Journal of Informetrics 9 (2015) 809–825

Contents lists available at ScienceDirect

Journal of Informetrics

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

How to become an important player in scientific collaboration networks?

Ashkan Ebadi[a,*], Andrea Schiffauerova[a,b]

[a] Concordia Institute for Information Systems Engineering (CIISE), Concordia University, 1515 Ste-Catherine Street West, Montreal, QC, Canada H3G 2W1

[b] Department of Engineering Systems and Management, Masdar Institute of Science and Technology, Masdar City, Abu Dhabi, United Arab Emirates

ARTICLE INFO

Article history:
Received 13 January 2015
Received in revised form 30 July 2015
Accepted 4 August 2015
Available online 3 September 2015

Keywords:
Network structure
Collaboration
Statistical analysis
NSERC
Canada

ABSTRACT

Scientific collaboration is one of the important drivers of research progress that supports researchers in the generation of novel ideas. Collaboration networks and their impact on scientific activities thus already attracted some attention in the research community, but no work so far studied possible factors which can influence the network positions of the researchers at the individual level. The objective of this paper is to investigate various characteristics and roles of the researchers occupying important positions in the collaboration network. For this purpose, we focus on the collaboration network among Canadian researchers during the period of 1996 to 2010 and employ multiple regression models to estimate the impact on network structure variables. Results highlight the crucial role of past productivity of the researchers along with their available funding in determining and improving their position in the co-authorship network. It is shown that researchers who have great influence on their local community do not necessarily publish high quality works. We also find that highly productive researchers not only have more important connections but also play a critical role in connecting other researchers. Moreover, although mid-career scientists tend to collaborate more in knit groups and on average have higher influence on their local community, our results specifically highlight the important role of young researchers who occupy mediatory positions in the network which enable them to connect different communities and fuel information transmission through the network.

1. Introduction

Recent progress in information technologies has cut the world-wide distances enabling researchers to get in contact easier. Hence, nowadays no specific border can be defined for scientific activities in a way that researchers have formed a global community aiming to advance the level of knowledge. Concurrently, the nature of the science has become more complex and inter-disciplinary which encourages scientists to be more collaborative in order to increase their scientific productivity, to get access to new knowledge and financial resources, *etc.* Katz and Martin (1997) define scientific collaboration as the process through which researchers with a common goal work together to produce new scientific knowledge. The importance of collaborative research is now acknowledged in scientific communities (Brad Wray, 2006). Through collaboration researchers get access to an often informal network of scientists that may facilitate knowledge and skill diffusion (Tijssen, van Leeuwen,

* Corresponding author. Tel.: +1 352 745 4468.
E-mail address: a_ebad@ciise.concordia.ca (A. Ebadi).

http://dx.doi.org/10.1016/j.joi.2015.08.002
1751-1577/

& Korevaar, 1996; Tijssen, 2004). Although it is not easy to quantify scientific collaboration, co-authorship has become the standard way of measuring collaboration since it is considered as a better sign of mutual scientific activity (De Solla Price, 1963; Ubfal & Maffioli, 2011). Co-authorship networks, as one of the main forms of scientific collaboration (Abbasi, Altmann, & Hwang, 2010), evolve over time. This evolution might reflect the growth/decay of a research subject, community or even a scientific field (Huang, Zhuang, Li, & Giles, 2008). This evolution and changes can be also seen at the nodes level (*i.e.* researchers in the co-authorship networks) where researchers' positions and their importance within their community and/or the whole collaboration network might also change over time. Position of a node in a network depends both on its direct and indirect connections with the other nodes (Mattsson & Johanson, 1992).

