Journal of Informetrics 5 (2011) 668–677

Contents lists available at ScienceDirect

Journal of Informetrics

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

h-Degree as a basic measure in weighted networks

Star X. Zhao [a], Ronald Rousseau [b,c,d], Fred Y. Ye [a,*]

[a] Department of Information Resource Management, Zhejiang University, 38 Zheda Road, Hangzhou, Zhejiang 310027, China
[b] KHBO (Association K.U. Leuven), Industrial Sciences and Technology, Zeedijk 101, B-8400 Oostende, Belgium
[c] K.U. Leuven, Department of Mathematics, Celestijnenlaan 200B, B-3000 Leuven (Heverlee), Belgium
[d] Universiteit Antwerpen, IBW, Venusstraat 35, B-2000 Antwerpen, Belgium

ARTICLE INFO

Article history:
Received 25 May 2011
Accepted 22 June 2011

Keywords:
h-Degree
Node strength
Network analysis
Weighted networks
Informetric networks

ABSTRACT

We introduce the h-degree of a node as a basic indicator for weighted networks. The h-degree (d_h) of a node is the number d_h if this node has at least d_h links with other nodes and the strength of each of these links is greater than or equal to d_h. Based on the notion of h-degree other notions are developed such as h-centrality and h-centralization, leading to a new set of indicators characterizing nodes in a network.

1. Introduction

Network science (Börner, Sanyal, & Vespignani, 2007; Fang, 2010; Watts, 2004) is currently gaining more and more impact in informetrics. Many kinds of typical informetric networks, such as citation and co-citation networks (Chen & Redner, 2010; Ding, Yan, Frazho, & Caverlee, 2009; Egghe & Rousseau, 2002), co-author networks (Liu, Bollen, Nelson, & Van de Sompel, 2005; Rodriguez & Pepe, 2008), keywords networks (Su & Lee, 2010), and patent networks (Liu & Shih, 2011) are being studied. Providing quantitative methods to express relational data and to resolve the structure of relations, network analysis is a useful tool for investigating networks, including those with informetric content. These methods can be used to study different forms of heterogeneous relations.

After the introduction of the h-index for scientists in 2005 (Hirsch, 2005) its theoretical aspects have been thoroughly discussed (Egghe & Rousseau, 2006; Glänzel, 2006; Schubert & Glänzel, 2007; Ye, 2011), while its meaning has been extended to various other source-item relations. One of the important reasons for the success of the h-index is that it measures the key part of a dataset in a relatively natural way. This feature is consistent with the 80/20 rule and the power-law distribution: common phenomena studied in information science (Egghe, 2005) as well as in network studies (Barabási & Albert, 1999; Clauset, Shalizi, & Newman, 2009).

Hence, combining network analysis and the h-index became an interesting idea. In 2009, Schubert, Korn and Telcs (Korn, Schubert, & Telcs, 2009; Schubert, Korn, & Telcs, 2009) introduced the lobby index as a centrality parameter for nodes and the h-index of a network as an indicator for complete networks (Schubert & Soos, 2010). The lobby index of node n is defined as the largest integer k such that n has at least k neighbors with a degree, $d(n)$, of at least k. The h-index of a network is defined as, "A network's h-index is h if not more than h of its nodes have a degree not less than h" (Schubert, 2010). These

* Corresponding author. Tel.: +86 571 87996929; fax: +86 571 87953814.
E-mail addresses: ronald.rousseau@khbo.be (R. Rousseau), yye@zju.edu.cn (F.Y. Ye).

doi:10.1016/j.joi.2011.06.005

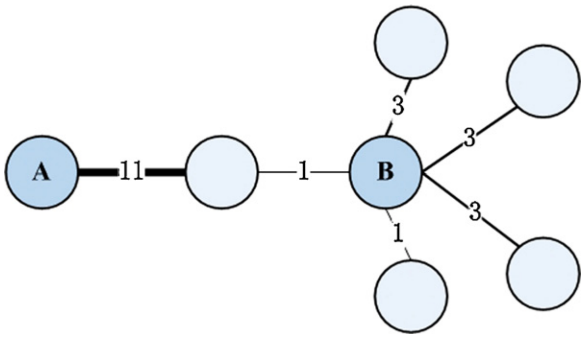

Fig. 1. Same node strength (nodes A and B) but different structural properties.

indicators can be further developed leading to enriched network analysis methods in numerous applications (Rousseau & Ye, 2011).

Moreover network analysis can be said to be a scientific method based on links and nodes as fundamental units. These units lead to the notion of the degree of a node which is an essential parameter in network analysis. Complex relations in networks are reflected by different numbers of links and weights, as most informetric networks are indeed weighted networks. For example, journals in journal citation networks are characterized by the number of citations they give or receive, author pairs in co-author networks are characterized by the number of times they collaborate, and the number of co-occurrences are essential in the study of co-word networks. Based on the idea of the h-index, we claim that there may be another basic degree which is suitable for the analysis of weighted networks. This idea is developed in this contribution. We use the terms graph and network as synonyms.

