Jointly published by Akadémiai Kiadó, Budapest and Springer, Dordrecht
Scientometrics, Vol. 78, No. 2 (2009) 375–382
DOI: 10.1007/s11192-008-2218-1

Hirsch-type indices for characterizing networks

ANDRÁS SCHUBERT,[a] ANDRÁS KORN,[b] ANDRÁS TELCS[c]

[a] *Hungarian Academy of Sciences, Institute for Research Policy Studies, Budapest, Hungary*
[b] *Budapest University of Technology and Economics, Department of Telecommunications and Media Informatics, Budapest, Hungary*
[c] *Budapest University of Technology and Economics, Department of Computer Science and Information Theory, Budapest, Hungary*

Hirsch-type indices are devised for characterizing networks and network elements. Their actual use is demonstrated on scientometric examples, and the potential value of the concept on a practically unlimited range of networks is suggested.

Introduction

In his original article, Hirsch [1] introduced the h-index as a performance measure. Most of the literature derived from this paper (nearly 200 citing items until the end of 2008; for a review, see e.g. [2]) studied the concept in the same or similar context. Even the predecessor of the index – the story goes back several decades as far as to Sir Arthur Eddington [3] – was intended to measure performance; as it happened, not in scientific achievements but in cycling prowess.

Among the rare exceptions, Glänzel [4] analyzed the mathematical properties of the h-index irrespective of its utility aspects. Using Gumbel's extreme value theory, he concluded that in the class of distributions obeying an asymptotical power law ("asymptotically Paretian distributions" – the most typical class of distributions in scientometrics, among other fields) the h-index can be approximated by a power function of the sample size and the sample mean. A refined version of this theoretical result gained later empirical support based on the example of h-indices of journals [5, 6].

Received October 8, 2008

Address for correspondence:
ANDRÁS SCHUBERT
E-mail: schuba@iif.hu

0138–9130/US $ 20.00
Copyright © 2008 Akadémiai Kiadó, Budapest
All rights reserved

Barabási [7] introduced the concept of *preferential attachment* to describe the evolution of a class of networks. The concept proved to be rather fruitful in modeling physical, biological and social (even scientometric [8]) networks. One of the key properties of the model is that the degree distribution has a power-law upper tail, i. e., the distribution is asymptotically Paretian.

In the present paper Hirsch-type indices are proposed as network indicators. Just like the "classical" h-index does with the number of publications and the citation rate, the network h-index is expected to characterize both the size and the density of the network in a particularly balanced way. Presuming that the networks under study represent finite stages of the evolution of "Barabási type" infinite networks, we expect that the network h-index will be a power function of the size and the density of the network.

Hirsch-type indices as network indicators

A *network* is a set of *nodes* optionally connected by *edges*. The *size* of the network can be measured by the number of nodes, n. The *degree* of a node, d, is the number of edges that are adjacent to the node. The average degree of nodes, $\rho = e/n$ (e is the total number of edges) is a measure of *network density* (or connectivity). Several characteristic features of a network (diameter, centrality, clustering properties, etc.) are determined by the *degree distribution*, of which ρ is the simplest statistic. A well-chosen supplementary statistic may significantly deepen our insight into the nature of a network.

Definition: A network has a *degree h-index* of h, if not more than h of its nodes have a degree of not less than h.

The degree h-index measures, so to say, the *influential weight* of a network determined by its size and density (similarly as Hirsch's original h-index was intended to measure the professional weight of an author through the number and citedness of his/her publications). A smaller but more connected network may thus 'outweigh' a larger but sparser one.

It is intuitively clear that the degree h-index might be related to the *centralization* ("global centrality") of the network.* Given the size and the density of a network, the degree h-index takes its maximum somewhere between the completely centralized and the completely decentralized cases.

We think that such an indicator can be effectively used in various applications, most particularly (but by no means exclusively) in the case of social networks. In what follows, two illustrative examples of its application in scientometrics will be presented.

---

* A similar concept for node centrality was introduced as "lobby index" in Ref. [9].

Networks of papers in journals

Papers published in a journal can be considered nodes of a network with links (edges) between papers that share at least one common author.

