

Research Article

High-Order Degree and Combined Degree in Complex Networks

Shudong Wang,¹ Xinzeng Wang ,² Qifang Song,² and Yuanyuan Zhang³

¹College of Computer and Communication Engineering, China University of Petroleum, Qingdao 266580, China

²College of Mathematics and Systems Science, Shandong University of Science and Technology, Qingdao 266510, China

³Computer Engineering Institute, Qingdao University of Technology, Qingdao 266520, China

Correspondence should be addressed to Xinzeng Wang; wanglxz@126.com

Received 30 January 2018; Revised 22 May 2018; Accepted 3 June 2018; Published 27 June 2018

Academic Editor: Ester Zumpano

Copyright © 2018 Shudong Wang et al. 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.

We define several novel centrality metrics: the high-order degree and combined degree of undirected network, the high-order out-degree and in-degree and combined out-degree and in-degree of directed network. Those are the measurement of node importance with respect to the number of the node neighbors. We also explore those centrality metrics in the context of several best-known networks. We prove that both the degree centrality and eigenvector centrality are the special cases of the high-order degree of undirected network, and both the in-degree and PageRank algorithm without damping factor are the special cases of the high-order in-degree of directed network. Finally, we also discuss the significance of high-order out-degree of directed network. Our centrality metrics work better in distinguishing nodes than degree and reduce the computation load compared with either eigenvector centrality or PageRank algorithm.

1. Introduction

The theory of network has gone through rapid development since the late 1990s. One of the hottest points is the research on the attributes of network. Node degree has always been considered as one of the most important and fundamental attributes. Many relative researches have defined other attributes of network based on node degree [1, 2], such as degree distribution [3, 4], clustering coefficient [5, 6], the characteristic path length [6], and so on. As early as 1960s, Rapoport [3, 4] emphasized the importance of the degree distribution in all kinds of real networks. Wasserman and Faust [5] introduced fraction of transitive triples in social network in 1994. In order to describe cliquishness of a typical neighborhood, Watts and Strogatz [6] defined clustering coefficient of general complex network in 1998 based on the fraction of transitive triples. Watts and Strogatz [6] also defined the characteristic path length to measure the typical separation between two nodes in the network. In more recent years, numerous researches paid attention to degree distribution [7–10]. Early researches believed that degree distribution followed Poisson Distribution just as the random network theory has described [11, 12]. Recent researches have

found out that the degree distribution for a large number of networks, such as the World Wide Web [13], the Internet [14], the metabolic networks [15], genome-wide disruption networks for yeast [16], and the network of interregional direct investment stocks across Europe [17], have a power-law tail. Such networks are called scale-free [18].

When a practical problem is transformed into a complex network model, people tend to use node centrality to describe the importance and influence of a node, or people need to sort these nodes [19]. Node degree (or degree centrality) is one of the basic methods of sorting the nodes [20–29]. Other common methods are also based on node degree [1, 2]. As long ago as 1948, Bavelas [30] studied the center of social network. Sabidussi et. al. [31] defined what it means for a network node to approach closeness. Freeman [32] used node degree and betweenness to define two kinds of node centrality, and, furthermore, he used node centrality to define graph centrality. Network eigenvector centrality [33, 34] is often used to describe the importance of nodes in social network. In 1998, Brin and Page [35, 36] simplified the eigenvector centrality for undirected network into PageRank algorithm (PR), which is widely used in searching engine Google [36] and many other directed networks [37–42]. By

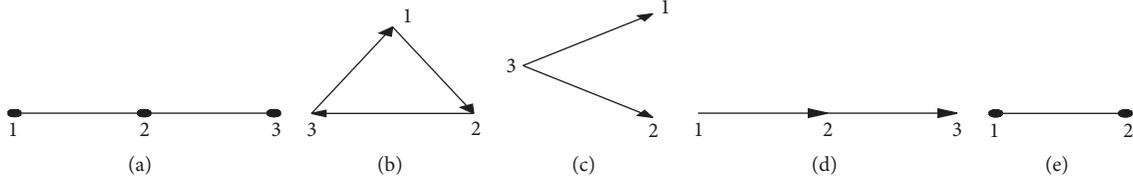


FIGURE 1: Diagram of random walk. Choose the origin node 1: (a) the process of splitting and random walking could be repeated an infinite number of times, because origin node 1 is connected to node 2; (b) the ball can walk randomly an infinite number of times without split since the origin node is one of the directed loop; (c) the ball split and walk randomly only once; (d) the ball can only walk twice without splitting; (e) the ball can walk an infinite number of times without split.

using the degree distribution of neighbor nodes, Ai [43] gave the definition of the neighbor vector centrality. Zeng [44] implemented the Mixed Degree Decomposition (*MDD*) procedure by using coreness centrality, and he defined the mixed degree of nodes. Bae [45] defined the kernel centrality of a node by using the kernel of the neighbor nodes.

The degree centrality used for sorting nodes has the advantages of simple calculation, but the results are not accurate enough. Therefore, it may require verification by other methods [20] or other network attributes [27, 28]. Closeness centrality and betweenness centrality are often given with degree centrality for comparison purpose [27, 28]. The computation of both closeness centrality and betweenness centrality are so complicated that big networks often need fast approximate algorithm [46–48]. Even though the eigenvector centrality on undirected network and the PageRank on directed work can give satisfying results in nodes sorting, those two methods often involve expensive computations, such as iterations [42]. In the following paragraphs, we will define several novel centrality metrics, which are cheaper in computation compared with closeness centrality, betweenness centrality, eigenvector centrality on undirected network, and PageRank on directed network. We will show that node degree, the eigenvector centrality on undirected network and in-degree, and PageRank on directed network are all special cases of (or equivalent to) one of our novel centrality metrics (see Sections 4.1 and 4.2). It should be pointed out that the combined degree defined in our paper is different from the mixed degree in Zeng [44], which is the Mixed Degree Decomposition of coreness centrality (see Section 4.4).

2. Methods

Random walk has always been one of the most important methods in the research of complex network [49–53]. Noh et. al. [49] derived the mean first passage time (*MFPT*) between any two nodes by using random walk of complex network. Tejedor [50] computed the *MFPT* of a network based on a broad class of random walk. Rosvall et. al. [51] used the probability flow of random walks on a network as a proxy for information flow in the real system. Saramäki et. al. [52] generated scale-free networks based on selecting parent nodes by using random walk. Weng [53] used the newly defined mean first traverse distance (*MFTD*) to describe anomalous random walks. Now let us think about an ideal

random walk of a simple network (could be an undirected network or a directed network): choose a node of the network, which we call the origin node, put a ball at the origin node, and the ball obeys the following rules to split repeatedly and to walk randomly.

- (i) Splitting: the ball splits up into $d_G(v)$ (or $d_G^+(v)$) balls at the node v , where $d_G(v)$ is the degree of v in undirected network and $d_G^+(v)$ is the out-degree of v in directed network.
- (ii) Random walk: after every split, the balls move to the adjacent nodes along the $d_G(v)$ edges; in directed network they can only move along $d_G^+(v)$ outgoing edges.
- (iii) Disappearance: balls that can no longer walk would disappear, for example, at the isolated nodes or dead nodes of directed network.

The number of repeats of the random walk would be different depending on the selection of network and/or the selection of origin node, such as in Figure 1.

In the undirected network, the number of balls after s^{th} split is defined as the s -order degree of the origin node v , denoted by $d_G^s(v)$. Clearly, if $s = 1$, $d_G^1(v)$ is the degree $d_G(v)$ of v (see Theorem 2, Section 4.1). In the directed network, the number of balls after s^{th} split is called the s -order out-degree of the origin node v , denoted by $d_G^{+s}(v)$. Also, when $s = 1$, $d_G^{+1}(v)$ is the out-degree $d_G^+(v)$ of v (see Theorem 3, Section 4.2).

