



Network Reconstruction in Terms of the Priori Structure Information

Jia-Qi Fu¹, Qiang Guo¹, Kai Yang² and Jian-Guo Liu^{3,4*}

¹Research Center of Complex Systems Science, University of Shanghai for Science and Technology, Shanghai, China, ²College of Information Engineering, Yangzhou University, Yangzhou, China, ³Institute of Accounting and Finance, Shanghai University of Finance and Economics, Shanghai, China, ⁴Shanghai Engineering Research Center of Finance Intelligence, Shanghai University of Finance and Economics, Shanghai, China

In this paper, we investigate the reconstruction of networks based on priori structure information by the Element Elimination Method (EEM). We firstly generate four types of synthetic networks as small-world networks, random networks, regular networks and Apollonian networks. Then, we randomly delete a fraction of links in the original networks. Finally, we employ EEM, the resource allocation (RA) and the structural perturbation method (SPM) to reconstruct four types of synthetic networks with 90% priori structure information. The experimental results show that, comparing with RA and SPM, EEM has higher indices of reconstruction accuracy on four types of synthetic networks. We also compare the reconstruction performance of EEM with RA and SPM on four empirical networks. Higher reconstruction accuracy, measured by local indices of success rates, could be achieved by EEM, which are improved by 64.11 and 47.81%, respectively.

OPEN ACCESS

Edited by:

Mahdi Jalili,
RMIT University, Australia

Reviewed by:

Francisco Wellington Lima,
Federal University of Piauí, Brazil
Ke Hu,
Xiangtan University, China

*Correspondence:

Jian-Guo Liu
liujg004@ustc.edu.cn

Specialty section:

This article was submitted to
Social Physics,
a section of the journal
Frontiers in Physics

Received: 29 June 2021

Accepted: 27 July 2021

Published: 11 August 2021

Citation:

Fu J-Q, Guo Q, Yang K and Liu J-G
(2021) Network Reconstruction in
Terms of the Priori
Structure Information.
Front. Phys. 9:732835.
doi: 10.3389/fphy.2021.732835

Keywords: network reconstruction, element elimination method, priori structure information, time-series information, evolutionary game

1 INTRODUCTION

Reconstructing a network based on priori structure information has attracted lots of attention for the network science [1]. Prior information about the connectivity patterns or potential interactions of the networks are accessible via public database [2, 3], high-throughput experiments [4], or data mining of interaction knowledge [5–7]. A wide diversity of methods based on priori structure information have been developed for the problem of network reconstruction [1, 8, 9]. Among various models, a few reconstruction models would provide a reliable estimate of a network's structure with priori structure information. Link prediction is a typical method which uses accessible structure to estimate the likelihood of existence of unobserved links or identifies spurious links in a network [10, 11]. The unknown structure of a network is then reconstructed by link prediction. A few link prediction models are validated in both synthetic networks and empirical networks, which are local similarity indices [12–14], maximum likelihood methods [11, 15] and methods based on predictability [16, 17].

The other method uses accessible structure information to reconstruct a class of networks with evolutionary games [18, 19]. Such model, known as compressive sensing reconstruction model (CSR), is initially proposed to solve the problems of global network reconstruction [20–22]. The CSR method provides theoretical framework to dealing with networks purely from measured time-series information. To reconstruct a network with N nodes, the CSR method reconstructs the adjacent matrix column by column and each column is a vector with N elements [23, 24]. Contrary to the CSR method, the adjacent matrix is reconstructed by the Element Elimination Method (EEM) in a similar fashion, but the number of elements in different column might be N_i ($N_i \leq N$, $i = 1, 2, \dots, N$) because EEM initially eliminates coupling nodes based on priori structure information. Exploiting the natural

sparsity of the vectors, the pioneering work has applied EEM to achieve a successful reconstruction in scale-free networks with a small fraction of hubs [25]. However, in many cases, examples of real-world networks are not characterized by scale-free [26], i.e., the collaboration network of film actors [27, 28], the neural network of the worm *Caenorhabditis elegans* [26], the power grid of the western United States [29, 30], and drug trafficking network [31], et al. In addition, unique structure could be observed in world airline networks [32, 33] and Apollonian networks [34–36], which are characterized by scale-free and also satisfies basic features of small-world. EEM for reconstructing networks characterized by other features has not been fully explored. We are interested in, to achieve a successful reconstruction, the detailed amount of time-series information required for EEM in spite of the priori structure information. This motivates us to investigate the application of EEM to other networks characterized by different features.

In this paper, we investigate the reconstruction of general networks, which are characterized by four types of synthetic networks as small-world networks, random networks, regular networks and Apollonian networks. Typically, the reconstruction accuracy of EEM is evaluated on four types of networks. We will show the performance of EEM, characterized by low information requirements and high reconstruction accuracy. Experiments on four synthetic networks demonstrate that comparing with the resource allocation (RA) [12] and the structural perturbation method (SPM) [16], EEM can effectively enhance the reconstruction accuracy. Further, three local indices of success rates demonstrate that the reconstruction accuracy obtained by EEM when reconstructing three separately local structure in a network is close. In addition, experiments on four empirical networks demonstrate that EEM outperforms RA and SPM. Compared with RA and SPM, EEM has higher reconstruction accuracy, measured by local indices of success rates, which are improved by 64.11 and 47.81%, respectively.

2 METHODS AND MODELS

2.1 The Procedure of the Network Reconstruction

Uncovering a network's structure has many potential applications so that we can assess the system's resilience [37–39], understand the dynamical mechanisms [40], identify significant nodes in a network [41, 42], detect community structure [43], locate diffusion sources Hu et al. [44, 45], and analyze the networks' properties [46–48]. In this paper, an Element Elimination Method (EEM) [25] is employed to reconstruct the structure of networks. We then give the illustration of the procedures of employing EEM to reconstruct synthetic networks: 1) Generate synthetic networks. 2) Extract time-series information from observed data. 3) Reconstruct the networks with EEM. Noting that the adjacent relationships between nodes in the network are sparse and would not change over time, we could explore the casual relationships between nodes' time-series information. Consequently, we could uncover the unknown link set E^P of the networks by EEM based on priori link set E^T .

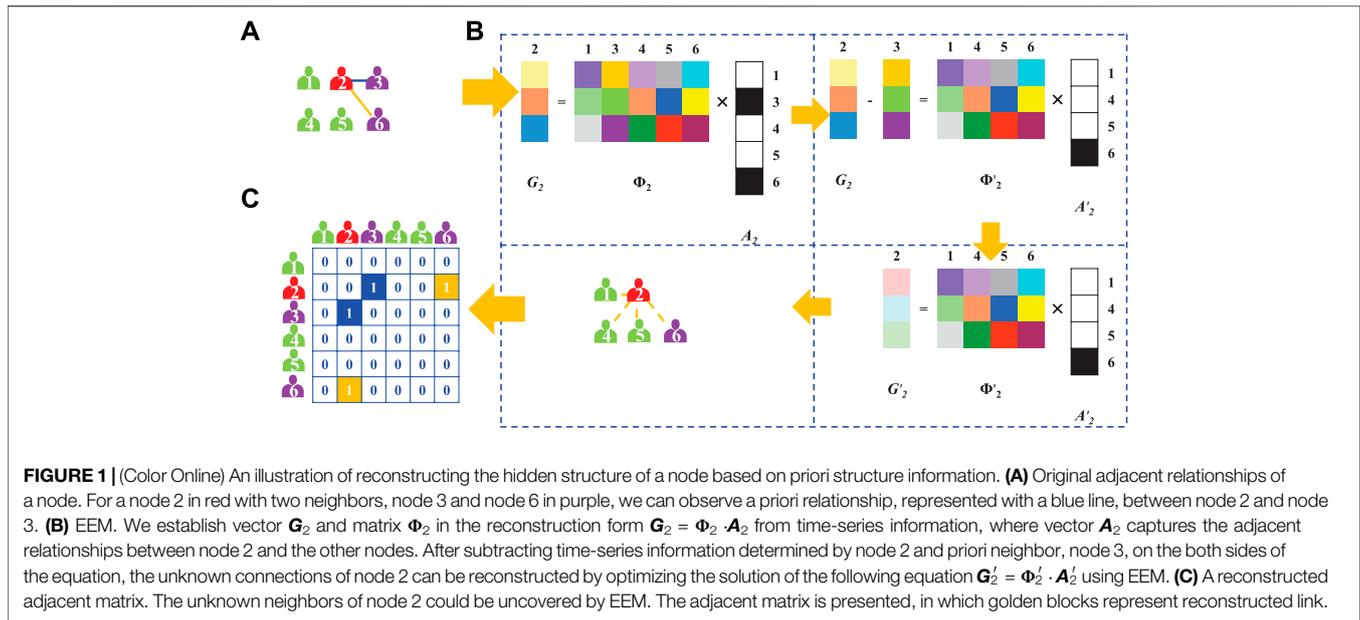
As illustrated in **Figure 1**, a procedure of network reconstruction is presented. Supposing the relationships between node 2 and other 5 nodes should be reconstructed, and only one adjacent relationship (a blue line in **Figure 1A**) is known. However, we are confused about which one is the original network from vastly different networks with possible connective relationships. Simultaneously, the network is evolving over the time, and a few time-series information of nodes' strategies and payoffs could be obtained. We then build a model to bridge node 2's strategies and its payoffs, as **Figure 1B** illustrated. Consequently, we can use EEM to reconstruct the network's structure and obtain the adjacent relationships as shown in **Figure 1C**.

