction index NSILR that takes into account the intralayer and interlayer information is proposed for a predicted layer. According to the experimental results on seven multiplex networks, for the NSILR index, its prediction performance that depends only on the intralayer information from the predicted layer can be significantly improved by the interlayer information from other layers. Meanwhile, the performance of NSILR index is compared based on different baseline measures.

The results suggest that the interlayer information from other layers is robust in the prediction procedure. In most cases, the interlayer information from other layers is more valuable than the additional information from the single predicted layer. Through analyzing the influence of layer relevance on the prediction performance, we find that the layer relevance plays a significant role in link prediction. The more relevant the layers are, the higher the performance is improved.

For the NSILR index, the methods of GOR and PCC are exploited to measure the layer relevance in multiplex networks. The result suggests that PCC has more advantages than GOR in identifying the low and negative relevance between layers. However, in order to improve the prediction performance, there may be some other ways to measure the relevance between layers. It is an interesting extension to validate the performance via different layer relevance methods for link prediction in multiplex networks.



## Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant No. 21503101), the Natural Science Foundation of Gansu Province, China (Grant No. 1506RJZA223) and the Project-sponsored by SRF for ROCS, SEM (Grant No. SEM[2015]311).

## Appendix A. Baseline Measures

For the NSILR index, we adopt five baseline measures to calculate the node similarity in the single-layer network. Here, we give a brief introduction of these measures as follows.

(1) Common Neighbor (CN) Index.[8] As the simplest measure, this index directly counts the number of overlap neighbors between two nodes $x$ and $y$. It is based on the basic assumption that nodes tend to connect if they have more common neighbors, and is defined as

$$\mathrm{sim}_{xy} = |\Gamma(x) \cap \Gamma(y)|, \tag{A.1}$$

where $\Gamma(x)$ is the set of neighbors of node $x$.

(2) Adamic–Adar (AA) Index.[9] The common neighbors that have high degree are assigned lower weights by this index. It is defined as

$$\mathrm{sim}_{xy} = \sum_{z \in \Gamma(x) \cap \Gamma(y)} \frac{1}{\log k(z)}. \tag{A.2}$$

(3) Resource Allocation (RA) Index.[10] Motivated by the resource allocation process in networks, the high-degree common neighbors is punished more severely than

AA index by this index. It is defined as

$$\mathrm{sim}_{xy} = \sum_{z \in \Gamma(x) \cap \Gamma(y)} \frac{1}{k(z)}. \tag{A.3}$$

(4) Katz Index.[12] As a representative global structural similarity measure, this index sums all paths connecting two nodes and assigns less weights to the longer paths. It is defined as

$$\mathrm{sim}_{xy} = \sum_{l=1}^{\infty} \beta^l \cdot |\mathrm{paths}_{xy}^{\langle l \rangle}| = \beta A_{xy} + \beta^2 (A^2)_{xy} + \beta^3 (A^3)_{xy} + \cdots, \tag{A.4}$$

where $|\mathrm{paths}_{xy}^{\langle l \rangle}|$ is the set of all paths with length $l$ connecting nodes $x$ and $y$, $\beta$ is a free parameter to adjust the weights of different length paths. When $\beta$ is lower than the reciprocal of the maximum of the eigenvalues of adjacent matrix $A$, this index can be written as $S = (I - \beta A)^{-1} - I$.

(5) Local Path (LP) Index.[13] A typical local structural similarity measure, this index takes into account local paths to tradeoff prediction accuracy and computational complexity. It is defined as

$$\mathrm{sim}_{xy} = A^2 + \epsilon A^3, \tag{A.5}$$

where $\epsilon$ is a free parameter, $A^3$ equals the number of different paths with length 3 connecting nodes $x$ and $y$.

## References


1. S. Boccaletti, V. Latora, Y. Moreno, M. Chavez and D.-U. Hwang, *Phys. Rep.* **424**, 175 (2006).
2. D. Liben-Nowell and J. Kleinberg, *J. Am. Soc. Inf. Sci. Technol.* **58**, 1019 (2007).
3. L. Lü and T. Zhou, *Phys. A, Stat. Mech. Appl.* **390**, 1150 (2011).
4. C. V. Cannistraci, G. Alanis-Lobato and T. Ravasi, *Sci. Rep.* **3**(4), 1613 (2013).
5. L. Miao, Q.-M. Zhang, D.-C. Nie and S.-M. Cai, *Phys. A, Stat. Mech. Appl.* **419**, 301 (2015).
6. F. Xie, Z. Chen, J. Shang, X. Feng and J. Li, *Knowl.-Based Syst.* **81**, 148 (2015).
7. D. Lin, An information-theoretic definition of similarity, *Fifteenth International Conference on Machine Learning*, Vol. 98 (Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1998), pp. 296–304.
8. M. E. Newman, *Phys. Rev. E* **64**, 025102 (2001) APS.
9. L. A. Adamic and E. Adar, *Soc. Netw.* **25**, 211 (2003).
10. T. Zhou, L. Lü and Y.-C. Zhang, *Eur. Phys. J. B* **71**, 623 (2009).
11. Z. Liu, Q.-M. Zhang, L. Lü and T. Zhou, *Europhys. Lett.* **96**, 48007 (2011).
12. L. Katz, *Psychometrika* **18**, 39 (1953).
13. L. Lü, C.-H. Jin and T. Zhou, *Phys. Rev. E* **80**, 046122 (2009).
14. P. Zhang, J. Li, E. Dong and Q. Liu, A method of link prediction based on betweenness, in *Computational Social Networks* (Springer, NY, 2015), pp. 228–235.
15. J. Ding, L. Jiao, J. Wu and F. Liu, *Knowl.-Based Syst.* **98**, 200 (2016).
16. B. Yan and S. Gregory, *Phys. Rev. E* **85**, 056112 (2012).