This network of papers has a degree h-index h_P, if h_P is the greatest number of papers in the network that have a degree of at least h_P.

Networks of authors in journals

Authors publishing in a journal can be considered nodes of a network with links (edges) between authors who co-authored at least one paper in the journal.

This network of authors has a degree h-index h_A, if h_A is the greatest number of authors in the network who have a degree of at least h_A.

Results

As an empirical sample, a set of 36 journals in the field of Dentistry & Oral Medicine has been taken. The bibliographic data of their articles published in 1999 and the citations received by these papers in the period 1999–2001 have been retrieved from the Science Citation Index database of Thomson–Reuter's Web of Science. Table 1 contains the basic bibliometric data of the 36 journals of the sample: the number of papers and authors, n_P and n_A, the density of the paper and the author network, ρ_P and ρ_A, the citation-based h-index, h_{cit} (cf. [5, 6]), the h_P and h_A indices and the degree centralization indices of the networks of papers and authors, c_{dP} and c_{dA}. The degree centralization index, c_d, is defined [10] as

$$c_d = \frac{\sum_{i=1}^{n}(d_{\max} - d_i)}{(n-1)(n-2)},$$

where $d_{\max}$ is the maximum degree of a node in the network, d_i is the degree of the i-th node and n is the number of nodes.

The dependence of the degree h-index on the size and density of the network

As mentioned in the Introduction, Glänzel [4] constructed a theoretical model of the dependence of the citation h-index on the sample size and the sample's mean citation rate, provided that the citation distribution follows an asymptotical inverse power law. In the simplest case this model results in the formula: $h = cn^{1/3}x^{2/3}$, where h is the degree h-index, n is the sample size, x is the mean citation rate and c is a positive constant. This formula gained empirical support, among others, in extensive samples of journals [6].

Table 1. Basic bibliometric data of 36 Dentistry & Oral Medicine journals

Journal title	n_P	n_A	ρ_P	ρ_A	h_{cit}	h_P	h_A	c_{dP}	c_{dA}
Acta Odontologica Scandinavica	66	161	0.818	2.994	4	3	7	0.0663	0.0507
Am. J. Dentistry	67	192	0.925	3.292	8	4	8	0.0793	0.0672
Am. J. Orthodont. & Dentofacial Orthoped.	239	478	1.264	2.699	6	7	7	0.0243	0.0175
Angle Orthodontist	104	214	1.462	2.907	4	7	7	0.0548	0.0241
Archives of Oral Biol.	142	492	0.761	3.988	7	6	11	0.0377	0.0307
British Dental J.	374	500	2.144	2.004	6	18	9	0.0481	0.0221
British J. Oral & Maxillofacial Surg.	84	223	0.524	3.390	4	3	8	0.0306	0.0210
Caries Research	61	202	0.689	3.752	7	2	7	0.0571	0.0314
Cleft Palate – Craniofacial J.	102	316	0.490	3.791	4	3	11	0.0253	0.0422
Clinical Oral Implants Research	52	207	0.346	3.855	6	2	7	0.0337	0.0301
Community Dentistry & Oral Epidemiol.	63	155	0.413	2.542	5	2	5	0.0264	0.0227
Critical Reviews in Oral Biol. & Med.	13	28	0.000	1.857	6	0	3	0.0000	0.0456
Eur. J. Oral Sciences	70	239	0.314	3.565	6	2	8	0.0251	0.0188
Eur. J. Orthodont.	61	167	0.459	2.539	4	2	5	0.0610	0.0272
Intl. J. Oral & Maxillofacial Implants	104	327	0.981	3.884	7	4	12	0.0299	0.0436
Intl. J. Oral & Maxillofacial Surg.	105	326	0.648	3.405	6	3	8	0.0329	0.0204
Intl. J. Periodont. & Restorative Dentistry	60	150	0.433	3.067	6	2	7	0.0450	0.0268
J. Am. Dental Association	308	404	2.649	2.401	7	14	11	0.0372	0.0214
J. Clinical Periodontol.	133	467	1.609	4.994	8	8	17	0.0645	0.0259
J. Cranio-Maxillofacial Surgery	56	185	0.500	3.773	6	3	8	0.0660	0.0507
J. Dental Research	1856	4666	2.875	4.069	10	14	19	0.0092	0.0049
J. Dentistry	78	244	1.103	4.328	8	3	15	0.0786	0.0650
J. Dentistry for Children	75	158	2.027	2.646	4	10	5	0.1107	0.0345
J. Endodont.	183	484	1.388	3.521	5	6	9	0.0312	0.0259
J. Oral & Maxillofacial Surgery	364	735	1.236	2.438	6	13	7	0.0326	0.0131
J. Oral Pathol. & Medicine	87	390	0.575	4.805	8	3	9	0.0289	0.0238
J. Oral Rehabilitation	147	412	0.925	3.316	5	4	8	0.0214	0.0188
J. Orofacial Pain	43	72	0.791	2.500	7	3	6	0.0552	0.0942
J. Periodontal Research	55	211	0.582	4.066	7	3	7	0.0849	0.0429
J. Periodontol.	201	739	0.985	6.005	11	5	34	0.0253	0.0571
J. Prosthetic Dentistry	240	592	0.983	3.084	8	5	10	0.0380	0.0287
J. Public Health Dentistry	50	124	1.400	3.790	4	4	11	0.0765	0.1257
Operative Dentistry	64	156	0.781	2.910	6	3	6	0.0527	0.0463
Oral Microbiol. & Immunol.	60	205	0.667	3.629	5	3	6	0.0409	0.0266
Oral Oncol.	101	427	0.574	4.885	6	3	10	0.0452	0.0262
Oral Surg. Med. Pathol. Radiol. & Endod.	293	908	0.758	3.844	8	5	11	0.0181	0.0112

