Physica A 392 (2013) 6107–6117

Contents lists available at ScienceDirect

Physica A

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

A centrality measure for communication ability in weighted network

Li Zhai [a,b], Xiangbin Yan [a,*], Guojing Zhang [c]

[a] School of Management, Harbin Institute of Technology, Harbin, 150001, China
[b] School of Applied Science, Harbin University of Science and Technology, Harbin, 150080, China
[c] School of Mathematics and Statistics, Northeast Normal University, Changchun, 130024, China

HIGHLIGHTS

- We proposed a new node centrality measurement in weighted network.
- We investigated the properties of the communication centrality.
- It is superior to other centrality measures in the use of information.
- It contains a well-balanced mix of other centrality measures.

ARTICLE INFO

Article history:
Received 6 March 2013
Received in revised form 27 June 2013
Available online 7 August 2013

Keywords:
Centrality measure
Weighted network
H-index
Lobby index
H-degree

ABSTRACT

This paper proposes a new node centrality measurement in a weighted network, the communication centrality, which is inspired by Hirsch's *h*-index. We investigated the properties of the communication centrality, and proved that the distribution of the communication centrality has the power-law upper tail in weighted scale-free networks. Relevant measures for node and network are discussed as extensions. A case study of a scientific collaboration network indicates that the communication centrality is different from other common centrality measures and other *h*-type indexes. Communication centrality displays moderate correlation with other indexes, and contains a well-balanced mix of other centrality measures and cannot be replaced by any of them.

1. Introduction

Efficient communication means high impact (wide access or high reach) and low cost, whether in communication networks, or in social and biological networks [1]. In a complex network, the roles, positions, influence and centrality of nodes is usually expressed by node degree centrality [2–4], node strength [5], closeness centrality [6,7], betweenness centrality [2] and eigenvector centrality [8] etc. Nonetheless, none of these indexes can accurately capture the communication ability of the nodes [1].

This paper brings forth a new concept of communication centrality in the weighted network to measure the communication ability of the node. We principally considered three major influencing factors. First, node degree, namely the number of neighbor nodes, is the most direct perceived factor exerting effects on the communication ability of the node. Clearly, more neighbor nodes indicate more effective communication of a node since it can pass on and receive information through more channels. Second, the stronger communication ability of the neighbor nodes of a node demonstrates the greater influence of its communication since the neighbor nodes boast a tremendous capacity to communicate. For instance in the network

* Corresponding author. Tel.: +86 451 86414017.
E-mail address: xbyan@hit.edu.cn (X. Yan).

0378-4371/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.physa.2013.07.056

of lobbyists (or diplomats), the persuasion of a lobbyist who is stronger in lobby ability can produce better results than that of a weaker lobbyist. Third, the edge weight of a node has a remarkable influence on the communication ability of this node. Without loss of generality, we suppose that greater edge weight represents faster arrival or lower communication costs or higher trust. Therefore greater edge weight means the stronger ability of this node to communicate with its neighbor nodes. For example, in a collaboration network, greater edge weight indicates more frequent collaboration, and then the bilateral communication is obviously highly efficient. For another example, in the network of friends, if the edge weight represents the intimacy degree or the trust, then the edge weight will exert influence on the communication between friends and further influence the communication ability of a person.

Hirsch [9] proposed the h-index, which integrates the amount of papers and citation times of papers to measure the academic achievements or influence of scholars. "A scientist has h-index h if h of his or her N_p papers have at least h citations each and the other $(N_p - h)$ papers have $\leq h$ citations each" [9]. The H-index simply and effectively measures the key part of a dataset in a relatively natural way [10]. Since its introduction, the h-index and some related bibliometric indices have received a lot of attention from the scientific community in the last few years [11].

Scholars have applied the h-index and some other indexes in the network to measure the centrality of the nodes. Zhao et al. [10] stated that the h-degree in the weighted network can be used as the centrality measure of nodes. "The h-degree of node x in a weighted network is equal to k if k is the largest natural number such that x has at least k links each with strength at least equal to k" [10]. Furthermore, Zhao et al. [12] promoted the concept of h-degree to the directional weighted network and introduced the directed h-degree [12]. But h-degree does not consider the influence of neighbor nodes upon the centrality of this node. Schubert [13] proposed the partnership ability index φ [13], where φ is a special case of the h-degree [14]. In 2009, Korn et al. put forward the lobby index for the non-weighted network to describe communication ability [1]. "The lobby index of a node x is the largest integer k such that x has at least k neighbors with a degree of at least k" [1]. Zhao et al. [10] promoted the lobby index to be the w-lobby index in the weighted network, stating that "the w-lobby index of a node x is the largest integer k such that x has at least k neighbors with node strength at least k" [10]. Consistent with [5,10], in this paper we define the node strength of a node in a weighted network as the sum of weights of all its edges. Campiteli et al. studied the nature of lobby index combining a actual non-weighted biological network and a linguistic network [15]. The lobby index and w-lobby index integrate node degree and degree (strength) of neighbor nodes. They are superior to indexes measuring node centrality without using information of neighbor nodes for communication ability, such as node degree, node strength and h-degree. However, lobby index and w-lobby index neglect an important factor: different edge weights of this node mean different communication ability. Additionally, the lobby index displays strong correlation with degree centrality [1].

Based on the idea of h-index, this paper proposes a communication centrality to measure node centrality reflecting a communication ability which is suitable for the analysis of weighted undirected networks. The structure of the paper is as follows: Section 2 gives the definition of communication centrality and discusses its theoretical properties. Section 3 documents a case study conducted to comprehensively evaluate the communication centrality in a large co-author network, along with other well-known centrality indexes. Section 4 concludes the paper.

