

## Research Article

# Roles of Clustering Coefficient for the Network Reconstruction

Jingti Han<sup>1,2</sup> and Changmei Mao<sup>1,2</sup> 

<sup>1</sup>School of Information Management and Engineering, Shanghai University of Finance and Economics, Shanghai 200433, China

<sup>2</sup>Institute of Fintech, Shanghai University of Finance and Economics, Shanghai 200433, China

Correspondence should be addressed to Changmei Mao; changmei146@163.com

Received 4 March 2018; Revised 19 July 2018; Accepted 9 October 2018; Published 24 October 2018

Guest Editor: Ricardo Soto

Copyright © 2018 Jingti Han and Changmei Mao. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

It is important to establish relations between the network reconstruction and the topological dynamical structure of networks. In this article, we quantify the effect for two types of network topologies on the performance of network reconstruction. First, we generate two network modes with variable clustering coefficient based on Holme-Kim model and Newman-Watts small-world model, then we reconstruct the artificial networks by using a novel framework called  $L_1$ -norm minimization algorithm based on a theory called compressive sensing (CS), a framework for recovering sparse signals. The results of the simulation experiment show that the accuracy rate for the network reconstruction is a monotonically increasing function of the clustering coefficient in Holme-Kim model, whereas the opposite occurs in Newman-Watts small-world network. And this yet demonstrates that the larger the network size, the higher the accuracy rate. Moreover, we compare the results of CS with orthogonal matching pursuit (OMP), a greedy algorithm. The results show that the accuracy rate of  $L_1$ -norm minimization method is 10% higher than that of OMP, and OMP yields 1.2 times the computation speed of  $L_1$ -norm minimization. Our work demonstrates that the topological structure of network has influence on the accurate reconstruction and it is helpful for offering proper method for the network reconstruction.

## 1. Introduction

Network reconstruction has attracted much attention for various collective dynamical behaviors [1–3]. The concept of complex network reconstruction is proposed for the first time by Guimerà [4]. It is an inverted problem that can rebuild the existing network to yield estimates of the true network properties and plays significant role in various aspects, such as locating the source of disease [5, 6], identifying hidden information, preventing virus transmission [7, 8], and predicting risk spreading in financial networks.

When dealing with the network reconstruction problem, we should think about the reliability of network data since the accessible data may be fragmentary and limited in consideration of network size. Efficient approaches to solve the reconstruction problem with low data requirements are mainly obtained from methods as Link prediction [9], Bayesian reasoning [10], ODE [11], and so on. Recently, a theory called compressive sensing (CS), used for recovering sparse signal [12, 13], such as image reconstruction,

large-scale sensor-network data processing, and coupled-oscillator networks reconstruction [14], either with time series (continuous-time) data, which has been outstanding in nonlinear dynamics with interactions among nodes [15], or discrete time series, which remains deeper researching [16], can reconstruct a propagation network using small data set information extracted from experiments or observations [17–19].  $L_1$ -norm minimization (Basis Pursuit), which can solve sparse problem, is often used as one kind of algorithm for CS. Under stochastic dynamical process, Shen et al. reconstruct networks from limited time series. Wang et al. [20] uncover interaction networks from small amount of data based on CS. Indeed, such studies proposed a wide range of issues related to accurate network reconstruction for diverse fields. Most of them studied sparse networks in applications from gene-regulatory networks [21], coupled oscillator networks, to social networks [22]. In real world network, human Brain Network has been investigated under CS [23]; in [24] the authors took into account three directed real-world networks, Football, USTop500, and *C. elegans*, and tested two methods

on the reconstruction for the three networks. In [25], CS is exploited for reducing the memory cost in switches and routers. In [26], the authors gather and spread information across a large P2P overlay network efficiently by CS.

Prior studies about network reconstruction mainly focused on analyzing the collective dynamics complex networks with evolutionary game data. Santos et al. [27] study the mechanism of cooperative evolution with data from game theory. Rong et al. [28] investigate, in the networked prisoner's dilemma game, how the degree-mixing schema influences the cooperation. Besides the network games [29, 30], the network synchronization [31], the topological structure [32, 33], and so on are popular research topics in the field of complex system. Specifically, analyzing the dynamic properties of complex system is of great significance. The topological characteristics play important role on the network controllability [34]. Guo et al. [35] investigate the role of the assortative coefficient for the network reconstruction and find that when the assortative coefficient is positive, the  $L_1$ -norm minimization method could regenerate the social network more accurately. However, whether network topology can affect the reconstruction accuracy and choose proper method is important for network reconstruction. Inspired by this, we investigate the effect on the accuracy rate of the network reconstruction by  $L_1$ -norm minimization method with different topological structure of networks measured by variable clustering coefficient among two type network models. Newman and Watts raise small-world network, highlighting their high clustering and shortest average path length [36]. Barabási and Albert (BA) established a network with both growth mechanism and preferential attachment mechanism [37, 38]; the nodes' degree of the system follows the principle of the power-law distribution. The existing classical network model that incorporates both properties is called Holme–Kim model [39]. It is built based on BA model and possesses the scale-free characteristic. In this article, we control the parameter of the model by adjusting the change of the clustering coefficient. To verify the reconstructing algorithm we have mentioned, we model various network structures with variable clustering coefficient and investigate the relation between accuracy and clustering coefficient. At last, we also investigate the performance by using orthogonal matching pursuit (OMP), a greedy algorithm, which shows difference with  $L_1$ -norm minimization.

In this paper, first we generate a series networks by Holme–Kim model and Newman-Watts small-world model, then we reveal the structure of network with evolutionary-game data by using  $L_1$ -norm minimization. Furthermore, we compare our framework with OMP, another method for network reconstruction. The results show that, with the network topological structure changing, our method is 10% higher than that of OMP in terms of success rate (SR), which measures the accuracy of network reconstruction. Specially, Holme–Kim model acts distinctly different among the two algorithms. Meanwhile, OMP has 1.2 times the convergence rate of  $L_1$ -norm minimization. Our work shows the accuracy of prediction for two types of network topologies, with the two methods performing differently, and choosing the proper method is important.

## 2. Methods

We address the mechanism of uncovering two types network topologies with evolutionary PD game time series data based on  $L_1$ -norm minimization, one kind of compressive sensing method. It is carried out through exploring two evolutionary games generating in Holme–Kim networks and Newman-Watts small-world networks; individual interactive information is accessible, then we make an attempt to decode the network structure from measurable data.

