Constructing bibliometric networks: A comparison between full and fractional counting ☆

Antonio Perianes-Rodriguez [a], Ludo Waltman [b], Nees Jan van Eck [b,*]

[a] *SCImago Research Group, Departamento de Biblioteconomia y Documentacion, Universidad Carlos III, Getafe, Madrid, Spain*
[b] *Centre for Science and Technology Studies, Leiden University, Leiden, The Netherlands*

A B S T R A C T

The analysis of bibliometric networks, such as co-authorship, bibliographic coupling, and co-citation networks, has received a considerable amount of attention. Much less attention has been paid to the construction of these networks. We point out that different approaches can be taken to construct a bibliometric network. Normally the full counting approach is used, but we propose an alternative fractional counting approach. The basic idea of the fractional counting approach is that each action, such as co-authoring or citing a publication, should have equal weight, regardless of for instance the number of authors, citations, or references of a publication. We present two empirical analyses in which the full and fractional counting approaches yield very different results. These analyses deal with co-authorship networks of universities and bibliographic coupling networks of journals. Based on theoretical considerations and on the empirical analyses, we conclude that for many purposes the fractional counting approach is preferable over the full counting one.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

The study of bibliometric networks, such as co-authorship, bibliographic coupling, and co-citation networks, has a long history in the field of bibliometrics, with early work dating back to the 1960s and 1970s (e.g., De Solla Price, 1965; Kessler, 1963; Small, 1973). Many different methods for analyzing and visualizing bibliometric networks have been studied by bibliometricians (e.g., Börner, Chen, & Boyack, 2003; Milojević, 2014; Van Eck & Waltman, 2014; Zhao & Strotmann, 2015). However, before bibliometric networks can be analyzed and visualized, they first need to be constructed. The construction of bibliometric networks has received remarkably little attention in the literature (for important exceptions, see Batagelj & Cerinšek, 2013; Park, Yoon, & Leydesdorff, 2016). It seems that the construction of bibliometric networks is typically seen as a more or less trivial step that does not need any special consideration. In this paper, we argue that this step is far from trivial. We point out that different approaches can be taken to construct bibliometric networks. Our aim is to draw attention to the existence of different approaches for constructing bibliometric networks, to clarify the conceptual differences between these approaches, and to show that these approaches may yield very different results.

A well-known problem in the field of bibliometrics is the issue of assigning co-authored publications to individual authors. For instance, when a publication is co-authored by three researchers, how should the publication be counted for each

individual researcher? In the context of the calculation of bibliometric indicators, many different approaches have been proposed to this problem (for overviews, see Gauffriau, Larsen, Maye, Roulin-Perriard, & Von Ins, 2007; Waltman, 2016). The most popular approaches are the full counting method (also known as the whole counting method) and the fractional counting method (e.g., Aksnes, Schneider, & Gunnarsson, 2012; Waltman & Van Eck, 2015). In the case of the full counting method, a publication co-authored by three researchers is assigned to each researcher with a full weight of one. On the other hand, in the case of the fractional counting method, the publication is assigned to each researcher with a fractional weight of 1/3.

In this paper, we show how the distinction between full and fractional counting, which has been studied extensively in the context of the calculation of bibliometric indicators, can be translated to the context of the construction of bibliometric networks. Consider for instance the construction of a co-authorship network. Suppose researcher X has co-authored a publication with five other researchers. In the conventional approach to the construction of bibliometric networks, this yields five co-authorship links with a weight of one for researcher X. We refer to this approach as the full counting method. An alternative approach is to assign a weight of 1/5 to each of the five co-authorship links. In this approach, which we refer to as the fractional counting method, the total weight of the co-authorship links that a researcher obtains because of co-authoring a publication equals one. This total weight of one is distributed equally over the individual co-authorship links.

To construct bibliometric networks, researchers have traditionally used the full counting method. To the best of our knowledge, the fractional counting method has hardly been used in the literature (for the only exception that we are aware of, see Newman, 2001c), although some related ideas have been proposed (Batagelj & Cerinšek, 2013; Cerinšek & Batagelj, 2015; Park et al., 2016; Persson, 1994, 2010).[1] In this paper, we carefully define the full and fractional counting methods. Our focus is on three popular types of bibliometric networks, namely co-authorship, bibliographic coupling, and co-citation networks, but our ideas extend to other types of bibliometric networks as well. We also provide two examples of situations in which the choice between the full and fractional counting methods makes a big difference. One example is about co-authorship networks of universities. The other example deals with bibliographic coupling networks of journals. In both examples, we argue that the fractional counting method is preferable over the full counting method.

We note that the full and fractional counting methods are both available in the VOSviewer software (www.vosviewer.com; Van Eck & Waltman, 2010, 2014) for constructing and visualizing bibliometric networks. The VOSviewer software can be used to construct bibliometric networks based on data downloaded from bibliographic databases such as Web of Science and Scopus. The software requests the user to choose between the use of the full and the fractional counting method. The information provided in this paper should help VOSviewer users in choosing the most appropriate counting method for their analyses.

This paper is organized as follows. Formal definitions of the full and fractional counting methods in the context of the construction of bibliometric networks are provided in Section 2. An empirical comparison between the two counting methods is reported in Section 3. We present our conclusions in Section 4.

2. Constructing bibliometric networks

In this section, we provide a detailed discussion of the full and fractional counting methods for constructing bibliometric networks. We first discuss in general terms the difference between full and fractional counting. We then focus specifically on co-authorship networks, followed by bibliographic coupling and co-citation networks. We focus on these three types of bibliometric networks because they seem to be the types of bibliometric networks that receive most attention in the literature. However, we emphasize that our ideas apply to other types of bibliometric networks as well. For an overview of the literature on different types of bibliometric networks, we refer to Van Eck and Waltman (2014, Subsection 2.1).

2.1. Full counting vs. fractional counting

In the context of the calculation of bibliometric indicators, the concepts of a publication and a co-author play a key role in the distinction between full and fractional counting. Full counting means that a co-authored publication is counted with a full weight of one for each co-author, which implies that the overall weight of a publication is equal to the number of authors of the publication. Fractional counting means that a co-authored publication is assigned fractionally to each of the co-authors, with the overall weight of the publication being equal to one. Hence, in the case of fractional counting, each publication has the same overall weight.

In the context of the construction of bibliometric networks, a similar distinction between full and fractional counting can be made. However, in order to do so, the concepts of a publication and a co-author need to be replaced by appropriate network-related concepts. We replace the concept of a publication by the concept of an action. The concept of a co-author is replaced by the concept of a link. For specific types of bibliometric networks, the concepts of an action and a link can be given a more concrete interpretation. For instance, in the case of a co-authorship network, co-authoring a publication with other

---

[1] Small and Sweeney (1985) also use a fractional counting approach in the context of the construction of a bibliometric network. However, they do not use fractional counting in the actual construction of the network, but instead they use fractional counting to select the publications to be included in the network.