2. Methodology

2.1. The communication centrality

In this section, we will define communication centrality and discuss its property. As mentioned, the communication ability of a node's neighbor nodes in the weighted network stands as the important factor influencing its communication ability. Then, how is the communication ability of the neighbor nodes reflected? For a given node, larger degree means more neighbor nodes and higher communication ability in the network, and greater edge weight with neighbor nodes indicates more frequent contact between this node and its neighbor nodes, which means less contact cost and higher communication ability. In addition, the node degree and node strength cannot be replaced by each other. Lobby index and w-lobby index measure the communication ability of neighbor nodes only through their degree or node strength. However, node degree and edge weight both should be considered in measuring the communication ability of neighbor nodes. Fig. 1 gives an example of a weighted network. The width of the line shows the strength of weight with the specific numerical values. In Fig. 1, the node degrees of node B and E are both 3 but their node strengths are respectively 11 and 5; the node strengths of node A and E are both 5 with different node degrees, 1 and 3 respectively. Thus, in a bid to combine the information of node degree and node strength, we adopt h-degree [10] to measure the communication ability of neighbor nodes since it can mirror the two factors simultaneously.

Considering factors of degree, edge weight of a node and communication ability of its neighbor node simultaneously, and based on the algorithm of the h-index, this paper proposes a new communication centrality for measuring the communication ability of a node. For a node, each edge and linked neighbor node can be regarded as a communication channel. To measure the communication ability of the channel, the communication centrality calculates the product of edge weight and h-degree of the linked neighbor node. Then, communication centrality applies the h-index to balance the communication ability of these channels and the number of communication channels. Here, the number of communication channels is node degree; it is equivalent to the number of papers when using the h-index to calculate the academic

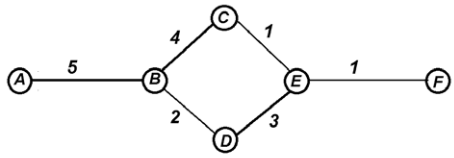

Fig. 1. Example of the communication centrality in a weighted network.

Serial number		Product($w_{Ey}\times d_h(y)$)
1	<	6(3×2)
2	>	1(1×1)
3	>	1(1×1)

Fig. 2. Calculation of communication centrality of node E of Fig. 1. Here w_{Ey} indicates the edge weight value of node E and node y, $d_h(y)$ means the h-degree of node y.

achievement of a scholar, and the communication ability of every communication channel is equivalent to the citations of every paper.

Note that there are many weighted networks with natural numbers as edge weights in complex networks, such as the collaboration networks with collaboration frequency as the edge weight and the microblog users networks with forwarding frequency as the edge weights. So, to probe, we consider the weighted networks whose edge weights are of natural numbers in this paper.

Definition 1 (*Communication Centrality*)**.** The communication centrality $c(x)$ of node x is the largest integer k such that the node x has at least k neighbor nodes satisfying the product of each node's h-degree and the weight of the edge linked with node x is no fewer than k.

This is to say, if the h-degree of M neighbor nodes $y_1, y_2, \ldots, y_M$ of node x are marked as $d_h(y_1), d_h(y_2), \ldots, d_h(y_M)$, and the edge weights of node x and $y_1, y_2, \ldots, y_M$ are marked as $w_1, w_2, \ldots, w_M$, then their product sequence is

$$w_1 d_h(y_1), w_2 d_h(y_2), \ldots, w_M d_h(y_M).$$

Here we can suppose

$$w_1 d_h(y_1) \geq w_2 d_h(y_2) \geq \cdots \geq w_M d_h(y_M).$$

Then

$$c(x) = \max\{k : w_k d_h(k) \geq k\}.$$

Obviously, the communication centrality can measure the communication ability of the node. Higher edge weights indicate more frequent communications and lower communication cost. Greater h-degree of neighbor nodes indicates more powerful communication ability and greater influence. Communication centrality comprehensively considers the two factors and calculates their product and applies the h-index to balance the products and node degree. It is thus seen that communication centrality is suitable for describing the communication ability of nodes.

The definition of the communication centrality tells that in a weighted network, if a node has many neighbor nodes with which it has greater edge weight or some of the neighbor nodes have strong communication ability (here the strong communication ability refers to the node that has many neighbors boasting greater edge weight, namely the h-degree), then its communication centrality will be higher. That is, communication centrality is the monotone nondecreasing function of edge weights and node degree. Here the node refers to both node x and its neighbor nodes; the edge both refers to the edge of node x and edge of its neighbor nodes.

We calculate the communication centrality of node E in Fig. 1 as an example. E has three neighbor nodes, namely C, D and F, which have respectively edge weights 1, 3, 1 with E. The h-degrees of C, D and F are 1, 2 and 1. Specifically, node F has only one edge. So, its h-degree is 1. Node D has 2 edges and their edge weights are 3 and 2 in descending order. So the h-degree of node D is 2, that is to say, node D has at least 2 edges whose weights are no lower than 2 and no weight of 3 edges is lower than 3. Similarly, the h-degree of node C is 1. Then we calculate the products of h-degrees of neighbor node and their edge weight with E respectively, and rank the products in descending order (shown in Fig. 2). It can be seen that E has at least 1 neighbor node whose product is no less than 2. So the communication centrality of E is 1. In addition, the node degree, node strength, communication centrality, lobby index, w-lobby index and h-degree of the 6 nodes can be calculated according to Fig. 1 with the results shown in Table 1. The results indicate that sequencing of node centralities according to different indexes leads to different results. Such differences rest in the fact that they measure the importance of nodes from diverse perspectives. Sequencing the communication ability of the node according to the value of communication centrality in descending order, the result is $B > D > A = C = E = F$.

Table 1
Degree centrality, node strength, h-degree, lobby index, w-lobby index and communication centrality of Fig. 1.

Index	Node					
	A	B	C	D	E	F
Degree centrality	1	3	2	2	3	1
Node strength	5	11	5	5	5	1
h-degree	1	2	1	2	1	1
Lobby index	1	2	2	2	2	1
w-lobby index	1	3	2	2	2	1
Communication centrality	1	3	1	2	1	1

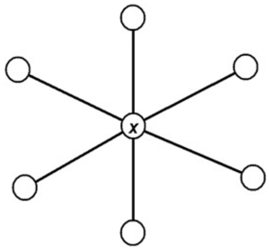

Fig. 3. Example of a star network.