Consider a more complicated random walk of a directed network: select a network node v , which we call the sink node. Then we put balls at every node in the network (including the sink node v), and all of the balls follow the above rules to split repeatedly and to walk randomly, the number of balls at the sink node after the s^{th} random walk is defined as the s -order in-degree of the sink node v , denoted by $d_G^{-s}(v)$. Clearly, if $s = 1$, $d_G^{-1}(v)$ is the in-degree $d_G^-(v)$ of v (see Theorem 3, Section 4.2).

2.1. 2-Order Degree. As for the undirected network, according to the above definition, 2-order degree $d_G^2(v)$ of the node v is the number of two-edge paths connected to v . The two-edge paths may overlap. For example, Figure 1(e), both nodes have an overlap two-edge path, then the 2-order degree of both nodes is 1. Note that the 2-order degree is not necessarily equal to the number of neighbors of a node's neighbors,

TABLE 1: The 2-order out-/in-degree of Figure 2.

Name of each node		1	2	3	4	5	6	7	8	9	10	Mean
Figure 2(a)	Degree	1	1	1	4	2	5	1	1	1	1	1.9
(Undirected network)	2-order degree	4	4	4	5	9	6	5	5	5	5	5.2
	Out-degree	2	2	1	1	2	2	1	1	1	1	1.4
Figure 2(b)	In-degree	0	1	2	2	1	2	1	2	1	2	1.8
(Directed network)	2-order out-degree	3	3	1	2	3	2	1	1	1	1	1.4
	2-order in-degree	0	0	2	1	1	3	2	3	2	4	1.8

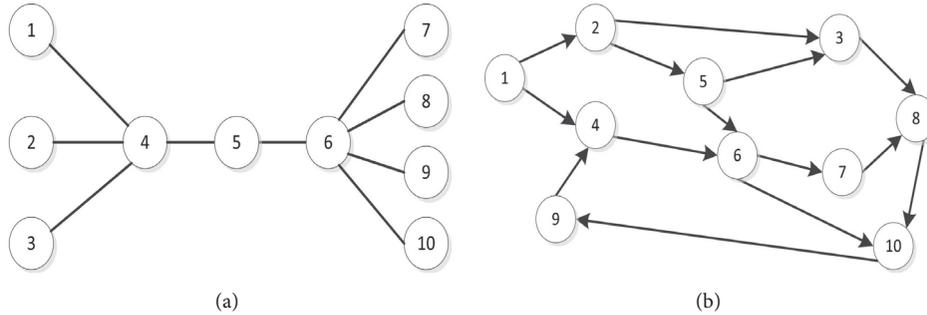


FIGURE 2: The degree and 2-order degree of a network.

since there may be more than one two-edge path between any two nodes (here the two nodes may be the same one). The computation of node 2-order degree $d_G^2(v)$ is relatively simple. Suppose matrix $A = (a_{ij})$ is the adjacency matrix of the undirected network G , we call $A_2 = A^2$ the 2-order adjacency matrix, and then the sum of elements in each row (or each column) of A_2 is the 2-order degree of the corresponding node. Use $\{A_2\}_i$ to denote the sum of the i -th row of matrix A_2 , then the 2-order degree of the i -th node v_i of the undirected network G is $d_G^2(v_i) = \{A_2\}_i$.

A directed graph has both a 2-order in-degree and a 2-order out-degree for each node, which are the numbers of incoming and outgoing two-edge paths, respectively. Still denote the adjacency matrix as $A = (a_{ij})$, and denote $\{A_2\}_i$ and $\{A_2\}^i$ as the sum of the i -th row and i -th column of the 2-order adjacency matrix $A_2 = A^2$, respectively. Then the 2-order out-degree and the 2-order in-degree of the i -th node v_i in this directed network are $d_G^{+2}(v_i) = \{A_2\}_i$ and $d_G^{-2}(v_i) = \{A_2\}^i$, respectively.

If G has vertices v_1, v_2, \dots, v_n , the sequence $(d(v_1), d(v_2), \dots, d(v_n))^T$ is called the **degree sequence** of G [54]. As we all know, the mean value and the distribution of degree sequence guide the definition of mean degree and degree distribution. We can obtain the mean 2-order degree and the 2-order degree distribution of the network by using the definition of 2-order degree.

In Figure 1, (a) the 2-order degree of the three nodes all is 2; (b) the 2-order out-degree and 2-order in-degree of the three nodes all are 1; (c) the 2-order out-degree and 2-order in-degree of the three nodes all are 0; (d) the 2-order out-degree of the three nodes is 1, 0, 0, and the 2-order in-degree

of the three nodes is 0, 0, 1; (e) the 2-order degree of the two nodes is 1.

Figure 2 is a relatively complicated undirected network and a relatively complicated directed network. Table 1 shows the 2-order (in-/out-)degree of the all nodes and both mean 2-order degrees for the two networks in Figure 2. As we can see from Table 1, node 5 has the highest 2-order degree in Figure 2(a), which is the same as the common sense that node 5 is the most important node in Figure 2(a). However this conclusion cannot be obtained simply by calculating node degree. Of course we can obtain such conclusion by calculating node betweenness centrality or eigenvector centrality, and so forth. Another thing we can tell from Table 1 is that nodes having the highest 2-order out-degree are 1, 2, and 5, and the node having the highest 2-order in-degree is 10. Therefore, we could consider that nodes 1, 2, and 5 are the source nodes of (b), and the node 10 could be thought as the collection node of (b). In general, we need to run the PageRank algorithm of the entire network in order to get such conclusions.

2.2. s -Order Degree and Its Computation. Same as 2-order degree, suppose matrix A is an adjacency matrix of a network G , we denote $A_s = A^s (s \in N^*)$ as the s -order adjacency matrix of G . The matrix A_s was used to calculate the number of walk with length s between nodes [54–56]. Denote $\{A_s\}_i$ and $\{A_s\}^i$ as the sum of the i -th row and i -th column respectively in matrix A_s ; then the s -order degree of node v_i in undirected network G is $\{A_s\}_i$ and $\{A_s\}^i = \{A_s\}_i$ (if G is a directed network, then the s -order out-degree is $\{A_s\}_i$, and the s -order in-degree is $\{A_s\}^i$). It is easy to define mean s -order (out-/in-)degree and the s -order (out-/in-)degree distribution, too.

The computation of s -order degree can make use of adjacency matrix, but this involves the exponentiation computation of matrix. The order of matrix A equals the number of nodes in the network, and there could be thousands and tens of thousands of nodes in a network. Therefore even though the computation of s -order adjacency matrix only involves the exponentiation computation of matrix and additions of integers, it would be a high requirement on the computer's CPU and memory. A simple method to tackle this problem is to firstly compute the lower order degree sequence of the network and then compute the higher-order degree sequence. We can prove Theorem 1 by the method of mathematical induction.