17. A. Cardillo, J. Gómez-Gardeñes, M. Zanin, M. Romance, D. Papo, F. del Pozo and S. Boccaletti, *Sci. Rep.* **3**(2), 1344 (2013).
18. V. Nicosia, G. Bianconi, V. Latora and M. Barthelemy, *Phys. Rev. Lett.* **111**, 058701 (2013).
19. M. Szell, R. Lambiotte and S. Thurner, *Proc. Nat. Acad. Sci.* **107**, 13636 (2010).
20. S. Boccaletti, G. Bianconi, R. Criado, C. I. Del Genio, J. Gómez-Gardeñes, M. Romance, I. Sendiña-Nadal, Z. Wang and M. Zanin, *Phys. Rep.* **544**, 1 (2014).
21. M. Kivelä, A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno and M. A. Porter, *J. Complex Netw.* **2**, 203 (2014).
22. K.-M. Lee, B. Min and K.-I. Goh, *Eur. Phys. J. B* **88**, 1 (2015).
23. B. Zhu and Y. Xia, *Sci. Rep.* **5**, 13707 (2015).
24. S. Gerard and J. M. Michael, *Introduction to Modern Information Retrieval* (McGraw-Hill, NY, 1983).
25. P. Jaccard, *Etude Comparative de la Distribution Florale dans une Portion des Alpes et du Jura* (Impr. Corbaz, 1901).
26. E. A. Leicht, P. Holme and M. E. Newman, *Phys. Rev. E* **73**, 026120 (2006).
27. A. Arenas, A. Diaz-Guilera and C. J. Pérez-Vicente, *Phys. D, Nonlinear Phenom.* **224**, 27 (2006).
28. H. Liao, A. Zeng and Y.-C. Zhang, *Phys. A, Stat. Mech. Appl.* **436**, 216 (2015).
29. S. Zeng, *Phys. A, Stat. Mech. Appl.* **443**, 537 (2016).
30. A.-L. Barabási and R. Albert, *Science* **286**, 509 (1999).
31. F. Li, J. He, G. Huang, Y. Zhang, Y. Shi and R. Zhou, *Knowl.-Based Syst.* **89**, 669 (2015).
32. M. Pujari and R. Kanawati, *NHM* **10**, 17 (2015).
33. D. Hristova, A. Noulas, C. Brown, M. Musolesi and C. Mascolo, arXiv:1508.07876.
34. S. Sharma and A. Singh, An efficient method for link prediction in complex multiplex networks, in *2015 11th Int. Conf. Signal-Image Technology & Internet-Based Systems (SITIS)*, 2015.
35. S. Sharma and A. Singh, *Comput. Soc. Netw.* **3**, 7 (2016).
36. G. Bianconi, *Phys. Rev. E* **87**, 062806 (2013).
37. F. Battiston, V. Nicosia and V. Latora, *Phys. Rev. E* **89**, 032804 (2014).
38. V. Nicosia and V. Latora, *Phys. Rev. E* **92**, 032805 (2015).
39. V. Gemmetto and D. Garlaschelli, *Scientific Rep.* **5**, 9120 (2015).
40. K.-M. Lee, J. Y. Kim, W.-k. Cho, K.-I. Goh and I. Kim, *New J. Phys.* **14**, 033027 (2012).
41. M. Vickers and S. Chan, *Melbourne: Victoria Institute of Secondary Education* (1981).
42. J. Coleman, E. Katz and H. Menzel, *Sociometry* **20**, 253 (1957).
43. E. Lazega, *The Collegial Phenomenon: The Social Mechanisms of Cooperation among Peers in a Corporate Law Partnership* (Oxford University Press on Demand, 2001).
44. T. A. Snijders, P. E. Pattison, G. L. Robins and M. S. Handcock, *Sociol. Methodol.* **36**, 99 (2006).
45. M. Magnani, B. Micenkova and L. Rossi, arXiv:1303.4986.
46. B. L. Chen, D. H. Hall and D. B. Chklovskii, *Proc. Nat. Acad. Sci. USA* **103**, 4723 (2006).
47. M. De Domenico, M. A. Porter and A. Arenas, *J. Complex Netw.* **3**(2), 159 (2015), Oxford University Press, 10.1093/comnet/cnu038.
48. M. Costanzo *et al.*, *Science* **327**, 425 (2010).
49. J. A. Hanley and B. J. McNeil, *Radiology* **143**, 29 (1982).