2.2. Comments of the communication centrality

2.2.1. Properties of communication centrality

Generally speaking, if we compare communication centrality with other indexes in a weighted network, we can have the following observations. Supposing G is any weighted network whose edge weights are of natural numbers, let N denote the total number of nodes in network G. $d(x)$, $s(x)$, $l(x)$, $wl(x)$, $d_h(x)$ and $c(x)$ respectively represent node degree, node strength, lobby index, w-lobby index, h-degree and communication centrality of node x. The following properties show relations between the communication centrality and degree centrality, lobby index, w-lobby index and h-degree.

Property 1. *For any isolated node x,* $c(x) = 0$*; and for any non-isolated node x,* $1 \le c(x) \le d(x) \le N - 1 < N$.

Proof. According to the definition of communication centrality, the Property 1 clearly is correct.

According to the definitions of lobby index and w-lobby index and the Property 1, the following Properties 2 and 3 clearly are correct.

Property 2. *For any node x with degree* 1, $c(x) = l(x) = wl(x) = d(x) = 1$.

Property 3. *For any node x in the globally coupled network which consists of M nodes network,* $c(x) \le l(x) = wl(x) = d(x) = M - 1$.

Property 3 manifests that lobby index and w-lobby index are equal to degree in the global coupling network. Visibly, lobby index and w-lobby index equivalence to node degree in some special circumstances will boost its correlation with the node degree.

Property 4. *If x is in the center of a star network (e.g. Fig. 3), then*

(1) $c(x) = d_h(x)$;
(2) *the necessary and sufficient condition of* $c(x) = d(x)$ *is that the edge weights of x are larger or equal to* $d(x)$*;*
(3) *if* $c(x) = d(x)$*, then* $wl(x) = c(x) = d_h(x) = d(x)$.

The proof is provided in Appendix A.1.

Property 5. *For any node x in the ring network (e.g. Fig. 4), it holds that* $c(x) = d_h(x)$.

The proof is provided in Appendix A.2.

2.2.2. Distribution of communication centrality

In recent years, an important discovery made in the field of complex network research is that the degree distribution has a power-law form of many networks, including the Internet, WWW, metabolic network etc. This kind of network is the called scale-free network [16]. Barabási et al. [17] proposed the scale free network model which has a key characteristic that the degree distribution has a power-law upper tail and the node degrees are independent [17].

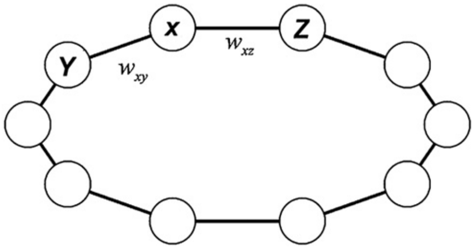

Fig. 4. Example of a ring network.

In this section, we will discuss the distribution of communication centrality, assuming the network is a weighted scale free network, and edge weights are of natural numbers for the simplicity and distribution of edge weights has a power-law upper tail. In what follows, some properties of the communication centrality are investigated in weighted scale free networks.

The H-index measures a scholar's academic achievement based on the set of the scholar's papers. Let x be a randomly chosen scholar, $n(x)$ is the number of papers of this scholar, y is a certain paper of this scholar and $cit(y)$ is the citation times of the paper y.

Theorem 1 (*Distribution of the H-index [18]*)**.** *Assume that the productivity has a α-fat tail: $\mathrm{P}(n(x) \geq n) \approx cn^{-\alpha}$, and the citation score has a β-fat tail $\mathrm{P}(c(y) \geq t) \approx ct^{-\beta}$ and assume that all the publication and citation scores are independent, then $\mathrm{P}(h(x) \geq k) \cong k^{-\alpha(\beta+1)}$.*

Note: in Theorem 1 and the following, c represents an arbitrary positive constant unless specified otherwise. Its value may change from occurrence to occurrence, and $\alpha > 1$, $\beta > 1$. Let $a_n \approx b_n$ mean that $a_n/b_n \to 1$ as $n \to \infty$ and $a_n \cong b_n$ mean that there is a $M > 1$ such that $1/M \leq a_n/b_n \leq M$ for all n.

Because h-degree and h-index have the same calculation principal, so by Theorem 1, it is easy to get the following theorem about h-degree.

Theorem 2 (*Distribution of the H-degree*)**.** *In the weighted undirected network G, the weight of any edge w has distribution $\mathrm{P}(w \geq k) \approx ck^{-\beta}$ and the distribution of degree of any node x is $\mathrm{P}(d(x) \geq k) \approx ck^{-\alpha}$. When the weights of all edges of node x and degree of node x are independent, then $\mathrm{P}(d_h(x) \geq k) \cong k^{-\alpha(\beta+1)}$.*

Next we will explain the distribution of communication centrality.

Theorem 3 (*Distribution of the Communication Centrality*)**.** *In the weighted undirected network G, the weight of any edge w has distribution $\mathrm{P}(w \geq k) \approx ck^{-\beta}$ and the distribution of degree of any node x is $\mathrm{P}(d(x) \geq k) \approx ck^{-\alpha}$. When the weights of all edges and degrees of all nodes are independent, then $\mathrm{P}(c(x) \geq k) \cong k^{-\alpha(\beta+1)}$.*

The proof is provided in Appendix A.3.

Suppose the communication centrality of a node is k, if its one neighbor node satisfies product of neighbor node's h-degree and the weight of the edge linked with the node is no fewer than k, naturally, the neighbor node is the core node that affects its communication ability.

Definition 2 (*C-core*)**.** C-core ($C(x)$) of any node x refers to the set of all the nodes y satisfying $w_{xy}d_h(y) \geq c(x)$, namely,

$$C(x) = \{y | w_{xy}d_h(y) \geq c(x)\}.$$

Obviously, the number of elements in the set $C(x)$ is no less than $c(x)$. It can be noticed that for y in c-core, it is because the edge weight w_{xy} is large or the h-degree $d_h(y)$ of y is also large, such that $w_{xy}d_h(y) \geq c(x)$. When w_{xy} is large and the h-degree $d_h(y)$ of y is small, y might also enter c-core. In this case, if the node degree or node strength is adopted to measure the importance of a node in the network, the importance of y might not be high. However, since its edge weight w_{xy} with node x is larger, then y is important for x. Hence its entry into c-core.

