Journal of Informetrics 10 (2016) 750–761

Contents lists available at ScienceDirect

Journal of Informetrics

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

# Mapping the evolution of scientific fields based on cross-field authors

Xiaoling Sun*, Kun Ding, Yuan Lin

*Institute of Science of Science and S.&T. Management, Dalian University of Technology, China*

ARTICLE INFO

*Article history:*
Received 24 December 2015
Received in revised form 23 April 2016
Accepted 23 April 2016

*Keywords:*
Scientific evolution
Network analysis
Interdisciplinary research

ABSTRACT

Mapping the evolution of scientific fields has drawn much attention in recent years. Researchers have proposed various methods to describe, explain and predict different aspects of science. Network-based analysis has been widely used for knowledge networks, in order to track the changes of research topics and the spread of scientific ideas. Here we propose a novel approach for mapping the science from the perspective of cross-field authors. Computer science is selected based on its interdisciplinary applications. We build a scientific network consisting of computer science conferences as nodes, and two conferences are linked if there exist authors that published papers on both conferences. The scientific fields are identified by community detection algorithm. The results suggest the proposed method based on author overlaps across fields are effective in mapping the science.



## 1. Introduction

Understanding the evolution of scientific fields is important for planning toward positive science policy and societal impact. Researchers have made great efforts to study different aspects of science in recent years (Börner, Glänzel, Scharnhorst, & den Besselaar, 2011; Börner & Scharnhorst, 2009; Scharnhorst, Börner, & Besselaar, 2012; Shi, Foster, & Evans, 2015). The studies of scientific evolution are based broadly on two philosophies (Sun, Kaur, Milojević, Flammini, & Menczer, 2013): the *cognitive* view and the *social* view. In the cognitive view, researchers put more emphasize on the importance of shared knowledge (Chavalarias & Cointet, 2013; Milojević, 2015; Yun, Kim, & Jeong, 2015). In the social view, studies offered qualitative descriptions of science evolution as stages of social group formation (Crane, 1972; Wagner, 2008).

Network-based analysis has been a popular way to study the evolution of science. Newman took an important step towards applying network ideas to study scientific collaboration networks (Newman, 2001). Later on, network analysis have been applied to various knowledge networks (Chang, Huang, & Lin, 2015), such as collaboration networks (Sun, Lin, Xu, & Ding, 2015), citation networks (Boyack, Klavans, & Börner, 2005; Chang et al., 2015).

In the context of scientific map, different level of maps has been proposed to reveal the scientific evolution. At the paper level, science mapping studies including document co-citation analyses (Boyack & Klavans, 2014; Chen, Ibekwe-SanJuan, & Hou, 2010) and word co-occurrence analyses (Chavalarias & Cointet, 2013). At the journal level, journal citations have been largely used as a way for mapping science (Boyack et al., 2005; Leydesdorff & Rafols, 2012; Leydesdorff, Rafols, & Chen,

* Corresponding author at: International Convention Center C513, Dalian University of Technology, Dalian 116024, China. Tel.: +86 15842488360.
*E-mail address:* xlsun@dlut.edu.cn (X. Sun).

http://dx.doi.org/10.1016/j.joi.2016.04.016
1751-1577/

2013). Web-of-Science categories are also employed to generate a global map of science (Leydesdorff, Carley, & Rafols, 2013). Different levels of maps have different characteristics, and could reveal scientific evolution from different aspects.

In this paper, we propose a novel method to map the evolution of scientific fields via the overlapping authors across fields. In recent years, interdisciplinary research is becoming more and more popular, as it has been recognized for the ability to encourage and accelerate innovative development. Many researchers evolve in multiple disciplines, changing research interests over time. Therefore, cross-field authors could capture the relationships between disciplines. We select computer science area to study the scientific evolution based on its interdisciplinary applications, as well as its fast-growing speed with many subareas. Here we build a scientific network consisting of computer science conferences as nodes, and two conferences are related and linked if there exist authors that published papers on both conferences. The scientific fields are identified by the community detection algorithm. In contrast to the previous approaches based on word co-occurrence networks and co-citation networks, here we used a methodology solely based on the overlapping authors across fields. The results suggest that cross-field research indeed play an important role in shaping the evolution of scientific fields.

The rest of this paper is organized as follows: first, a brief background of the study is given in Section 2; Second, Section 3 introduces the proposed method including how to build the conference network and identify scientific fields in computer science; Next, Section 4 presents the results and shows the performance of our method; then we conclude in Section 5.

## 2. Related work

There are basically three levels of scientific maps: paper-level, journal-level and category-level scientific maps (Börner, Chen, & Boyack, 2003; Boyack, 2008).

### 2.1. Paper-level maps

For paper-level maps, citation data is widely used to draw scientific maps. Griffith, Small, Stonehill, and Dey (1974) used co-citation method to create the maps of the scientific literature. Klavans and Boyack (2006) presented a method for generating maps directly from the data on the relationships between hundreds of thousands of documents. Chen (2006) designed as a tool CiteSpace for progressive knowledge domain visualization using co-citation data.

Word co-occurrence analysis is another way to map the science at paper-level. Chavalarias and Cointet (2013) proposed a bottom-up reconstruction of the dynamics of scientific fields based on co-word analysis. They also proposed an asymmetric paradigmatic proximity metric between terms which provide insight into hierarchical structure of scientific activity (Chavalarias & Cointet, 2008). Cui et al. (2011) presented TextFlow, a visual topic analysis system to help users explore and understand topic evolutions.

### 2.2. Journal-level maps

Researchers also have made many efforts to study journal-level maps, which provide a lesser resolution than paper-level maps. Leydesdorff and Rafols (2012), Leydesdorff, Rafols, et al. (2013) used scientometric overlay mapping techniques to build map of science based on aggregated journal–journal citation relations. Boyack et al. (2005) built the backbone of science using journal citation data. Rosvall and Bergstrom (2010) proposed a method for significance clustering and mapping change in science. The mapping visualization could be very impressive (Börner, 2010; van Eck & Waltman, 2009).

Boyack, Börner, and Klavans (2009) combined paper-level and journal-level maps, which used the ISI Proceedings database with the Science and Social Science Indexes at the paper level and generated a disciplinary map of 7227 journals and 671 journal clusters to study the structure and evolution of chemistry.

### 2.3. Category-level maps

There are also other researches using category information to build the science map. For example, Herrera, Roberts, and Gulbahce (2010) offered a top-down scheme for categorizing physics literature described through PACS numbers. The ISI subject categories classify journals included in the Science Citation Index (SCI). The aggregated journal–journal citation matrix can be aggregated on the basis of these categories (Leydesdorff & Rafols, 2009).

From the previous research, we could see that citations, terms and category information are commonly used in scientific maps. In computer science area, authors have also drawn maps by leveraging these informations. Fried and Kobourov (2014) proposed a map of computer science using the words and phrases co-occurrence in the paper titles. Pham, Klamma, and Jarke (2011) combined DBLP and CiteSeerX databases, and constructed knowledge network based on journal/conference paper citations. These maps reveal the evolution of scientific fields at different resolutions. Here we provide an alternative and innovative way from the perspective of authors. The idea is inspired by the observation that, there are more and more researchers evolving in multiple disciplines and changing research interests over time. A full fabric of science is created because many authors move from discipline to discipline (Klavans & Boyack, 2010). Therefore, we use overlapping authors to capture the relationships between conferences.