EasyChair Preprint

№ 8021

Scientific Communication Networks: Tracking Victorian and Twenty-First Century Communication with Social Network Analysis

Bertrum MacDonald, Anatoliy Gruzd and Valerie Collins

May 22, 2022

Work-in-Progress Paper

Title: Scientific Communication Networks: Tracking Victorian and Twenty-First Century Communication with Social Network Analysis

ABSTRACT

Background

Communication is a deeply rooted characteristic of science. Beginning with the launch of scientific periodicals in London and Paris in the seventeenth century, scientists have given a high priority to the exchange of information and widely disclosing the results of their research. This compelling motivation led William Garvey to declare that communication is the "essence of science" (Garvey, 1979). While information sharing is central to scientific pursuits, much about the networks in which scientific information circulates remains to be explored by communication scholars, particularly historical systems. In a 2004 essay, entitled "Knowledge in transit," James Secord, Director of the Darwin Correspondence Project, stated that "the narrative frameworks used by historians of science need to come to terms with diversity of understanding science as a form of communication" (Secord, 2004, p. 654). Today's movement of massive quantities of information around the globe, made possible by advances in networking and computing technologies, has triggered the development of powerful digital tools and methodologies to probe communication networks in new ways. These techniques open up the possibility of studying both historical and current-day communication networks to determine similarities and differences in information exchange activities among scientists. While the large volume of information flowing through today's communication networks is receiving extensive study (e.g., Cvitanovic et al., 2017; Haythornthwaite et al., 2018), much less attention has been focused on gaining an understanding of the substantial movement of scientific information through the postal networks of the nineteenth and twentieth centuries. The difficulty in working with print-based records has inhibited detailed analyses of historical networks. However, the growing body of historical records being made available in digitized format recently offers the potential of applying digital tools to draw out characteristics of communication networks in the print-dominated periods, which can provide evidence to compare with current day networks.

Objective

Communication fosters relationships among scientists, allows for the exchange of ideas and research findings, and creates avenues for scientists to seek and obtain information by formal and informal means. The act of transmitting scientific findings directly to other individuals creates ties among the individuals, regardless of the transmission medium. Letters, tweets, email messages, and posts to a variety of other platforms are all means by which networks of scientific exchange can be created. Scientific work is not placeless; who is involved, where the activity occurs, when research is conducted, and how information is shared matters (Livingston, 2003; 2005). Thus, the communication of scientific information needs to be examined in more than local

settings of time and place, even though such study originates in particular places. This paper begins with applying social network analysis methods to gain an understanding of communication patterns in the correspondence records of Edwin Gilpin, a Victorian era government-based scientist employed in a geological and mining engineering position (Duggan & MacDonald, 2008; 2011), which is then compared to the Twitter activity of selected twenty-first scientists in similar employment. The objective is to demonstrate that social network analysis methodology can uncover prominent characteristics of communication networks in both periods.

Methods

As an example of many thousands of letters circulating within scientific networks in the Victorian period (for example, see Goldstein, 1994; 2008, MacDonald, 2012), the surviving correspondence (numbering more than 1,300 letters) of Edwin Gilpin Jr. was read and coded for network analysis. A database of three types of metadata was created, namely information about the author and recipient of each letter, the subject content of each letter, and the connections or links between individuals identified in the letters. Visualizations of the coded data were prepared with ORA software (Organizational Risk Analyzer) developed by the Center for Computational Analysis of Social and Organizational Systems at Carnegie Mellon University. This network analysis and visualization software provides techniques for studying the "design structure of an organization…[e.g.] the relationship among its personnel, knowledge resources, and tasks entities" (Carley et al., 2011).