The following theorem shows the truncated distribution of product of h-degree of y in c-core and the edge weight w_{xy}.

Theorem 4. *The condition is same as Theorem 3, if node $y \in C(x)$, then the truncated distribution of the product $w_{xy}d_h(y)$ is $\mathrm{P}(w_{xy}d_h(y) \geq k | y \in C(x), c(x) = s) \approx c(\frac{k}{s})^{-\beta}$ for $k > s \geq 1$.*

Proof. For $k > s \geq 1$, $\mathrm{P}(w_{xy}d_h(y) \geq k | y \in E(x), c(x) = s) = \mathrm{P}(w_{xy}d_h(y) \geq k | w_{xy}d_h(y) \geq s) = \frac{\mathrm{P}(w_{xy}d_h(y) \geq k)}{\mathrm{P}(w_{xy}d_h(y) \geq s)} \approx c(\frac{k}{s})^{-\beta}$. □

Theorem 4 manifests that for node y in c-core, the product of its h-degree and the edge weight of x and y has the property of power-law distribution and compared with original distribution $\mathrm{P}(w_{xy} \cdot d_h(y) \geq k) \approx ck^{-\beta}$, this truncated distribution has a higher mathematics expectation.

2.3. Other communication centrality measures

Similar to degree centrality, node strength and h-degree, based on the communication centrality, relevant measures for nodes, and networks as a whole are proposed in this subsection.

2.3.1. C_c-centrality measures for nodes

To compare the node centralities in different networks, the centrality indexes of nodes in different networks should go through standardized definition to gain the standardized centrality measurement. The standardized form can also be given to communication centrality, called C_c-centrality.

Definition 3. In a weighted network with N nodes, the c_c-centrality, C_c, of node x is defined as: $C_c(x) = c(x)/(N-1)$, here $N-1$ is the maximum value attained by communication centrality $c(x)$.

For example, for the network in Fig. 1, $N = 6$. So, c_c-centrality of nodes A–F are sequentially 1/5, 3/5, 1/5, 2/5, 1/5, 1/5. One property of c_c-centrality is given below.

Property 6. *For non-isolated nodes in weighted network with N nodes, c_c-centrality ($C_c(x)$) of node x always satisfies the following inequality:*

$$0 < \frac{1}{N-1} \le C_c(x) \le \frac{d(x)}{N-1} \le 1.$$

Proof. By Property 1, we have $1/(N-1) \le c(x)/(N-1) \le d(x)/(N-1) \le (N-1)/(N-1)$. So Property 6 is correct. □

2.3.2. C_c-centralization for whole networks

To define c_c-centrality of the whole network consistent with [10], this paper also adopts Freeman's centralization procedure [2]. Specifically, given a node centrality index F, a centralization index F_1 for the whole network G with N nodes is defined as $F_1(G) = \frac{\sum_{i=1}^{N}(\mathrm{Max}(G)-F(x_i))}{\mathrm{Max}(N)}$, where Max($G$) is the maximum value attained by F in the network G and Max(N) is the maximum value attained by the numerator in all possible graphs with N nodes. Based on this procedure we define the c_c-centralization of a weighted network based on the communication centrality.

Definition 4. In a weighted network G with N nodes, the c_c-centralization, $C_c(G)$ of this network is

$$C_c(G) = \frac{\sum_{i=1}^{N}(\mathrm{Max}(G) - c(x_i))}{(N-1)(N-2)}. \tag{1}$$

Here $c(x_i)$ represents the communication centrality of node x_i. The denominator of Eq. (1) is obtained as follows. The largest possible value for Max(G) is $N-1$, which can be reached in a star network. Then, there are $N-1$ nodes with communication centrality equal to 1, and hence a difference with the largest value of $N-2$, leading to a denominator equal to $(N-1)(N-2)$.

C_c-centralization describes the distribution characteristic of communication centrality in a weighted network. Using Fig. 1 as an example to calculation c_c-centralization, it contains six nodes, $N = 6$. Communication centralities of nodes A–F are sequentially 1, 3, 1, 2, 1 and 1 (see Table 1). Hence, for the network Max(G) = 3. The numerator of formula (1) is 9, while the denominator is $(6-1)(5-1) = 20$. Hence $C_c(G) = 0.45$.

3. A case study

3.1. Data

We choose eight top academic journals in the field of information systems (see Table 2) as the data source to construct a co-author network. We retrieved data from the Web of Science databases on September 1, 2012. 3457 articles with 4322 authors recorded by "Article" are downloaded from the nine journals for the period of January 1, 1981–September 1, 2012. The co-author network takes the scholar as the node, constructs the edge according to the collaboration relationship and makes collaboration frequency as the edge weight. That is to say, if two scholars have collaborated in n ($n \ge 1$) papers, then they have linked edges and their edge weight is n. Otherwise they have no linked edge [19].

3.2. Results and analysis

First, we analyzed the basic features of the co-author network and examined the distribution of node degree (d), edge weight and node strength (s). After the log–log transformation of cumulative probability distribution of degree and node strength, we observed that their distributions all follow power-law distributions approximately. So the co-author network can be regarded as a weighted scale free network [17].

Table 2
List of eight academic journals.*

Serial number	Journal
1	European Journal of Information Systems
2	Information Systems Journal
3	Information Systems Research
4	Journal of Information Technology
5	Journal of Management Information Systems
6	Journal of Strategic Information Systems
7	Journal of the Association for Information Systems
8	MIS Quarterly

* http://www.vvenkatesh.com/isranking/.

