# Applying social network analysis to explore Twitter diffusion patterns

Juan Pablo Alperin[1,2]
*juan@alperin.ca*
[1] School of Publishing, Simon Fraser University, Vancouver (Canada)
[2] Public Knowledge Project

Stefanie Haustein[3,4]
*stefanie.haustein@uottawa.ca*
[4] School of Information Studies, University of Ottawa, Ottawa (Canada)
[4] Observatoire des Sciences et des Technologies (OST), Centre Interuniversitaire de Recherche sur la Science et la Technologie (CIRST), Université du Québec à Montréal, Montréal (Canada)

Social media have changed communication by enabling direct interactions between those who were traditionally referred to as the sender and receiver of information. Although scholarly communication still heavily relies on authors publishing in journal articles, social media platforms enable public discussions of scientific knowledge in real-time, when readers share their opinion online. Platforms like Twitter, which are used by both researchers and the general public, provide a means for scientists to engage directly with interested members of the public and vice versa, which is why tweets to scientific papers have been one of the most promising sources of altmetric data to capture societal impact. However, studies have shown that tweets to scientific papers are created mostly by academics (Alperin, 2015; Tsou, Bowman, Ghazinejad, & Sugimoto, 2015), implying that public engagement is rare. Moreover, a significant amount of tweets are automatically generated by Twitter bots (Haustein et al., 2016), which further questions the ability of the number of tweets to reflect impact outside the scholarly community.

This paper focuses on a set of seven highly tweeted articles published in the open access journal *BMC Biology* in 2014 as a case study to explore whether social network analysis can be used to differentiate between papers that diffused among well-connected users and those that reached more disparate groups and individuals. The study relies on Altmetric data to identify the seven articles in the sample and on the Twitter API to collect the follower information of each of the users who shared a paper. It then applies social network analysis on the follower networks to identify diffusion patterns that can, in turn, be used to understand the various users and uses of scientific papers on Twitter.

Several network indicators and Twitter statistics such as percentage of retweets and tweet half-life are calculated for each of the seven networks and summarized in Table 1 below. Although several of the papers were tweeted a similar number of times, network indicators point to very different patterns of diffusion. In studying these indicators as well as the follower network graphs, we uncover that network characteristics can serve as a valuable source of information to detect when research articles disseminate beyond a single Twitter community. For example, even though tweeted by almost exactly the same number of users, *Biol1* (196 users) and *Biol7* (190 users) exhibit the highest and lowest modularity scores, which suggests that *Biol1* reached a diverse group of Twitter users, while *Biol7* was tweeted by a homogeneous, well-connected community. When examining the Twitter accounts of users tweeting about *Biol1*, it reveals penetration to the popular nutrition and fitness authors. In contrast, *Biol7* only has 68.5% of its users in the largest component, indicating that around a third of the users it reached were not part of a core group. In fact, five of the seven networks (*Biol2-5* and *Biol7*) were shared mostly within established and well-connected communities, with 89% or more of the users in a single component whose modularity is below .25. These structures seem to confirm above mentioned studies that research on Twitter is shared primarily among academic communities of users who are already well-connected outside of Twitter. However, the analysis of follower networks also reveal that certain publications are able to gain the attention of more diverse communities and disconnected users. These examples show that Tweet counts alone mask very different structures in the follower networks of those who tweeted.

This study signals the potential of using social network analysis on social media networks to differentiate and improve current altmetric indicators. This small sample indicates that metrics derived from the follower networks on Twitter may be useful for detecting which papers break out of academia and move into the general public. Such indicators would provide greater insight into the relationship between academia and the public. Future research includes validation of network indicators with qualitative methods, triangulation of network-based user information with different classification approaches as well as large-scale application of social network analysis in the context of altmetrics.

Table 1. Tweet and network indicators for 7 BMC Biology articles shared on Twitter

| **Name** | **Num Twitter users/ nodes** | **% of users in largest component** | **Mean shortest diffusion path** | **Infomap modularity of largest component** | **Density** | **% of Retweets** | **Tweet half-life** |
|---|---|---|---|---|---|---|---|
| Biol1 | 196 | 81.6 | 4.9 | 0.57 | 0.01 | 65.3 | 56.3 |
| Biol2 | 153 | 90.9 | 3.9 | 0.25 | 0.04 | 59.0 | 6.0 |
| Biol3 | 44 | 90.9 | 3.8 | 0.17 | 0.07 | 53.7 | 9.1 |
| Biol4 | 50 | 96.0 | 3.4 | 0.02 | 0.10 | 63.6 | 1.4 |
| Biol5 | 70 | 97.1 | 3.2 | 0.01 | 0.14 | 64.8 | 2.2 |
| Biol6 | 168 | 68.5 | 3.9 | 0.37 | 0.02 | 53.6 | 8.5 |
| Biol7 | 190 | 89.5 | 3.3 | 0.00 | 0.07 | 59.8 | 1.5 |

## References

Alperin, J. P. (2015). Moving beyond counts: A method for surveying Twitter users. In *altmetrics15: 5 years in, what do we know?* Amsterdam, The Netherlands. Retrieved from http://altmetrics.org/wp-content/uploads/2015/09/altmetrics15_paper_3.pdf

Haustein, S., Bowman, T. D., Holmberg, K., Tsou, A., Sugimoto, C. R., & Larivière, V. (2016). Tweets as impact indicators: Examining the implications of automated "bot" accounts on Twitter. *Journal of the Association for Information Science and Technology*, *67*(1), 232–238. https://doi.org/10.1002/asi.23456

Tsou, A., Bowman, T. D., Ghazinejad, A., & Sugimoto, C. R. (2015). Who tweets about science? In *Proceedings of the 2015 International Society for Scientometrics and Informetrics* (pp. 95–100). Istanbul, Turkey. Retrieved from https://pdfs.semanticscholar.org/81fe/8b63188cf25648a7c592bc6b5457fee3c101.pdf?_ga=1.184338726.1264550827.1478885332