2. Methodology

2.1. The h-degree

In unweighted, undirected networks two nodes are either linked or not. This may be represented by the value 1, indicating that a link exists, or 0, indicating that there is no link. However, in many real networks, the strength of links is an important parameter, leading to the notion of weighted (or valued) networks.

The degree of a node, denoted as $d(n)$, is one of the most basic characteristics in network studies. Many other parameters such as those based on shortest paths and fundamental properties of networks such as scale-free phenomena (Barabási & Albert, 1999; Barabási, Albert, & Jeong, 2000) are based on this notion.

In this article we define the node strength of a node in a weighted network as the sum of the strengths (or weights) of all its links (Barrat, Barthélemy, Pastor-Satorras, & Vespignani, 2004). Often the term node degree is used in the case of unweighted networks as well as in weighted ones. In order to make a clear distinction we use the term node degree (degree of a node) only for unweighted networks. In weighted networks we use the term node strength, denoted as $d_s(n)$. Weights in a network can, moreover, be of similarity type (as in co-authorship strength) or of dissimilarity type (as distances in a road map). Moreover, in a network of cities where links are existing roads and the corresponding weight is the distance between two cities measured along roads, this definition certainly does not seem to be meaningful. In such a graph the weight is related to the reachability of a city (the larger the weights, the smaller the reachability), but, for example, when there are two different roads between cities then this increases reachability so that it makes no sense to add distances (weights). Hence, in our opinion the standard definition of node strength is not always appropriate. Consequently, our analysis is restricted to cases where this definition can be used. For simplicity we assume in this article that weights are natural numbers.

The lobby index mentioned in the introduction can be adapted to a weighted network as follows: the w-lobby index (weighted network lobby index) of node n, denoted as $l(n)$ is defined as the largest integer k such that node n has at least k neighbors with node strength at least k.

Fig. 1 shows that node strength as defined above does not distinguish between obviously different situations. Node A in Fig. 1 only has one link with one other node; node B has 5 links with 5 other nodes. The node strengths of A and B are the same and equal to 11, although their position and role in the network are significantly different.

This example illustrates that more suitable basic parameters are needed to describe structural properties in a weighted network. Definition 1 introduces a new type of degree, which takes a step in this direction. It is called the h-degree, denoted as d_h, and is defined as follows.

Definition 1. The h-degree (d_h) of node n in a weighted network is equal to $d_h(n)$ if $d_h(n)$ is the largest natural number such that n has at least $d_h(n)$ links each with strength at least equal to $d_h(n)$.

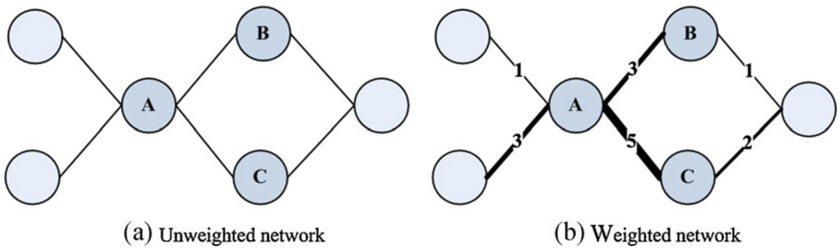

Fig. 2. Illustration of the h-degree in a weighted network.

Table 1
Degree centralities of Fig. 2.

	Underlying unweighted network node degree	Node strength	h-Degree	w-Lobby index
A	4	12	3	3
B	2	4	1	2
C	2	7	2	3

2.1.1. Comments

In directed networks one can define an in-h-degree and an out-h-degree. For simplicity these are not used in this article.The h-degree of node A in Fig. 1 is 1, while the h-degree of node B is 3, illustrating the fact that h-degree can – sometimes – make a distinction in circumstances where node strength cannot.From Definition 1 it follows that a node with h-degree d_h may have more than d_h links, but then these other ones have a lower strength. This is illustrated in Fig. 1 where node B has 5 links, but only three have a weight equal to or larger than 3, so that node B's h-degree is 3.In a weighted network with N nodes the highest possible h-degree is $N-1$. This happens in a star network where the centre is linked to the other nodes with a weight equal to (at least) $N-1$.In a ring where each weight is at least 2, each node has h-degree 2.In unweighted networks the h-degree of a node is either 1 (a non-isolated node) or 0 (an isolated node). Also in weighted networks isolated nodes have $d_h = 0$.Another solution for the problem illustrated in Fig. 1 was proposed by Opsahl, Agneessens, and Skvoretz (2010). They make use of a tuning parameter to find a balance between the number of links and the weights attached to them in a weighted network.

Fig. 2 gives another illustration of the difference between node degree (in the underlying unweighted network), node strength and the h-degree in a weighted network. Degree centralities are given in Table 1.