Theorem 1. *For s -order degree sequence, the following conclusions are true:*

- (i) *If d^s is a s -order degree sequence of an undirected network G , A is the adjacency matrix, then the $(s + 1)$ -order degree sequence of G is $A * d^s$; particularly, if d^1 is the degree sequence of an undirected network G , then the $(s + 1)$ -order degree sequence of G is $A^s * d^1$.*
- (ii) *Similarly, if d^{+s} and d^{-s} are the s -order out-degree sequence and s -order in-degree sequence of directed network respectively, A is the adjacency matrix, then the $(s + 1)$ -order out-degree sequence and the $(s + 1)$ -order in-degree sequence are $A * d^{+s}$ and $A^T * d^{-s}$ respectively; particularly, if d^{+1} and d^{-1} are the out-degree sequence and in-degree sequence of the directed network respectively, then the $(s + 1)$ -order out-degree sequence and the $(s + 1)$ -order in-degree sequence are $A^s * d^{+1}$ and $(A^T)^s * d^{-1}$, respectively.*

2.3. Combined Degree. Based on the s -order degree defined as above, the following gives the definition of combined degree of a node v_i of an undirected network:

$$d_G^{Com}(v_i) = \alpha_1 \cdot d_G^1(v_i) + \alpha_2 \cdot d_G^2(v_i) + \cdots + \alpha_s \cdot d_G^s(v_i) + \cdots, \quad (1)$$

where $d_G^m(v_i)$ ($m = 1, 2, \dots, s, \dots$) denotes the various order degrees of node v_i , and constants α_m ($m = 1, 2, \dots, s, \dots$) are all nonnegative real numbers, with a sum of 1. If $\alpha_1 = 1$ and $\alpha_m = 0$ ($m = 2, 3, \dots, s, \dots$), the combined degree is the node degree in common sense, if $\alpha_2 = 1$ and $\alpha_m = 0$ ($m = 1, 3, \dots, s, \dots$), the combined degree is the 2-order degree, and so forth. For the values of parameters, we usually consider the case $\alpha_1 \geq \alpha_2 \geq \alpha_3 \geq \cdots$, where $\alpha_1 \geq \alpha_2$ indicates that a single neighbor's influence to v_i is no less than a single neighbor's influence. Notice that the value of constant α_m ($m = 1, 2, \dots, s, \dots$) may not be integers; therefore the combined degree $d_G^{Com}(v_i)$ usually not be integer. Since the combined degree is the combination of regular degree and various high-order degree, we need not discuss the mean combined degree and combined degree distribution. Same as above, we can define combined in-degree and combined out-degree of a directed network. According to Theorem 1, we could give the following formula of combined degree

sequence. Here we omit the formulas of combined in-degree sequence and combined out-degree sequence of directed networks.

$$d_G^{Com} = (\alpha_1 \cdot E_n + \alpha_2 \cdot A + \cdots + \alpha_s \cdot A^{s-1} + \cdots) d^1, \quad (2)$$

where E_n is the identity matrix of order n .

3. Results

In order to compare the behavior of degree and high-order degree and/or combined degree, we apply these attributes to a couple of best-known networks. We believe that the attributes of network with better behavior should be better to discover the differences among different nodes. By the definition of the differences of a given set of data, standard deviation (std) is the most important metric parameter. The greater the standard deviation is, the more diversities the data has, and the better the discrimination is, the better the attributes behavior is. As a complement to std , we define a novel parameter and call it as overflow ratio (OR_α), which denotes the ratio of the number of elements outside $[\alpha \cdot mean, 1/\alpha \cdot mean]$ to the number of elements in the given set of data, where $mean$ denotes the mean value of the given set of data and $0 < \alpha < 1$. The greater the overflow ratio is, the less the possibility that data cluster around the mean value of the given set is, and the better the discrimination is, the better the attributes behavior is. Therefore, these two parameters both reflect the diversity of a given set of data to some extent. Clearly, if one of the network attributes shows more diversity, it is easier to distinguish or sort the nodes.

3.1. Random Network. Random network was put forward for the first time by Paul Erdős and Alfred Rényi in 1960 (which is called *ER* Random Network) [11]. Here we generate a random network with n nodes, and we investigate the influence of node connection probability p taking different values on degree and high-order degree. Figure 3 shows the simulation results of 1000 times for node degree distribution, 2-order degree distribution, 4-order degree distribution, and combined degree distribution with $n = 10000$ and $p = 0.05$. We can tell from the results that the two high-order degrees and combined degree of random network still preserve the property of degree distribution, which is similar to Poisson Distribution (or Normal Distribution) with mean degree/mean high-order degree as the peak value.

Table 2 gives the discrimination of some attributes with $n = 10000$ and $p = 0.05$. Since the node degree distribution and high-order degree/combined degree distribution are all similar to Poisson Distribution (or Normal Distribution), we take big values for the overflow ratios. We can also tell from Table 2 that high-order degree has better discrimination compared with regular node degree, which is also true when n and p change (see Supplementary Table 1).

3.2. Small World Model. Watts [57] gave the basic attributes of Small World Model in 1999. Watts and Strogatz [6] proposed the construction of small world network by edge redistribution, which is called the *WS* small world network. Monasson

TABLE 2: Random Network: the discrimination parameters of 10000 simulations for different stages in Random Network, with number of nodes $n = 1000$ and connection probability $p = 0.05$.

Node degree	Standard Deviation	overflow ratio				
		$\alpha = 0.9$	$\alpha = 0.91$	$\alpha = 0.93$	$\alpha = 0.95$	$\alpha = 0.98$
degree	$2.80 * 10e1$	1.5%	2.73%	9.1%	23.08%	63.01%
2-order degree	$1.093 * 10e4$	1.55%	2.86%	9.1%	23.86%	64.48%
4-order degree	$2.747 * 10e9$	1.57%	2.9%	9.21%	23.96%	64.54%
combined degree	$3.440 * 10e8$	1.57%	2.9%	9.21%	23.96%	64.54%