Since the degree distribution of "preferential attachment" networks (to which co-author networks presumably belong) is supposed to follow an asymptotical inverse power law [7, 8], the degree h-indices, h_P and h_A, are expected to obey similar relations:

$$h_P = c_P n_P^{1/3} \rho_P^{2/3},$$

$$h_A = c_A n_A^{1/3} \rho_A^{2/3}.$$

Figure 1 shows the fit of the empirical degree h-indices to the theoretical model.

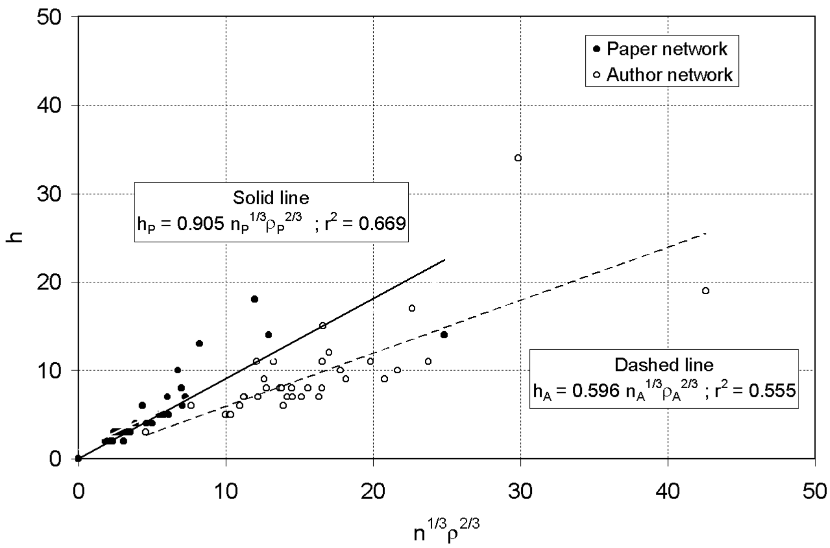

Figure 1. Fit of the degree h-indices to the theoretical model

The fairly good fit strongly supports the the Paretian model of degree distribution in the paper and author networks and, indirectly, the underlying "preferential attachment" model of their evolution.