*2.1. Reconstruction Processes.* CS is first used for finding solutions to underdetermine linear systems through processing signal to acquire and remodeling signal. The advantage of CS [40–42] is that it can reconstruct a small number of linear measurements of the signal that contain enough information. It often happens on the fact that abundant signals are compressible or sparse when they are expressed in the proper basis and frame. The purpose of CS is mainly aimed at reconstructing vector  $\mathbf{X}$  from  $\mathbf{G}$  ( $\mathbf{X} \in \mathbb{R}^N$ ,  $\mathbf{G}$  is linear measurements); convex optimization problem can be expressed as follows:

$$\begin{aligned} \min \quad & \|\mathbf{X}\|_1 \\ \text{s.t.} \quad & \mathbf{G} = \Phi \cdot \mathbf{X}, \end{aligned} \quad (1)$$

from which the vector  $\mathbf{G} \in \mathbb{R}^M$  denotes the measurement vector and measurement matrix  $\Phi \in \mathbb{R}^{M \times N}$  whose rows have unit length projects data.  $\|\mathbf{X}\|_1 = \sum_{i=1}^N \|\mathbf{X}_i\|$  is the  $L_1$ -norm of vector  $\mathbf{X}$ . As a kind of compression estimates, the calculation speed of  $L_1$ -norm method is slow, the measurement data required is small, and the precision is high. Optimum solution to the convex optimization can be accessible [43]; it has been widely used for network reconstruction problem easily [44]. Considering that the length of an unknown vector is much more than the observable measurements ( $N \gg M$ ) and the number of non-zero elements is less than  $M$ , all these are main advantages of CS. The measurement vector  $\mathbf{G}$  and matrix  $\Phi$  are generated from the evolutionary game progress where accessible data about strategy and payoff for each agent can be got from the dynamic interaction. In general, two game theories are used to investigate cooperation in social networks; one is prisoner's dilemma (PD) game [45], and the other is snowdrift game (SG) [46]. Recently, the prisoner's dilemma (PD) game has been used combined with spatial or topological dimension [47, 48]. PD game has been a framework to exploit cooperation for complex social network among nodes [49]; it can serve for network reconstruction. In fact, the structure of network often can affect the style of cooperation; much research has been done about it. In terms of mechanism for PD game, there exist many ways to deal with cooperation; in [50], it is demonstrated that costly punishment may decrease cooperation sometimes. In general, two strategies  $\mathbf{S}$  belong to an agent: cooperation (C) or defection (D), the strategy matrices for cooperation is  $\mathbf{S}(C) = (1, 0)^T$  and strategy matrices for defection is  $\mathbf{S}(D) = (0, 1)^T$ ; for the PD game, we can simplify the payoff matrices as follows: we implement Eq. (1) based on evolutionary

game data with numerical computations through complex networks. The payoff matrices  $\mathbf{P}$  are

$$\mathbf{P} = \begin{pmatrix} 1 & 0 \\ 1+b & 0 \end{pmatrix}, \quad (2)$$

where  $b(0 < b < 1)$  is the parameter that can indicate the temptation to defect; each agent plays the PD game with its neighbors and gains payoffs step by step. For example, as for agent  $i$ , after one round, the payoff  $\mathbf{Z}_i$  is

$$\mathbf{Z}_i = \sum_{j \in \Gamma_i} \mathbf{S}_i^T \cdot \mathbf{P} \cdot \mathbf{S}_j, \quad (3)$$

in Eq. (3)  $\mathbf{S}_i$  denotes the strategies of agent  $i$  and  $\mathbf{S}_j$  denotes the strategies of agent  $j$  respectively, and  $\mathbf{Z}_i$  is total payoff over all the neighbor set  $\Gamma_i$  of  $i$ . After each round of the game, the agent will adjust its own strategy according to the its payoff and the neighbors', allowing them to achieve the desired benefits in subsequent rounds, thus forming game evolution process. Common strategy evolution rules have imitated the best Copy dynamics, Fermi dynamics, and so on. The experimental part of this paper will use the Fermi dynamics rule to carry out the PD game evolution. The rule of the dynamics of Fermi: the probability of the node  $i$  next round learning node  $j$ 's strategy is [51]

$$\mathbf{Q}(\mathbf{S}_i \leftarrow \mathbf{S}_j) = \frac{1}{1 + \exp[(\mathbf{Z}_i - \mathbf{Z}_j)/\kappa]}. \quad (4)$$

The parameter  $\kappa$  denotes the rationality of the player in the game; when it comes close to zero, players will only learn the strategy, the revenue of which is higher than its own in this round. When  $\kappa$  increases, the probability of a player learning a low yield neighbor strategy will increase, which can be expressed by  $\kappa$  being 0 if  $\mathbf{Z}_j < \mathbf{Z}_i$  and 1 if  $\mathbf{Z}_i < \mathbf{Z}_j$ . When  $\kappa$  approaches  $\infty$ , this means that the player chooses completely random decision making.

Each agent possesses one node in the evolutionary game. Assume that the number of nodes in the game network is  $N$ ; the links between nodes can be represented by an adjacency matrix  $\mathbf{A}$  with  $N$  order. If node  $j$  is the neighbor of node  $i$ , the matrix element  $a_{ij} = 1$ ; otherwise  $a_{ij} = 0$ . The total payoff for node  $i$  can be expressed as follows:

$$\begin{aligned} \mathbf{G}_i(t) &= a_{i1} \mathbf{S}_i^T(t) \cdot \mathbf{P} \cdot \mathbf{S}_1(t) + \cdots + a_{i,i-1} \mathbf{S}_i^T(t) \cdot \mathbf{P} \\ &\cdot \mathbf{S}_{i-1}(t) + a_{i,i+1} \mathbf{S}_i^T(t) \cdot \mathbf{P} \cdot \mathbf{S}_{i+1}(t) + a_{iN} \mathbf{S}_i^T(t) \\ &\cdot \mathbf{P} \cdot \mathbf{S}_N(t), \end{aligned} \quad (5)$$

in which  $a_{ij}(j = 1, \dots, i-1, i+1, \dots, N)$  denotes connection between agent  $i$  and its neighbor  $j$ , and  $a_{ij} \mathbf{S}_i^T(t) \cdot \mathbf{P} \cdot \mathbf{S}_j(t)(j = 1, \dots, i-1, i+1, \dots, N)$  denotes the total payoff of node  $i$  from the PD game with  $j$ , and  $t$  denotes the the total rounds that all

agents implement PD game with their own neighbors. And  $\Phi_i$  is the payoff matrix for node  $i$ , expressed as follows:

$$\Phi_i = \begin{pmatrix} \mathbf{F}_{i1}(t_1) & \mathbf{F}_{i,2}(t_1) & \cdots & \mathbf{F}_{in}(t_1) \\ \mathbf{F}_{i1}(t_2) & \mathbf{F}_{i,2}(t_2) & \cdots & \mathbf{F}_{in}(t_2) \\ \mathbf{F}_{i1}(t_3) & \mathbf{F}_{i,2}(t_3) & \cdots & \mathbf{F}_{in}(t_3) \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{F}_{i1}(t_m) & \mathbf{F}_{i,2}(t_m) & \cdots & \mathbf{F}_{i,n}(t_m) \end{pmatrix}, \quad (6)$$

then we solve the data information

$$\mathbf{A}_i = (a_{i1}, \dots, a_i, a_{i-1}, a_{i+1}, \dots, a_{iN})^T, \quad (7)$$

and

$$\mathbf{G}_x = (\mathbf{G}_i(t_1), \mathbf{G}_i(t_2), \dots, \mathbf{G}_i(t_m))^T, \quad (8)$$

where  $\mathbf{F}_{ij}(t_i) = \mathbf{S}_i^T \cdot \mathbf{P} \cdot \mathbf{S}_j(t_i)$  and Eq. (6)–(8) satisfy the following equation:

$$\mathbf{G}_i = \Phi_i \cdot \mathbf{A}_i. \quad (9)$$

In a similar fashion, through compressive sensing method, the remaining agents yield payoff from their neighbor-connection; the overall network adjacency matrix can be expressed as  $\mathbf{A} = (\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_N)$ . Due to the sparsity of  $\mathbf{A}_j$ , it is assured that it can be solved by using CS theory to get accurate results from partial time series data ( $\eta \equiv m/n$ , where  $\eta$  denotes the length of time series data, and  $m$  denotes the number of moment: for example,  $\eta = 60\%$  means that the number of used measurements equals  $0.6N$ ).