2.2 Generation of Synthetic Network

In order to evaluate the reconstruction performance of EEM in small-world networks and networks characterized by other features, we generate four types of synthetic networks. Noting that small-world network is a model of network that can be tuned between random network and regular network [26], we also consider the networks when their connection topology is assumed to be completely regular or completely random. Besides, the performance on the Apollonian networks by EEM has seldom been evaluated. Then, we generate four types of synthetic networks which are small-world networks, random networks, regular networks and Apollonian networks. The precedent findings indicate that the assortative coefficient has a direct influence on the accuracy of network reconstruction [49]. Therefore, some statistical properties have to be tuned when the networks are generated.

Supposing a network is composed of N nodes and $|E|$ links. To minimize the influence from different network structure, we fix a default mean assortative coefficient $\langle r \rangle$ for three types of synthetic networks, excluding Apollonian networks. Given wiring rules between nodes, we could generate vastly different networks with the given number of nodes N . Initially, the generated synthetic networks should have sufficient links that the total number of links of the network should exceed the number of links $|E|$. Then we randomly delete some of the links so that the number of the residual links is equal to $|E|$. In this way, the generated synthetic networks would have N nodes and $|E|$ links. We select one network from the synthetic network set whose mean assortative coefficient is close to the value of default $\langle r \rangle$ (the absolute error is less than 10^{-3}). The other types of synthetic networks are generated by another wiring rules in a similar way. Actually, synthetic networks generated whose statistical properties are close to default value are limited. On the other hand, the generation procedure of the regular network and the Apollonian network results in merely one realization of the synthetic networks. In this paper, each synthetic network has performed only one realization for the experiments.

Due to privacy or confidentiality issues, the complete structure of a network is not accessible. In addition, it is an impossible mission for us to record nodes' complete time-series information. In spite of the difficulties, some priori information about the adjacent relationships between a few nodes, and discrete records of nodes' time-series information might be available. Despite the



limited information, the connective relationships between nodes has a direct effect on the individual node, which contributes to node's attitude or selection in the next time. The dependence from the network's structure on nodes' interactions provide information for us to utilize the time-series information of nodes to describe the adjacent relationships behind them [24, 50].

2.3 The Model of the Evolutionary Game

The main challenge lies in that the structure of the network is inaccessible, also in that merely limited nodes' time-series information is available. Since the time-series information is closely related to the connective relationships between nodes, we can reconstruct the unknown structure from the limited time-series information.

We use an evolutionary game model, the Prisoner Dilemma Game (PDG) model, to describe the nodes' dynamics [51–53]. In each round of the game, the nodes usually weigh the benefits against the risks and selects a strategy. Here, we use $\mathbf{SY}_i(t)$ to define the strategy of node i . We denote vector $\mathbf{SY}_i(t) = (1,0)^T$ to represent a cooperation strategy, while we denote $\mathbf{SY}_i(t) = (0,1)^T$ to represent a defection strategy. Here, T stands for 'transpose'.

When node i and node j trigger a game, the payoff of node i is dependent on both two nodes' strategies and a uniform payoff matrix \mathbf{P} , which is defined as:

$$\mathbf{P} = \begin{pmatrix} 1 & 0 \\ b & 0 \end{pmatrix} \quad (1)$$

where b ($1 < b < 2$) is a parameter characterizing the volume of payoff when node i select a defection strategy. In the t round, node i would play with all its different neighbors with the same strategy. When node i encounters a neighbor j , node i would gain payoff from node j as:

$$F_{ij}(t) = \mathbf{SY}_i^T(t) \cdot \mathbf{P} \cdot \mathbf{SY}_j(t). \quad (2)$$

In the same round, node i 's total payoffs \mathbf{G}_i would be calculated, and it is the sum of the payoffs from all node i 's neighbors.

In a new round, node i would attempt to maximize its payoffs by updating its strategy. According to Fermi rule [54], node i randomly select a node j from its neighbors after t round. In $t + 1$ round, node i would then adopt node j 's strategy with the probability

$$W(\mathbf{SY}_i(t+1) \leftarrow \mathbf{SY}_j(t)) = \frac{1}{1 + \exp[(TG_i(t) - TG_j(t))/\kappa]}, \quad (3)$$

where $TG_i(t)$ is node i 's cumulative payoffs from 1 to t round. $TG_j(t)$ is similarly defined. Parameter κ characterizes node's rationality when it update strategies. Parameter $\kappa = 0$ corresponds to rational selection behavior of nodes.

Since game occurs among connected nodes, the information of the adjacent relationships between nodes are hidden in their dynamical records of strategies or payoffs in the game. Then we can utilize the information to uncover a networks' structure when we collect the time-series information about the strategies and payoffs of nodes. When we reconstruct a certain network, the limited time-series information is usually presented in a random sample of sufficient time-series information.

2.4 Element Elimination Method

Given limited time-series information of nodes, an EEM could be applied to reconstruct a network based on priori structure information. EEM is a variant of the CSR method, which utilizes priori structure information to exclude the priori connective relationships before reconstruction. Suppose that the relationships between nodes in a certain network can be represented by an adjacency matrix \mathbf{A} with dimensions $N \times N$, where N is the number of nodes in the network. EEM decomposes the process of reconstructing the entire network into many subnetwork recovery

TABLE 1 | The statistical properties of four synthetic networks.

Networks	N	$ E $	$\langle k \rangle$	$\langle r \rangle$	$\langle C \rangle$	$\langle D \rangle$
WS network	120	480	8	-0.05	0.50	3.38
RM network	120	480	8	-0.05	0.10	2.53
RG network	120	480	8	NAN	0.64	7.94
AP network	124	366	5.90	-0.27	0.81	2.57
WS network	250	1,000	8	-0.05	0.50	4.07
RM network	250	1,000	8	-0.05	0.10	Inf
RG network	250	1,000	8	NAN	0.64	16.06
AP network	367	1,095	5.97	-0.21	0.82	2.96

problems, and the network structure, namely, the adjacency matrix \mathbf{A} , is reconstructed column by column [55,56]. An adjacency vector \mathbf{A}_i of a node is used to describe the adjacent relationships between node i ($i = 1, 2, \dots, N$) and the other $N - 1$ nodes in the network, which contains no loop. The adjacency vector $\mathbf{A}_i = (a_{i1}, a_{i2}, \dots, a_{i,i-1}, a_{i,i+1}, \dots, a_{iN})^T$ with element $a_{ij} = 1$ when node i and node j are connected, and $a_{ij} = 0$ otherwise. Suppose that N_i ($N_i \leq N - 1$) nodes in the adjacency vector \mathbf{A}_i have undetermined relationships with node i . EEM is employed to find out node i 's ($i = 1, 2, \dots, N$) direct neighbors from N_i possible nodes, namely a shorter adjacency vector $\mathbf{A}_i = (a'_{i1}, a'_{i2}, \dots, a'_{iN_i})^T$ of node i ($i = 1, 2, \dots, N$).