The h-index and the measures of network centralization

The relation of the degree h-index to measures of network centralization is somewhat deceptive. Obviously, the degree h-index of an *n*-element star network (maximum centralization) is 1 and that of an *n*-element complete network (minimum centralization) is $n-1$. This fact would suggest an inverse relation between the degree h-index and the centralization measures. In general, however, the normalized degree h-index or h-fraction, *h*/*n*, the number of nodes in the 'h-core' (nodes with degree equal to or above *h*) is not in direct functional relation with the measures of network centralization.

The networks studied in our empirical samples are rather sparse, i.e., the great majority of nodes have zero or very few neighbors. In extremely sparse networks both the degree h-index and the centralization is very low. During the evolution of such networks, as newer and newer links are formed among the nodes, the h-index and centralization increase in parallel. It seems that our samples clearly reflect this evolutionary phase. In Figure 2, the degree centralization turns out to be practically equal to the normalized h-indices.

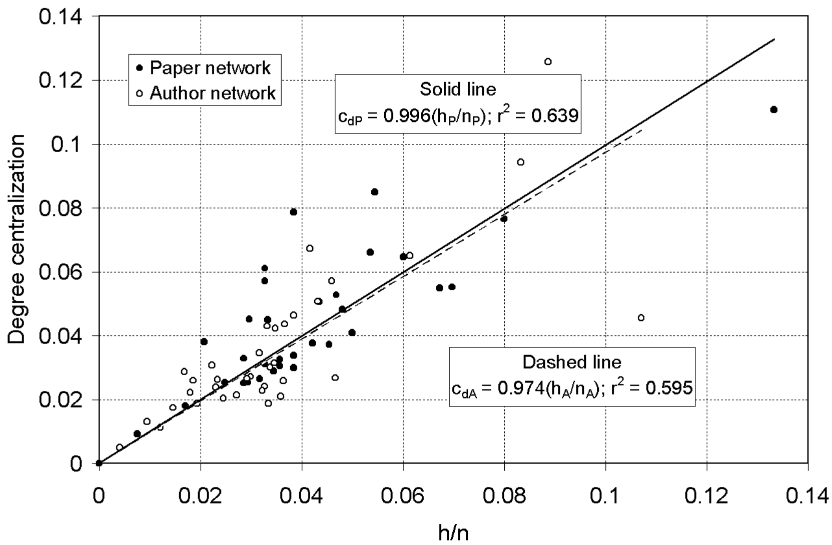

Figure 2. Correlation of degree centralization with the degree h-indices

The normalized degree h-index thus appears useful as an empirical indicator of centralization, although the theoretical background of its behavior is still waiting for elaboration.

The influence of the degree h-index on the citation features of a journal

The relation of the degree h-index with other network measures results from the internal structure of the networks concerned. A subsequent step might be to find relations between the network structure and some 'external' features.

The degree h-index was said to represent a kind of 'influential weight' of the network. One may wonder whether this kind of weight of the paper or the author network of a journal has any influence on a substantially independent feature of the journal, such as, for example, its citedness. Figure 3 shows the correlation between the degree h-indices and the citation-based journal h-index (cf. [5, 6]).