**2.2. Network Model.** A network can be denoted as an adjacent matrix  $\mathbf{W} = (a_{ij})_{N \times N}$ , which contains  $N$  nodes and  $E$  edges. In general,  $k_i$  represents the degree of node  $i$ , and  $C_i$  represents the clustering coefficients of node  $i$ ; it represents the ratio between the number of closed triplets and the total number of connected triplets of vertices; it can be calculated as follows:

$$C_i = \frac{1}{k_i(k_i - 1)} \sum_{j,k=1}^N a_{ij} a_{jk} a_{ki}, \quad (10)$$

where  $a_{ij} = 1$  denotes that there exists a partnership between node  $i$  and  $j$ , so when  $a_{ij} a_{jk} a_{ki} = 1$ , then a triangle forms among nodes  $i, j, k$ . The whole network clustering coefficient should be expressed as

$$C = \frac{1}{N} \sum_{i=1}^N C_i. \quad (11)$$

We validated our method by using PD games data occurring in Holme–Kim networks. Contrary to BA model, Holme–Kim model adds Triad Formation process (TF process), in order to change the principle when a new node is attached to an existing node by the principle called Preferential Attachment process (PA process). By this way, it

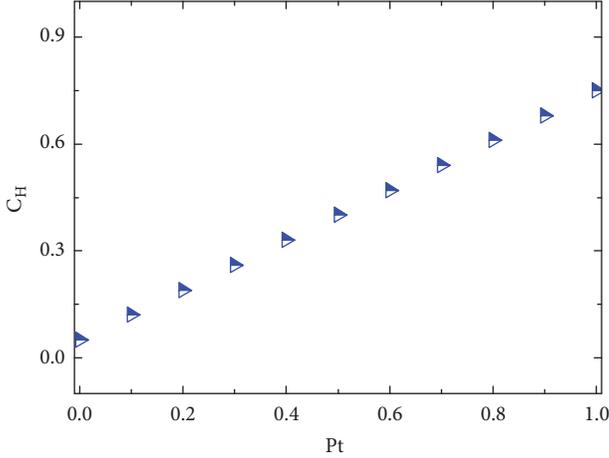


FIGURE 1: Correlation between the  $C_H$  and  $P_t$  in Holme–Kim network.

not only makes the network growth mode more flexible, but also increases the clustering coefficient. In the Holme–Kim model, the evolution process of network has three driver factors: growth, preference attachment, and triad formation. The algorithm of the Holme–Kim model is shown as follows [52]:

*Growth.* At the beginning, the network has  $m_0$  nodes; on each time step, one new node  $v$  with  $m$  edges connects with the existing nodes ( $m \leq m_0$ ).

*Preferential Attachment.* There is preference when the new node  $v$  chooses other nodes to connect. The ratio of choosing node  $i$  depends on the degree of this node. It can be expressed as follows:

$$\Pi_i = \frac{k_i}{\sum_j k_j}. \quad (12)$$

*Triad Formation.* After adding an edge between new node  $v$  and  $i$ 's neighbor in the previous PA step, another neighbor node of  $i$  is chosen to connect with node  $v$  with probability  $P_t$ . The network structure varies as  $P_t$  changes; Figure 1 shows the relationship between probabilities  $P_t$  and clustering coefficients  $C_H$  of network. It is obvious that as  $P_t$  increases,  $C_H$  grows as well, which implies that the network gets more closely.

Another PD game data occurring in Newman-Watts small-world network (homogeneous small-world network) exhibits a homogeneous connectivity distribution, in the sense that the number of connections for all nodes is the same; the algorithm steps can be described as

- (1) First, initializing a network with  $N$  nodes, it is ring-shaped; each node has  $2k$  neighbors,  $k > 0$ ,  $k$  is an integer (usually small).
- (2) Then adding an edge between the unconnected nodes with probability  $P_N$  ( $0 < P_N \ll 1$ ).

During the process, there will be multiple edges between any pair of nodes. All nodes will have no self-loops; the clustering coefficients of the network are denoted as  $C_N$ .

### 3. Experimental Results

*3.1. Generating Artificial Networks with Tunable Clustering Coefficient.* The PD game is simulated on two types of networks, Holme–Kim networks and Newman-Watts small-world. To test the efficiency in reconstructing a network with our method, we first generate an artificial scale-free network with 100 nodes by using Holme–Kim model and vary the  $P_t$  value from 0 to 1. The clustering coefficients ( $C_H$ ) are also incorporated into this range correspondingly. Then, without loss of generality, we set  $m_0=m=5$ ,  $N=100$ . We, especially, study the relationship between SR, which measures the accuracy of network reconstruction, and the degree  $k$  in Holme–Kim networks. Each piece of data is averaged, performed on 10 networks by 10 runs. In order to reconfirm the performance of the algorithm, we set different network sizes to further investigate the performance. Analogously, we also reproduce a group of Newman-Watts small-world networks by using compressive sensing method to adjust their clustering coefficients into the same range. In order to uncover the topology structure of the evolution network, we record the strategies and payoffs in the form of time series as the system tends to the steady state.

*3.2. Evaluation Criterion.* PD game is implemented to investigate the performance of the network diffusion dynamics and structures; after recording the measurement matrix  $G_i$  and measurement vector  $\Phi_i$ , we introduce the SR to measure the performance; the accuracy rate can be expressed as follows:

$$SR = \frac{1}{N} \sum_{i=1}^N \left( \frac{Q_{ir} \cap Q_{io}}{Q_{io}} \right), \quad (13)$$

in which  $Q_{io}$  denotes neighbors of node  $i$  during test set;  $Q_{ir}$  denotes the neighbors of node  $i$  during training set. For a single player, SR denotes the ratio of the successfully predicted number of neighbor connections to the actual number of neighbors. At last, we get the average value for the whole network.

