Journal of Informetrics 5 (2011) 146–166

Contents lists available at ScienceDirect

Journal of Informetrics

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

An approach for detecting, quantifying, and visualizing the evolution of a research field: A practical application to the Fuzzy Sets Theory field

M.J. Cobo*, A.G. López-Herrera, E. Herrera-Viedma, F. Herrera

Dept. Computer Science and Artificial Intelligence, CITIC-UGR (Research Center on Information and Communications Technology), University of Granada, E-18071-Granada, Spain

ARTICLE INFO

Article history:
Received 26 April 2010
Received in revised form 3 October 2010
Accepted 6 October 2010

Keywords:
Science mapping
Co-word analysis
Bibliometric studies
Fuzzy Sets Theory
Thematic evolution
h-Index

ABSTRACT

This paper presents an approach to analyze the thematic evolution of a given research field. This approach combines performance analysis and science mapping for detecting and visualizing conceptual subdomains (particular themes or general thematic areas). It allows us to quantify and visualize the thematic evolution of a given research field. To do this, co-word analysis is used in a longitudinal framework in order to detect the different themes treated by the research field across the given time period. The performance analysis uses different bibliometric measures, including the h-index, with the purpose of measuring the impact of both the detected themes and thematic areas. The presented approach includes a visualization method for showing the thematic evolution of the studied field.

Then, as an example, the thematic evolution of the Fuzzy Sets Theory field is analyzed using the two most important journals in the topic: *Fuzzy Sets and Systems* and *IEEE Transactions on Fuzzy Systems*.

1. Introduction

Bibliometrics is usually used for the quantitative research assessment of academic output, and it is starting to be used for practice based research (for more information see Callon, Courtial, & Laville, 1991; Coulter, Monarch, & Konda, 1998; Henderson, Shurville, & Fernstrom, 2009; Ramos-Rodrguez & Ruz-Navarro, 2004; van Raan, 2005a). Concretely, bibliometrics is a set of methods used to study or measure texts and information, especially in big datasets. Many research fields use bibliometric methods to explore the impact of their field, the impact of a set of researchers, or the impact of a particular paper (Henderson et al., 2009; van Raan, 2005a).

In bibliometrics, there are two main procedures: performance analysis and science mapping (Noyons, Moed, & Luwel, 1999; van Raan, 2005a). Performance analysis aims at evaluating groups of scientific *actors* (countries, universities, departments, researchers) and the impact of their activity (Noyons, Moed, & van Raan, 1999; van Raan, 2005a) on the basis of bibliographic data. Science mapping aims at displaying the structural and dynamic aspects of scientific research (Börner, Chen, & Boyack, 2003; Noyons, Moed, & Luwel, 1999). A science map is used to represent the cognitive structure of a research field.

Various types of techniques have been developed to build a science map (Small, 2006), the most commonly used being documents co-citation (Small, 1973) and co-word analysis (Callon, Courtial, Turner, & Bauin, 1983). Moreover, different

* Corresponding author.
E-mail addresses: mjcobo@decsai.ugr.es (M.J. Cobo), lopez-herrera@decsai.ugr.es (A.G. López-Herrera), viedma@decsai.ugr.es (E. Herrera-Viedma), herrera@decsai.ugr.es (F. Herrera).

1751-1577/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.joi.2010.10.002

methods have been proposed to address the problem of delimiting a research field, and quantifying and visualizing the detected subfields by means of co-word or co-citation analysis (Börner et al., 2003; Callon et al., 1991; Chen, Ibekwe-SanJuan, & Hou, 2010; Coulter et al., 1998; Courtial & Michelet, 1994; Courtial, 1990; Kandylas, Upham, & Ungar, 2010; Leydesdorff & Rafols, 2009; Rip & Courtial, 1984; Small & Upham, 2009; Small, 1977, 2006; Upham & Small, 2010). The majority of these methods are mainly focused on measuring the performance of the scientific actors and little research has been carried out in order to measure the performance of given research fields in a conceptual way (specific themes or whole thematic areas). A performance analysis of specific themes or whole thematic areas can measure (quantitatively and qualitatively) the relative contribution of these themes and thematic areas to the whole research field, detecting the most prominent, productive, and highest-impact subfields.