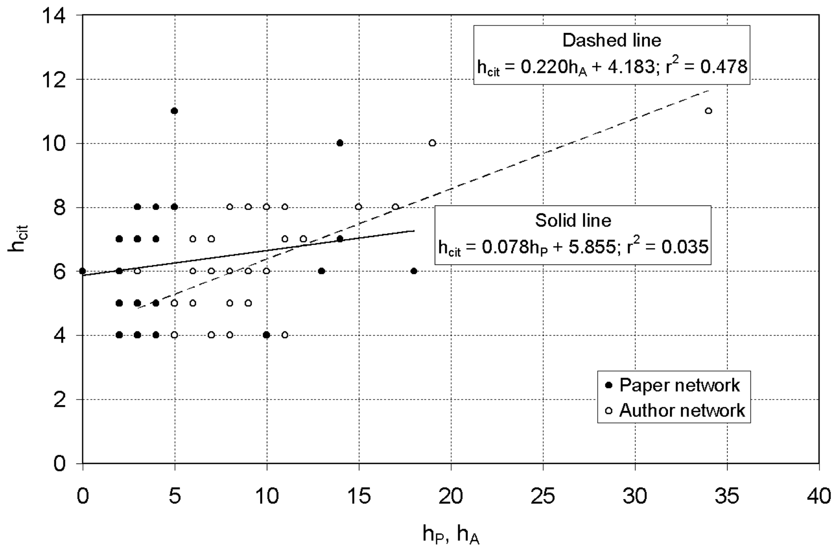

Figure 3. Correlation of the citation h-index with the degree h-indices

Apparently, the degree h-index of the paper network has no observable correlation with the citation h-index; that of the author network, however, exhibits significant positive correlation. This is exactly what could be expected: the paper network is a largely artificial construction without any 'organic' association, whereas the author network reflects an existing organized community. The influence of the latter on citation behavior (partly through 'condemnable' practices such as self-citation or cross-citation) is easily comprehensible.

Discussion and conclusions

A new indicator, the degree h-index was devised for characterizing networks. The examples considered gave empirical support to some of its features supposed to hold for scientometric networks of papers and authors.

(i) The degree h-index depends on the size and the density of the paper and author networks as predicted by Glänzel's model [4, 6]. This indirectly substantiates the Barabási-type 'preferential attachment' model for their evolution [7, 8].

(ii) The normalized degree h-index (h-fraction) is in an intricate relation with the centralization of the network. In the realm of journal paper and author networks under study, the degree h-index is approximately equal to the degree centralization of the network.

(iii) The degree h-index represents a kind of 'influential weight' of networks, which may affect the behavior of the network in external relations. As an example, it was found that the degree h-index of the author networks is in positive correlation with the citation h-index of the journals. The paper network exhibited no such correlation.

The examples of this paper were taken from the field of scientometrics, but the concept of the degree h-index can be applied and its relations with other network properties and indicators can be sought for in any of the vast areas of network studies. Social, as well as electric, electronic, information or transportation networks might be interesting and promising fields of application and, of course, the Internet may not only serve as a boundless source of data for investigation, but its inherent structure recently became the primary testbed of methods for exploring huge networks.

Among the diverse directions of possible future research, a more detailed understanding of the relation between the degree h-index and the degree centrality (as well as other centrality measures) and the connection of the h-index concept and the 'rich-club phenomena' might be stressed.

The authors hope that the initial steps taken in this paper will be followed by others to deepen and widen our knowledge in this intriguing topic.

References

1. J. E. HIRSCH, An index to quantify an individual's scientific output, *Proceedings of the National Academy of Sciences of the United States of America*, 102 (2005) 16569–6572.
2. L. BORNMANN, H. D. DANIEL, What do we know about the h index? *Journal of the American Society for Information Science and Technology*, 58 (9) (2007) 1381–1385.
3. A. W. F. EDWARDS, System to rank scientists was pedalled by Jeffreys, *Nature*, 437 (2005) 951.
4. W. GLÄNZEL, On the h-index – A mathematical approach to a new measure of publication activity and citation impact, *Scientometrics*, 67 (2) (2006) 315–321
5. T. BRAUN, W. GLÄNZEL, A. SCHUBERT, A Hirsch-type index for journals, *Scientometrics*, 69 (1) (2006) 169–173.
6. A. SCHUBERT, W. GLÄNZEL, A systematic analysis of Hirsch-type indices for journals, *Journal of Informetrics*, 1 (3) (2007) 179–184.
7. A. L. BARABÁSI, *Linked: How Everything is Connected to Everything Else and What It Means for Business, Science, and Everyday Life*. Plume, New York, 2003.
8. A. L. BARABÁSI, H. JEONG, Z. NÉDA, E. RAVASZ, A. SCHUBERT, T. VICSEK, Evolution of the social network of scientific collaborations, *Physica A*, 311 (2002) 590–614.
9. A. KORN, A. SCHUBERT, A. TELCS, Lobby index in networks, submitted to *Physical Reviews E.*
10. L. C. FREEMAN, Centrality in social networks. I. Conceptual clarification, *Social Networks*, 1 (1979) 215–239.