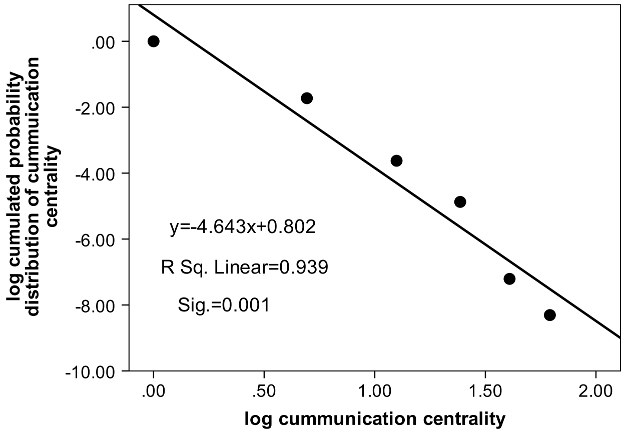

Fig. 5. The distribution of communication centrality.

Table 3
Top 10 scholars in the sequencing of communication centrality.

AU	*c*	*d*	*s*	*cl*	*bw*	*ev*	d_h	*l*	*wl*
Whinston, AB	6	39	63	0.121	0.031	0.004	3	8	10
Benbasat, I	5	48	82	0.134	0.043	0.012	3	9	12
Gupta, A	5	33	51	0.112	0.015	0.002	4	7	7
Agarwal, R	4	33	45	0.135	0.030	0.049	2	9	9
Barki, H	4	15	27	0.115	0.008	0.003	3	5	6
Barua, A	4	13	22	0.111	0.006	0.002	2	7	8
Briggs, RO	4	8	22	0.109	0.001	0.054	3	4	6
Brown, SA	4	19	31	0.116	0.005	0.014	2	9	10
Chan, FKY	4	5	9	0.110	0.000	0.005	2	5	5
Chen, HC	4	14	21	0.103	0.001	0.011	2	6	7

Next, we calculate the communication centrality (c) of all nodes in the co-author network; make a log–log transformation of cumulated probability distribution of the communication centrality. Linear regression of the transformed data indicates that the distribution of communication centrality follows the power-law distribution approximately. The model and coefficients passed the significance test with the results shown in Fig. 5.

Fig. 6 displays the sub-network of co-authors composed of 31 authors whose communication centralities are greater than or equal to 4. In the figure, the wider edge indicates the greater weight of linked nodes. The larger point means the greater communication centrality. It can be seen that most of the nodes boasting high communication ability are not isolated. Five isolated nodes account for 16.13% and the other 26 (83.87%) nodes form 3 connected components. Besides, the edge weight of the sub-network averages out to be 2.62, much larger than the average edge weight 1.14 of the whole collaboration network. It demonstrates that scholars with larger communication centrality have a more frequent collaboration with others.

We calculate degree, node strength, closeness centrality (cl), betweenness centrality (bw), eigenvector centrality (ev) and other h-type indexes of the nodes, specifically lobby index (l), w-lobby index (wl) and h-degree (d_h). Table 3 lists the top 10 scholars of communication centrality and their related indexes. If communication centralities of two scholars are identical, the sequence of scholars is listed by alphabetical order of their names. From Table 3, we can know that communication centrality is different from other indexes and that scholars of larger communication centrality might have other smaller indexes, and vice versa.

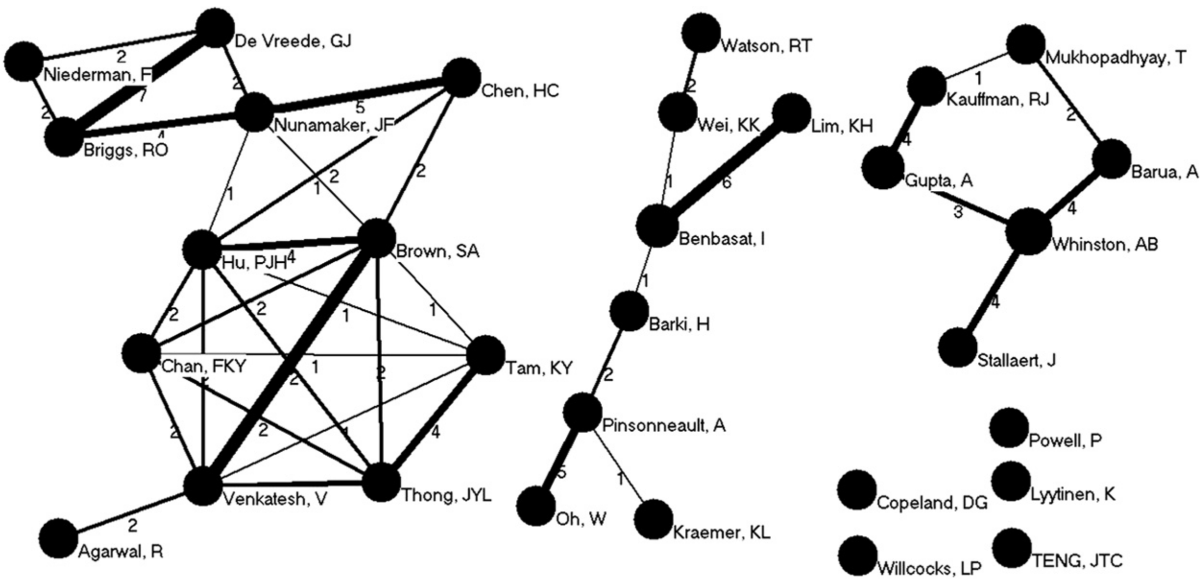

Fig. 6. Sub-network of co-authors (the top 31 scholars of communication centrality).

Table 4
Spearman rank correlation of indexes.

Spearman's rho	*c*	*d*	*s*	*cl*	*bw*	*ev*	d_h	*l*
d	.485**							
s	.543**	**.972****						
cl	.467**	.540**	.550**					
bw	.495**	.656**	.664**	.444**				
ev	.415**	.315**	.333**	.527**	.265**			
d_h	.656**	.344**	.430**	.259**	.405**	.263**		
l	.451**	**.940****	*.905***	.515**	.441**	.312**	.298**	
wl	.468**	**.946****	**.914****	.521**	.461**	.314**	.319**	**.995****

** Correlation is significant at the 0.01 level (2-tailed).

Although these indexes are different, they display the tendency of consistency to a certain extent. To measure such consistency, we made a Spearman rank correlation analysis of all the indexes of 4051 non-isolated scholars, and the results are shown in Table 4. We can see that all the indexes are in a positive relationship.