For comparison of the historical case with current day information sharing, Twitter was selected because the communication traffic is publicly accessible for analytical purposes. Twitter serves much like a letter of earlier decades since information about the senders, recipients (followers), subjects, other connections, and attachments are all available. While differences exist between Twitter posts and written letters — the openly public nature of tweets contrasts to the limited access of most letters to sender and recipient, and the brief message length of tweets compared to letters (many letters were brief, however) — information was (is) transmitted between scientists in both media. A sample of scientists communicating via Twitter was selected, via the online application FollowerWonk, to reflect Gilpin's field of activity, i.e., scientists who work in earth science subjects and work within a government-based context. Netlytic, a tool designed for analyzing social networks (netlytic.org) and employed in studies of online conversations, was used to capture Twitter traffic of the selected scientists over the period of a month. Based on the search terms used, Netlytic only collects and displays tweets directed at another user. Thus, only tweets that are part of a conversation or clearly labeled as directed messages are collected and analyzed. General tweets from a user, without an @ tag, are not comparable to a letter, as they are similar to announcements or diary entries that happen to be public. The collected tweets were subjected to content analysis following a framework for science discourse on Twitter developed by Holmberg and Thelwall (2014), and the findings of Gruzd and Goertzen (2013) grounded in the

Uses and Gratifications theory regarding reasons that some social scientists use social media to disseminate information, socialize, and gather information.

Results

For this research, we use a mixed-method approach that relies on both network and content analysis. The network analysis of scholarly communication networks helped us examine a variety of relationships among numerous actors in the networks, and the content analysis of the correspondence offers insights about information exchanges between local, national, and international nodes in the networks. While the analysis of the data is currently being finalized, below are some preliminary observations.

With regard to the communication activities of the selected current-day scientists, the methodology applied in this study generated datasets that are appropriate for assessing the similarity and differences with communication patterns in the Victorian period. The coding framework developed for current-day communication via Twitter can be applied to the correspondence records of earlier decades to identify similarities and differences in information sharing through the networks in each period. The points of comparison show that geography and distance may not be markedly different, while timing of information transfer is shaped by the communication technology used in the network. Moreover, network analysis methods can reveal characteristics of historical periods in a manner similar to communication through digital platforms today (see, for example, Figure 1). These analytical techniques will aid in understanding the importance of communication networks that grew over time.

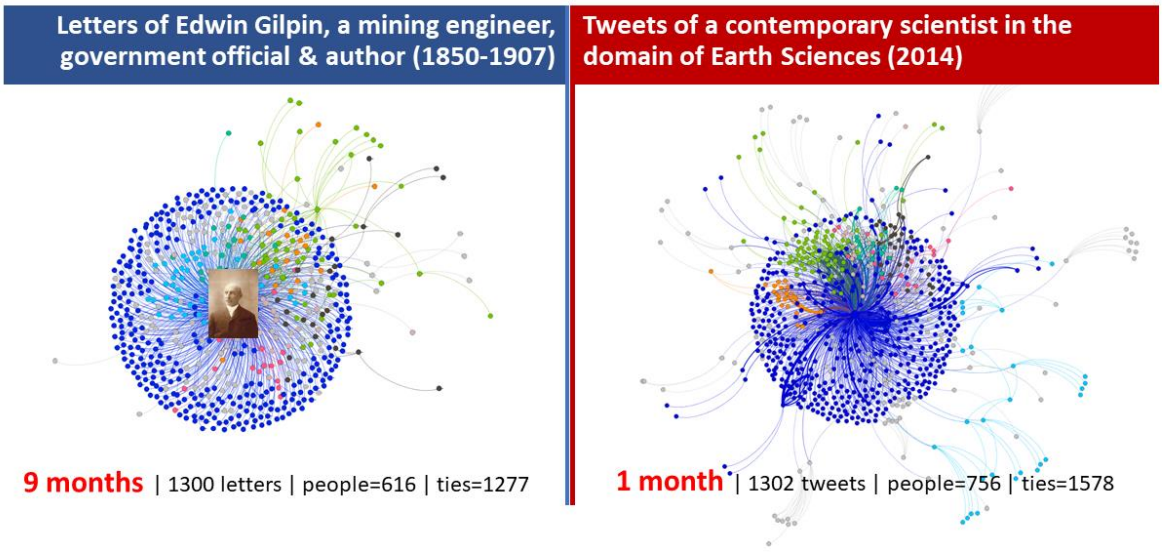

Figure 1. Scholarly Communication: Then and Now

In "Landscapes of Knowledge," D. N. Livingston stated: "scientific knowledge is a geographical phenomenon. It is acquired in specific sites, it circulates from location to location, it transforms the world" (Livingston, 2010, p. 18). That information networks facilitate the circulation of scientific knowledge is well established. How

