Journal of Informetrics 7 (2013) 129–137

Contents lists available at SciVerse ScienceDirect

Journal of Informetrics

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

A probe into dynamic measures for h-core and h-tail

Dar-zen Chen[a], Mu-hsuan Huang[b], Fred Y. Ye[c,d,*]

[a] Department of Mechanical Engineering and Institute of Industrial Engineering, National Taiwan University, Taipei, Taiwan
[b] Department of Library and Information Science, National Taiwan University, Taipei, Taiwan
[c] School of Information Management, Nanjing University, Nanjing, China
[d] Department of Information Resource Management, Zhejiang University, Hangzhou, China

ARTICLE INFO

Article history:
Received 27 July 2012
Received in revised form
28 September 2012
Accepted 3 October 2012

Keywords:
h-core
h-tail
Shape centroid
Shape descriptor
h-index
k-index

ABSTRACT

Using the concepts of h-core and h-tail, shape descriptors and shape centroids, *k*-index and *k′*-index, dynamic measures are probed, with practical data in the fields of Physics and sociology. It is revealed that there are obvious differences between natural sciences (Physics, particles & fields) and social sciences (sociology) when *c*-descriptor, h-core centroid and *k*-index are applied as dynamic measures, while few differences exist when using *t*-descriptor, h-tail centroid and *k′*-index, following a time span from 1 to 10 years.

1. Introduction

Since the *h*-index was introduced in 2005 (Hirsch, 2005), it has been applied as an academic measure (Alonso, Cabrerizo, Herrera-Viedma, & Herrera, 2009; Egghe, 2010) and has led to a simple and meaningful unification of publications and citations (Ye, 2009, 2011).

The *h*-index is regarded as a robust indicator of measuring both the impact and output of publications, which link with quality and quantity respectively. Mostly, the mistaken data on citations are easily caused in the long-tail part of low-citation publications when we count citations in a database; therefore the *h*-index has high accuracy in academic assessment (Vanclay, 2007). On the other hand, the *h*-index can be applied to forecast the future academic performance of scholars (Hirsch, 2007). However, the *h*-index is presented in an integral in which similar *h*-indices in scholars or institutions readily exist, making it ineffective indifferentiating academic performances (Nair & Turlach, 2012). Huang and Chi (2010) also compared three different indices for the institution level research evaluation. Hence, some h-type indices have been introduced to improve the *h*-index, and several related indicators have been observed (Egghe, 2006; Glanzel, 2006; Jin, Liang, Rousseau, & Egghe, 2007; Kuan, Huang & Chen, 2011a, 2011b, 2012a, 2012b; Ye, 2010).

Meanwhile, the *h*-index ignores its research target's number of papers and citation distribution, which may result in *h*-inconsistency in some cases (Waltman & van Eck, 2012). Some scholars consider that the number of papers and citation distribution in h-core and h-tail should be looked to for a more correct academic assessment (Rousseau, 2006; Ye & Rousseau, 2010). Kuan et al. (2011a) suggested the two indicators, *c*-descriptor and *t*-descriptor, for analyzing patent performance of assignees according to their rank-citation curves based on practical data. Since *c*-descriptor and *t*-descriptor are not able to

* Corresponding author. Tel.: +86 571 87996929; fax: +86 571 87951670.
E-mail addresses: yye@nju.edu.cn, yye@zju.edu.cn (F.Y. Ye).