We analyzed the reason why degree, node strength, lobby index and w-lobby index have a strong correlation. We analyzed the network and found that 89.9% of the edge weights of this co-author network is 1, which is the main reason for the strong correlation between degree centrality and node strength. Because the lobby index values of 76.28% of the 4051 nodes in the network have equal degree and $l(x) \leq wl(x) \leq d(x)$, the correlations between any two of w-lobby index, lobby index and degree are in strong correlation. Another reason for the strong correlation between the w-lobby index and lobby index is that 89.9% of the edge weight is 1, while for node x whose edge weight of the neighbor nodes is 1, it is that $wl(x) = l(x)$. Furthermore, due to strong correlation between degree and node strength, it is easily understood that lobby index and w-lobby index are also in strong correlation with node strength.

The correlation between communication centrality and all the other indexes are moderate. The correlation between communication centrality and h-degree is the largest, at 0.656, and that with eigenvector centricity is the smallest, to be 0.415. It can be deduced that communication centrality perfectly achieves the balance of the characteristics of all indexes while it cannot be replaced by any other index. This case demonstrates that communication centrality surmounts the strong correlation between lobby index and w-lobby index with degree.

However, we should also discuss the dynamic range and ranking procedure, comparing to the other indexes in the co-author network. Fig. 7 is the frequency histograms of each centrality measure of 4322 scholars in the co-author network. The maximum, mean, coefficient of variation (CV, CV = (Standard deviation)/Mean) of distributions of each centrality measure are given in Fig. 7. Because the minimum value is 0 for all indexes, the maximum values mean the dynamic range of indexes. The CVs indicate the degree of variation of all indexes. For one index, the smaller CV and dynamic range are disadvantageous for ranking the nodes.

From Fig. 7 we can see that the dynamic ranges of lobby index, w-lobby index, degree centrality and node strength are growing, and it is also this case for the mean and the CV. The dynamic ranges of h-degree, communication centrality, degree centrality and node strength are also growing, and for the mean and the CV it is also this case. This is due to

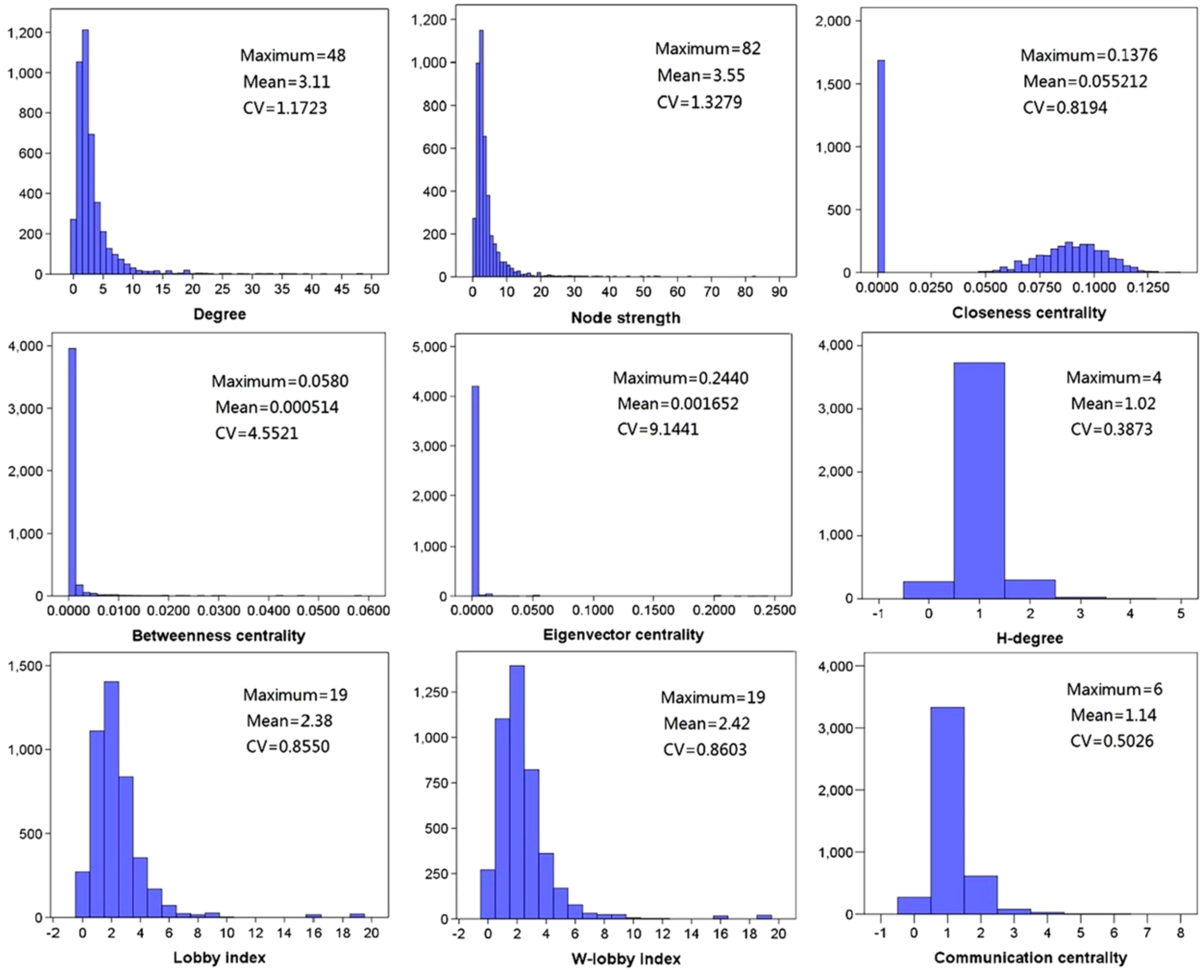

Fig. 7. Frequency histograms of centrality measures.