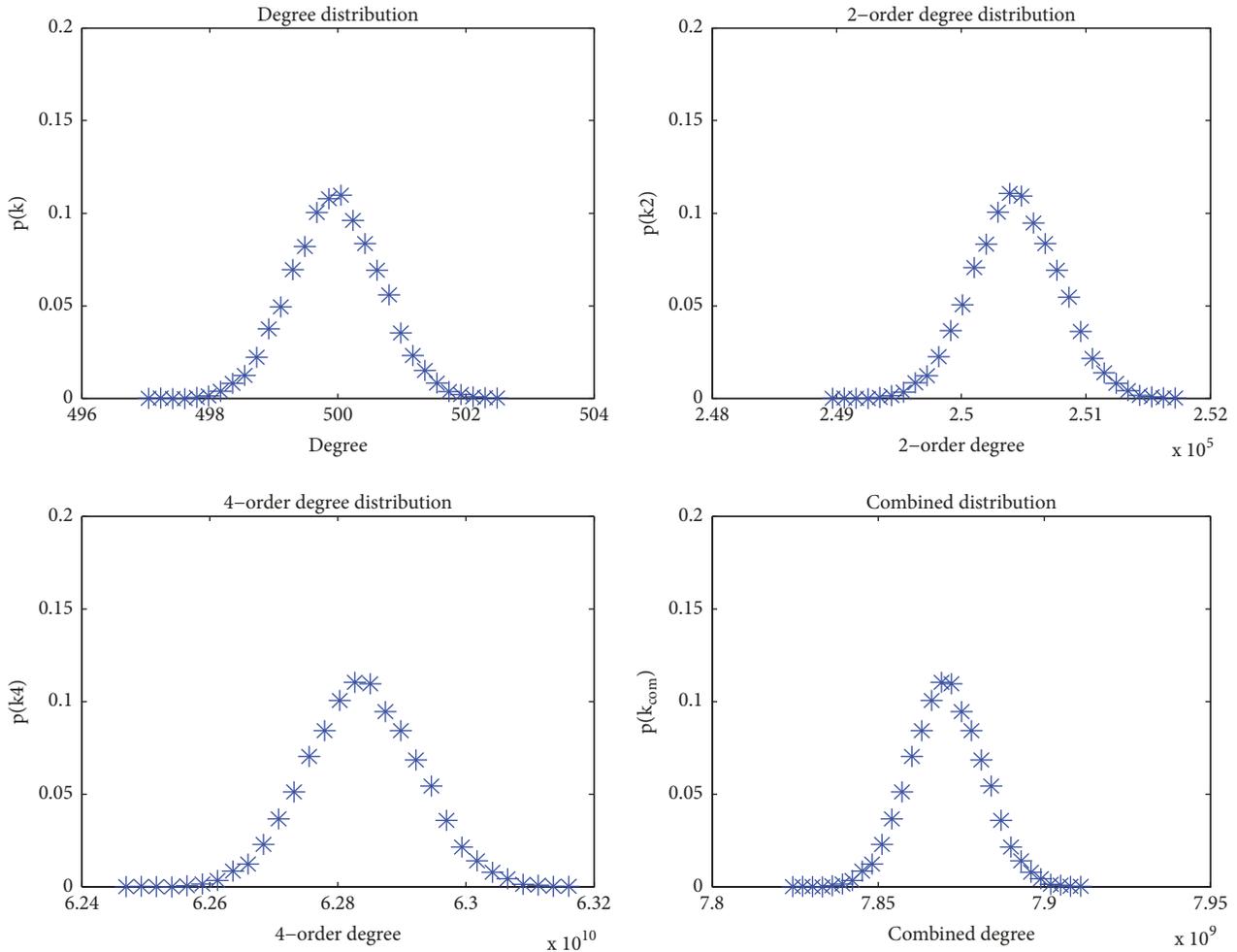


FIGURE 3: Random network: simulation results of 1000 times for node degree (mean) distributions with number of nodes $n = 10000$ and edge augment connection probability $p = 0.05$, taking $\alpha_1 = 1/2, \alpha_2 = 1/4, \alpha_3 = \alpha_4 = 1/8, \alpha_5 = \dots = 0$ for combined degree distribution.

[58] and Newman et al. [59] initiated the construction of small world network through edge augment, which is called the NW small world network. In this subsection we construct a small world network by edge augment, and investigate the behaviors of high-order degrees in small world network. Firstly, we construct a regular network with n nodes, and with each node connected to the nearest d nodes (where $n \gg d \gg \ln n \gg 1$). For a given probability p , we stochastic add $pnd/2$ edges to produce the small world network. We can tell from Figure 4 that the scatter plot of 2-order degree

distribution, 3-order degree distribution and 4-order degree distribution are all similar to Poisson Distribution (or Normal Distribution), which is the same as the degree distribution of the small world network.

For small world network, similar to random network, we are more concerned with the discrimination of small world network. Table 3 shows the mean discrimination of 1000 simulations for degree, 2-order degree, 4-order degree and 8-order degree, with $n = 800$, initial edge connection number $d = 30$, and connection probability $p = 0.05$. Of all

TABLE 3: Small world network: the discrimination parameters of 1000 simulations for different stages in small world network with $n = 5000$, $p = 0.05$, and $d = 20$.

order	Standard	overflow ratio			
	Deviation	$\alpha = 0.85$	$\alpha = 0.9$	$\alpha = 0.93$	$\alpha = 0.95$
degree	1.2476	0.05%	0.6102%	7.2873%	7.2954%
2-order degree	41.321	0.15%	1.57%	5.9382%	5.9425%
4-order degree	45674	0.21%	1.8%	9.87%	9.76%
8-order degree	5.268e + 010	0.4%	4%	16.8%	32.875%

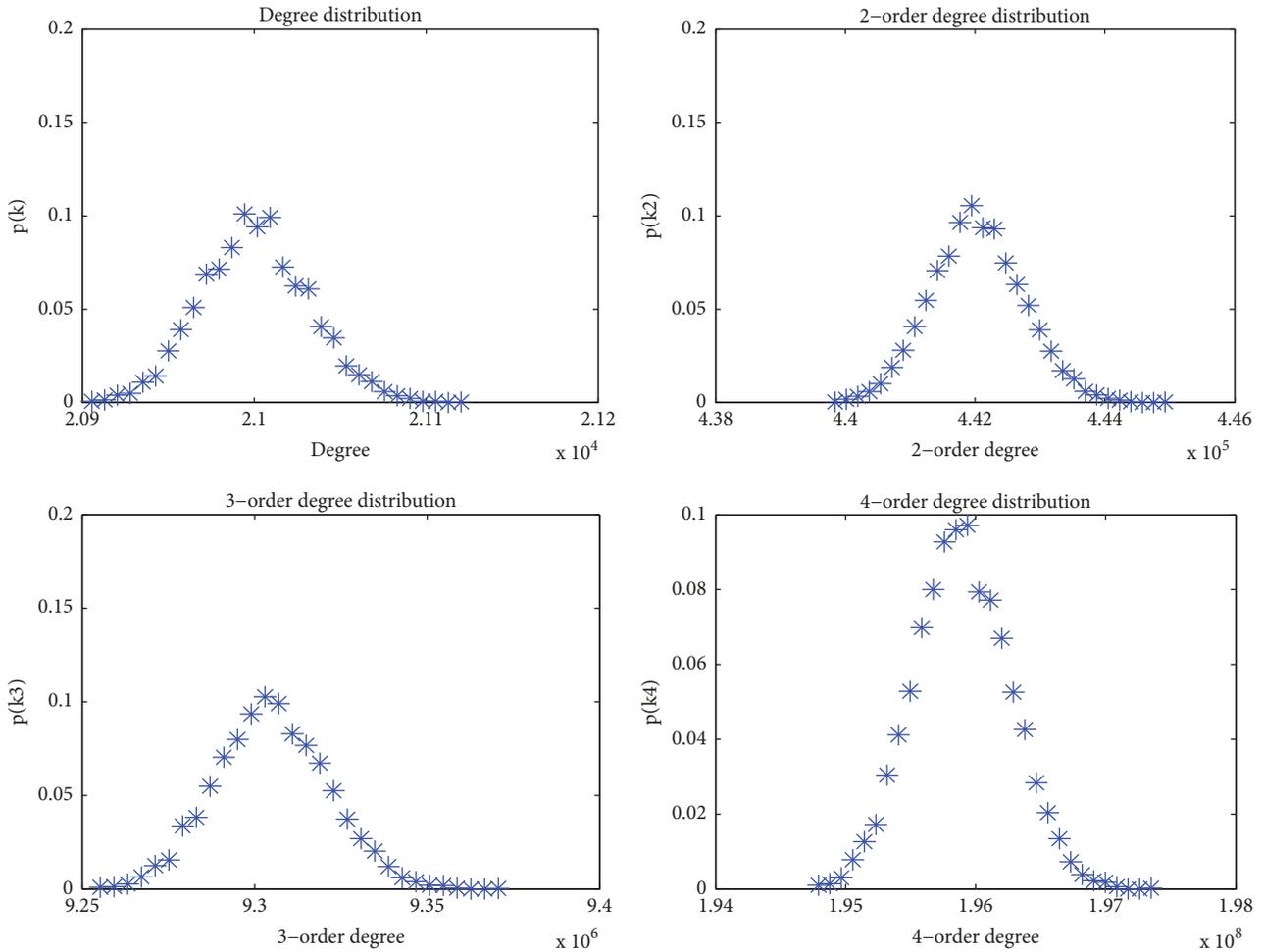


FIGURE 4: Small world network: simulation results of 1000 times for node degree distributions with number of nodes $n = 5000$ and given probability $p = 0.05$ and with each node connected to the nearest $d = 20$ nodes initially.

the situations, 4-order degree and 8-order degree show better discrimination results than the others. But sometimes the 2-order degree does not show good results in discrimination than degree, which is probably caused by the smallness of the value of parameter p . In Supplementary Table 2, we increase the value of p , and show the discrimination of degree, 2-order degree, and 4-order degree, with the number of nodes n and initial connection edges d change over 54 cases. Among all of the cases, the standard deviation discrimination gets bigger (better), 4-order degree shows smaller overflow ratio twice than that of node degree, and 2-order degree shows smaller

overflow ratio four times than that of node degree, which suggests that 4-order degree and 2-order degree are better than node degree in the aspect of overflow ratio. Moreover, as the order of high-order degree increases, the overflow ratio shows better results in discriminating the nodes.