*3.3. Results.* Figure 2(a) shows the SR of the network construction by means of  $L_1$ -norm minimization, as shown in the picture; different  $C_H$  indicate different topological structures of networks with tunable clustering coefficients; judging from transverse direction, the SR would increase along with the increase of  $C_H$  value; the SR, especially, would reach the highest and would not get larger obviously when  $C_H$  value gets close to 1. From the vertical perspective, the value of SR increases with the increase of  $\eta$ , for instance, the SR with which, when the length of data was 60%, is higher than that of 50%, and so forth. It demonstrates that we can use  $L_1$ -norm minimization to reconstruct network effectively. Moreover, it is necessary to detail the progress of whole network reconstruction to infer local characteristic for each single node; we investigate the SR among single nodes, as

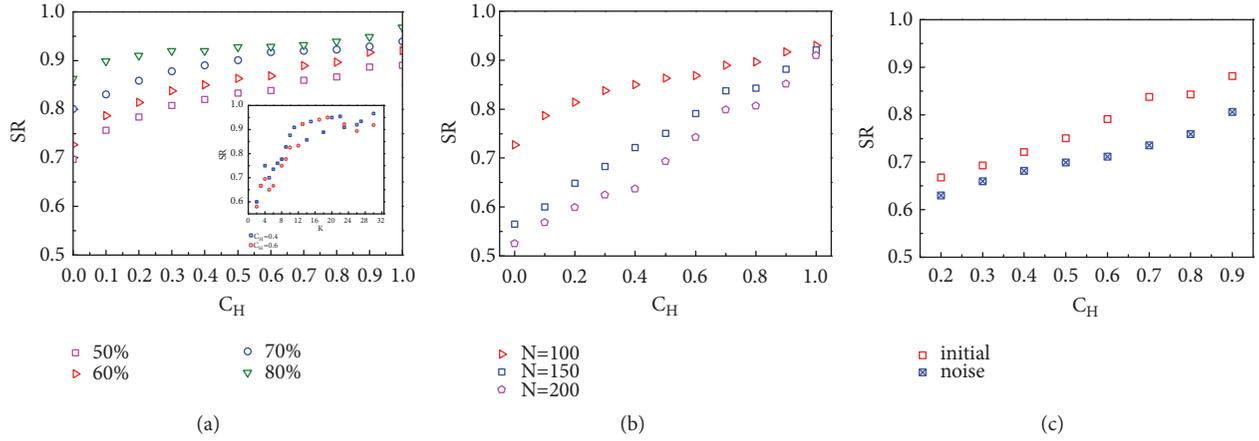


FIGURE 2: Success rate (SR) of networks with tunable clustering coefficients  $C_H$  in Holme–Kim model by means of  $L_1$ -norm minimization. (a) The size is set to  $N = 100$ ,  $\eta$  is the length of data information,  $\kappa = 0.1$ . As shown in the picture, when clustering coefficients  $C_H$  increases, the accuracy SR increases. The tendency also applies in the variable value of  $\eta$ . Each piece of data is averaged, performed on 10 networks by 10 runs. The subgraph shows the relation between the SR and the degree of each game player, in which the network size is 100,  $\eta = 50\%$ . It shows that the bigger the degree of the node, the higher the SR. (b) is the performance of the SR for different network sizes, with  $N = 100, 150, 200$  respectively and  $\eta$  is fixed to 60%. When  $C_H$  value increases, the SR increases correspondingly. SR decreases as the network size increases. (c) tests the SR with noise compared with that of original data; the Gaussian noise is defined as  $\varepsilon \sim [\mathcal{N}(0, 0.1^2)]$ . The size of the network is  $N = 150$ ,  $\eta = 60\%$ . Each piece of data is averaged, performed on 10 networks by 10 runs.

shown in the subgraph in Figure 2(a). The bigger the degree for each node, the higher the SR. It indicates that the hub node has higher SR, which manifests that the node, which has the maximum degree, contributes to the accurate reconstruction.

Furthermore, in order to validate  $L_1$ -norm minimization algorithm, we set different network sizes to further investigate the performance; we generate another two kinds of network size, which incorporates 150 nodes and 200 nodes, respectively. As shown in Figure 2(b), when the  $C_H$  value increases, the SR increases correspondingly, but gets lower when the network size increases. From this point of view, we can conclude that the network size can influence the result of network reconstruction. For instance, when the  $C_H = 0.5$ , network size decreases from 150 to 100, then SR increases from 0.85 to 0.9, and so forth. The relationship between SR and network size is inversely related. In reality, noise exists everywhere in physical systems; the data get from complex systems are more or less influenced by noise. In order to further improve the experiment, Gaussian noises  $[\mathcal{N}(0, 0.1^2)]$  are added in the experiment ( $Q' = Q + \varepsilon$ ,  $Q'$  is the noise parameter) to investigate the stability of the result. Figure 2(c) indicates that the network size  $N$  is 150 and  $\eta = 60\%$ . We find that, compared with the data after adding the Gaussian noise, the accuracy of initial data is higher, although the CS method could reconstruct the missed links in the network.

As a control group, we investigate the relationship between SR and clustering coefficient  $C_N$  in Newman-Watts small-world network by means of  $L_1$ -norm minimization, in which  $N = 20, 50, 100$  respectively. As shown in Figure 3(a) we can find that SR decreases as the clustering coefficients  $C_N$  of the network increase, which shows adverse tendency compared with Holme–Kim model. The main reason may be originating from the fact that the degree distribution in Newman-Watts small-world network is homogeneous; the

nodes which have homogeneous degrees need similar data to reconstruct the network. The subgraph in Figure 3(b) shows the relation between the SR and degree distribution of each game player, in which the network size is  $N = 100$  and  $\eta = 80\%$ . It shows that the bigger the degree of the node, the lower the SR.

The methods used to solve sparse approximation problems are available in a variety of ways. As we present our  $L_1$ -norm minimization algorithm in the above context of networks, we would like to compare the performance of our framework for network reconstruction with OMP [53], a greedy algorithm, which can be expressed as

$$\begin{aligned} \min \quad & \|\mathbf{X}\|_0 \\ \text{s.t.} \quad & \mathbf{G} = \Phi \cdot \mathbf{X}, \end{aligned} \quad (14)$$

compared with  $L_1$ -norm minimization. OMP needs much more data and has lower accuracy [54]. The number of iterations required to find the correct support set determines the complexity of the algorithm. For the comparison, we carry out all experiments. The results are shown in Figures 4 and 5. The overall tendency is implemented by OMP in accordance with  $L_1$ -norm minimization. In addition, the two methods are simultaneously implemented on scale-free networks and small-world network, with data occurring in evolutionary games. As shown in Figures 6 and 7, we can find that the former proposed method is higher than or equal to OMP; for the Holme–Kim model network, the resulting SR in our method is higher than OMP; for the Newman-Watts small-world model, the resulting SR is nearly equal to OMP. The main reason may result from the fact that both  $L_1$ -norm minimization and OMP can solve the problem of sparse signal recovery [55]. Both the  $L_1$ -norm and orthogonal matching pursuit are the methods used to solve sparse