$l(x) \leq wl(x) \leq d(x) \leq s(x)$ and $d_h(x) \leq c(x) \leq d(x) \leq s(x)$ for any node x in weighted networks with natural number as the edge weight.

Also from Fig. 7 we can see the maximum, mean and CV of communication centrality are less than the ones of lobby index respectively, but the propositions that "$c(x) < l(x)$" and "dynamic range of $c(x)$ is smaller than the range of $l(x)$" are not always right. This is very easy to prove according to their definitions. In the co-author network, the phenomenon that the dynamic range of communication centrality is smaller than the dynamic range of lobby index is caused mainly by the fact that 89.9% of the edge weights mentioned in this paper are 1, the average of edge weights is 1.14, which is small, close to the minimum value 1. Thus by definitions of communication centrality and lobby index, we can know that in the above case, the dynamic range of communication centrality would tend to be smaller than the dynamic range of lobby index, and from the definition we can see if the edge weights are larger, then the range of communication centrality will also be larger.

4. Conclusions

Built upon well-known scientific measurement index h-index and h-degree in network, this paper proposes the communication centrality to measure the communication ability of a node in a weighted undirected network. Communication centrality overcomes the strong correlation of lobby index and w-lobby index with node degree, adds the information of edge weight of the node that influences communication ability in the calculation and effectuates more accurate measurement. Nevertheless, the lobby index does not make use of the information of node edge weight. Our comparison of communication centrality with other well-known centrality indexes reveals that in the use of information it is apparently superior to the degree centrality (only calculation of the number of neighbor nodes), node strength (only calculation of the sum of edge weights), h-degree (without utilizing the communication ability of the neighbor nodes), lobby index (without utilizing edge weights) and w-lobby index (without utilizing the edge weights of this node and calculating the importance of neighbor nodes only through node strength).

A case study of a co-author network shows that the communication centrality proposed in this paper is different from other common centrality measures (degree centrality, betweenness centrality, closeness centrality, eigenvector centrality and node strength) and other h-type indexes (lobby index, w-lobby index and h-degree). In addition, communication centrality displays moderate correlation with these indexes, which illustrates that communication centrality balances all indexes and cannot be replaced by any other index.

Acknowledgments

This work is partly supported by the National Natural Science Foundation of PRC (No. 71171067, 71328103, 71171068) and Postdoctoral Science-research Developmental Foundation of Heilongjiang province (No. LBH-Q11114).

Appendix A

A.1

(1) For the star network in Fig. 3, no matter how the edge weights of the central node x are, the h-degrees of the neighbor nodes are equal to 1. So the product of edge weights and the h-degree of corresponding neighbor nodes are equal to the value of the edge weight. By the definitions of h-degree and communication centrality, they are equivalent to each other at this time, hence $c(x) = d_h(x)$.
(2) Since $c(x) = d_h(x)$, if $c(x) = d(x)$, then x has $d(x)$ edges (namely all the edges) whose weights are no smaller than $d(x)$; conversely, if each edge weight of x is no smaller than $d(x)$, then for x, $d_h(x) = d(x)$. Hence $c(x) = d_h(x)$, $c(x) = d(x)$. Consequently (2) is proved.
(3) If $c(x) = d(x)$, then each edge weight of x is no smaller than $d(x)$, then the node strengths of $d(x)$ neighbor nodes of x are no smaller than $d(x)$, hence $wl(x) = d(x)$. Summing up the above, $c(x) = d_h(x) = wl(x) = d(x)$.

A.2

Any node x in the ring network in Fig. 4 has two neighbor nodes. Let us suppose y and z are its neighbor nodes. We will respectively discuss the different circumstances of edge weights w_{xy} and w_{xz} of x with y and z.

(1) If $w_{xy} \geq 2$ and $w_{xz} \geq 2$, then $d_h(x) = 2$. As $d_h(y) \geq 1$ and $d_h(z) \geq 1$, hence $w_{xy}d_h(y) \geq 2$, $w_{xz}d_h(z) \geq 2$, then $e(x) = 2$, consequently $c(x) = d_h(x)$.
(2) If $w_{xy} = w_{xz} = 1$, then $d_h(x) = d_h(y) = d_h(z) = 1$, and then $w_{xy}d_h(y) = w_{xz}d_h(z) = 1$, so $c(x) = 1$ and $c(x) = d_h(x)$.
(3) If one of w_{xy} and w_{xz} is equal to 1 and the other is greater than or equal to 2, assuming $w_{xy} = 1$, $w_{xz} \geq 2$ without generality, then $d_h(y) = 1$ and $d_h(z) \geq 1$, so $w_{xy}d_h(y) = 1$ and $w_{xz}d_z(y) \geq 2$. Consequently $c(x) = 1 = d_h(x)$.

The above discussion shows that for any node x in the ring network then $c(x) = d_h(x)$ is always correct.

A.3

(1) Calculate the distribution of the product of the weight of any edge of node x and the h-degree of the neighboring y linked with it first. Suppose w_{xy} is the weight of the edge of node x and y, and $d_h(y)$ is the h-degree of node y. For any $k \geq 1$, have

$$\begin{aligned} P(w_{xy} \cdot d_h(y) \geq k) &= \sum_{s=1}^{+\infty} P(w_{xy} \cdot d_h(y) \geq k | d_h(y) = s) P(d_h(y) = s) \\ &= \sum_{s=1}^{k-1} P\left(w_{xy} \geq \frac{k}{s} | d_h(y) = s\right) P(d_h(y) = s) + \sum_{s=k}^{+\infty} P(d_h(y) = s) \\ &= A + B \end{aligned}$$

here $B = P(d_h(y) \geq k) \approx ck^{-\alpha(\beta+1)}$ (according to the Theorem 2),

$$\begin{aligned} A &= \sum_{s=1}^{k-1} P\left(w_{xy} \geq \frac{k}{s}\right) P(d_h(y) = s) = \sum_{s=1}^{k-1} c \left]\frac{k}{s}\right[^{-\beta} s^{-\alpha(\beta+1)-1} \\ &\approx \sum_{s=1}^{[\sqrt{k}]} c \left]\frac{k}{s}\right[^{-\beta} s^{-\alpha(\beta+1)-1} + \sum_{s=[\sqrt{k}]+1}^{k-1} c \left]\frac{k}{s}\right[^{-\beta} s^{-\alpha(\beta+1)-1} = A_1 + A_2. \end{aligned}$$

Here the operator $[r]$ is the largest integer not more than r, the operator $]r[$ is the smallest integer not less than r for any real number r.