3.3. Undirected Scale-Free Network. Barabási and Albert [18, 60, 61] proposed scale-free network (which is called BA scale-free network) with node degree distribution following the power-law distribution. In this subsection, we construct an undirected, scale-free network to investigate the behavior

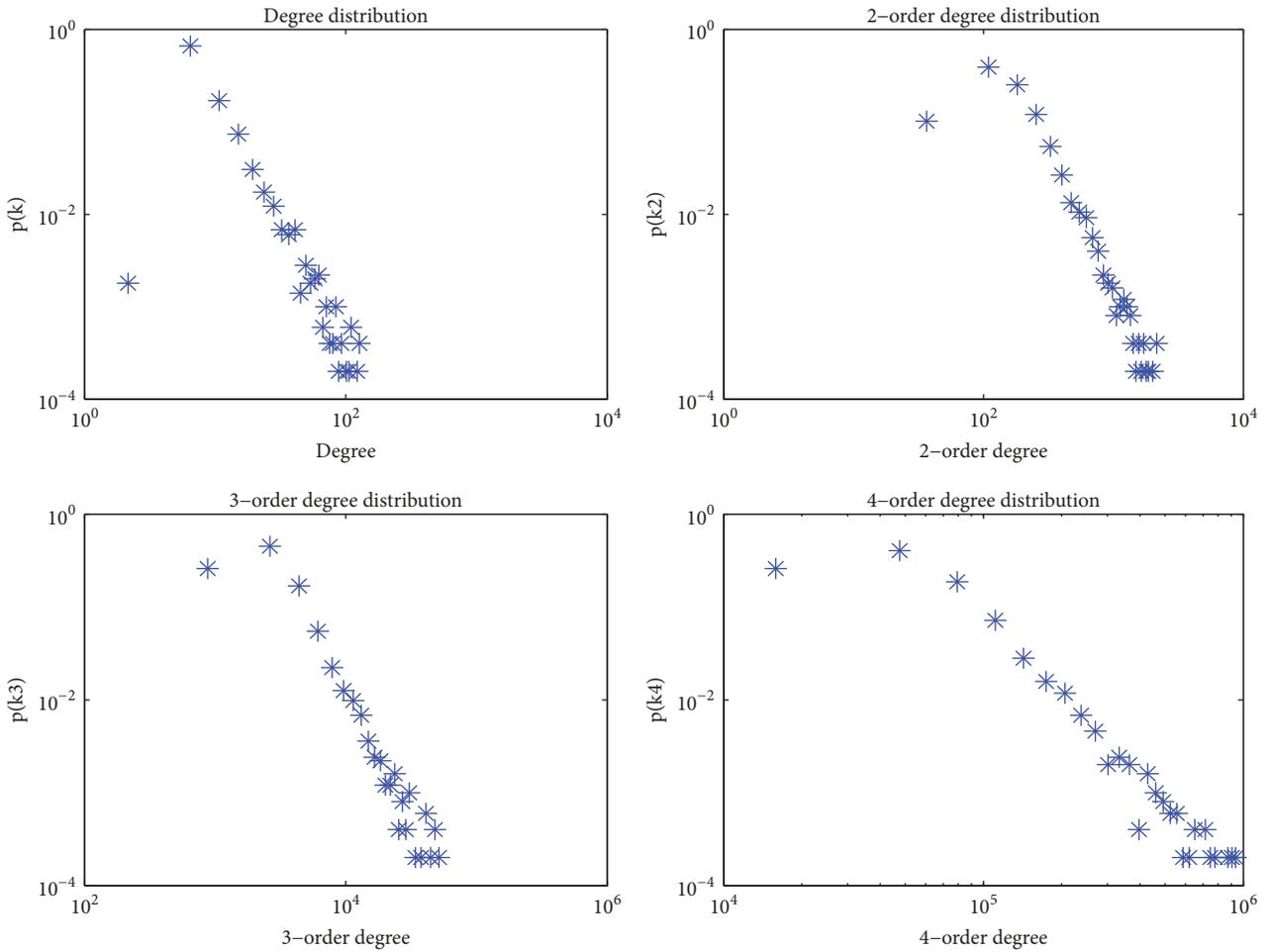


FIGURE 5: Undirected scale-free network: the node degree distribution and high-order degree distribution of a scale-free network, with $n = 100$, $p = 0.05$, $m = 5$, $t = 4900$.

of various orders of degree. Firstly, we produce a random network with the number of nodes n , stochastic connection probability p . Then we add one node each time whose degree equals m by the preferential attachment mechanism. Preferential attachment means that the more connected a node is, the more likely it is to receive new links. Nodes with higher degree have stronger ability to grab links added to the network. In our paper, the probability that new node connects to node i is proportional to the degree of i . Repeating this process for t times, we get an undirected scale-free network with $n + t$ nodes and about $np/2 + mt$ edges.

The high-order degree distribution and combined degree distribution are also power-law distribution; see Figure 5. Supplementary Table 3 gives the discrimination results of various orders of degree, with initial number of nodes n , initial node connection probability p , and running times t , and the new node added to the network each time has a degree of m . As a matter of fact, only if $0.3 \leq \alpha \leq 0.5$, the higher-order degree shows better discrimination results. The reason maybe that the range of node degree sequence is relatively large for scale-free network (compared with

random network and small world network), which coincides with our knowledge [18, 60, 61].

3.4. Directed Scale-Free Network. In this subsection, we construct a directed scale-free network. Firstly, produce a random network with nodes number n and connection probability p . Secondly assign the direction of each existing edge at random (with equal probability) to establish a directed network. Thirdly, run this step t times, add a new node at each time, and add m_1 in-edges by out-degree priority mechanism and m_2 out-edges by in-degree priority mechanism; there we usually choose $m_2 \geq m_1 \geq 0$. Therefore we obtain a directed network with $n + t$ nodes and $np/2 + (m_1 + m_2)t$ edges. Figure 6 shows that both the in-degree distribution and the out-degree distribution in the directed network we developed follow power-law distribution; therefore it is a directed scale-free network [56, 57]. But there are variations at the beginning of the graphs for high-order in-degree and high-order out-degree (especially the out-degree), even though the general shapes of these graphs still resemble the power-law distributions; see Figure 6. We suspect that this may be