Due to the growing large number of researchers and their co-authorship links, scientific collaboration networks are among the complex ones (Abbasi, Hossain, & Leydesdorff, 2012). Role of a researcher (node) in a network can bring some advantages to the researcher (*e.g.* better access to knowledge sources, political factors, awareness of potential projects, *etc.*), and the surrounding community. This becomes more interesting as one notes that the roles of nodes in a network might change over time (Abbasi *et al.*, 2012). Barabási and Albert (1999) showed that a new node in a network will be linked to the other nodes with large number of connections (higher degree centrality) with a higher probability. This indicates the importance of the highly connected nodes in a network. This is also confirmed by Moody (2004) who showed that authors who are new in a scientific network are more likely to get connected to highly reputable authors with many collaborators thus making the surrounding community of the reputable researcher denser. On the other hand, there exist studies indicating that getting connected to high performing nodes (researchers, organizations, *etc.*) can affect the performance of the connecting node. For example, Mote (2005) analyzed the impact of inter-organizational complexity on the research output of 20 projects in national labs and found that groups that were connected to prolific organizations also showed higher performance. All of this highlights the importance of structural collaboration network positions in scientific and technological activities. Thus this paper specifically focuses on researchers' roles in their collaboration networks and assesses the impact of influencing factors.

The remainder of the paper proceeds as follows: Section 2 discusses the gaps in the literature and objectives of the research; Section 3 presents the data, methodology and the models; Section 4 presents the empirical results and interpretations; Section 5 concludes; and Section 6 discusses the limitations.

2. Research motivation and objectives

Scientific collaboration is more and more attracting the attention of researchers as the science is evolving toward a more complex and highly inter-disciplinary nature. The continuous growing trend of collaboration in terms of the number of co-authored papers has been widely confirmed in bibliometric studies (*e.g.* Grossman, 2002; Cronin, 2005). In addition, it has been studied in a vast number of different disciplines such as computer science, sociology, research policy, and philosophy (Sonnenwald, 2007), focusing on different aspects of collaboration. In a quite different study, Jiang (2008) presented an algorithm for detecting active researchers in scientific communities which is based on an abstract definition of collaboration cost and number of interactions between researchers. Their assumption of considering active researchers to be more attractive for collaboration partially confirms the importance of collaboration in scientific communities. Abbasi *et al.* (2010) used the three measures of researchers' collaboration network structure, number of collaborations and productivity of co-authors to quantify the collaboration activities of researchers. They proposed two indices, namely researchers collaboration (RC-index) and community collaboration (CC-index), which can be also used for detecting the best partners for a research project.

It is argued that the structure of the network can affect the collaboration patterns and scientific output (Ebadi & Schiffauerova, 2015a). Several studies assessed the impact of collaboration patterns and network positions on scientific activities and performance of researchers (*e.g.* Eslami, Ebadi, & Schiffauerova, 2013; Beaudry & Allaoui, 2012; Abbasi, Altmann, & Hossain, 2011) as well as their level of funding (*e.g.* Ebadi & Schiffauerova, 2015b) and found a positive relation in most of the cases. For example, Abbasi *et al.* (2011) focused on the impact of four network indicators (*i.e.* degree centrality, closeness centrality, betweenness centrality and eigenvector centrality) along with some other factors on the citation-based performance of researchers who were active in information systems field and found a positive relation between eigenvector and degree centralities and the performance of the target scholars. In another study, Abbasi *et al.* (2012) analyzed the impact of possessing various roles in co-authorship network in observing new researchers for collaboration. Their results suggest the higher importance of betweenness centrality as well as the degree centrality of an existing researcher in attracting new entrants. In addition, there are a number of studies that evaluated the impact of several influencing factors (*e.g.* funding, gender, scientific fields) on scientific collaboration and its patterns (*e.g.* Bozeman & Corley, 2004; Adams, Black, Clemmons, & Stephan, 2005; Gulbrandsen & Smeby, 2005; Rosenzweig *et al.*, 2008; Defazio, Lockett, & Wright, 2009). For more information, see the critical review of the literature by Ebadi and Schiffauerova (2013).

Although there are some studies that confirms the importance of structural network positions and relationships in business and scientific communities (*e.g.* Håkansson & Ford, 2002), to the best of our knowledge no study shows how one can possess such network positions by analyzing the impact of influencing factors on different network positions in scientific collaboration networks. In other words, network structure variables have been so far considered at the right hand-side of the equations, estimating their impact on various scientific activities or performance of the researchers, *etc.*

Apart from performance related factors and financial power, we hypothesize that career age and affiliation type of a researcher might help him/her to possess more influential network positions. We define an *influential researcher* as a highly