Here $A_1 = \sum_{s=1}^{[\sqrt{k}]} c]\frac{k}{s}[^{-\beta} s^{-\alpha(\beta+1)-1} = ck^{-\beta}\sum_{s=1}^{[\sqrt{k}]} s^{-\alpha(\beta+1)-1+\beta} + ck^{-\beta}\sum_{s=1}^{[\sqrt{k}]} r_{k,s}s^{-\alpha(\beta+1)-1+\beta}$, where $r_{k,s} = \frac{]\frac{k}{s}[^{-\beta}}{(\frac{k}{s})^{-\beta}} - 1$. For all $0 \le s \le [\sqrt{k}]$, $|r_{k,s}| \le 1 - \frac{(\frac{k}{s}+1)^{-\beta}}{(\frac{k}{s})^{-\beta}} \le ck^{-\frac{1}{2}}$. So $A_1 \approx ck^{-\beta}$.

$$
\begin{aligned}
|A_2| &= \left| \sum_{s=[\sqrt{k}]+1}^{k-1} c \left]\frac{k}{s}\right[^{-\beta} s^{-\alpha(\beta+1)-1} \right| \le \sum_{s=[\sqrt{k}]+1}^{k-1} c \left(\frac{k}{s}\right)^{-\beta} s^{-\alpha(\beta+1)-1} \\
&\le ck^{-\beta}(\sqrt{k})^{-\beta(\alpha-1)-1} \sum_{s=[\sqrt{k}]+1}^{k-1} s^{-\alpha} \\
&= ck^{-\beta} \cdot k^{-\frac{1}{2}\beta(\alpha-1)-\frac{1}{2}} \cdot o(1).
\end{aligned}
$$

Here $o(1)$ means that a higher order infinitesimal (when $k \to +\infty$). So, $A = A_1 + A_2 \approx ck^{-\beta}$, and then

$$\mathrm{P}(w_{xy} \cdot d_h(y) \ge k) = A + B \approx ck^{-\beta} + ck^{-\alpha(\beta+1)} \approx ck^{-\beta}.$$

(2) Similar with the proof of Theorem 2, as the weights of all edges and degrees of all nodes are independent in network G, node degree $d(x)$ of x and $w_{xy}d_h(y)$ of any neighbor node y are independent. The presupposition shows $\mathrm{P}(d(x) \ge k) \approx ck^{-\alpha}$, and (1) indicates $\mathrm{P}(w_{xy} \cdot d_h(y) \ge k) \approx ck^{-\beta}$, so Theorem 1 shows that $\mathrm{P}(c(x) \ge k) \cong k^{-\alpha(\beta+1)}$.

References

[1] A. Korn, A. Schubert, A. Telcs, Lobby index in networks, Physica A 388 (2009) 2221–2226.
[2] L.C. Freeman, Centrality in social networks: conceptual clarification, Social Networks 1 (1979) 215–239.
[3] J. Nieminen, On the centrality in a graph, Scandinavian Journal of Psychology 15 (1974) 332–336.
[4] M.E. Shaw, Group structure and the behavior of individuals in small groups, Journal of Psychology: Interdisciplinary and Applied 38 (1954) 139–149.
[5] A. Barrat, M. Barthélemy, R. Pastor-Satorras, A. Vespignani, The architecture of complex weighted networks, Proceedings of the National Academy of Sciences of the United States of America 101 (2004) 3747–3752.
[6] A. Bavelas, Communication patterns in task oriented groups, Journal of the Acoustical Society of America 22 (1950) 725–730.
[7] G. Sabidussi, The centrality index of a graph, Psychomatrika 31 (1966) 581–603.
[8] P. Bonacich, Factoring and weighting approaches to status scores and clique identification, Journal of Mathematical Sociology 2 (1972) 113–120.
[9] J.E. Hirsch, An index to quantify an individual's scientific research output, Proceedings of the National Academy of Sciences of America 102 (2005) 16569–16572.
[10] S.X. Zhao, R. Rousseau, F.Y. Ye, *H*-degree as a basic measure in weighted networks, Journal of Informetrics 5 (2011) 668–677.
[11] S. Alonso, F.J. Cabrerizo, E. Herrera-Viedma, F. Herrera, *H*-index: a review focused in its variants, computation and standardization for different scientific fields, Journal of Informetrics 3 (2009) 273–289.
[12] S.X. Zhao, F.Y. Ye, Exploring the directed *h*-degree in directed weighted networks, Journal of Informetrics 6 (2012) 619–630.
[13] A. Schubert, A Hirsch-type index of co-author partnership ability, Scientometrics 91 (2012) 303–308.
[14] R. Rousseau, Comments on "A Hirsch-type index of co-author partnership ability", Scientometrics 91 (2012) 309–310.
[15] M.G. Campiteli, A.J. Holanda, P.R.C. Soles, L.H.D. Soares, O. Kinouchi, Hirsch index as a network centrality measure, http://arxiv.org/abs/1005.4803, (2010).
[16] X.F. Wang, X. Li, G.R. Chen, The Complex Network Theory and its Application, Tsinghua University Press, 2006.
[17] A.L. Barabási, R. Albert, Emergence of scaling in random networks, Science 5439 (1999) 509–512.
[18] K. Barcza, A. Telcs, Paretian publication patterns imply Paretian Hirsch index, Scientometrics 81 (2009) 513–519.
[19] M.E.J. Newman, The structure of scientific collaboration networks, Proceedings of the National Academy of Sciences 98 (2001) 404–409.