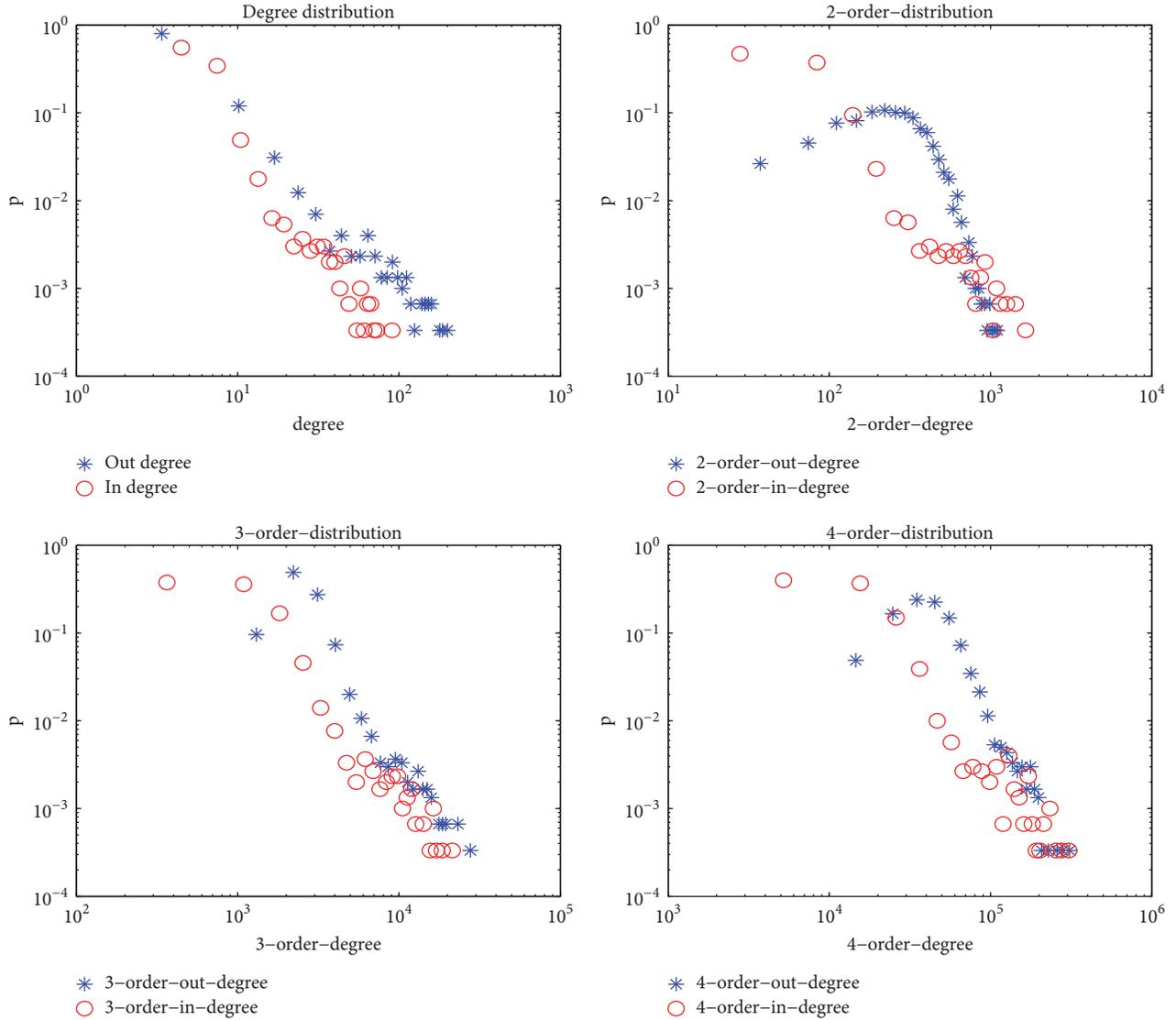


FIGURE 6: Directed scale-free network: node degree distribution and high-order distribution in the directed scale-free network with initial nodes $n = 100$, initial connection probability $p = 0.1$, the direction of an edge is assigned randomly, and $m_1 = 2, m_2 = 5$ in-edges/out-edges are added to network by the priority mechanism, running times $t = 2900$.

caused by the fact that we always add in-edges and out-edges at each time when we construct the network, and the numbers of edges maintain constants rather than random numbers. Since the discrimination comparison between high-order degree and node degree in directed scale-free network is similar to that in undirected scale-free network, we will not demonstrate the results here.

4. Discussion

The high-order degree in a network considers the influence of different path distances on a node. In a social network, consider a node is a person, and an edge is a friendship. Node degree of node v_a indicates the influence of v_a 's friends on v_a . The 2-order degree of v_a indicates the influence of v_a 's friends' friends on v_a . An interesting question is that, if v_a and v_b are

friends, there is an undirected edge connecting v_a and v_b , then v_a will influence v_b , which will in return have influence on v_a itself also. If v_a has many friends, v_a will have influence on all of its friends, which in return will affect v_a many times.

We know from Section 3 that high-order degrees and combined degree are superior to degree in discriminating nodes. However the computation of degree is simpler than high-order degrees and combined degree. Compared with the eigenvector centrality in undirected network, and PageRank in directed network, the high-order (out-/in-)degree and/or combined (out-/in-)degree we defined in this paper do not need iterations. The computation of our method only involves multiplications of matrices and vectors, which is clearly easier than eigenvector centrality or PageRank.

TABLE 4: Different centralities of Figure 2.

	Name of each node	1	2	3	4	5	6	7	8	9	10
Figure 2(a)	4-order degree	21	21	21	26	50	35	29	29	29	29
	Betweenness	0	0	0	21	20	26	0	0	0	0
	Eigenvector centrality	0.0547	0.0547	0.0547	0.1296	0.1431	0.2096	0.0884	0.0884	0.0884	0.0884
Figure 2(b)	3-order in-degree	0	0	1	2	0	2	3	4	4	6
	PageRank*	0	0	0	0.1795	0	0.1154	0.1795	0.1154	0.2308	0.1795
	PageRank**	0.0150	0.0318	0.0610	0.1255	0.0505	0.1134	0.1417	0.1283	0.1827	0.1501

* PageRank without damping factor.

** PageRank with damping factor equal to 0.85.

4.1. *The High-Order Degree and Eigenvector Centrality/Betweenness in Undirected Network.* In the undirected network, node degree is used to measure the influence of all the nodes connected to a node. 2-order degree considers the influence of a node's neighbor's neighbor on the node. 3-order degree considers the influence of a node's neighbors' neighbors' neighbor on that node. As a result, the lower the order of the high-order degree is, the closer it is to the node degree. On the other hand, the higher the order of the high-order degree is, the closer it is to eigenvector centrality.

Theorem 2. *For s -order degree in undirected network, if $s = 1$, it is the node degree of the network. As s approaches to infinity, s -order degree is equivalent to eigenvector centrality.*

We only need to show that when s approaches to infinity, s -order degree is equivalent to eigenvector centrality. Suppose A is the adjacency matrix in an undirected network and d^1 is the degree sequence. Then by Theorem 1, the s -order degree sequence is $A^{s-1} * d^1$, which is consistent with calculating the largest eigenvalue and the corresponding eigenvector by the method of power rule [62]. Therefore, the normalized result of $A^{s-1} * d^1 (s \rightarrow +\infty)$ is the node eigenvector centrality. If the eigenvector centrality is deemed as the most accurate method in ranking nodes, then from node degree to high-order degree, and to eigenvector centrality, the accuracies get higher and higher, and the computation complexity gets higher and higher at the same time. Therefore, if there is not high requirement on ranking accuracy and computation complexity, high-order degree is a relatively good choice.

In fact, high-order degrees are not simple alternative to eigenvector centrality. Sometimes high-order degrees may be more intuitive than some methods including eigenvector centrality. For example, Figure 2(a), the top three nodes sorted by both 2-order degree and 4-order degree are 5, 6, 4, which we can see from Tables 1 and 4; it is the same as our supposition. Otherwise the top three nodes sorted by both betweenness and eigenvector centrality are 6, 5, 4.

4.2. *The High-Order Degree and PageRank for Directed Network*

Theorem 3. *For s -order out-degree in a directed network, if $s = 1$, it is the out-degree of the directed network. For s -order in-degree in a directed network, if $s = 1$, it is the in-degree of*

the directed network. When s approaches to infinity, the s -order in-degree is equivalent to PageRank (without damping factor).

Theorem 3 is clearly true because PageRank is the simplification of eigenvector centrality in directed network [63]. Moreover, we can reach similar conclusions that when ranking network nodes, high-order in-degree is a good choice if one has certain but not high requirements on the accuracy and computation complexity.

Same as Section 4.2, high in-order degrees are not simple alternative to PageRank. Sometimes high in-order degrees may be more intuitive than PageRank. For example Figure 2(b), the top node sorted by both 2-order in-degree and 3-order in-degree is 9, which we can see from Tables 1 and 4, it is the same with our supposition. Otherwise the top node sorted by both PageRank without damping factor and PageRank with damping factor is 9.