networks operated historically is not yet well understood, however. This paper shows that the tools developed to characterize and understand the complexities of digitally-enabled networks of the present day can be applied to networks of earlier periods. While information exchanges were affected by the technologies of the time, determining the attributes of earlier scientific networks can identify communication patterns that persist regardless of platform and those that do not.

References

Carley, K. M., Reminga, J., Storrick, J., & Columbus, D. (2011, June). *ORA user's guide 2011*. Pittsburgh: Institute for Software Research, School of Computer Science, Carnegie Mellon University.

Cvitanovic, C., Cunningham, R., Dowd, A.-M., Howden, S. M., & van Putten, E. I. (2017). Using social network analysis to monitor and assess the effectiveness of knowledge brokers at connecting scientists and decision-makers: An Australian case study. *Environmental Policy and Governance, 27*(3), 256–269. https://doi.org/10.1002/eet.1752

Duggan, L. J., & MacDonald, B. H. (2008). Reading scientific and technical literature: The case of the nineteenth-century Nova Scotian mining engineer, Edwin Gilpin. *Journal of the Royal Nova Scotia Historical Society, 11,* 146-167.

Duggan, L. J., & MacDonald, B. H. (2011). Reading science: Evidence from the career of Edwin Gilpin, mining engineer. In S. Towhead, & W. R. Owens (Eds.). *The history of reading. Volume 1: International perspectives, c. 1500-1990* (pp. 117-132). London: Palgrave Macmillan. https://epdf.pub/the-history-of-reading-volume-1-international-perspectives-c-1500-1990.html

Garvey, W. D. (1979). *Communication: The essence of science*. Oxford: Pergamon Press.

Goldstein, D. (1994). "Yours for science": The Smithsonian Institution's correspondents and the shape of scientific community in nineteenth-century America. *Isis, 85*, 573-599. https://www.jstor.org/stable/235279

Goldstein, D. (2008). The knowledge trade and the expansion of scientific community in post-Civil War America. *Isis, 99*, 519-546. https://www.jstor.org/stable/10.1086/591711

Gruzd, A., & Goertzen, M. (2013). Wired academia: Why social science scholars are using social media. The 46th Hawaii International Conference on System Sciences (HICSS): 3332-3341, doi: 10.1109/HICSS.2013.614

Haythornthwaite, C., Kumar, P., Gruzd, A., Gilbert, S., Esteve Del Valle, M., & Paulin, D. (2018). Learning in the wild: Coding for learning and practice on Reddit. *Learning, Media and Technology* 43(3): 219-235. https://doi.org/10.1080/17439884.2018.1498356

Livingstone, D. N. (2003). *Putting science in its place: Geographies of scientific knowledge*. Chicago: University of Chicago Press.

Livingstone, D. N. (2005). Science, text and space: Thoughts on the geography of reading, *Transactions of the Institute of British Geographers, n.s. 30*(4), 391-401. https://doi.org/10.1111/j.1475-5661.2005.00179.x

Livingstone, D. N. (2010). Landscapes of knowledge. In P. Meusburger, et al. (Eds.). *Geographies of science, Knowledge and Space 3* (pp. 3-22). Berlin: Springer Science+Business Media B.V.

MacDonald, B. H. (2012). Crossing borders. The Smithsonian Institution and nineteenth-century diffusion of scientific information between the United States and Canada. In R. D. Apple, G J. Downey, & S. L. Vaughn (Eds.). *Science in print. Essays on the history of science and the culture of print* (pp. 87-106). Madison, WI: The University of Wisconsin Press.

Secord, J. (2004). Knowledge in transit. *Isis, 95*, 654-672.