



Breadth and depth of citation distribution



Siluo Yang*, Ruizhen Han

School of Information Management, Wuhan University, Wuhan 430072, China

ARTICLE INFO

Article history:

Received 18 October 2013

Received in revised form 24 September 2014

Accepted 16 December 2014

Available online 7 January 2015

Keywords:

Citation analysis

Bibliometrics

Citation distribution

X index

Y index

XY index

ABSTRACT

This study proposes a new 4D (i.e., spatial, temporal, breadth, and depth) framework for citation distribution analysis. The importance and differences in the breadth and depth of citation distribution are analyzed. Easily computable indices, X , Y , and XY , are proposed, which provide estimates of the breadth and depth of citation distribution. A knowledge unit can be an article, author, institution, journal, or a set of something. Index X , which represents the breadth of citation distribution, is the number of different knowledge units that cite special knowledge units. Index Y , which represents the depth of citation distribution, is the maximum number of citations among several knowledge units that refer to specific knowledge units. Index XY , which synthetically represents Indices X and Y , the feature and focus impacts of a knowledge unit, is index X divided by index Y . We analyze empirically the citation and reference distributions of 84 journals from the “Information science and library science” category of the Journal Citation Reports (2012) at the journal-to-journal level. Indices X , Y , and XY reflect the actual breadth and depth of citation distribution. Differences exist among Indices X , Y , and XY . Differences also exist between these indices and other bibliometric indicators. These indices cannot be replaced by existing bibliometric indicators. Specifically, the absolute values of indices X and Y are good supplements to existing bibliometric indicators. However, index XY and the relative values of Indices X and Y represent new aspects of bibliometric indicators.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

The first scientific journals appeared in the 16th century, but the systematic practice of citation was popularized later (Nicolaisen, 2007). With the rapid development of science and technology, researchers have begun to pay more attention to citations and references. In the 20th century, the practice of citing other works is almost second nature to anyone writing a scholarly or scientific paper (Kaplan, 1965). In the 1950s, Johns Hopkins University Librarian Eugene Garfield was inspired by an existing law database to create the Science Citation Index (SCI), which enables researchers to analyze millions of citations automatically and conveniently. Thereafter, some important contents and concepts of citation analysis were proposed, such as “Bibliographic Coupling” (Kessler, 1963), “Science Citation Network” (Price, 1965), “Co-Citation” (Small, 1973), and “Citation Visualization” (White & McCain, 1998). In 1998, the advent of the network version database Web of Science further promoted the popularization of citation analysis. In the network environment, Webometrics (Almind & Ingwersen, 1997) and Altmetrics (Wouters & Costas, 2012) were proposed, and a new era of the development of citation analysis began.

* Corresponding author. Tel.: +86 18040599879.

E-mail address: slyang2005@aliyun.com (S. Yang).

Currently, citation analysis is widely used in scientific evaluation, scholarly communication reveal, academic behavior analysis, and information retrieval (Garfield, 1983; Hammarfelt, 2011; Ketzler & Zimmermann, 2013).