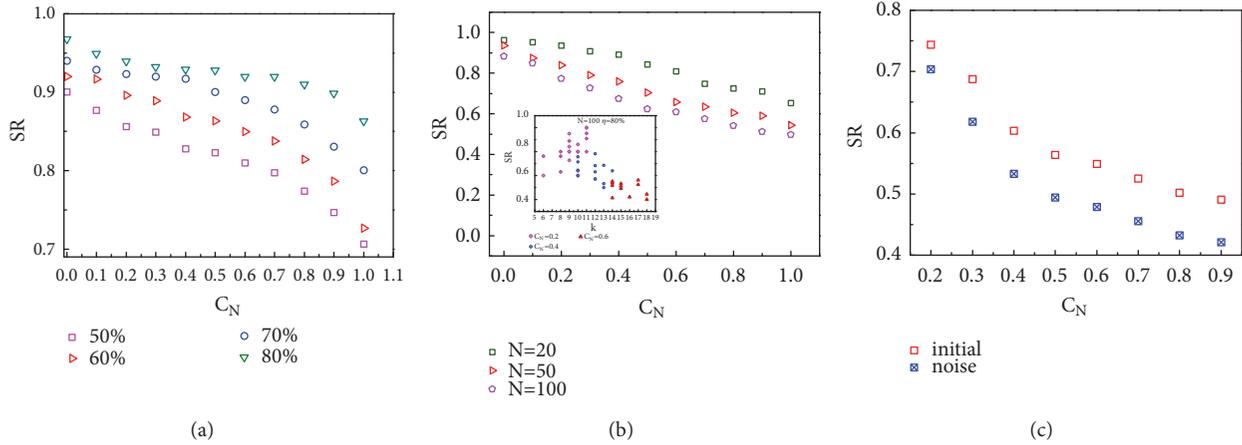


FIGURE 3: (a) Correlation between the clustering coefficients  $C_N$  of the Newman-Watts small-world networks and SR by means of  $L_1$ -norm minimization.  $N = 50, \eta = 50\%, 60\%, 70\%, 80\%$ . (b)  $N = 20, 50, 100, \eta$  is fixed to 60%. From the picture we can see that SR decreases as  $C_N$  increases. The same with Holme–Kim model, when network size increases, SR tends to be lower. Each piece of data is averaged, performed on 10 networks by 10 runs. It shows that the bigger the degree of the node, the lower the SR. (c) tests the SR with noise compared with that of original data; the Gaussian noise is defined as  $\varepsilon \sim [\mathcal{N}(0, 0.1^2)]$ . The size of the network  $N = 100, \eta = 60\%$ . Each piece of data is averaged, performed on 10 networks by 10 runs.

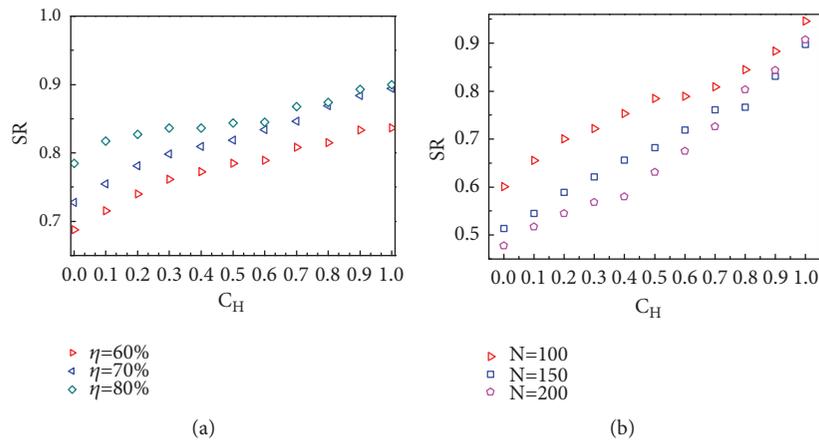


FIGURE 4: (a) SR of networks with tunable clustering coefficients  $C_H$  in Holme–Kim model by OMP method,  $N = 100, \kappa = 0.1$ . It is obvious that SR behaves in accordance with  $L_1$ -norm minimization. Each piece of data is averaged, performed on 10 networks by 10 runs. (b) is the performance of the SR for different network size, with  $N = 100, 150, 200$  respectively and  $\eta$  is fixed to 60%. When network size increases, SR tends to be lower. Each piece of data is averaged, performed on 10 networks by 10 runs.

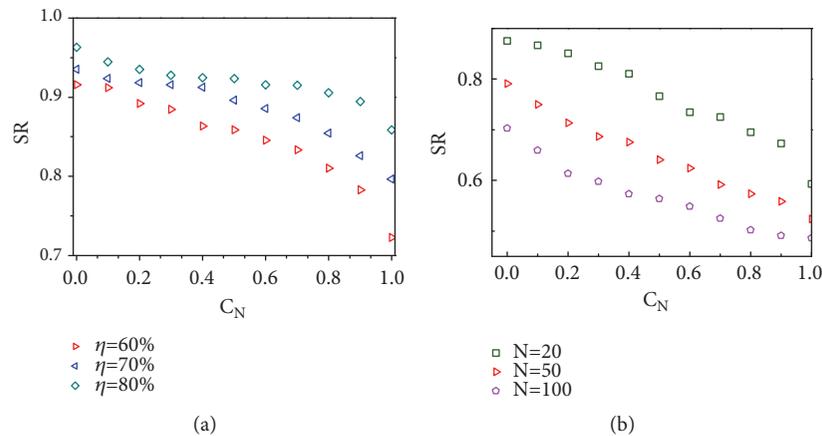


FIGURE 5: Correlation between the clustering coefficients  $C_N$  of the Newman-Watts small-world networks and SR by OMP method. (a)  $N = 50, \eta = 60\%, 70\%, 80\%$  (b)  $N = 20, 50, 100, \eta$  is fixed to 60%. Each piece of data is averaged, performed on 10 networks by 10 runs.

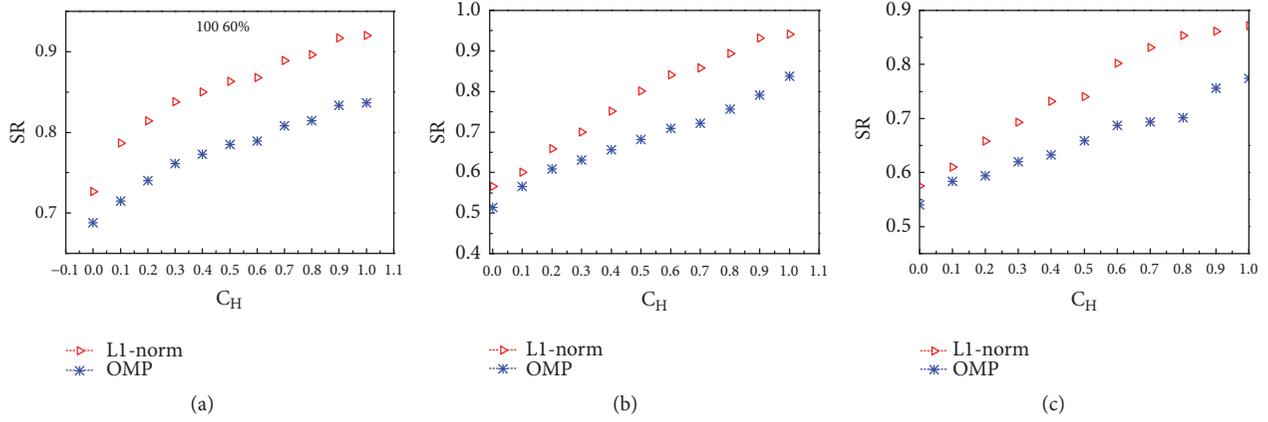


FIGURE 6: Performance comparison of  $L_1$ -norm minimization with OMP on scale-free network, considering different values of  $\eta$  parameter and  $N = 100, 150, 200$ . (a)  $N = 100, \eta = 60\%$  (b)  $N = 150, \eta = 60\%$  (c)  $N = 200, \eta = 80\%$ . Each piece of data is averaged, performed on 10 networks by 10 runs.