According to Definition 1 the h-degree can be considered as a compromise between the strength of links with adjacent nodes and the amount of adjacent nodes. In a sense, the difference between the h-degree and the node strength degree in weighted networks is similar to the difference between the h-index and the total number of citations in citation analysis. Compared with the node strength, another advantage of the h-degree is that it includes more information about local structural characteristics. As shown in Fig. 3, if we know that node A has node strength 3, we still cannot determine the links' structure: is it like A1, A2 or A3? However, if we know a node B with h-degree = 3, this fact contains more information. First, the main part of links' structure for node B is fixed. Second, we know that this node has 3 main links and the strength or weight of each main link is not less than 3. Of course, nodes with same h-degree may still have very different structural properties. For instance, one node may have no other links besides the d_h ones determining its h-degree, while another node with the same d_h-value may have hundreds of other links with weight d_h. As the h-degree is a summary indicator it is to be expected that information is lost.

The following properties establish a relation between the h-degree and the w-lobby index.

Proposition 1. *A node's h-degree is always smaller than or equal than its w-lobby index:* $d_h(n) \le l(n)$.

Proof. If node n's h-degree is k then node n has at least k neighbors (as it has k links). These neighbors' node strength is at least k (the link with n alone yields at least a value equal to k). Hence $l(n)$ is at least k. □

Proposition 2. *If node n has exactly* $d_h(n)$ *links then* $d_h(n) = l(n)$.

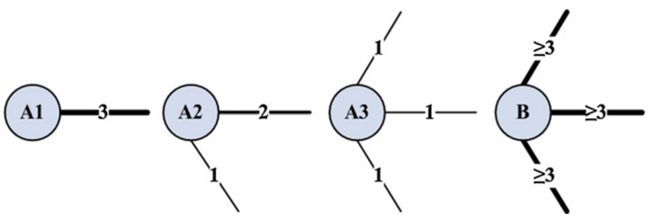

Fig. 3. Examples of h-degree.

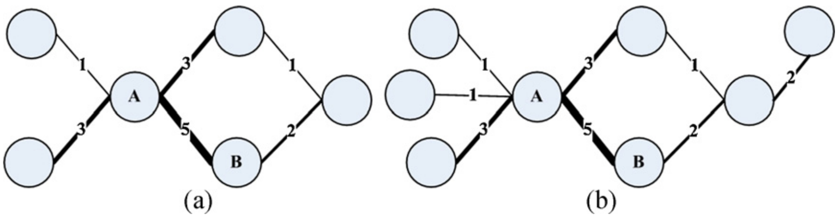

Fig. 4. h-Degree, w-lobby index and h-centrality.

Proof. If n has exactly $d_h(n)$ links then its w-lobby index can at most be equal to $d_h(n)$. However, by Proposition 1 $d_h(n) \le l(n)$. This proves the equality. □

Let N denotes the total number of nodes in a network and let N_a be the number of adjacent nodes of a given node. Clearly, $N_a \le N - 1 < N$. Then, the following basic inequalities hold.

Proposition 3. *For non-isolated nodes in a weighted network, the following inequality involving the h-degree* (d_h) *always holds*:

$$1 \le d_h \le N_a < N \tag{1}$$

Proof. For non-isolated nodes, there is at least one adjacent node linked to it, and the strength of this link is larger than or equal to 1, hence $1 \le d_h$. N_a is the total number of links of this node, leading to: $d_h \le N_a < N$. □

Proposition 4. *If node n has node strength* $d_s(n)$ *and h-degree* $d_h(n)$ *then*

$$d_s(n) \ge (d_h(n))^2 \tag{2}$$

Proof. When a node has h-degree d_h, then the minimum number of links is d_h, and each of these links has at least a strength equal to d_h. Hence $d_s(n) \ge (d_h(n))^2$. □

Similar to node degree in unweighted networks and node strength in weighted networks, the h-degree is a basic parameter for network analysis. Based on the h-degree, relevant measures for nodes, and networks as a whole are proposed in the next sections.

2.2. Other h-centrality measures for nodes

The use of centrality measures, originating from social network analysis (Butts, 2008; Scott, 2000) has led to valuable methods in all types of networks (Bollen, Van de Sompel, Hagberg, & Chute, 2009; Borgatti, Mehra, Brass, & Labianca, 2009; Otte & Rousseau, 2002). Although these methods are mainly applied to identity and characterize key nodes in a network, various centrality measures focus on different roles played by nodes in a network. The best known centrality measures are degree centrality, based on node degree, closeness centrality, based on distance to other nodes, and betweenness centrality, based on the ability of nodes to control flows in a network (Everett, Sinclair, & Dankelmann, 2004; Freeman, 1979). These measures were originally designed for unweighted networks. For example, when calculating the closeness centrality of nodes in a weighted network, the strength of links is often ignored and the weighted network is actually converted to an unweighted network.

Degree centrality of node n in an unweighted network, denoted as $C_d(n)$, is defined as (Freeman, 1979):

$$C_d(n) = \frac{d(n)}{N-1} \tag{3}$$

Adapting this definition to the context of the h-degree we propose the following definition.