The training set E^T sheds light on the priori neighbor set Γ_i^K of node i , which contains $(N - N_i - 1)$ nodes. Then we could calculate the sum of payoffs $\mathbf{G}'_{\Gamma_i^K}$ of node i obtained from the priori neighbors in neighbor set Γ_i^K according to Eq. 2. Subtracting payoffs $\mathbf{G}'_{\Gamma_i^K}$ from \mathbf{G}_i , we obtain payoffs \mathbf{G}'_i of node i . The payoffs \mathbf{G}'_i implies the hidden adjacent relationships between node i and N_i other nodes because node i gains payoffs merely from its neighbors.

Most real-world networks are characterized by natural sparsity and the adjacency vector \mathbf{A}_i of node i is sparse, which refers to vector \mathbf{A}_i has only a few nonzero elements (i.e. $a_{ij} = 1$). Noting that the value of each element in node i 's priori adjacency vector $\mathbf{A}_{\Gamma_i^K}$ is 1, vector \mathbf{A}'_i would still be sparse because the number of zero elements has not been changed but the number of nonzero elements has decreased when we remove the priori adjacency vector $\mathbf{A}_{\Gamma_i^K}$ from vector \mathbf{A}_i . The sparsity of \mathbf{A}'_i makes EEM applicable. Initially, the nodes' strategies and payoffs are recorded in discrete round t_1, t_2, \dots, t_M . Since new payoffs are obtained from the game between node i and N_i nodes, we can build a model as Eq. 4. The sparse vector \mathbf{A}'_i then can be reconstructed by solving the following convex optimization problem [57, 58]:

$$\begin{aligned} \min & \|\mathbf{A}'_i\|_1 \\ \text{s. t. } & \mathbf{G}'_i = \Phi'_i \cdot \mathbf{A}'_i, \end{aligned} \quad (4)$$

where $\|\mathbf{A}'_i\|_1 = \sum_{j=1}^{N_i} |a'_{ij}|$ is the L_1 norm of vector \mathbf{A}'_i . The available dynamical payoffs of node i can be expressed by $\mathbf{G}'_i = (G'_i(t_1), G'_i(t_2), \dots, G'_i(t_M))^T$. The payoffs of node i obtained from the corresponding nodes in limited rounds can be expressed by an $M \times N_i$ sensing matrix Φ'_i ($M \ll N_i$). In particular, we write $\Phi'_i =$

$$\begin{pmatrix} F_{i1}(t_1) & F_{i2}(t_1) & \dots & F_{iN_i}(t_1) \\ F_{i1}(t_2) & F_{i2}(t_2) & \dots & F_{iN_i}(t_2) \\ \vdots & \vdots & \dots & \vdots \\ F_{i1}(t_M) & F_{i2}(t_M) & \dots & F_{iN_i}(t_M) \end{pmatrix}$$

TABLE 2 | The statistical properties of four empirical networks.

Networks	N	$ E $	$\langle k \rangle$	$\langle r \rangle$	$\langle C \rangle$	$\langle D \rangle$
FWMW [59]	97	1,446	29.81144	-0.1506	0.4683	1.6929
FWW [59]	128	2075	32.4219	-0.1117	0.3346	1.7763
Jazz musicians [60]	198	2,742	27.6970	0.0202	0.6175	2.2530
<i>C. elegans</i> [26]	297	2,148	14.4646	-0.1632	0.2924	2.4553

The elements in matrix Φ'_i could be calculated using the formula shown in Eq. 2. According to Eq. 4, we could obtain adjacency vector $\mathbf{A}'_i = (a'_{i1}, a'_{i2}, \dots, a'_{iN_i})^T$ by solving the convex optimization problem. We could obtain the complete adjacency vector $\mathbf{A}_i = (a_{i1}, a_{i2}, \dots, a_{iN})^T$ by combining the reconstructed vector \mathbf{A}'_i and the priori neighbor set $\mathbf{A}_{\Gamma_i^K}$ of node i . In a similar fashion, the neighbor-connection vectors of all the other nodes can be obtained, yielding the network's adjacency matrix $\mathbf{A} = (\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_N)$.

3 EXPERIMENTAL RESULTS

3.1 Datasets

In order to understanding the performance of EEM in reconstructing the synthetic networks, the experiments are conducted in four types of networks. The basic statistical properties of the synthetic networks are presented in Table 1. N and $|E|$ are the number of nodes and links. $\langle k \rangle$ is the mean degree, $\langle r \rangle$ is the mean assortative coefficient, $\langle C \rangle$ is the mean clustering coefficient, and $\langle D \rangle$ is the mean shortest distance. Here, we use abbreviation WS, RM, RG and AP to represent small-world networks, random networks, regular networks and Apollonian networks, respectively.

We assume that the strategies and payoffs of each node in a certain round t is one piece of time-series information. In the experiments, we use M pieces of accessible time-series information obtained from discrete round t_1 to round t_M to reconstruct different networks. In this paper, we set N , namely the number of nodes in the network, as the maximum value of M . Then we use an index of information sufficiency η ($\eta \equiv M/N$) to represent the size of the time-series information used in the network reconstruction. Intuitively, the time-series information is sufficient when the pieces of the accessible time-series information $M = N$, while the time-series information is insufficient when $0 < M < N$. Correspondingly, the accessible time-series information is sufficient when the index of information sufficiency $\eta = 1$ and the accessible time-series information is insufficient when $0 < \eta < 1$. The reconstruction models are also applied to reconstruct networks with different priori information of the structure, measured by a probability P_s ($0 \leq P_s \leq 1$).

In addition, the performance of EEM is also evaluated in reconstructing the empirical networks. Table 2 shows the basic statistical properties of all four networks. These networks are chosen because they are characterized by large clustering coefficient and short distance.

3.2 Metrics

To test the EEM's accuracy, the original existent link set, \mathbf{E} , are randomly divided into two parts: the priori set E^T , and the probe

set E^P . Clearly, $E = E^T \cup E^P$ and $E^T \cap E^P = \emptyset$. In this paper, the priori set always contains P_s of links, and the remaining $1 - P_s$ of links constitute the probe set. We apply four standard indices to quantify the reconstruction accuracy: the success rates of existent links SR , the success rates of nonexistent links SN [24], precision PRE [61, 62] and the area under the receiver operating characteristic curve AUC [63] are applied. In addition, we apply local indices of success rates in the experiments.

Both the success rates of existent links SR and the success rates of nonexistent links SN estimate the similarity of the reconstructed networks and the original networks. The success rates of existent links SR denotes the ratio of the number of links reconstructed by the reconstruction models to the number of real existent links in the network. The success rates of nonexistent links SN denotes the ratio of the number of nonexistent links distinguished by the reconstruction models to the number of real nonexistent links in the network. We obtain

$$SR = \frac{1}{N} \sum_{i=1}^N \frac{|\Gamma_{io} \cap \Gamma_{ir}|}{|\Gamma_{io}|} \quad (5)$$

$$SN = \frac{\frac{1}{N} \sum_{i=1}^N |\bar{\Gamma}_{io} \cap \bar{\Gamma}_{ir}|}{|\bar{\Gamma}_{io}|} \quad (6)$$

where Γ_{io} and Γ_{ir} denote real neighbor set of node i and neighbor set of node i reconstructed by the reconstruction models, respectively. $|\cdot|$ denotes the number of elements in a set \cdot . $\bar{\Gamma}_{io}$ and $\bar{\Gamma}_{ir}$ are the supplementary set of set Γ_{io} and Γ_{ir} . Each node in set $\bar{\Gamma}_{io}$ is not adjacent to node i . Correspondingly, each node in reconstructed set $\bar{\Gamma}_{ir}$ is not adjacent to node i . A successful reconstruction is achieved when the success rates of existent links SR ($0 \leq SR \leq 1$) and the success rates of nonexistent links SN ($0 \leq SN \leq 1$) are close to the value of 1.