Citation analysis is the method of examining the frequency, patterns, and graphs of citations in articles and books (Garfield, 1983). However, the development of citation analysis was accompanied by a controversy regarding its effectiveness and reliability. Citation analysis received criticisms that pertain to weak basic theory on citation, unknown citing motivation, deficient citing process, disadvantageous citation analysis methods and data, and limited citation application (Bornmann & Daniel, 2008; MacRoberts & MacRoberts, 2010; Nicolaisen, 2007). Frequency and distribution are key points in citation analysis. Most researchers focus on citation frequency analyses (Egghe, Bornmann, & Guns, 2011), such as the simple statistical analysis of citation frequency, citation weighted analysis involving Eigenfactors, and a combination of other factors, such as H-index combining citation frequency and paper number (Bergstrom, 2007; Price, 1965; Yang, Han, Ding, & Song, 2012). Specialized research on citation distribution consists of four main aspects. First, it consists of the analysis of the mechanism and curves of the citation distribution model in general (Redner, 1998; Rodríguez-Navarro, 2011; Sangwal, 2013). Second, it consists of spatial and temporal distribution (Larivière, Archambault, & Gingras, 2008), for example, impact factor is a result of combining spatial and temporal analyses. Ding, Liu, Guo, and Cronin (2013) studied specific location distribution of citations in citing literature context. Third, several scholars have studied information diffusion by analyzing citation networks or citation patterns (Chatterjee & Chowdhury, 2008; Shi, Tseng, & Adamic, 2009). The breadth of citations represent the diffusion of knowledge into other fields and from basic to applied research and development (R&D), which was initially suggested by W.F. Lancaster in his “issue management” studies (Lancaster & Lee, 1985). Wu (2013) studied the geographical knowledge diffusion and spatial diversity by exploring and investigating the spatial properties of citation distances and patterns. Fourth, some researchers have explored the breadth of citation distribution at the macro level. Evans (2008) argued that researchers cite recent papers and concentrate their citations on a few papers despite the availability of online research (older and recent). Larivière, Gingras, and Archambault (2009) challenged the conclusion of Evans by analyzing the changes in the concentration of citations received within 2 and 5 years by papers published between 1900 and 2005. Their results showed that the dispersion of citations is increasing. The conclusions of Yang, Ma, Song, and Qiu (2010) were consistent with those of Larivière et al. (2009) to a certain extent. However, few papers have studied systematically the breadth and depth of citation distribution at the micro level and lack indicators that measure the breadth and depth of citation. Thus, we should pay attention to these indicators to strengthen and validate citation distribution analysis (Yang et al., 2012).

2. 4D model of citation distribution

We argue that citation distribution can be divided into four dimensions, namely, spatial, temporal, breadth, and depth distribution. Citation analysis can also be studied in general regardless of the four dimensions, which integrate citations frequency statistics, citations weighted analysis, citations network analysis, and research citations related to the topic (see Fig. 1).

Spatial and temporal distribution are universal. Temporal distribution is the variation in distribution related to time. Everything has a life cycle, and year, month, and day are used as the units of analysis in bibliometrics. For example, the series of classic laws of literature is growing and aging. By contrast, spatial distribution is the variation in distribution related to position in space. In citation analysis, spatial distribution is not limited to geographical space and can be extended to the relationship between various knowledge units, such as country, organization, and journal. The distribution laws of

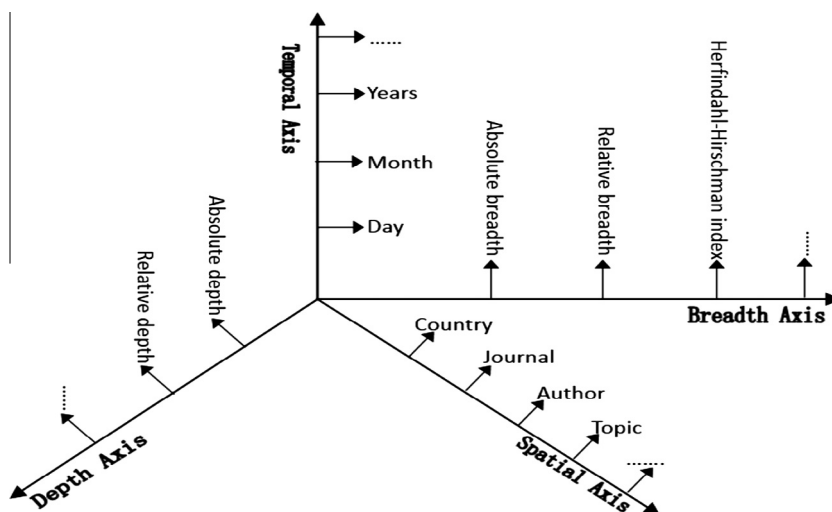


Fig. 1. 4D model of citation distribution.

Download English Version:

<https://daneshyari.com/en/article/515472>

Download Persian Version:

<https://daneshyari.com/article/515472>

[Daneshyari.com](https://daneshyari.com)