1751-1577/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.joi.2012.10.002

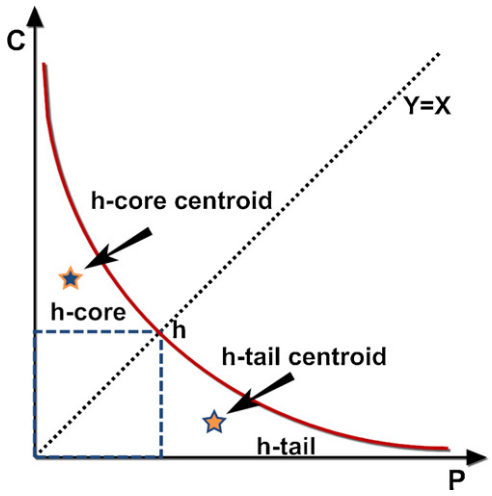

Fig. 1. h-core and h-tail in *P*–*C* plane.

instantly show the relative patent performance of all assignees, Kuan et al. (2011b) further introduced h-core centroid and h-tail centroid, which are located at the geometric centers of the h-core and h-tail areas following the rank-citation curves. Comparing with g-core (Egghe, 2006), pi-core (Vinkler, 2010), with 0.1, 0.01, and 1% of total papers, etc. (Radicchi, Fortunato, & Castellano, 2008), with the I3-index (Leydesdorff & Bornmann, 2011), or with CDS-index (Vinkler, 2011), h-core is simpler and easier so that we choose it as elite concept, even though the others also possess potential.

When Liang (2006) introduced the *h*-index sequence and *h*-index matrix to overcome the faults of the *h*-index in specific time spans, Rousseau and Ye (2008) also proposed dynamic the h-type index for measuring the dynamic non-linear properties. While Egghe (2009a, 2009b) set up the mathematical model for the *h*-index sequence, Nair and Turlach (2012) developed the stochastic *h*-index. Also, we mention dynamic h-measures with references (Egghe, 2007; Vinkler, 2010; Ye, 2012). All the studies show that scholars have paid attention to dynamic process of the *h*-index.

In the beginning, Hirsch found that there were differences among different fields with ten cases each in Physics, particles & fields and biology when he proposed the *h*-index (Hirsch, 2005). In recent years, many scholars have revealed the informetric differences in various fields (Batista, Campiteli, Kinouchi, & Martinez, 2006; Iglesias & Pecharroman, 2007; Lillquist & Green, 2010). Batista et al. (2006) set the fields of Physics, Chemistry, Biology/Biomedicine and Mathematics as targets for the analysis of the differences in co-author status among various fields and subjects, and applied h_I-index to correct the expansion that co-authors caused in the *h*-index. Iglesias and Pecharroman (2007) compared different fields with a simple method for scaling the *h*-index so that we could compare *h*-indices across fields. Lillquist and Green (2010) focused on several target institutions and collected scholars' paper data in the fields of Biology, Chemistry, Mathematics, Physics and Engineering. They further divided the Engineering field into Civil, Mechanical, Electrical and Chemical Engineering as major disciplines, and analyzed researchers' *h*-index performance in Biology, Chemistry, Mathematics, Physics, Engineering and the sub-fields of Engineering, to observe if diversity of the *h*-index was shown.

Based on above studies, we try to probe into the dynamic measures for h-core and h-tail, with indicators such as *c*-descriptor, *t*-descriptor, h-core centroid (c_x, c_y) and h-tail centroid (t_x, t_y), as well as *k* and *k*′ (Ye & Rousseau, 2010). On the basis of the data used, the time span was changed from one or two years to 10 years, the dynamic changes of publications, citations, and rank-citation curves in different fields during the 2001–2010 period have been studied, particularly in the fields of Physics (as one of the typical natural sciences) and Sociology (as one of the typical social sciences).

2. Methodology

2.1. Method

It is well known that publications (*P*) and citations (*C*) can be arranged into a diagram when ranked according to total citations of each publication from high to low, in which the *h*-index is always located on the *P*–*C* curve (as *R*–*C*, rank-citation). In the *P*–*C* plane, the h-core and h-tail are distributed as shown in Fig. 1.

We are interested in the h-core and h-tail portions, and the difference between natural science and social science, for which the *e*-index (Zhang, 2009) and *k*-index (Ye & Rousseau, 2010) could be defined as:

$$C_H = h^2 + e^2 \tag{1}$$

$$k = \frac{C/P}{C_T/C_H} \tag{2}$$