Precision PRE is defined as the ratio of existent links reconstructed by models to the number of the whole unknown existent links. In our case, to calculate precision we need to rank all the unknown links in decreasing order according to existent possibilities computed by reconstruction models. Then we focus on the top- L (here $L = |E^P|$) links. If there are H links successfully reconstructed, then

$$PRE = \frac{H}{L} \quad (7)$$

The area under the receiver operating characteristic curve AUC evaluates the reconstruction models' performance according to the whole unknown link list. Provided the existent possibility of all unknown links, AUC can be interpreted as the probability that a randomly chosen unknown existent link is given a higher existent possibility than a randomly chosen nonexistent link. In the implementation, the value of AUC is calculated with a function `perfcurve` by Matlab.

Clearly, a higher value of the success rates of existent links SR , the success rates of nonexistent links SN , precision PRE or the area under the receiver operating characteristic curve AUC means a higher reconstruction accuracy. We conduct 50 times independent simulation for averaging the indices of reconstruction accuracy as the mean success rates of existent

links $\langle SR \rangle$, the mean success rates of nonexistent links $\langle SN \rangle$, the mean precision $\langle PRE \rangle$ and the mean area under the receiver operating characteristic curve $\langle AUC \rangle$.

To understand the reconstruction performance of EEM when reconstructing local structure of the network divide the structure of each type of network into separately local structure. Supposing that the roles of nodes in the network are leaders, brokers and peripheral executors. We denote leaders are nodes with small degrees and the number of leaders in each type of network is 6. In addition, the subnetwork composed of leaders is a connected subgraph. Then brokers are nodes which are connected with leaders, and the residual nodes are peripheral executors. The sets of leaders, brokers and peripheral executors are not overlapped. We use letters L , B and P to represent the adjacent relationships between leaders, the adjacent relationships between leaders and brokers, and the adjacent relationships among peripheral executors and brokers, respectively. Then, we could obtain the success rates of existent links of each local structure normalized by the number of real existent links $|\Gamma_{io}|$ of the network.

$$SR_{Lr} = \frac{1}{N} \sum_{i=1}^N \frac{|\Gamma_{Lio} \cap \Gamma_{Lir}|}{|\Gamma_{io}|} \quad (8)$$

$$SR_{Br} = \frac{1}{N} \sum_{i=1}^N \frac{|\Gamma_{Bio} \cap \Gamma_{Bir}|}{|\Gamma_{io}|} \quad (9)$$

$$SR_{Pr} = \frac{1}{N} \sum_{i=1}^N \frac{|\Gamma_{Pio} \cap \Gamma_{Pir}|}{|\Gamma_{io}|} \quad (10)$$

The sum of three local success rates of existent links is equal the global success rates of existent links.

$$SR = SR_{Lr} + SR_{Br} + SR_{Pr} \quad (11)$$

Correspondingly, the maximum of three local success rates of existent links would be

$$SR_{Lo} = \frac{1}{N} \sum_{i=1}^N \frac{|\Gamma_{Lio}|}{|\Gamma_{io}|} \quad (12)$$

$$SR_{Bo} = \frac{1}{N} \sum_{i=1}^N \frac{|\Gamma_{Bio}|}{|\Gamma_{io}|} \quad (13)$$

$$SR_{Po} = \frac{1}{N} \sum_{i=1}^N \frac{|\Gamma_{Pio}|}{|\Gamma_{io}|} \quad (14)$$

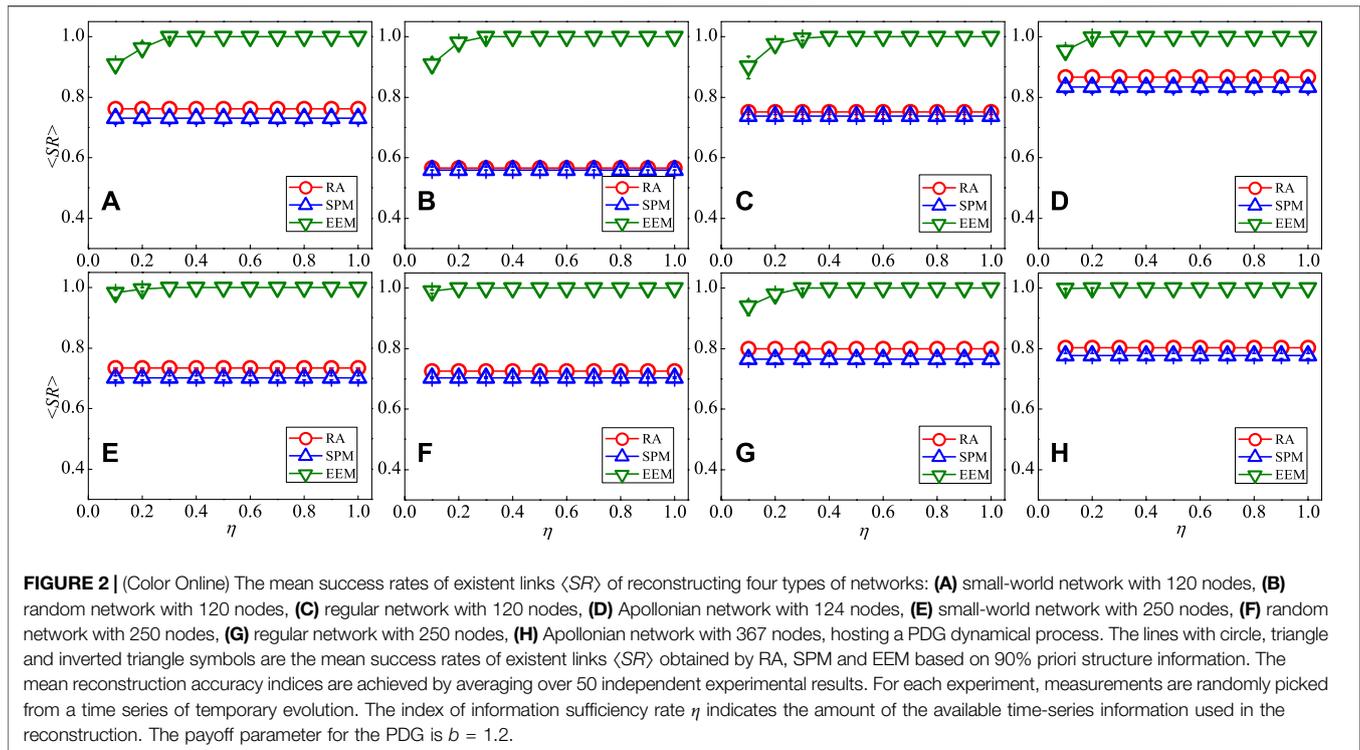
when the original network is successfully reconstructed. To quantify the success rates of three different local structure, we define local indices of success rates as follows:

$$APP_{SRL} = \frac{SR_{Lr}}{SR_{Lo}} \quad (15)$$

$$APP_{SRB} = \frac{SR_{Br}}{SR_{Bo}} \quad (16)$$

$$APP_{SRP} = \frac{SR_{Pr}}{SR_{Po}} \quad (17)$$

Similarly, a higher value of local index of success rates APP_{SRL} , APP_{SRB} , or APP_{SRP} means a higher reconstruction accuracy. We



conduct 50 times independent simulation for averaging the indices of success rates $\langle APP_{SRL} \rangle$, $\langle APP_{SRB} \rangle$ and $\langle APP_{SRP} \rangle$.

