Journal of Informetrics 11 (2017) 564–582

Contents lists available at ScienceDirect

Journal of Informetrics

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

Genealogical index: A metric to analyze advisor–advisee relationships

Luciano Rossi, Igor L. Freire, Jesús P. Mena-Chalco*

Center for Mathematics, Computation and Cognition, Federal University of ABC, Santo André, Brazil

ARTICLE INFO

Article history:
Received 14 October 2016
Received in revised form 1 April 2017
Accepted 2 April 2017

Keywords:
Genealogical index
Academic genealogy
Advisor–advisee relationship
Genealogical graph

ABSTRACT

Academic genealogy can be defined as the study of intellectual heritage that is undertaken through the relationship between a professor (advisor/mentor) and student (advisee) and on the basis of these ties, it establishes a social framework that is generally represented by an academic genealogy graph. Obtaining relevant knowledge of academic genealogy graphs makes it possible to analyse the academic training of scientific communities, and discover ancestors or forbears who had special skills and talents. The use of metrics for characterizing this kind of graph is an active form of knowledge extraction. In this paper, we set out a formal definition of a metric called 'genealogical index', which can be used to assess how far researchers have affected advisor–advisee relationships. This metric is based on the bibliometrics h-index and its definition can be broadened to measure the effect of researchers on several generations of scientists. A case study is employed that includes an academic genealogy graph consisting of more than 190,000 Ph.D.s registered in the Mathematics Genealogy Project. Additionally, we compare the genealogical indices obtained from both the Fields Medal and Wolf Prize winners, and found that the latter has had a greater impact than the former.

1. Introduction

Scientific evolution traces economic and social development that is carried out through scientific research, technological innovation and patenting. The University is the natural environment for scientific evolution since it involves academic mentoring that seeks to develop high quality human resources. The existence of this phenomenon is the driving-force behind the search for variables/metrics that can measure it. In this paper, we introduce a metric that is designed to assess the effects of academic mentoring on the achievements of the scientific community by means of academic genealogy graphs. The reason for applying this metric is to answer a fundamental research question: does academic supervision affect the performance of scientists?

Academic genealogy (AG) can be defined as a quantitative study of intellectual inheritance that has been perpetuated by generations of researchers through academic advising by mentors to their students (Sugimoto, 2014). The AG allows an analysis to be conducted about the dissemination of scientific knowledge and the progress made by scientific communities.

* Corresponding author.
E-mail addresses: luciano.rossi@ufabc.edu.br (L. Rossi), igor.freire@ufabc.edu.br (I.L. Freire), jesus.mena@ufabc.edu.br (J.P. Mena-Chalco).

http://dx.doi.org/10.1016/j.joi.2017.04.001
1751-1577/

The varied work of academic advisors influences the training programs of researchers and encourages future generations of researchers to continue their activities (Malmgren, Ottino, & Amaral, 2010). The AG provides a means of assessing and quantitatively analyzing the way these training schemes are conducted (Sugimoto, Ni, Russell, & Bychowski, 2011).

The analysis of academic and scientific communities has attracted a good deal of attention among researchers. Particular importance has been attached to the classification and identification of researchers involved in several knowledge areas. These include the adoption and improvement of quantitative metrics that support this analysis, as well as the study of the acquisition of scientific knowledge. The studies in this area mostly entail carrying out an in-depth analysis of academic publications, but there are a few studies devoted to evaluating research from the perspective of AG. In most cases, these studies are used to identify the forebears and descendants of an individual researcher (i.e., to compile an egocentric genealogy) simply to honor them (Kobayashi, 2015; van der Kruit, 2015).

A number of studies have been carried out to characterize the AG with the aim of analyzing specific knowledge areas, such as Neuroscience (David & Hayden, 2012), Chemistry (Andraos, 2005), Mathematics (Chang, 2011; Gargiulo, Caen, Lambiotte, & Carletti, 2016; Malmgren et al., 2010), Physiology (Bennett & Lowe, 2005; Jackson, 2011), Meteorology (Hart & Cossuth, 2013), Primatology (Kelley & Sussman, 2007), Bibliometry and Information Science (Russell & Sugimoto, 2009), and Protozoology (Elias, Floeter-Winter, & Mena-Chalco, 2016), and many others.

It should be noted that these studies converge insofar as they share the following common objectives: (i) prospecting, structuring and storing data about academic genealogy (i.e., historical records) (Andraos, 2005; Bennett & Lowe, 2005; Chang, 2011; Hart & Cossuth, 2013), (ii) characterizing knowledge areas and/or disciplines (egocentric and honorific genealogy), by analyzing genealogical frameworks using basic descriptive statistics (David & Hayden, 2012; Elias et al., 2016; Gargiulo et al., 2016; Malmgren et al., 2010; Russell & Sugimoto, 2009), and (iii) making the information available (i.e., publishing it) to members of the community and interested parties (David & Hayden, 2012; Hart & Cossuth, 2013). However, the majority of these studies does not focus their analysis in topological structures neither use metrics to find out key academic groups or individuals. Among the few studies that use metrics to characterize the topology of a network, it is worth highlighting the work of David and Hayden (2012) who employed fecundity metrics to characterize the neuroscientific community, Rossi and Mena-Chalco (2015) whose work examined the basis of the genealogical index and its main applications, and Lü, Zhou, Zhang, and Stanley (2016) where the h-index was used for the characterization of scientific networks.

There have also been initiatives that rely on web applications to document and share the academic genealogy of researchers across several fields, such as the following: The Mathematics Genealogy Project,[1] The Neurotree Project,[2] The Academic Family Tree,[3] and the Academic Genealogy Wiki.[4] These projects, which register and document the names of individuals, make it possible to study the influence of generations of researchers on the academic world, through academic mentoring.

According to Sugimoto (2014), the academic genealogy is mainly used by researchers interested in discovering and describing their own origins. These studies have tended to be neglected by those who are studying a branch of science from a historical, philosophical, sociological or scientific perspective. The real importance of academic genealogy is that it provides quantitative and qualitative inputs that can assist in measuring interactions at different levels. The academic genealogy allows science to be analyzed from the standpoint of a transmission of scientific knowledge through generations of researchers.

The academic genealogy can be analyzed by topological metrics that represent different features and provide useful information on the training of academic communities as well as by revealing the names of key researchers who have made a significant contribution to areas of knowledge. In this paper, a topological metric is used, called genealogical index, which can be applied to measure the academic influence of researchers by means of academic genealogy graphs. We use the term 'rank' to define the limited coverage of the generations that must be included for the calculation of the genealogical index.

This approach makes a formal adjustment to the bibliometric h-index in the academic genealogy project. This is a system to measure the influence of a researcher quantitatively from the perspective of human resources training, rather than simply concentrating on publications, citations, coauthorships, or research projects.

Finally, it should be emphasized that this work is aligned with the epistemology of big data analysis (Big Data), in the form of data-driven science and seeks to discover knowledge about universally-accepted scientific theories, as described by Frické (2014).

2. Method

2.1. Academic genealogy graphs

The advisor–advisee relationships can be represented in the form of a graph that can be useful for the study of academic genealogy. The structure used is called a genealogical tree. In fact, the structures built from academic genealogy data cannot

[1] http://genealogy.math.ndsu.nodak.edu, last accessed on January 15, 2017.
[2] http://neurotree.org/neurotree, last accessed on January 18, 2017.
[3] http://academictree.org/physics, last accessed on January 20, 2017.
[4] http://phdtree.org, last accessed on February 02, 2017.