Altmetrics: A Case Analysis of PLoS Article Level Metrics (ALM)

Müge Akbulut
Hacettepe University, Department of Information Management, 06800 Beytepe, Ankara, Turkey. mugeakbulut@hacettepe.edu.tr

Abstract: *The possibility of creating and easily distributing something via a digital platform creates an enormous material volume. With the increase of scientific publications it becomes harder for academicians to choose the most related and significant resources from the others (Henning & Gunn, 2012). Bibliometrics measures the impact of academicians' works through citations in the scholarly literature, but the impact is not limited by this. Altmetrics measures the impact of articles on the web, alongside bibliometric data. The aim of this study is to investigate the probable relationships between traditional metrics and altmetrics by analysing the PLoS Article-Level Metrics (ALM) dataset.*

Keywords*: Bibliometrics, altmetrics, Article Level Metrics (ALM).*

Introduction

The number of papers published in scientific journals is gradually increasing in parallel with developments in science and technology. Scientists engaged in research need to follow more journals and papers every day. A web environment which is globally accessible to everyone leads to benefits from a shared intelligence obtained through the contributions of users. In this environment, in which users gradually become participants, the media allows a more flexible and open system for cooperation and sharing instead of focusing only on the call to consume (O'Reilly, 2005; Shirky, 2008). At some point, mass amateurisation will cause a filtration problem which is bigger than that found in traditional environments. Therefore, solutions used before might be inadequate (Shirky, 2002). Academicians are making use of filtration to follow academic literature. However, traditional filtration methods are becoming gradually more dysfunctional as the environment continues to diversify. Traditional performance measurement systems are desirably adequate in comparison with current technology. Now, other ways of handling this problem should be focused and a number of different sources should be taken into account. The recent increase in online academic indexing systems has enabled new filtration opportunities. The introduction of altmetrics is significant, as it could rapidly impact on academic filtration within the digital ecosystem.

Academicians carry out their daily tasks on Web. According to some studies, there are more than 40 million papers listed on some online reference managers like Zotero and Mendeley (more universal than PubMed). Today, conversations about an invention are pondered and discussed on blogs and within social media (Mollett, Moran & Dunleavy, 2011). Nearly one third of academicians are also Twitter users (Priem & Costello, 2010). Besides, the number of scientific citations attributed to articles on Twitter is more than 58,000. Researchers are following other researchers whom they think to be important opinion leaders via Twitter instead of reading a bulk of pages by various authors in a peer-reviewed journal.

All these interactions are reflected in scientific communication processes. Articles which are dog-eared and have not been quoted can be found and recounted in some online environments like Mendeley, Zotero and CiteULike (Howard, 2012).

Individuals have an opportunity to get access to all kinds of online environments in which they can share a variety of articles, images, videos and so on. They can make use of such opportunities for research and cooperation. Thus, these environments may provide a useful measure of the impact of scientific research or performance.

Scientific communication has developed along with digital technologies. As communications have moved to an electronic environment, not only articles but also a number of new structures (data sets, analysis, reference managers, blogs, social networks, social marking, discussion lists and so on) have begun to appear. Hence, the homogeneity of channels has decreased while the diversification of channels has increased. Informal communication sparks significant debates in scientific communication. Researchers share their own studies with each other within online environments, potentially changing the direction of the research. More diverse environments mean more debates and discussions. If there are many ideas, then there are many individuals who discuss these with each other. Informal systems, which have been at the centre of research communications in the past, have now migrated to and are being followed online.

There are a great many more amateurs than professionals, as consumers also become producers who can easily communicate with each other in online environments (Surowiecki, 2004). This activity could be utilized to measure the impact of scientific research or performance.

This study will focus on correlations between altmetrics and traditional metrics. It will also discuss how altmetrics could affect traditional filters and whether it could be employed for forecasting.

Literature Review

As a result of the increasing opportunities provided by digital environments, researchers have begun to make use of different methods to measure scientific impact. One of the most significant of these methods is webometrics. Webometrics, first put forward by Almind and Ingwersen, applies infometric methods on the World Wide Web (Almind & Ingwersen, 1997, p. 404). There have been a number of studies about webometrics usage and usage records of online article access (Bollen, Van de Sompel, Hagberg, & Chute, 2009). However, webometric data are only manually collected periodically, as they are limited by automatic mining. This method is not functional to a large extent because publishers are not willing to open their sources for an extensive usage (Haustein & Siebenlist, 2011). Therefore, usage metrics are only automatically collected on a very small scale.