3.3 Experimental Results on Synthetic and Empirical Networks

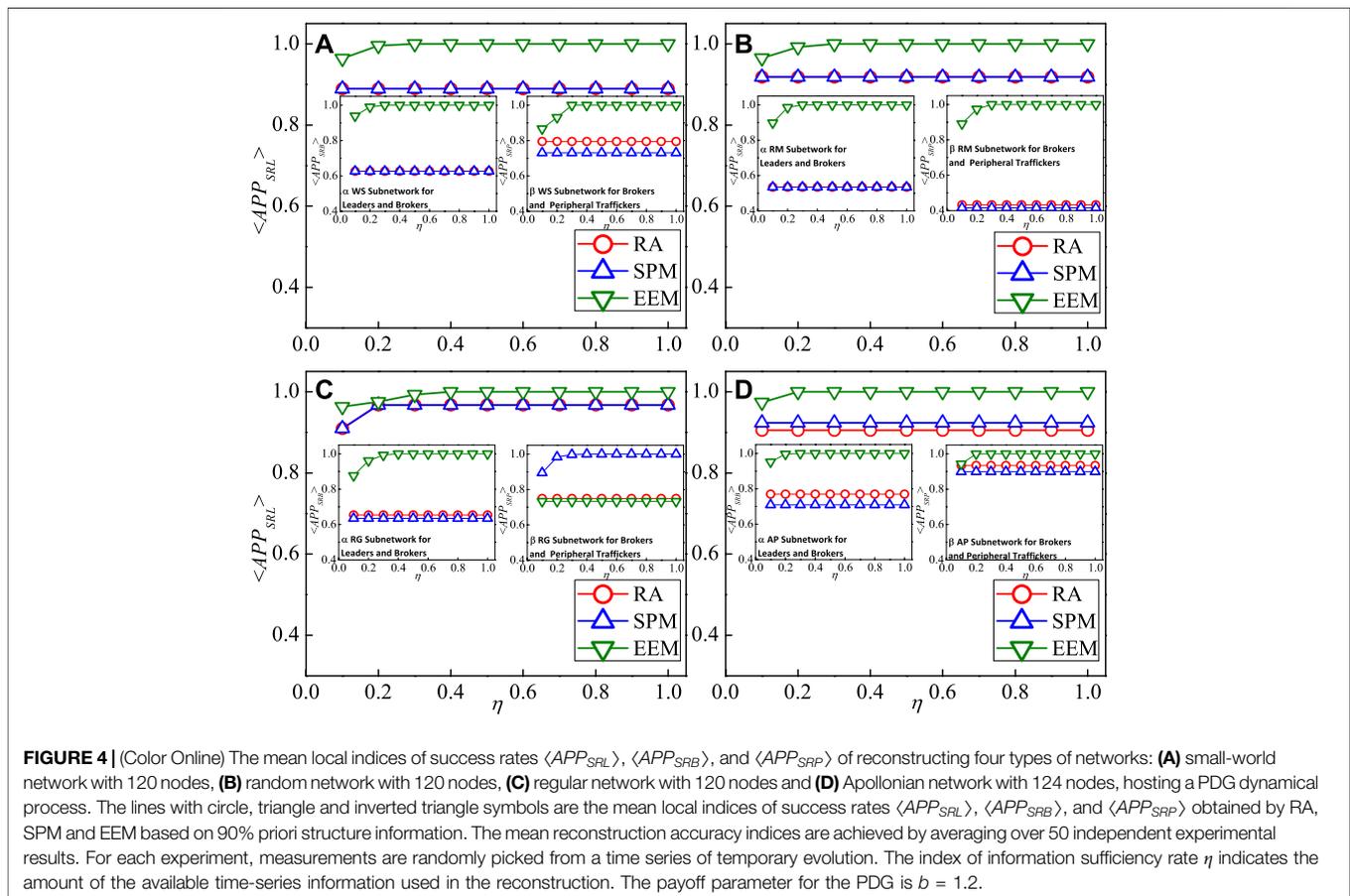
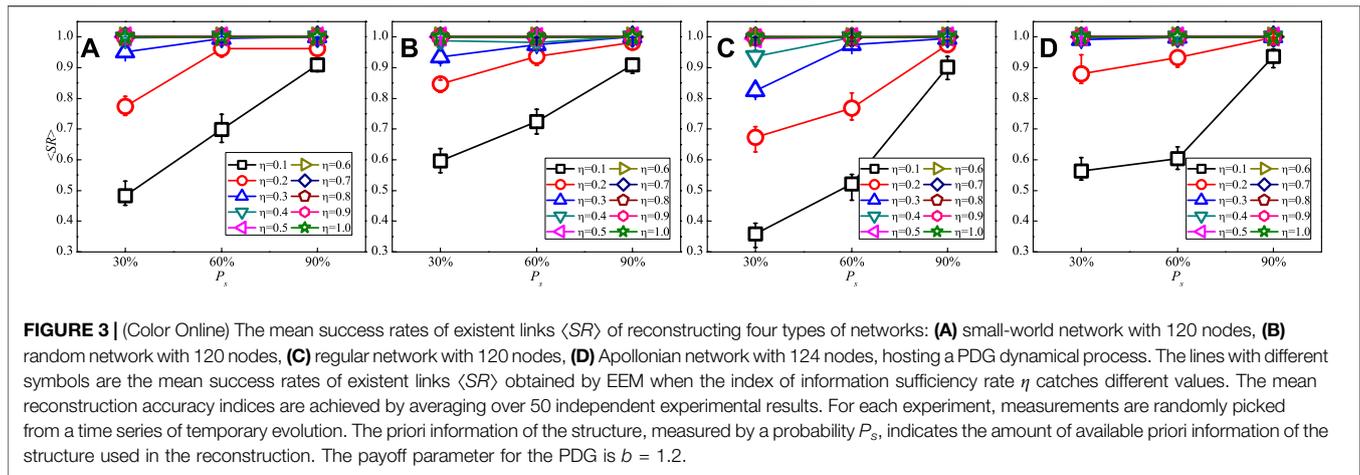
In order to understand the performance of EEM, four types of synthetic networks hosting a PDG dynamical process are considered in our paper. **Figure 2A** depicts the index of reconstruction accuracy for a synthetic small-world network, measured by the mean success rates of existing links $\langle SR \rangle$, based on 90% priori structure information. The mean success rates of existing links $\langle SR \rangle$ increases monotonously when the index of information sufficiency η is varying from 0.1 to 0.4. Especially the mean reconstruction accuracy $\langle SR \rangle$ reaches the maximum value of 1 when the index of information sufficiency $\eta = 0.4$. The increment rate of the mean reconstruction accuracy $\langle SR \rangle$ is 9.97%. Then the mean reconstruction accuracy $\langle SR \rangle$ keeps the value of 1 when the index of information sufficiency η is larger than 0.4. As shown in **Figures 2B–H**, the mean reconstruction accuracy $\langle SR \rangle$ increases monotonously when the index of information sufficiency η is less than 0.4. In addition, the mean reconstruction accuracy $\langle SR \rangle$ reaches 1 for the different types of synthetic networks when the index of information sufficiency η exceeds 0.4.

Moreover, we compare the experimental results between EEM and two link prediction models which are the resource allocation (RA) and the structural perturbation method (SPM). **Figures 2A–H** show that when the index of information sufficiency η is low (i.e., $\eta = 0.1$), the mean success rates of existing links $\langle SR \rangle$

obtained by EEM on small-world networks, random networks, regular networks and Apollonian networks reaches 0.9093, 0.9085, 0.9021, 0.9361, 0.9823, 0.9897, 0.9402 and 0.9982, respectively. Compared with RA and SPM, EEM's mean success rates of existing links $\langle SR \rangle$ are higher, which is improved by at least 8.07 and 12.22% on the networks with 120–124 nodes, respectively. Compared with RA and SPM, EEM's mean success rates of existing links $\langle SR \rangle$ are higher, which is improved by at least 17.53 and 22.81% on the networks with 250–367 nodes, respectively. The experimental results of **Figure 2** indicate that EEM has a well tradeoff that provides high quality reconstruction accuracy while requiring less time-series information.