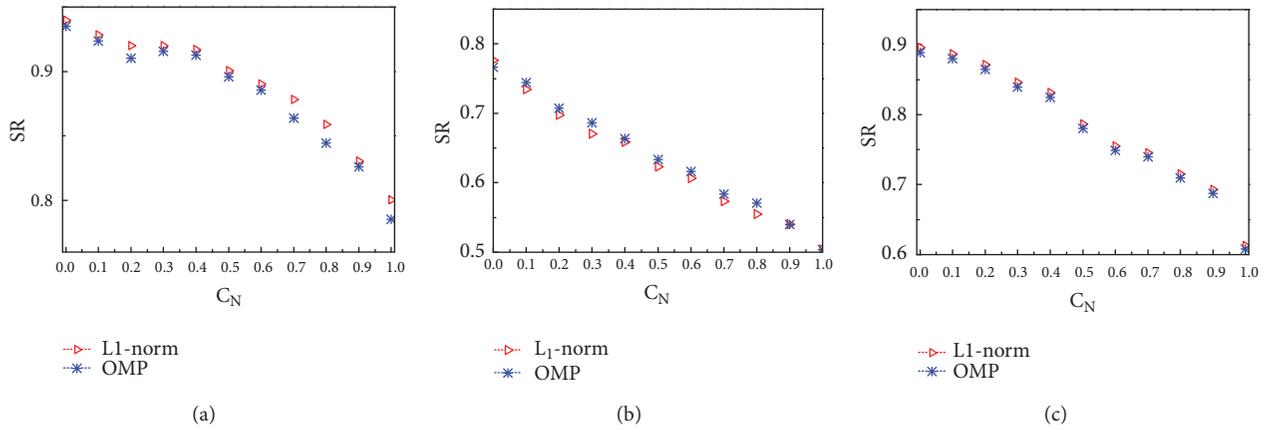


FIGURE 7: Performance comparison of  $L_1$ -norm minimization with OMP on small-world network, considering different values of  $\eta$  parameter and  $N = 50, 70, 100$ . (a)  $N = 50, \eta = 60\%$  (b)  $N = 70, \eta = 60\%$ . (c)  $N = 100, \eta = 90\%$ . Each piece of data is averaged, performed on 10 networks by 10 runs.

approximation problems; by using  $L_1$ -norm minimization method, it can get optimal solution to the problem of network reconstruction [56]. However, Newman-Watts small-world behaves much more sparse [57], as described in Table 1. These results demonstrate that two type networks can be efficiently inferred from limited measurements using  $L_1$ -norm minimization method. In addition, we compare the two algorithms on the performance of convergence rate between two models as shown in Tables 2, 3, 4, and 5. In order to achieve high accuracy with different data length, the convergence rate is different. Compared with the  $L_1$ -norm minimization, the OMP yields a faster computation speed; overall OMP is 1.2 times the computation speed of  $L_1$ -norm minimization. Inspired by this, we may consider using OMP in small-world network. All the above demonstrate that the network topological structure influences the accuracy of network reconstruction. We should use a proper method to deal with the complex problem of network reconstruction.

TABLE 1: The parameter for different network.

Type	clustering coefficient	modularity
B-A	big	big
W-S	bigger	small

## 4. Conclusion

In our study, we investigate the effect on the accuracy of the network reconstruction by  $L_1$ -norm minimization method with different topological structures of networks, measured by variable clustering coefficients among two types of network models: one is scale-free network and the other is small-world network. First, we construct two types of network models with variable clustering coefficients, then we reconstruct the network based on the evolutionary-game data by  $L_1$ -norm minimization method. We find that the success rate of the two kinds of network performs differently; for

TABLE 2: To achieve high accuracy (85%) the convergence rate was performed on scale-free network with two methods.

data length	time/s							
	N=100		N=150		N=200		N=250	
	$l_1$	OMP	$l_1$	OMP	$l_1$	OMP	$l_1$	OMP
$\eta = 50\%$	784	653	1075	853	1283	1058	1462	1107
$\eta = 60\%$	923	842	1382	1191	1596	1227	1686	1340
$\eta = 70\%$	1080	818	1652	1458	1789	1503	2074	1773
$\eta = 80\%$	1425	1096	1893	1578	2056	1631	2261	1916

TABLE 3: To achieve high accuracy (90%) the convergence rate was performed on scale-free network with two methods.

data length	time/s							
	N=100		N=150		N=200		N=250	
	$l_1$	OMP	$l_1$	OMP	$l_1$	OMP	$l_1$	OMP
$\eta = 50\%$	803	676	1104	876	1309	1087	1497	1130
$\eta = 60\%$	945	869	1398	1209	1616	1250	1686	1378
$\eta = 70\%$	1109	840	1678	1483	1809	1528	2098	1799
$\eta = 80\%$	1445	1108	1914	1596	2076	1651	2279	1934

TABLE 4: To achieve high accuracy (80%) the convergence rate was performed on small-world network with two methods.

data length	time/s							
	N=20		N=50		N=70		N=100	
	$l_1$	OMP	$l_1$	OMP	$l_1$	OMP	$l_1$	OMP
$\eta = 50\%$	976	820	1180	984	1375	1100	1498	1247
$\eta = 60\%$	1195	1012	1416	1151	1564	1348	1207	1040
$\eta = 70\%$	1401	1120	1692	1421	1689	1407	1002	871
$\eta = 80\%$	1672	1416	1795	1504	1879	1527	864	701

TABLE 5: To achieve high accuracy (85%) the convergence rate was performed on small-world network with two methods.

data length	time/s							
	N=20		N=50		N=70		N=100	
	$l_1$	OMP	$l_1$	OMP	$l_1$	OMP	$l_1$	OMP
$\eta = 50\%$	998	832	1203	997	1397	1117	1512	1264
$\eta = 60\%$	1215	1032	1431	1171	1584	1360	1221	1062
$\eta = 70\%$	1420	1141	1709	1445	1700	1427	1029	901
$\eta = 80\%$	1692	1438	1815	1528	1899	1557	896	728

Holme–Kim (scale-free) network, the success rate increases correspondingly when the clustering coefficients  $C_H$  increase but decreases when the clustering coefficients  $C_N$  increase in the Newman-Watts (small-world network). The explanation can be derived from the fact that, in Holme–Kim model, the bigger the degree for each node, the lower the success rate, while the small-world behaves opposite. Meanwhile, both of the success rates for the two network models decrease as the network size increases. At last, we adopt OMP method to reconstruct network. The results show that the SR carried out by  $L_1$ -norm minimization method is higher than or at least equal to OMP, and the accuracy rate of  $L_1$ -norm minimization method is 10% higher than that of OMP. Meanwhile, OMP is 1.2 times the convergence rate of  $L_1$ -norm minimization. From the above, we can conclude that

the network structure has influence on the accuracy of the network reconstruction by compressive sensing method.