4.3. *The Significance of Network High-Order Out-Degree.* PageRank (high-order in-degree) indicates that the PR value of a node v_a is larger when there are more nodes pointing to v_a in a directed network (quantity hypothesis) and/or the nodes pointing v_a have larger PR values (quality hypothesis) [35, 36]. Reversely, we call high-order out-degree Reverse PageRank (RPR). The RPR value of a node v_b in a directed network is determined by how many nodes pointed by v_b (reverse quantity hypothesis) and/or how large RPR value that nodes pointed by v_b have (reverse quality hypothesis). For example, we can say that the World Wide Web navigation websites have large RPR value. When someone surfs the Internet, if he does not know which website is worth visiting (with larger PR value), he should start with the navigation website (with larger RPR value). There is more research on the high-order in-degree (PR) for directed network, but there is no research on high-order out-degree (RPR) as far as we know.

Figure 2(b), although nodes 1, 2, 5 have the same 2-order out-degree, only node 1 has the largest n -order out-degree ($n \geq 3$). Node 1 is the only one that can reach any nodes of this graph.

4.4. *The Combined Degree We Defined and the Mixed Degree Zeng Defined.* Zeng [44] used the weighted sum of both the residual degree and the exhausted degree to define the mixed degree of the node of a network. On one hand, Zeng's definition involves Mixed Degree Decomposition (MDD) to

the whole network by using coreness centrality, while the combined degree we defined is based on a linear combination of various high-order degrees. Even though our method involves more terms in the weighted sum, it does not need to consider network decomposition. Therefore it is cheaper in computation than Zeng's definition. On the other hand, Zeng divides the nodes that are connected to a node into two classes according to *MDD*, while we divide the nodes that are connected to a node into finite number of classes according to path distance, which is more delicate in classification and has more accurate results.

5. Conclusion

In this paper, we define several novel centrality metrics: the high-order (out-/in-)degree and combined degree. For the values of combined degree's parameters, we usually consider the case $\alpha_1 \geq \alpha_2 \geq \alpha_3 \geq \dots \geq 0$ and $\sum_i \alpha_i = 1$. We prove that both the degree centrality and eigenvector centrality are the special cases of the high-order degree of undirected network, and both the in-degree and PageRank algorithm without damping factor are the special cases of the high-order in-degree of directed network. We present several experiments to discuss the performance of our novel centrality metrics. It can be seen from the experiments that the centrality metrics we defined are easy to calculate and perform better than degree centrality. In a large-scale complex network study, our centrality metrics will be an effective alternative to the eigenvector centrality/PageRank algorithm. The manuscript is only limited in introducing the definition of new metrics. We hope to discuss their efficacy and computational cost in the further works.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

Authors' Contributions

All authors worked together to produce the results and read and approved the final manuscript.

Acknowledgments

The research is supported by the National Natural Science Foundation of China (61572522 and 11371230) and Shandong Provincial Natural Science Foundation (ZR2018PF004).

Supplementary Materials

Supplementary 1. Supplementary Table 1: in the random network, we increase the value of p and show the discrimination of degree, 2-order degree, and 4-order degree, with the

number of nodes n and initial connection edges d changing over 27 cases.

Supplementary 2. Supplementary Table 2: in the small world network, we increase the value of p and show the discrimination of degree, 2-order degree, and 4-order degree, with the number of nodes n and initial connection edges d changing over 54 cases.

Supplementary 3. Supplementary Table 3: in the undirected scale-free network, the discrimination results of various orders of degree, with initial number of nodes n , initial node connection probability p , running times t , and the new node added to the network each time have a degree of m over 90 cases.

References

- [1] M. E. Newman, "The structure and function of complex networks," *SIAM Review*, vol. 45, no. 2, pp. 167–256, 2003.
- [2] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Reviews of Modern Physics*, vol. 74, no. 1, pp. 47–97, 2002.
- [3] A. Rapoport, "Contribution to the theory of random and biased nets," *Bulletin of Mathematical Biology*, vol. 19, no. 4, pp. 257–277, 1957.
- [4] A. Rapoport and W. J. Horvath, "A study of a large sociogram," *Behavioural Science*, vol. 6, pp. 279–291, 1961.
- [5] S. Wasserman and K. Faust, *Social Network Analysis: Methods and Applications*, Cambridge University Press, Cambridge, UK, 1994.
- [6] D. J. Watts and S. H. Strogatz, "Collective dynamics of small-world networks," *Nature*, vol. 393, pp. 440–442, 1998.
- [7] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, "Size-dependent degree distribution of a scale-free growing network," *Physical Review E: Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, vol. 63, no. 6, 2001.
- [8] J. Lee, C. Heaukulani, Z. Ghahramani, L. F. James, and S. Choi, "Bayesian inference on random simple graphs with power law degree distributions," in *Proceedings of the International Conference on Machine Learning*, pp. 2004–2013, 2017.
- [9] F. Mohd-Zaid, C. M. S. Kabban, R. F. Deckro, and E. D. White, "Parameter specification for the degree distribution of simulated Barabási-Albert graphs," *Physica A: Statistical Mechanics and Its Applications*, vol. 465, pp. 141–152, 2017.
- [10] X. Wang, S. Trajanovski, R. E. Kooij, and P. Van Mieghem, "Degree distribution and assortativity in line graphs of complex networks," *Physica A: Statistical Mechanics and Its Applications*, vol. 445, pp. 343–356, 2016.
- [11] P. Erdős and A. Rényi, "On the evolution of random graphs," *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, vol. 38, pp. 17–61, 2012.
- [12] P. Erdős, "On extremal problems of graphs and generalized graphs," *Israel Journal of Mathematics*, vol. 2, pp. 183–190, 1964.
- [13] R. Albert, H. Jeong, and A.-L. Barabási, "Internet: diameter of the World-Wide Web," *Nature*, vol. 401, no. 6749, pp. 130–131, 1999.
- [14] V. M. Eguiluz and K. Klemm, "Epidemic threshold in structured scale-free networks," *Physical Review Letters*, vol. 89, no. 10, Article ID 108701, 2002.