Definition 2. In a weighted network with N nodes, the h-centrality, C_h, of node n is defined as:

$$C_h(n) = \frac{d_h(n)}{N-1} \tag{4}$$

where $d_h(n)$ is the h-degree of node n. h-Centrality is just a normalized form of the h-degree. The difference between the h-degree, the w-lobby index and h-centrality is illustrated in Fig. 4; values are given in Table 2.

Table 2
h-Degree, w-lobby index and h-centrality of Fig. 4.

Node	h-Degree	w-Lobby index	h-Centrality
Fig. 4a: A	3	3	3/5
Fig. 4a: B	2	2	2/5
Fig. 4b: A	3	3	3/7
Fig. 4b: B	2	2	2/7

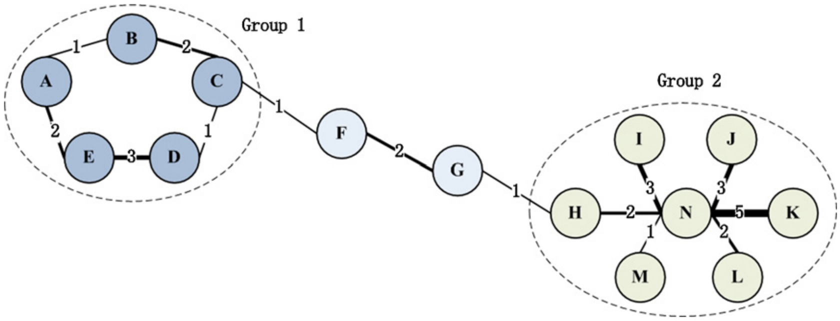

Fig. 5. An example of h-centralization: a network G and two subnetworks (G_1 and G_2).

Nodes A in Fig. 4a and b have the same h-degree and w-lobby index, but different h-centrality; the same observation holds for node B.

h-Centrality measures the importance of a node in a weighted network. Moreover, because it takes the size of the network into account, this measure makes nodes from different networks comparable. When considering a node, one difference between the w-lobby index and h-centrality is that the w-lobby index of a node is based on its neighboring nodes' degree while h-centrality is based on the number of links and their strength. Some basic properties of h-centrality are given below.

Proposition 5. *For non-isolated nodes in weighted networks, h-centrality (C_h) always satisfies the following inequality*

$$0 < \frac{1}{N-1} \le C_h \le \frac{N_a}{N-1} \le 1 \tag{5}$$

Proof. By Proposition 3, we have $1/(N-1) \le d_{\mathrm{h}}/(N-1) \le N_{\mathrm{a}}/(N-1) \le (N-1)/(N-1)$. As $d_{\mathrm{h}}/(N-1) = C_{\mathrm{h}}$ inequality (5) follows immediately. □

Proposition 6. *If a node has node strength d and h-degree d_h then the following inequality holds*:

$$C_h \le \frac{\sqrt{d}}{N-1} \tag{6}$$

Proof. By Proposition 4 we know that $d_{\mathrm{h}} \le \sqrt{d}$. Dividing by $N-1$ yields inequality (6). □

2.3. h-Centralization for whole networks

Freeman (1977) observes that there are two meanings for the notion of centrality of a network: it could mean the extent to which all nodes are central, or it could refer to the dominance of a single point. Applying Freeman's centralization procedure (Everett et al., 2004; Freeman, 1979) we take the second approach. Given a vertex centrality index F a centralization index F_1 for the whole graph G with N nodes is defined as:

$$F_1(G) = \frac{\sum_{i=1}^{N} [\mathrm{MAX}(G) - F(n_i)]}{\mathrm{MAX}(N)} \tag{7}$$

where MAX(G) is the maximum value attained by F in the graph G and MAX(N) is the maximum value attained by the numerator in all possible graphs with N nodes.

Based on this principle we define the h-centralization of a network based on the h-degree.

Definition 3. In a weighted network G with N nodes, the h-degree centralization, $C_{\mathrm{h}}(G)$ of this network is

$$C_{\mathrm{h}}(G) = \frac{\sum_{i=1}^{N} [\mathrm{MAX}(G) - d_{\mathrm{h}}(n_i)]}{(N-1)(N-2)} \tag{8}$$

The denominator of Eq. (8) 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 $d_{\mathrm{h}} = 1$, and hence a difference with the largest value of $N-2$, leading to a denominator equal to $(N-1)(N-2)$.

h-Centralization describes the distribution of weights in a network. A network has a high h-centralization if, compared to the node with largest h-centrality, the h-centrality of the other nodes is low. In those circumstances the links or weights in this high h-centralization network are concentrated in the central node. The distribution of weights in this network is unbalanced.

Definition 3 can also be applied to a subnetwork. An example is given based on Fig. 5 where we consider two subnetworks, which we compare among themselves and with the network as a whole.

Group 1 (G_1) contains five nodes ($N = 5$); four have h-degree equal to 1 and one, namely node E, has h-degree equal to 2. Hence for this subnetwork MAX(G_1) = 2. The numerator of formula (8) is 4, while the denominator is 12. Hence $C_{\mathrm{d}}(G_1) = 0.33$.