Intuitively, a network's structure would be accurately reconstructed when more priori information about the structure of the network are presented. **Figure 3** shows the dependence of the values of $\langle SR \rangle$ on probability P_s , the priori information of the structure, where we see that, in the cases of lower index of information sufficiency η ($\eta \leq 0.4$), $\langle SR \rangle$ increases monotonously when the probability P_s increases. On the other hand, the mean success rates of existing links $\langle SR \rangle$ approaches the maximum value of 1 when the index of information sufficiency η is larger than 0.4. In terms of the probability P_s , the highest performance is achieved for the highest P_s . The intuitive reason for the relatively superior performance with the four synthetic networks lies in the sufficiency of the available information of the networks' structure.

In the following, we verify the performance of EEM in local structure of the networks. We divide the structure of each type of network into three separately local structure with subscript L, B, P



for them. **Figure 4A** depicts reconstruction success rate of a small-world network, measured by the mean local index of success rates $\langle APP_{SRL} \rangle$, $\langle APP_{SRB} \rangle$, and $\langle APP_{SRP} \rangle$, based on 90% priori structure information.

As illustrated in the main graph in **Figure 4A**, the mean local index of success rates $\langle APP_{SRL} \rangle$ obtained by EEM is higher than RA or SPM. Especially the mean local index of success rates

$\langle APP_{SRL} \rangle$ obtained by EEM reaches 96.41% when the index of information sufficiency $\eta = 0.1$, while the mean local index of success rates $\langle APP_{SRL} \rangle$ obtained by RA and SPM are both 88.95%. The mean local index of success rates $\langle APP_{SRB} \rangle$ and $\langle APP_{SRP} \rangle$ obtained by EEM are 93.95 and 86.68% when the index of information sufficiency $\eta = 0.1$, as shown in the subgraph (α)-(β) in **Figure 4A**. Correspondingly, the mean local index of

success rates $\langle APP_{SRB} \rangle$ and $\langle APP_{SRP} \rangle$ obtained by RA are 62.67 and 79.51%, $\langle APP_{SRB} \rangle$ and $\langle APP_{SRP} \rangle$ obtained by SPM are 62.67 and 73.06%. The similar experimental results could also be found in the cases of random network, regular network and Apollonian network in **Figures 4B–D**, which indicate that EEM can achieve higher reconstruction accuracy with low time-series information than RA or SPM.

The underlying reason that EEM could obtain higher reconstruction accuracy than RA or SPM might be twofold. Firstly, EEM is applicable to reconstruct networks with sparse connective relationships because Wang et al. developed a paradigm [19, 24, 25] to address the network reconstruction problems and Candès et al. provided the theoretical framework for this paradigm [57, 58]. Both EEM and two link prediction models utilize the identical priori structure information of the network to obtain direct information of the unknown structure. In addition, EEM bridges the relationships between the nodes' payoffs and strategies by virtue of time-series information because the payoffs can merely be obtained from each node's neighbors. Then EEM could extract indirect information of the unknown structure from the above relationships which strengthens the reliability of the experimental results. RA and SPM could also extract valuable indirect information of the unknown structure, but the valuable information still originates from the priori structure information of the network due to lack of a universal theoretical framework.

Secondly, both the reconstruction accuracy of the local structure and the reconstruction accuracy of the global structure obtained by EEM highly consist. As illustrated in **Figure 4**, the absolute error between three mean local index of success rates $\langle APP_{SRL} \rangle$, $\langle APP_{SRB} \rangle$ and $\langle APP_{SRP} \rangle$ obtained by EEM on each network is less than 0.1, which indicates that the reconstruction accuracy on three separate local structure obtained by EEM is almost the same. Consequently, the global reconstruction accuracy and the local reconstruction accuracy highly consist because the global reconstruction accuracy is the linear combination of three mean local index of success rates as: $\langle SR \rangle = SR_{Lo} \cdot \langle APP_{SRL} \rangle + SR_{Bo} \cdot \langle APP_{SRB} \rangle + SR_{Po} \cdot \langle APP_{SRP} \rangle$, where SR_{Lo} , SR_{Bo} and SR_{Po} are constant for each network. The high reconstruction accuracy of three separately local structure contribute to a high reconstruction accuracy of the global structure. We also observe that the reconstruction accuracy on three separate local structure obtained by RA or SPM fluctuates. Especially in the reconstruction experiments on synthetic random networks, the maximum absolute error between three mean local index of success rates obtained by RA or SPM reaches 0.3837. The experimental results indicate that the reconstruction accuracy obtained by RA and SPM is largely dependent on the priori structure information of the network. The reconstruction accuracy of RA or SPM would be high when the local priori structure is consistent with the global structure, and the reconstruction accuracy would be low otherwise.

Finally, we test the results for four empirical networks. As shown in **Table 3**, we reconstruct the network structure by EEM, RA and SPM with 90% priori structure information. The empirical results indicate that four indices of reconstruction accuracy obtained by EEM are higher than RA and SPM for

TABLE 3 | The value of four indices of reconstruction accuracy for four empirical networks.

Network	Accuracy	RA	SPM	EEM	
				$\eta = 0.1$	$\eta = 0.6$
FWMW	SR	0.1833	0.1833	0.8351	0.9996
	SN	0.0016	0.0016	0.9577	0.9996
	PRE	0	0.0003	0.5730	0.9989
	AUC	0.6786	0.6968	0.8836	1
FWFW	SR	0.1575	0.3450	0.8765	1
	SN	0.9742	0.9712	0.9567	0.9999
	PRE	0.0385	0.2043	0.6143	0.9999
	AUC	0.4191	0.7816	0.9375	1
Jazz musicians	SR	0.4488	0.5962	0.8813	1
	SN	0.9902	0.9894	0.9723	0.9999
	PRE	0.2291	0.3486	0.5544	0.9980
	AUC	0.9151	0.9085	0.9807	1
<i>C. elegans</i>	SR	0.4770	0.4621	0.7828	0.9998
	SN	0.9927	0.9948	0.9965	0.9993
	PRE	0.0512	0.0446	0.4834	0.9539
	AUC	0.7817	0.7598	0.9243	0.9999

four empirical networks when the index of information sufficiency rate $\eta = 0.1$. Four indices of reconstruction accuracy obtained by EEM are higher than RA and SPM. Compared with RA, EEM's reconstruction accuracy, measured by the mean success rates of existent links $\langle SR \rangle$, which are improved by 355.54, 456.38, 96.37 and 64.11%, corresponding to FWMW, FWFW, Jazz musicians, Neural network of *C. elegans*. Compared with SPM, EEM's reconstruction accuracy, measured by the mean success rates of existent links $\langle SR \rangle$, which are improved by 355.54, 154.07, 47.81 and 69.38%, corresponding to FWMW, FWFW, Jazz musicians, Neural network of *C. elegans*. Empirical results indicate that the empirical networks reconstructed by EEM are closer to the original networks than those reconstructed by RA and SPM.

3.4 CONCLUSION

In summary, we have investigated the performance of EEM for reconstructing synthetic networks, which are characterized by four types of networks as small-world networks, random networks, regular networks and Apollonian networks, based on priori structure information. The mean success rates of existent links $\langle SR \rangle$ obtained by EEM could achieve at least 0.9021 when the index of information sufficiency η is 0.1. Compared with RA and SPM, EEM has higher mean success rates of existent links $\langle SR \rangle$, which is improved by 8.07 and 12.22% on the networks with 120–124 nodes, respectively. Compared with RA and SPM, EEM has higher mean success rates of existent links $\langle SR \rangle$, which is improved by 17.53 and 22.81% on the networks with 250–367 nodes, respectively. The experimental results also indicate that separately local structure in each type of network could be accurately reconstructed by EEM. In addition, EEM's reconstruction accuracy is also evaluated on four empirical networks. Compared with RA and SPM, EEM has higher mean

success rates of existent links $\langle SR \rangle$, which is improved by 64.11 and 47.81%, respectively. The reason that EEM obtain higher reconstruction accuracy than RA or SPM might lie in that EEM could utilize time-series information to strengthen the reliability of the experimental results and EEM's capability to reconstruct the local structure and the global structure highly consist. The evaluation of EEM on both synthetic networks and empirical networks suggest that EEM is applicable for networks with sparsely connective relationships and it has high reconstruction accuracy by low information requirements.