The main aim of this paper is to present a general approach to analyze the thematic evolution of a given research field. This approach combines performance analysis and science mapping for detecting and visualizing conceptual subdomains (particular themes or general thematic areas). It also allows us to quantify and visualize the thematic evolution of the research field. To do this, co-word analysis is used in a longitudinal framework (Garfield, 1994). For a better interpretation of the results, strategic diagrams are used in order to categorize the detected themes. Furthermore, thematic areas are used to show conceptual evolution, proposing a visualization approach for graphically showing the thematic evolution of the studied field. Additionally, we develop a performance analysis using different basic bibliometric indicators (the number of published documents, the number of received citations, etc.,) and the h-index (Alonso, Cabrerizo, Herrera-Viedma, & Herrera, 2009; Cabrerizo, Alonso, Herrera-Viedma, & Herrera, 2010; Hirsch, 2005). As an example, the proposed approach is applied to analyze the thematic evolution of the Fuzzy Sets Theory (FST)[1] research field (Zadeh, 1965, 2008) by only considering the documents published in the two most important journals on the topic: *Fuzzy Sets and Systems* and *IEEE Transactions on Fuzzy Systems*.

This paper is organized as follows. Section 2 gives a brief overview of the science mapping and longitudinal studies. Section 3 introduces the approach to analyze the evolution of a research field. Section 4 uses the approach in order to analyze the FST research field. Finally, some conclusions are drawn in Section 5.

2. Science mapping and longitudinal studies

Science mapping or bibliometric mapping is a spatial representation of how disciplines, fields, specialities, and individual papers or authors are related to one another (Small, 1999). It is focused on monitoring a scientific field and delimiting research areas to determine its cognitive structure and its evolution (Noyons, Moed, & van Raan, 1999).

Various types of techniques have been developed to build a science map (Small, 2006), the most commonly used being documents co-citation and co-word analysis.

Co-citation analysis was proposed by Small (1973). This tool maps the structure of a research field through pairs of documents that are commonly cited together (Coulter et al., 1998). Co-citation has been used in the literature to delimit research areas (Small, 2006), discover knowledge communities (Kandylas et al., 2010), research fronts (Upham & Small, 2010) and invisible colleges (Noma, 1984), and also to study different research fields such as the absorptive capacity field (Calero-Medina & Noyons, 2008), the organic thin film transistors (Small & Upham, 2009), to analyze the Strategic Management Journal (Ramos-Rodrguez & Ruz-Navarro, 2004) or to study the marrow of science (Moya-Anegón et al., 2007), among other applications.

Co-word analysis was proposed by Callon et al. (1983) as a content analysis technique that is effective in mapping the strength of association between information items in textual data. It deals directly with sets of terms shared by documents, mapping the pertinent literature directly from the interactions of key terms. Co-word analysis has been used to analyze the interactions between basic and technological research (Callon et al., 1991), study the software engineering field (Coulter et al., 1998), the information research field (Ding, Chowdhury, & Foo, 2001), the scientific area of physical chemistry of surfactants (Bailón-Moreno, Jurado-Alameda, & Ruz-Baños, 2006), the Spanish FST field (López-Herrera et al., 2009), to study the hybridization of the FST field with other computational intelligence techniques (and fields) (López-Herrera, Cobo, Herrera-Viedma, & Herrera, 2010), among others applications.

At the end of the co-word or co-citation analysis, a set of clusters is returned which can be understood as conglomerates of different scientific aspects. In the case of co-citation analysis, the clusters represent groups of references that can be understood as the intellectual base of the different subfields. On the other hand, in the case of co-word analysis, the clusters represent groups of textual information that can be understood as semantic or conceptual groups of different topics treated by the research field. So, the detected clusters can be used with several purposes such as:

- To analyze their evolution through measuring continuance across consecutive subperiods.
- To quantify the research field by means of a performance analysis.

In some studies, co-citation and co-word analysis are used in a longitudinal framework (Garfield, 1994) in order to analyze and track the evolution of a research field along consecutive time periods. One of the first longitudinal studies was that carried

[1] The Fuzzy Sets Theory field was founded by Zadeh in 1965 (for more information see Zadeh, 1965, 2008).