Our method, mentioned in the paper, can be adopted to reconstruct sparse networks among the real-world networks for the reason that small-world phenomenon and scale-free characteristics are two typical complex network characteristics. Our contribution to the current literature can be summarized as follows: the method mentioned above needs time series data; usually the topology structure of the network cannot be obtained. From this point of view, if we do not know about the topology structure, we may first choose  $L_1$ -norm minimization method; if we know the topology structure of the network, the network has the same characteristics as Holme–Kim model (scale-free characteristics); we may be more likely to choose  $L_1$ -norm minimization; if the network

has the same properties as the small-world network, for the accuracy recovery by  $L_1$ -norm minimization is nearly equal to that of OMP, we can adopt both of these methods; however, OMP method has an advantage over  $L_1$ -norm minimization from a convergence speed perspective, so if we consider the speed of convergence, we may choose OMP method. Overall, our methods, in principle, are applicable in real-world such as protein network, Internet network. Meanwhile, the research also puts forward application of CS in other research fields, such as remote sensing and social networks.

Meanwhile, there exist some shortcomings in this paper, which may influence the further research for network reconstruction. First, the two methods mentioned in this paper are used to recover sparse network; better methods are worth exploring [58, 59]. Second, we only investigate the effect of compressive sensing depending on dynamic clustering coefficients of the macro network structure. More characteristics should be considered such as betweenness, closeness, and eigenvector centrality to verify our methods. In addition, we should also consider the micro network structure, such as Ring, Chain, and Star, which may affect the result [60]. Third, in this paper, we consider only two phenomena: one is the small-world, and the other is scale-free network, but in the actual network there are both scale-free network and small world network natures of the network, so it is possible that the two methods may not effectively reconstruct the network; we should find a better way. Last, during the process of simulation, all the data are observable. In fact, the measurement can be obtained partially [61, 62]; that is to say, some hidden nodes cannot be got in the system. How to implement our methods in these situations deserves to be studied further.

Generally speaking, this paper puts forward two methods to deal with the problem of network reconstruction through two different manners. It provides directions for us to reconstruct complex network; yet it is expected to make efforts to pursue better approaches. All these deserved to be explored.

## Data Availability

No data were used to support this study.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work is supported by the National Natural Science Foundation of China (Grant No. 71271126), doctoral Fund of Ministry of Education of China (No. 20120078110002), and Fundamental Research Funds for the Central Universities, supported by the GFSUFE (Grant No. CXJJ-2016-414).

## References

- [1] M. Timme, "Revealing Network Connectivity from Response Dynamics," *Physical Review Letters*, vol. 98, no. 22, 2007.
- [2] J. Bongard and H. Lipson, "Automated reverse engineering of nonlinear dynamical systems," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 24, pp. 9943–9948, 2007.
- [3] A. Clauset, C. Moore, and M. E. J. Newman, "Hierarchical structure and the prediction of missing links in networks," *Nature*, vol. 453, no. 7191, pp. 98–101, 2008.
- [4] R. Guimerà and M. Sales-Pardo, "Missing and spurious interactions and the reconstruction of complex networks," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 52, pp. 22073–22078, 2009.
- [5] Z. Levnajić and A. Pikovsky, "Network Reconstruction from Random Phase Resetting," *Physical Review Letters*, vol. 107, no. 3, 2011.
- [6] D. Marbach, R. J. Prill, T. Schaffter, C. Mattiussi, D. Floreano, and G. Stolovitzky, "Revealing strengths and weaknesses of methods for gene network inference," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 14, pp. 6286–6291, 2010.
- [7] G. Michailidis and F. D'Alché-Buc, "Autoregressive models for gene regulatory network inference: sparsity, stability and causality issues," *Mathematical Biosciences*, vol. 246, no. 2, pp. 326–334, 2013.
- [8] Z. Shen, W.-X. Wang, Y. Fan, Z. Di, and Y.-C. Lai, "Reconstructing propagation networks with natural diversity and identifying hidden sources," *Nature Communications*, vol. 5, no. 5, p. 4323, 2014.
- [9] L. Lü and T. Zhou, "Link prediction in complex networks: a survey," *Physica A: Statistical Mechanics and its Applications*, vol. 390, no. 6, pp. 1150–1170, 2011.
- [10] N. Dojer, A. Gambin, A. Mizera, B. Wilczyński, and J. Tiuryn, "Applying dynamic Bayesian networks to perturbed gene expression data," *BMC Bioinformatics*, vol. 7, article no. 249, 2006.
- [11] E. D. Sontag, "Network reconstruction based on steady-state data," *Essays in Biochemistry*, vol. 45, pp. 161–176, 2008.
- [12] E. J. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *Institute of Electrical and Electronics Engineers Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.
- [13] J. Jin, Y. Gu, and S. Mei, "An Introduction to Compressive Sampling and Its Applications," *Journal of Electronics Information & Technology*, vol. 32, no. 2, pp. 470–475, 2010.
- [14] J. Romberg, "Imaging via compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 14–20, 2008.
- [15] J. M. Smith, *Evolution and the Theory of Games*, Springer, Boston, MA, USA, 1993.
- [16] S. Hempel, A. Koseska, J. Kurths, and Z. Nikoloski, "Inner composition alignment for inferring directed networks from short time series," *Physical Review Letters*, vol. 107, no. 5, 2011.
- [17] W. Wang, R. Yang, Y. Lai, V. Kovanis, and C. Grebogi, "Predicting Catastrophes in Nonlinear Dynamical Systems by Compressive Sensing," *Physical Review Letters*, vol. 106, no. 15, 2011.
- [18] L. Xiao, P. Yang, Y. Xiao et al., "Efficient compressive sensing detectors for generalized spatial modulation systems," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 2, pp. 1284–1298, 2017.
- [19] L. D. Abreu and J. L. Romero, "MSE estimates for multitaper spectral estimation and off-grid compressive sensing," *Institute*