Table 3
Top 10 cited papers citing Hirsch's original h-index article.

No.	Author/Paper/Journal	Number of citations	Number of refs.
1	Egghe, L. (2006, April). Theory and practise of the g-index. *Scientometrics, 69*(1), 131–152	169	13
2	Van Raan, A. F. J. (2006, June). Comparison of the Hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups. *Scientometrics, 67*(3), 491–502	132	14
3	Hirsch, J. E. (2007, December 4). Does the h index have predictive power? *PNAS, 104*(49), 19193–19198	104	14
4	Meho, L. I., & Yang, K. (2007, November). Impact of data sources on citation counts and rankings of LIS faculty: Web of science versus Scopus and Google scholar. *JASIST, 58*(13), 2105–2125	97	61
5	Jin, B. H., Liang, L. M., Rousseau, R., & Egghe, L. (2007, March). The R-and AR-indices: Complementing the h-index. *Chinese Science Bulletin, 52*(6), 855–863	95	31
6	Batista, P. D., Campiteli, M. G., Kinouchi, O., & Martinez, A. S. (2006, July). Is it possible to compare researchers with different scientific interests? *Scientometrics, 68* (1), 179–189	89	13
7	Glanzel, W. (2006, May). On the h-index—A mathematical approach to a new measure of publication activity and citation impact. *Scientometrics, 67*(2), 315–321	84	13
8	Braun, T., Glanzel, W., & Schubert, A. (2006, April). A Hirsch-type index for journals. *Scientometrics, 69*(1), 169–173	83	7
9	Egghe, L., & Rousseau, R. (2006, April). An informetric model for the Hirsch-index. *Scientometrics, 69*(1), 121–129	82	11
10	Bornmann, L., & Daniel, H. D. (2007, July). What do we know about the h index? *JASIST, 58*(9), 1381–1385	80	43

Group 2 (G_2) contains seven nodes ($N=7$); it is a star subnetwork, with six nodes having $d_\mathrm{h}=1$ and one node having $d_\mathrm{h}=3$ (which is not the maximum possible for such a star network). Consequently, the numerator of formula (8) is here equal to 12, while the denominator is 30. $C_\mathrm{d}(G_2)=0.4$.

Finally, for the whole network we have $N=14$; $\mathrm{MAX}(G)=3$. The numerator is $12\times 2+1\times 1=25$; the denominator is $13\times 12=156$. Hence $C_\mathrm{d}(G)=0.16$. We note that the h-index of the whole network is 2.

2.4. Indicators for measuring the concentration of weights

Standard network indicators were initially designed for unweighted networks. In this contribution we introduced the h-degree, h-centrality and h-centralization for weighted networks. When taking the ratio of these two types of measures, we obtain new indicators related to the distribution of weights in a network, and this on the level of nodes as well as on the level of whole components. In this context the following indicator is proposed.

Definition 4. In a weighted network, the h-ratio index $R_\mathrm{h}(n)$ is defined as

$$R_\mathrm{h}(n)=\frac{d_\mathrm{h}(n)}{d(n)} \tag{9}$$

The h-ratio index R_h satisfies the following property.

Proposition 7. *The following inequalities always hold*

$$\frac{1}{d}\le R_h\le\frac{N_a}{d}<\frac{N}{d} \tag{10}$$

Proof. This follows immediately from Proposition 3 and the definition of R_h. □

3. A case study

3.1. Data collection and processing

Data were retrieved from the Web of Science (WoS) databases (SCI, SSCI and A & HCI) on March 23, 2011. A co-citation network derived from highly cited papers in h-index research was constructed.

3.1.1. Method

(1) We obtained the 49 most-cited articles citing Hirsch's original paper "An index to quantify an individual's scientific research output". These papers we each cited at least 20 times. We added Hirsch' original paper to this set, leading to 50 papers. The top most-cited ones, excluding Hirsch' original article, are shown in Table 3.

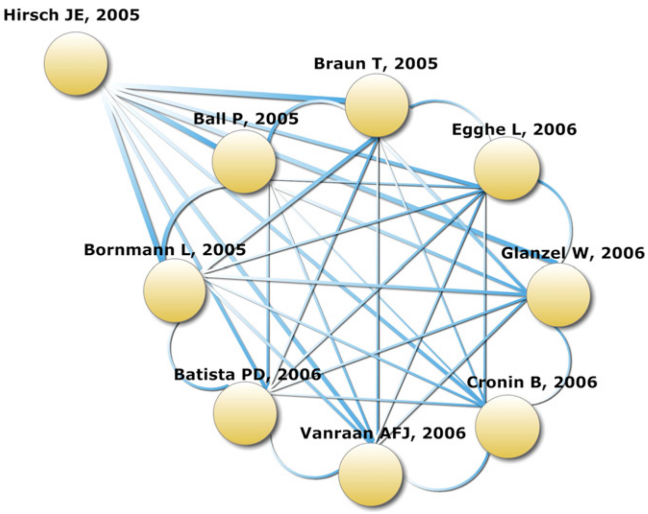

Fig. 6. Sub-network constituted by high h-degree nodes (h-degree ≥7).