Although the efficiency of EEM has been measured in reconstructing network's structure with both synthetic networks and empirical networks, there are still a lot of questions to be considered further. For example, the results show that EEM can give remarkably higher reconstruction accuracy on a network hosing a PDG dynamical process, but the performance of EEM has not been validated under another dynamical process. Although EEM could also be extended to cases with large-scale network, the computing time might increase exponentially. In addition, EEM's capability to identify spurious links has not been explored. Noting that EEM can well capture the adjacent relationships from limited information and thus give more accurate reconstruction, such features make EEM appealing to reconstructing general networks with extremely low data requirement. Despite underlying challenges, we will make attempt to continue our research referring to the problems of network reconstruction.

REFERENCES

- Liao JC, Boscolo R, Yang Y-L, Tran LM, Sabatti C, and Roychowdhury VP. Network Component Analysis: Reconstruction of Regulatory Signals in Biological Systems. *Proc Natl Acad Sci* (2003) 100:15522–7. doi:10.1073/pnas.2136632100
- Matys V, Fricke E, Geffers R, Gößling E, Haubrock M, and Hehl R. TRANSFAC(R): Transcriptional Regulation, from Patterns to Profiles. *Nucleic Acids Res* (2003) 31:374–8. doi:10.1093/nar/gkg108
- Keseler IM, Collado-Vides J, Gama-Castro S, Ingraham J, Paley S, and Paulsen IT. Ecocyc: A Comprehensive Database Resource for Escherichia Coli. *Nucleic Acids Res* (2004) 33:D334–D337. doi:10.1093/nar/gki108
- Lee TI, Rinaldi NJ, Robert F, Odom DT, Bar-Joseph Z, and Gerber GK. Transcriptional Regulatory Networks in Saccharomyces Cerevisiae. *Science* (2002) 298:799–804. doi:10.1126/science.1075090
- Dong GG, Wang F, Shekhtmane LM, Danziger MM, Fan JF, and Du RJ. Optimal Resilience of Modular Interacting Networks. *Proc Natl Acad Sci USA*. (2021) 118:e1922831118. doi:10.1073/pnas.1922831118
- Bussemaker HJ, Li H, and Siggia ED. Building a Dictionary for Genomes: Identification of Presumptive Regulatory Sites by Statistical Analysis. *Proc Natl Acad Sci* (2000) 97:10096–100. doi:10.1073/pnas.180265397
- Bussemaker HJ, Li H, and Siggia ED. Regulatory Element Detection Using Correlation With Expression. *Nat Genet* (2001) 27:167–71. doi:10.1038/84792
- Chang C, Ding Z, Hung YS, and Fung PCW. Fast Network Component Analysis (Fastnca) for Gene Regulatory Network Reconstruction From Microarray Data. *Bioinformatics* (2008) 24:1349–58. doi:10.1093/bioinformatics/btn131
- Cugueró-Escofet MÀ, Quevedo J, Alippi C, Roveri M, Puig V, and García D. Model- vs. Data-Based Approaches Applied to Fault Diagnosis in Potable

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

J-QF provided this topic and wrote the paper. QG, KY and J-GL guided, discussed and modified the manuscript. All authors contributed to manuscript and approved the submission version.

FUNDING

This work is supported by the National Natural Science Foundation of China (Grant Nos. 71771152 and 61773248), the National Social Science Fund of China (No.16BJY158), the Major Program of National Fund of Philosophy and Social Science of China (Nos. 20ZDA060 and 18ZDA088), and the Scientific Research Project of Shanghai Science and Technology Committee (Grant No. 19511102202).

ACKNOWLEDGMENTS

The authors acknowledge the valuable discussion with Huan-Mei Qin, Guang Liang, Ren-De Li, Hong-Yi Ding, Shao-Yong Han.

- Water Supply Networks. *J Hydroinformatics* (2016) 18:831–50. doi:10.2166/hydro.2016.218
- Lü L, and Zhou T. Link Prediction in Complex Networks: A Survey. *Physica A: Stat Mech Its Appl* (2011) 390:1150–70. doi:10.1016/j.physa.2010.11.027
- Clauset A, Moore C, and Newman MEJ. Hierarchical Structure and the Prediction of Missing Links in Networks. *Nature* (2008) 453:98–101. doi:10.1038/nature06830
- Zhou T, Lü L, and Zhang Y-C. Predicting Missing Links via Local Information. *Eur Phys J B* (2009) 71:623–30. doi:10.1140/epjb/e2009-00335-8
- Lü L, Jin C-H, and Zhou T. Similarity Index Based on Local Paths for Link Prediction of Complex Networks. *Phys Rev E* (2009) 80:046122. doi:10.1103/PhysRevE.80.046122
- Liben-Nowell D, and Kleinberg J. The Link-Prediction Problem for Social Networks. *J Am Soc Inf Sci* (2007) 58:1019–31. doi:10.1002/asi.20591
- Guimerà R, and Sales-Pardo M. Missing and Spurious Interactions and the Reconstruction of Complex Networks. *Proc Natl Acad Sci* (2009) 106:22073–8. doi:10.1073/pnas.0908366106
- Lü L, Pan L, Zhou T, Zhang Y-C, and Stanley HE. Toward Link Predictability of Complex Networks. *Proc Natl Acad Sci USA* (2015) 112:2325–30. doi:10.1073/pnas.1424644112
- Sun J, Feng L, Xie J, Ma X, Wang D, and Hu Y. Revealing the Predictability of Intrinsic Structure in Complex Networks. *Nat Commun* (2020) 11:1–10. doi:10.1038/s41467-020-14418-6
- Zhang H-F, and Wang W-X. Complex System Reconstruction. *Acta Physica Sinica* (2020) 69:088906. doi:10.7498/aps.69.20200001
- Wang W-X, Lai Y-C, and Grebogi C. Data Based Identification and Prediction of Nonlinear and Complex Dynamical Systems. *Phys Rep* (2016) 644:1–76. doi:10.1016/j.physrep.2016.06.004
- Xu M, Xu C-Y, Wang H, Li Y-K, Hu J-B, and Cao K-F. Global and Partitioned Reconstructions of Undirected Complex Networks. *Eur Phys J B* (2016) 89:1–6. doi:10.1140/epjb/e2016-60956-2