- [15] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A. L. Barabási, “The large-scale organization of metabolic networks,” *Nature*, vol. 407, pp. 651–654, 2002.
- [16] J. Rung, T. Schlitt, A. Brazma, K. Freivalds, and J. Vilo, “Building and analysing genome-wide gene disruption networks,” *Bioinformatics*, vol. 18, no. 2, pp. S202–S210, 2002.
- [17] S. Battiston, J. F. Rodrigues, and H. Zeytinoglu, “The network of inter-regional direct investment stocks across Europe,” *Advances in Complex Systems (ACS)*, vol. 10, no. 1, pp. 29–51, 2007.
- [18] A.-L. Barabási and R. Albert, “Emergence of scaling in random networks,” *American Association for the Advancement of Science: Science*, vol. 286, no. 5439, pp. 509–512, 1999.
- [19] L. D. F. Costa, O. N. Oliveira, G. Travieso et al., “Analyzing and modeling real-world phenomena with complex networks: a survey of applications,” *Advances in Physics*, vol. 60, no. 3, pp. 329–412, 2011.
- [20] L. Li, Y. Wei, C. To et al., “Integrated Omic analysis of lung cancer reveals metabolism proteome signatures with prognostic impact,” *Nature Communications*, vol. 5, article no. 5469, 2014.
- [21] Y. Zhang, M. Zhao, J. Su, X. Lu, and K. Lv, “Novel model for cascading failure based on degree strength and its application in directed gene logic networks,” *Computational & Mathematical Methods in Medicine*, vol. 2018, Article ID 8950794, 2018.
- [22] S. Wang, Y. Chen, Q. Wang, E. Li, Y. Su, and D. Meng, “Analysis for gene networks based on logic relationships,” *Journal of Systems Science & Complexity*, vol. 23, no. 5, pp. 999–1011, 2010.
- [23] H. Jeong, S. P. Mason, A. L. Barabási, and Z. N. Oltvai, “Lethality and centrality in protein networks,” *Nature*, vol. 411, no. 6833, pp. 41–42, 2001.
- [24] Y. Zhang, K. Lv, S. Wang, J. Su, and D. Meng, “Modeling gene networks in *saccharomyces cerevisiae* based on gene expression profiles,” *Computational and Mathematical Methods in Medicine*, vol. 2015, Article ID 621264, 10 pages, 2015.
- [25] M. C. González, C. A. Hidalgo, and A.-L. Barabási, “Understanding individual human mobility patterns,” *Nature*, vol. 453, no. 7196, pp. 779–782, 2008.
- [26] W. Wang and T. Zhang, “Caspase-1-mediated pyroptosis of the predominance for driving CD₄⁺ T cells death: a nonlocal spatial mathematical model,” *Bulletin of Mathematical Biology*, vol. 80, no. 3, pp. 540–582, 2018.
- [27] M. Rubinov and E. Bullmore, “Schizophrenia and abnormal brain network hubs,” *Dialogues in Clinical Neuroscience*, vol. 15, no. 3, pp. 339–349, 2013.
- [28] J. Bohannon, “Counterterrorism’s new tool: ‘metanetwork’ analysis,” *Science*, vol. 325, no. 5939, pp. 409–411, 2009.
- [29] T. Zhang, X. Meng, and T. Zhang, “Global dynamics of a virus dynamical model with cell-to-cell transmission and cure rate,” *Computational and Mathematical Methods in Medicine*, Article ID 758362, 8 pages, 2015.
- [30] D. W. Franks, J. Noble, P. Kaufmann, and S. Stagl, “Extremism propagation in social networks with hubs,” *Adaptive Behavior*, vol. 16, no. 4, pp. 264–274, 2008.
- [31] A. Bavelas, “A mathematical model for group structures,” *Human Organization*, vol. 7, no. 3, pp. 16–30, 1948.
- [32] G. Sabidussi, “The centrality index of a graph,” *Psychometrika*, vol. 31, pp. 581–603, 1966.
- [33] L. C. Freeman, “Centrality in social networks conceptual clarification,” *Social Networks*, vol. 1, no. 3, pp. 215–239, 1978.
- [34] P. Bonacich, “Technique for analyzing overlapping memberships,” *Sociological Methodology*, vol. 4, pp. 176–185, 1972.
- [35] E. Costenbader and T. W. Valente, “The stability of centrality measures when networks are sampled,” *Social Networks*, vol. 25, no. 4, pp. 283–307, 2003.
- [36] L. Page, *The PageRank citation ranking: bringing order to the web*, *Stanford Digital Libraries Working Paper*, vol. 9, 1998.
- [37] S. Brin and L. Page, “The anatomy of a large-scale hypertextual web search engine,” *Computer Networks*, vol. 56, no. 18, pp. 3825–3833, 2012.
- [38] M. A. Kłopotek, S. T. Wierzchoń, R. A. Kłopotek, and E. A. Kłopotek, “Traditional PageRank versus network capacity bound,” in *Social and Information Networks*, 2017.
- [39] T. H. Haveliwala, “Topic-sensitive PageRank,” in *Proceedings of the 11th International Conference on World Wide Web, WWW ’02*, pp. 517–526, May 2002.
- [40] A. N. Langville and C. D. Meyer, “Deeper inside PageRank,” *Internet Mathematics*, vol. 1, no. 3, pp. 335–380, 2004.
- [41] M. Bressan, E. Peserico, and L. Pretto, “The power of local information in pagerank,” in *Proceedings of the 22nd International Conference on World Wide Web, WWW 2013*, pp. 179–180, May 2013.
- [42] I. M. Kloumann, J. Ugander, and J. Kleinberg, “Block models and personalized PageRank,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 114, no. 1, pp. 33–38, 2017.
- [43] A. Altman and M. Tennenholtz, “Ranking systems: the PageRank axioms,” in *Proceedings of the EC’05: 6th ACM Conference on Electronic Commerce*, pp. 1–8, June 2005.
- [44] J. Ai, H. Zhao, K. M. Carley, Z. Su, and H. Li, “Neighbor vector centrality of complex networks based on neighbors degree distribution,” *The European Physical Journal B*, vol. 86, no. 4, 2013.
- [45] A. Zeng and C.-J. Zhang, “Ranking spreaders by decomposing complex networks,” *Physics Letters A*, vol. 377, no. 14, pp. 1031–1035, 2013.
- [46] J. Bae and S. Kim, “Identifying and ranking influential spreaders in complex networks by neighborhood coreness,” *Physica A: Statistical Mechanics and Its Applications*, vol. 395, pp. 549–559, 2014.
- [47] U. Brandes, “A faster algorithm for betweenness centrality,” *Journal of Mathematical Sociology*, vol. 25, no. 2, pp. 163–177, 2001.
- [48] K. Goel, R. R. Singh, S. Iyengar, and S. Gupta, “A faster algorithm to update betweenness centrality after node alteration,” *Internet Mathematics*, vol. 11, no. 4–5, pp. 403–420, 2015.
- [49] R. Guns, Y. X. Liu, and D. Mahbuba, “Q-measures and betweenness centrality in a collaboration network: a case study of the field of informetrics,” *Scientometrics*, vol. 87, no. 1, pp. 133–147, 2011.
- [50] J. D. Noh and H. Rieger, “Random walks on complex networks,” *Physical Review Letters*, vol. 92, no. 11, Article ID 118701, 2004.
- [51] V. Tejedor, O. Bénichou, and R. Voituriez, “Global mean first-passage times of random walks on complex networks,” *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, vol. 80, no. 6, 2009.
- [52] M. Rosvall and C. T. Bergstrom, “Maps of random walks on complex networks reveal community structure,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 4, pp. 1118–1123, 2008.
- [53] J. Saramäki and K. Kaski, “Scale-free networks generated by random walkers,” *Physica A: Statistical Mechanics and Its Applications*, vol. 341, no. 1–4, pp. 80–86, 2004.

- [54] T. Weng, J. Zhang, M. Khajehnejad, M. Small, R. Zheng, and P. Hui, "Navigation by anomalous random walks on complex networks," *Scientific Reports*, vol. 6, 2016.
- [55] E. K. Lloyd, J. A. Bondy, and U. S. Murty, "Graph theory with applications," *The Mathematical Gazette*, vol. 28, pp. 237-238, 1976.
- [56] E. Terzi and M. Winkler, "A spectral algorithm for computing social balance," in *Algorithms and Models for the Web Graph*, vol. 6732 of *Lecture Notes in Comput. Sci.*, pp. 1-13, Springer, Heidelberg, 2011.
- [57] J. Tang, Y. Chang, C. Aggarwal, and H. Liu, "A survey of signed network mining in social media," in *ACM Computing Surveys (CSUR)*, vol. 49, pp. 237-238, 2015.
- [58] D. J. Watts, *Small Worlds: The Dynamics of Networks between Order and Randomness*, Princeton University Press, Princeton, NJ, USA, 1999.
- [59] R. Monasson, "Diffusion, localization and dispersion relations on small-world lattices," *The European Physical Journal B*, vol. 12, pp. 555-567, 1999.
- [60] M. E. Newman and D. J. Watts, "Renormalization group analysis of the small-world network model," *Physics Letters A*, vol. 263, no. 4-6, pp. 341-346, 1999.
- [61] A. L. Barabási, R. Albert, and H. Jeong, "Mean-field theory for scale-free random networks," *Physica A Statistical Mechanics & Its Applications*, vol. 272, pp. 173-178, 1999.
- [62] A. Barabási, E. Ravasz, and T. Vicsek, "Deterministic scale-free networks," *Physica A: Statistical Mechanics and Its Applications*, vol. 299, no. 3-4, pp. 559-564, 2001.
- [63] M. E. J. Newman, "The mathematics of networks," in *The New Palgrave Encyclopedia of Economics*, vol. 208, pp. 1-12, 2008.



Hindawi

Submit your manuscripts at
www.hindawi.com