(2) We downloaded the details of these 50 papers, including all these papers' references, saved them as .txt, and then imported them into a software program Network Workbench (NWB) (Börner et al., 2010). In this way we obtained 1532 different papers (some of which are not published in WoS source journals). A co-citation network was derived based on these 1532 nodes.
(3) We wrote a computer program to transform this network into the format required by UCINET (Borgatti, Everett, & Freeman, 2002).

3.2. Results and analysis

The subnetwork of the co-citation network containing those nodes with h-degree larger than or equal to 7 is shown in Fig. 6. The main part of the co-citation network after the removal of the links whose strength is 1 is shown in Fig. 7.

All nodes in the sub-network shown in Fig. 6 have not less than 7 strong links (number of co-citations ≥7) with other nodes in the whole network. Based on the h-degree, all these nodes are important nodes in the whole network. They, moreover,

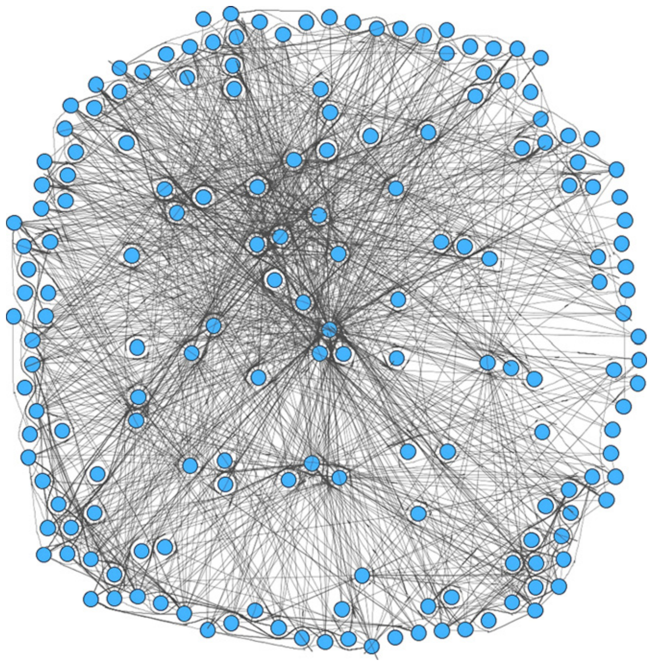

Fig. 7. The co-citation sub-network (after removing all links with strength one).

Table 4
Some parameters for nodes of h-degree ≥5 in the case study.

Nodes (author/year/journal)	Node strength	h-Degree	Degree centrality	h-Centrality ($\times 10^3$)	R_h ($\times 10^3$)
Hirsch, J. E. (2005). *PNAS*	1969	10	1.286	6.532	5.079
Egghe, L. (2006). *Scientometrics*	1047	8	0.684	5.225	7.641
Braun, T. (2005). *Scientist*	1025	8	0.669	5.225	7.805
Glänzel, W. (2006). *Scientometrics*	1063	8	0.694	5.225	7.526
Bornmann, L. (2005). *Scientometrics*	1092	8	0.713	5.225	7.326
van Raan, A. F. J. (2006). *Scientometrics*	992	8	0.648	5.225	8.065
Batista, P. D. (2006). *Scientometrics*	1020	8	0.666	5.225	7.843
Ball, P. (2005). *Nature*	472	7	0.308	4.572	14.831
Cronin, B. (2006). *JASIST*	1038	7	0.678	4.572	6.744
Egghe, L. (2006). *ISSI Newsletter*	306	6	0.200	3.919	19.608
Egghe, L. (2006). *Scientist*	229	6	0.150	3.919	26.201
Kelly, C. D, (2006). *Trends in Ecology & Evolution*	880	6	0.575	3.919	6.818
Banks, M. G. (2006). *Scientometrics*	849	6	0.555	3.919	7.067
Glänzel, W. (2006). *SCI Focus*	188	6	0.123	3.919	31.915
Glänzel, W. (2005). *ISSI Newsletter*	809	5	0.528	3.266	6.180
Egghe, L. (2006). *Scientometrics*	211	5	0.138	3.266	23.697
Braun, T. (2006). *Scientometrics*	967	5	0.632	3.266	5.171

Table 5
Wilcoxon signed ranks test of node strength and h-degree for all nodes in the co-citation network.

	h-Degree vs. node strength
Z	−33.818
Statistical significance (2-sided test)	0.000

have strong links with each other so that the sub-network has a high density and approaches a fully connected network. Obviously, Hirsch's original paper must be among these key nodes. Visually Fig. 7 exhibits a core/periphery structure (Borgatti & Everett, 1999). Some parameters for nodes of h-degree ≥5 are given in Table 4.