- of Electrical and Electronics Engineers Transactions on Information Theory*, vol. 63, no. 12, pp. 7770–7776, 2017.
- [20] W.-X. Wang, Y.-C. Lai, C. Grebogi, and J. Ye, “Network Reconstruction Based on Evolutionary-Game Data via Compressive Sensing,” *Physical Review X*, vol. 1, no. 2, Article ID 021021, pp. 1–7, 2011.
- [21] T. S. Gardner, D. Di Bernardo, D. Lorenz, and J. J. Collins, “Inferring genetic networks and identifying compound mode of action via expression profiling,” *Science*, vol. 301, no. 5629, pp. 102–105, 2003.
- [22] X. Han, Z. Shen, W. Wang, and Z. Di, “Robust Reconstruction of Complex Networks from Sparse Data,” *Physical Review Letters*, vol. 114, no. 2, 2015.
- [23] H. Lee, D. S. Lee, H. Kang, B.-N. Kim, and M. K. Chung, “Sparse brain network recovery under compressed sensing,” *IEEE Transactions on Medical Imaging*, vol. 30, no. 5, pp. 1154–1165, 2011.
- [24] P. Siyari, H. R. Rabiee, M. Salehi, and M. E. Mehdiabadi, “Network reconstruction under compressive sensing,” in *Proceedings of the 2012 ASE International Conference on Social Informatics, SocialInformatics 2012*, pp. 19–25, USA, December 2012.
- [25] Y. Lu, A. Montanari, B. Prabhakar, S. Dharmapurikar, and A. Kabbani, “Counter braids: A novel counter architecture for per-flow measurement,” in *Proceedings of the 2008 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems, SIGMETRICS’08*, pp. 121–132, USA, June 2008.
- [26] R. Gaeta, M. Grangetto, and M. Sereno, “Local access to sparse and large global information in P2P networks: A case for compressive sensing,” in *Proceedings of the 2010 IEEE 10th International Conference on Peer-to-Peer Computing, P2P 2010*, Netherlands, August 2010.
- [27] F. C. Santos and J. M. Pacheco, “Scale-free networks provide a unifying framework for the emergence of cooperation,” *Physical Review Letters*, vol. 95, no. 9, Article ID 098104, 2005.
- [28] Z. Rong, X. Li, and X. Wang, “Roles of mixing patterns in cooperation on a scale-free networked game,” *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, vol. 76, no. 2, 2007.
- [29] D.-H. Shin and Y.-J. Shin, “Why do people play social network games?” *Computers in Human Behavior*, vol. 27, no. 2, pp. 852–861, 2011.
- [30] M. O. Jackson, “Allocation rules for network games,” *Games and Economic Behavior*, vol. 51, no. 1, pp. 128–154, 2005.
- [31] T. Yang, Z. Meng, G. Shi, Y. Hong, and K. H. Johansson, “Network synchronization with nonlinear dynamics and switching interactions,” *Institute of Electrical and Electronics Engineers Transactions on Automatic Control*, vol. 61, no. 10, pp. 3103–3108, 2016.
- [32] M. P. van den Heuvel, R. C. W. Mandl, C. J. Stam, R. S. Kahn, and H. E. Hulshoff Pol, “Aberrant frontal and temporal complex network structure in schizophrenia: a graph theoretical analysis,” *The Journal of Neuroscience*, vol. 30, no. 47, pp. 15915–15926, 2010.
- [33] S. Boccaletti, G. Bianconi, and R. Criado, “The structure and dynamics of multilayer networks,” *Physics Reports*, vol. 544, no. 1, pp. 1–122, 2014.
- [34] M. Pósfai, Y. Liu, J. Slotine, and A. Barabási, “Effect of correlations on network controllability,” *Scientific Reports*, vol. 3, no. 1, 2013.
- [35] Q. Guo, G. Liang, J. Fu, J. Han, and J. Liu, “Roles of mixing patterns in the network reconstruction,” *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, vol. 94, no. 5, 2016.
- [36] D. J. Watts and S. H. Strogatz, “Collective dynamics of “small-world” networks,” *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.
- [37] A. L. Barabási, R. Albert, and H. Jeong, “Mean-field theory for scale-free random networks,” *Physica A: Statistical Mechanics and its Applications*, vol. 272, no. 1, pp. 173–187, 1999.
- [38] A. Barabasi and R. Albert, “Emergence of scaling in random networks,” *Science*, vol. 286, no. 5439, pp. 509–512, 1999.
- [39] P. Holme and B. J. Kim, “Growing scale-free networks with tunable clustering,” *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, vol. 65, no. 2, Article ID 026107, 2002.
- [40] D. L. Donoho, “Compressed sensing,” *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [41] E. J. Candès and M. B. Wakin, “An introduction to compressive sampling: a sensing/sampling paradigm that goes against the common knowledge in data acquisition,” *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, 2008.
- [42] E. J. Candès and T. Tao, “Near-optimal signal recovery from random projections: universal encoding strategies?” *Institute of Electrical and Electronics Engineers Transactions on Information Theory*, vol. 52, no. 12, pp. 5406–5425, 2006.
- [43] E. J. Candès, J. K. Romberg, and T. Tao, “Stable signal recovery from incomplete and inaccurate measurements,” *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, pp. 1207–1223, 2006.
- [44] S. G. Shandilya and M. Timme, “Inferring network topology from complex dynamics,” *New Journal of Physics*, vol. 13, no. 1, p. 013004, 2011.
- [45] K. Sigmund and M. A. Nowak, “Evolutionary game theory,” *Current Biology*, vol. 9, no. 14, pp. R503–R505, 1999.
- [46] C. Hauert and M. Doebell, “Spatial structure often inhibits the evolution of cooperation in the snowdrift game,” *Nature*, vol. 428, no. 6983, pp. 643–646, 2004.
- [47] G. Abramson and M. Kuperman, “Social games in a social network,” *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, vol. 63, Article ID 030901, 2001.
- [48] G. Szabó, J. Vukov, and A. Szolnoki, “Phase diagrams for an evolutionary prisoner’s dilemma game on two-dimensional lattices,” *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, vol. 72, no. 4, Article ID 047107, 2005.
- [49] M. A. Nowak and R. M. May, “Evolutionary games and spatial chaos,” *Nature*, vol. 359, no. 6398, pp. 826–829, 1992.
- [50] J.-J. Wu, B.-Y. Zhang, Z.-X. Zhou et al., “Costly punishment does not always increase cooperation,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 41, pp. 17448–17451, 2009.
- [51] G. Szabó and C. Tóke, “Evolutionary prisoner’s dilemma game on a square lattice,” *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, vol. 58, no. 1, pp. 69–73, 1998.
- [52] Z.-Y. Wang, J.-T. Han, and J. Zhao, “Identifying node spreading influence for tunable clustering coefficient networks,” *Physica A: Statistical Mechanics and its Applications*, vol. 486, pp. 242–250, 2017.
- [53] A. Kulkarni and T. Mohsenin, “Low Overhead Architectures for OMP Compressive Sensing Reconstruction Algorithm,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 64, no. 6, pp. 1468–1480, 2017.

- [54] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.
- [55] V. K. Goyal, A. K. Fletcher, and S. Rangan, "Compressive sampling and lossy compression: Do random measurements provide an efficient method of representing sparse signals?" *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 48–56, 2008.
- [56] E. J. Candès, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathématique*, vol. 346, no. 9-10, pp. 589–592, 2008.
- [57] Z.-J. Jiang, "The numbers of participants and nucleon-nucleon collisions in high-energy heavy-ion collisions," *Acta Physica Sinica*, vol. 56, no. 9, pp. 5191–5195, 2007 (Chinese).
- [58] Z. Xu, X. Chang, F. Xu, and H. Zhang, " $L_{1/2}$  regularization: a thresholding representation theory and a fast solver," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 23, no. 7, pp. 1013–1027, 2012.
- [59] L. Ma, X. Han, Z. Shen, W. Wang, Z. Di, and Z. Gao, "Efficient Reconstruction of Heterogeneous Networks from Time Series via Compressed Sensing," *PLoS ONE*, vol. 10, no. 11, p. e0142837, 2015.
- [60] S. Battiston, G. Caldarelli, R. M. May, T. Roukny, and J. E. Stiglitz, "The price of complexity in financial networks," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 113, no. 36, pp. 10031–10036, 2016.
- [61] Y. Yuan, G.-B. Stan, S. Warnick, and J. Goncalves, "Robust dynamical network structure reconstruction," *Automatica*, vol. 47, no. 6, pp. 1230–1235, 2011.
- [62] D. Hayden, Y. H. Chang, J. Goncalves, and C. Tomlin, *Compressed sensing for network reconstruction*, 2014.