21. Barranca VJ, and Zhou D. Compressive Sensing Inference of Neuronal Network Connectivity in Balanced Neuronal Dynamics. *Front Neurosci* (2019) 13:1101. doi:10.3389/fnins.2019.01101
22. Li R-D, Guo Q, Ma H-T, and Liu J-G. Network Reconstruction of Social Networks Based on the Public Information. *Chaos* (2021) 31:033123. doi:10.1063/5.0038816
23. Shen Z, Wang W-X, Fan Y, Di Z, and Lai Y-C. Reconstructing Propagation Networks With Natural Diversity and Identifying Hidden Sources. *Nat Commun* (2014) 5:1–10. doi:10.1038/ncomms5323
24. Wang W-X, Lai Y-C, Grebogi C, and Ye J. Network Reconstruction Based on Evolutionary-Game Data via Compressive Sensing. *Phys Rev X* (2011) 1: 021021. doi:10.1103/PhysRevX.1.021021
25. Ma L, Han X, Shen Z, Wang W-X, and Di Z. Efficient Reconstruction of Heterogeneous Networks From Time Series via Compressed Sensing. *PLoS One* (2015) 10:e0142837. doi:10.1371/journal.pone.0142837
26. Watts DJ, and Strogatz SH. Collective Dynamics of 'Small-World' Networks. *Nature* (1998) 393:440–2. doi:10.1038/30918
27. Amaral LAN, Scala A, Barthélemy M, and Stanley HE. Classes of Small-World Networks. *Proc Natl Acad Sci* (2000) 97:11149–52. doi:10.1073/pnas.200327197
28. Ravasz E, and Barabási A-L. Hierarchical Organization in Complex Networks. *Phys Rev E* (2003) 67. doi:10.1103/PhysRevE.67.026112
29. Albert R, Albert I, and Nakarado GL. Structural Vulnerability of the North American Power Grid. *Phys Rev E Stat Nonlin Soft Matter Phys* (2004) 69: 025103. doi:10.1103/PhysRevE.69.025103
30. Crucitti P, Latora V, and Marchiori M. A Topological Analysis of the Italian Electric Power Grid. *Physica A: Stat Mech Its Appl* (2004) 338:92–7. doi:10.1016/j.physa.2004.02.029
31. Bright D, Koskinen J, and Malm A. Illicit Network Dynamics: The Formation and Evolution of a Drug Trafficking Network. *J Quant Criminol* (2019) 35: 237–58. doi:10.1007/s10940-018-9379-8
32. Barrat A, Barthélemy M, Pastor-Satorras R, and Vespignani A. The Architecture of Complex Weighted Networks. *Proc Natl Acad Sci* (2004) 101:3747–52. doi:10.1073/pnas.0400087101
33. Verma T, Araújo NAM, Nagler J, Andrade JS, Jr, and Herrmann HJ. Model for the Growth of the World Airline Network. *Int J Mod Phys C* (2016) 27: 1650141. doi:10.1142/S0129183116501412
34. Soares DJB, Andrade JS, Jr, Herrmann HJ, and da Silva LR. Three-Dimensional Apollonian Networks. *Int J Mod Phys C* (2006) 17:1219–26. doi:10.1142/S0129183106009175
35. Andrade RFS, Andrade JS, Jr, and Herrmann HJ. Ising Model on the Apollonian Network With Node-Dependent Interactions. *Phys Rev E* (2009) 79:036105. doi:10.1103/PhysRevE.79.036105
36. Araújo NA, Andrade RFS, and Herrmann HJ. Q-State Potts Model on the Apollonian Network. *Phys Rev E* (2010) 82:046109. doi:10.1103/physreve.82.046109
37. Dong G, Wang F, Shekhtman LM, Danziger MM, Fan J, and Du R. Optimal Resilience of Modular Interacting Networks. *Proc Natl Acad Sci USA* (2021) 118:e1922831118. doi:10.1073/pnas.1922831118
38. Dong G, Fan J, Shekhtman LM, Shai S, Du R, and Tian L. Resilience of Networks With Community Structure Behaves as if Under an External Field. *Proc Natl Acad Sci USA* (2018) 115:6911–5. doi:10.1073/pnas.1801588115
39. Gao J, Barzel B, and Barabási A-L. Universal Resilience Patterns in Complex Networks. *Nature* (2016) 530:307–12. doi:10.1038/nature16948
40. Xu X, Zhang J, and Small M. Superfamily Phenomena and Motifs of Networks Induced From Time Series. *Proc Natl Acad Sci* (2008) 105:19601–5. doi:10.1073/pnas.0806082105
41. Ren Z-M. Age Preference of Metrics for Identifying Significant Nodes in Growing Citation Networks. *Physica A: Stat Mech its Appl* (2019) 513:325–32. doi:10.1016/j.physa.2018.09.001
42. Liu J-G, Ren Z-M, and Guo Q. Ranking the Spreading Influence in Complex Networks. *Physica A: Stat Mech its Appl* (2013) 392:4154–9. doi:10.1016/j.physa.2013.04.037
43. Pan Y, Li D-H, Liu J-G, and Liang J-Z. Detecting Community Structure in Complex Networks via Node Similarity. *Physica A: Stat Mech its Appl* (2010) 389:2849–57. doi:10.1016/j.physa.2010.03.006
44. Hu Z-L, Shen Z, Tang C-B, Xie B-B, and Lu J-F. Localization of Diffusion Sources in Complex Networks With Sparse Observations. *Phys Lett A* (2018) 382:931–7. doi:10.1016/j.physleta.2018.01.037
45. Hu Z-L, Han X, Lai Y-C, and Wang W-X. Optimal Localization of Diffusion Sources in Complex Networks. *R Soc Open Sci* (2017) 4:170091. doi:10.1098/rsos.170091
46. Wang X, Su J, Ma F, and Yao B. Mean First-Passage Time on Scale-free Networks Based on Rectangle Operation. *Front Phys* (2021) 9:238. doi:10.3389/fphys.2021.675833
47. Ren Z-M, Zeng A, and Zhang Y-C. Bridging Nestedness and Economic Complexity in Multilayer World Trade Networks. *Humanit Soc Sci Commun* (2020) 7:1–8. doi:10.1057/s41599-020-00651-3
48. Buldyrev SV, Parshani R, Paul G, Stanley HE, and Havlin S. Catastrophic cascade of Failures in Interdependent Networks. *Nature* (2010) 464:1025–8. doi:10.1038/nature08932
49. Guo Q, Liang G, Fu JQ, Han JT, and Liu JG. Roles of Mixing Patterns in the Network Reconstruction. *Phys Rev E* (2016) 94:052303. doi:10.1103/PhysRevE.94.052303
50. Han X, Shen Z, Wang W-X, Lai Y-C, and Grebogi C. Reconstructing Direct and Indirect Interactions in Networked Public Goods Game. *Sci Rep* (2016) 6: 1–12. doi:10.1038/srep30241
51. Nowak MA, and May RM. Evolutionary Games and Spatial Chaos. *Nature* (1992) 359:826–9. doi:10.1038/359826a0
52. Rong Z, Yang H-X, and Wang W-X. Feedback Reciprocity Mechanism Promotes the Cooperation of Highly Clustered Scale-free Networks. *Phys Rev E* (2010) 82:047101. doi:10.1103/PhysRevE.82.047101
53. Tang Y, Jing M, and Yu Y. Conditional Neutral Reward Promotes Cooperation in the Spatial Prisoner's Dilemma Game. *Front Phys* (2021) 9:79. doi:10.3389/fphys.2021.639252
54. Szabó G, and Tóke C. Evolutionary Prisoner's Dilemma Game on a Square Lattice. *Phys Rev E* (1998) 58:69–73. doi:10.1103/PhysRevE.58.69
55. Wang W-X, Yang R, Lai Y-C, Kovanis V, and Harrison MAF. Time-Series-Based Prediction of Complex Oscillator Networks via Compressive Sensing. *Epl (Europhysics Letters)* (2011) 94:48006. doi:10.1209/0295-5075/94/48006
56. Han X, Shen Z, Wang WX, and Di Z. Robust Reconstruction of Complex Networks From Sparse Data. *Phys Rev Lett* (2015) 114:028701. doi:10.1103/PhysRevLett.114.028701
57. Candès EJ, Romberg J, and Tao T. Robust Uncertainty Principles: Exact Signal Reconstruction from Highly Incomplete Frequency Information. *IEEE Trans Inform Theor* (2006) 52:489–509. doi:10.1109/TIT.2005.862083
58. Candès EJ, and Wakin MB. An Introduction to Compressive Sampling. *IEEE Signal Process Mag* (2008) 25:21–30. doi:10.1109/msp.2007.914731
59. Baird D, Luczkovich J, and Christian RR. Assessment of Spatial and Temporal Variability in Ecosystem Attributes of the St marks National Wildlife Refuge, Apalachee bay, florida. *Estuarine, Coastal Shelf Sci* (1998) 47:329–49. doi:10.1006/ecss.1998.0360
60. Gleiser PM, and Danon L. Community Structure in Jazz. *Adv Complex Syst* (2003) 06:565–73. doi:10.1142/S0219525903001067
61. Zhou T, Ren J, Medo M, and Zhang YC. Bipartite Network Projection and Personal Recommendation. *Phys Rev E Stat Nonlin Soft Matter Phys* (2007) 76: 046115. doi:10.1103/PhysRevE.76.046115
62. Liu JG, Shi K, and Guo Q. Solving the Accuracy-Diversity Dilemma via Directed Random Walks. *Phys Rev E* (2012) 85:016118. doi:10.1103/PhysRevE.85.016118
63. Hanley JA, and McNeil BJ. The Meaning and Use of the Area under a Receiver Operating Characteristic (Roc) Curve. *Radiology* (1982) 143:29–36. doi:10.1148/radiology.143.1.7063747

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Fu, Guo, Yang and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.