Table 4 clearly shows that the notions of node strength and h-degree are different. For instance: the two nodes with h-degree equal to 7 have node strengths that differ by a factor two. The node strengths of the five nodes with h-degree equal to six vary from 188 to 880. This difference can be measured by the R_h indicator. The node "Glänzel, W. (2006), SCI Focus" has the largest R_h-value in Table 4. This means that the weight of this node's links is concentrated (in a relative sense) in a few strong links.

For all nodes in our case network, as shown in Table 5, Wilcoxon's signed rank test shows the same result: the node degree and h-degree of nodes in this network are (statistically) significantly different.

One reason why network science became a hot topic during the end of last century was the fact that it was observed that links in the Internet follow a power-law distribution (Barabási, 2009; Barabási & Albert, 1999; Rousseau, 1997). Using the data of our case study, we observe that the distribution of the h-degree also follows (approximately) a power-law. This is illustrated in Fig. 8 using a relative size-frequency log–log plot.

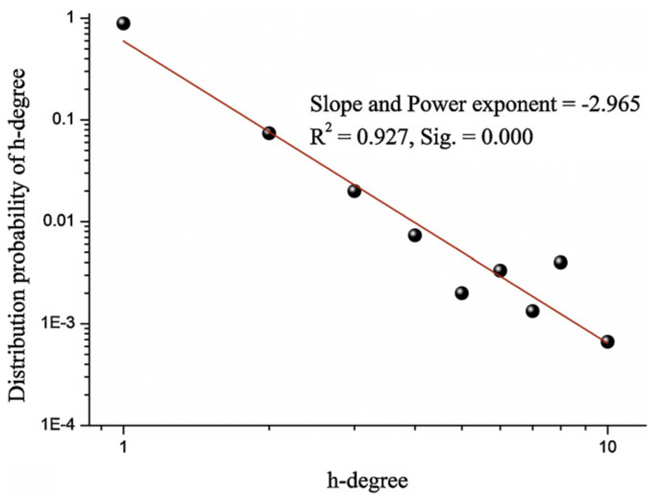

Fig. 8. Power-law distribution of h-degree in logarithmic coordinate axes.

Table 6
h-Centralization and other parameters for whole network.

	h-Index of network	Degree centralization ($\times10^{-2}$) (a)	h-Centralization ($\times10^{-6}$) (b)	b/a
Network case study	8	4.83	3.69	0.76×10^{-4}

The newly introduced network parameters for our case study are shown in Table 6. The main point made here is that the distribution of weights in our example is quite uneven.

4. Discussion and conclusion

The methodological section and the case study indicate that the h-degree and derived indicators have interesting features. First, they are more suitable for weighted networks than indicators based on the underlying unweighted network, they combine the number of links and the strength of links, and reflect more information about the links' strength and structure. Second, they focus on the most important nodes and are consistent with a power-law. Third, they inherit some merits of the h-index. For example, the h-degree is easy to calculate.

However, there are also some limits on the use of the h-degree and its derived measures. In some network studies, the numerical scale of links' weight is not suitable for ranking, for instance when weights belong to the interval [0, 1] and certainly when weights can be negative. At this point, a numerical transformation is needed (but then scientists should agree on the exact formulation of this transformation). Second, similar to the h-index, also the h-degree may be an indicator lacking discriminatory power.

The measures introduced in this paper may lead to further informetric network studies. In citation networks, the h-degree refers to the most cited nodes. In keyword networks, it points to core concepts. In co-author networks, it singles out the most collaborative authors. Using the notion of h-centralization different networks can be compared. Furthermore, the h-degree may be applied in any discipline, and in any weighted network.

We conclude by saying that the h-degree, h-centrality, h-centralization and other parameters based on the h-degree provide a set of useful measures for the study of weighted networks. Similar to node strength the h-degree is a basic measure, leading to new characterizations in networks, which may stimulate further studies.

Acknowledgements

The authors are supported by NSFC Grants (Nos. 70773101 and 7101017006) and thank Mr. Paul L. Zhang for assistance in data processing. The authors thank Raf Guns (Antwerp University) for correcting some errors in an earlier version.

References

Barabási, A.-L. (2009). Scale-free networks: A decade and beyond. *Science*, *325*(5939), 412–413.
Barabási, A.-L., Albert, R., & Jeong, H. (2000). Scale-free characteristics of random networks: The topology of the World-Wide Web. *Physica A*, *281*(1–4), 69–77.
Barabási, A. L., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, *286*(5439), 509–512.
Barrat, A., Barthélemy, M., Pastor-Satorras, R., & Vespignani, A. (2004). The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(11), 2747–3752.
Bollen, J., Van de Sompel, H., Hagberg, A., & Chute, R. (2009). A principal component analysis of 39 scientific impact measures. *PLoS One*, *4*(6), e6022.
Borgatti, S. P., Everett, M. G., & Freeman, L. C. (2002). *UCINET for Windows: Software for social network analysis*. Harvard, MA: Analytic Technologies.
Borgatti, S. P., Mehra, A., Brass, D. J., & Labianca, G. (2009). Network analysis in the social sciences. *Science*, *323*(5916), 892–895.
Borgatti, S. P., & Everett, M. G. (1999). Models of core/periphery structures. *Social Networks*, *21*(4), 375–395.
Börner, K., Huang, W. X., Linnemeier, M., Duhon, R. J., Phillips, P., Ma, N. L., et al. (2010). Rete-netzwerk-red: Analyzing and visualizing scholarly networks using the Network Workbench Tool. *Scientometrics*, *83*(3), 863–876.
Börner, K., Sanyal, S., & Vespignani, A. (2007). Network science. *Annual Review of Information Science and Technology*, *41*, 537–607.
Butts, C. T. (2008). Social network analysis: A methodological introduction. *Asian Journal of Social Psychology*, *11*(1), 13–41.
Chen, P., & Redner, S. (2010). Community structure of the physical review citation network. *Journal of Informetrics*, *4*(3), 278–290.
Clauset, A., Shalizi, C. R., & Newman, M. (2009). Power-law distributions in empirical data. *SIAM Review*, *51*(4), 661–703.
Ding, Y., Yan, E. J., Frazho, A., & Caverlee, J. (2009). PageRank for ranking authors in co-citation networks. *Journal of the American Society for Information Science and Technology*, *60*(11), 2229–2243.
Egghe, L. (2005). *Power laws in the information production process: Lotkaian informetrics*. Oxford (UK): Elsevier.
Egghe, L., & Rousseau, R. (2002). Co-citation, bibliographic coupling and a characterization of lattice citation networks. *Scientometrics*, *55*(3), 349–361.
Egghe, L., & Rousseau, R. (2006). An informetric model for the Hirsch-index. *Scientometrics*, *69*(1), 121–129.
Everett, M., Sinclair, P., & Dankelmann, P. (2004). Some centrality results new and old. *Journal of Mathematical Sociology*, *28*(4), 215–227.
Fang, J. Q. (2010). Network science—Theory and application. *Journal of the American Society for Information Science and Technology*, *61*(6), 1297–1298.
Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry*, *40*, 35–41.
Freeman, L. C. (1979). Centrality in social networks: Conceptual clarification. *Social Networks*, *1*(3), 215–239.
Glänzel, W. (2006). On the h-index—A mathematical approach to a new measure of publication activity and citation impact. *Scientometrics*, *67*(2), 315–321.
Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(46), 16569–16572.
Korn, A., Schubert, A., & Telcs, A. (2009). Lobby index in networks. *Physica A-Statistical Mechanics and its Applications*, *388*(11), 2221–2226.
Liu, D. R., & Shih, M. J. (2005). Hybrid-patent classification based on patent-network analysis. *Journal of the American Society for Information Science and Technology*, *62*(2), 246–256.

Liu, X. M., Bollen, J., Nelson, M. L., & Van de Sompel, H. (2005). Co-authorship networks in the digital library research community. *Information Processing & Management*, *41*(6), 1462–1480.
Opsahl, T., Agneessens, F., & Skvoretz, J. (2010). Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks*, *32*, 245–251.
Otte, E., & Rousseau, R. (2002). Social network analysis: A powerful strategy, also for the information sciences. *Journal of Information Science*, *28*(6), 441–453.
Rodriguez, M. A., & Pepe, A. (2008). On the relationship between the structural and socioacademic communities of a coauthorship network. *Journal of Informetrics*, *2*(3), 195–201.
Rousseau, R. (1997). Situations: An exploratory study. *Cybermetrics*, *1*(1) http://www.cindoc.csic.es/cybermetrics/v1i1p1.html
Rousseau, R., & Ye, F. Y. (2011). Subgraphs derived from the h-core in undirected, unweighted networks. *ISSI Newsletter*, *7*(1), 5–9.
Schubert, A. (2010). A reference-based Hirschian similarity measure for journals. *Scientometrics*, *84*(1), 133–147.
Schubert, A., & Glänzel, W. (2007). A systematic analysis of Hirsch-type indices for journals. *Journal of Informetrics*, *1*(3), 179–184.
Schubert, A., Korn, A., & Telcs, A. (2009). Hirsch-type indices for characterizing networks. *Scientometrics*, *78*(2), 375–382.
Schubert, A., & Soos, S. (2010). Mapping of science journals based on h-similarity. *Scientometrics*, *83*(2), 589–600.
Scott, J. (2000). *Social network analysis: A handbook*. London: Sage Publications.
Su, H. N., & Lee, P. C. (2010). Mapping knowledge structure by keyword co-occurrence: A first look at journal papers in Technology Foresight. *Scientometrics*, *85*(1), 65–79.
Watts, D. J. (2004). The “new” science of networks. *Annual Review of Sociology*, *30*, 243–270.
Ye, F. Y. (2011). A unification of three models for the h-index. *Journal of the American Society for Information Science and Technology*, *62*(1), 205–207.