According to the studies of Procter and others, 80% of academicians are estimated to have social media user accounts (Procter et al., 2010). There are also some studies about how much altmetrics data we have and how they are distributed (Priem & Hemminger, 2010). In studies in which comment systems based on journals and "rapid impacts" are focused, there are many asymmetric distributions. In a study of Schriger and others, prepress articles in the repository and rates of being tweeted are compared. According to the results of this study, it has been revealed that sample articles from the arXiv preprint repository were tweeted at the rate of 95% (Schriger, Chehrazi, Merchant, & Altman, 2011). Wardle points out that citations in Wikipedia do not coincide with the Journal Citation Reports, although there are a few more citations in more effective journals (Wardle, 2010).

Few studies have been found which examine how altmetric data analysis will affect traditional filters and whether it can be used for the purpose of forecasting.

Method

The objective of this study is to seek any correlations between traditional and altmetric metrics by analysing PLoS Article-Level Metric (ALM) data sets. This study also aims to test whether there are statistically significant correlations among the frequency of use of articles and their impact factors, total citation numbers and half-lives.

This paper addresses the following research questions:

- Is the impact of an article on the Internet sensitive enough to forecast an increase in citations?
- Do journals with high impact factors and total citation numbers (i.e., publishing high impact articles) also have higher altmetrics impact values?
- Do 18 metric statistics designating the altmetrics values of the articles show changes over the years?

Essential data sets were downloaded from the PLoS Article-Level Metric (ALM) web page (http://article-level-metrics.plos.org/plos-alm-data/) on January 2013. Data sets included 78,386 articles published in eight PLoS journals (PLOSBiology, PLOSClinicalTrials, PLOS Computational Biology, PLOSGenetics, PLOSMedicine, PLOS Neglected Tropical Diseases, PLOS ONE ve PLOSPathogens) between 2003 and 2013. Citation numbers were manually extracted from Web of Science (WoS) for all articles matching the journal "PLoS *". The detailed Web of Science (WoS) citation information was exported for the 22,000 PLoS articles, 500 articles at a time (maximum permitted by ISI website). For the impact factor, JCR (Journal Citation Reports) data sets were merged and transferred to Numbers, SPSS and Tableau programs to evaluate them after making necessary arrangements.

Citation numbers of the articles have been gained from Web of Science (WoS) database published by Institute for Scientific Information in USA. Journal Citation Reports (JCR) published by ISI has been used to find out impact factors of the articles along with citation numbers, self-citation rates and so on.

References

Almind, T. & Ingwersen, P. (1997). Informetric analysis on the World Wide Web: Methodological approaches to 'webometrics'. *Journal of Documentation,* 53(4), 404-426. Retrieved March 14, 2013 from http://comminfo.rutgers.edu/~muresan/IR/Docs/Articles/jdocAlmind1997.pdf

Bollen, J., Van de Sompel, H., Hagberg, A. & Chute, R. (2009). A principal component analysis of 39 scientific impact measures. *PLoS ONE* 4. doi:10.1371/journal.pone.0006022. Retrieved March 14, 2013 from http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0006022

Haustein, S. & Siebenlist, T. (2011). Applying social bookmarking data to evaluate journal usage, *Journal of Informetrics*, 5(3), 446-457, 10.1016/j.joi.2011.04.002. Retrieved March 14, 2013 from http://www.sciencedirect.com/science/article/pii/S1751157711000393

Henning, V. & Gunn, W. (2012). *Impact factor: Researchers should define the metrics that matter to them*. Retrieved March 12, 2013 from http://www.guardian.co.uk/higher-education-network/blog/2012/sep/06/mendeley-altmetricss-open-access-publishing

Howard, J. (March 12, 2012). *Tracking scholarly influence beyond the impact factor*. *Chronicle of Higher Education*, Retrieved March 12, 2013 from http://chronicle.com/blogs/wiredcampus/tracking-scholarly-influence-beyond-the-impact-factor/35565

Mollett, A., Moran, D. & Dunleavy, P. (2011). *Using Twitter in university research, teaching and impact activities: A guide for academics and researchers*. Retrieved March 12, 2013 from http://blogs.lse.ac.uk/impactofsocialsciences/files/2011/11/Published-Twitter_Guide_Sept_2011.pdf

O'Reilly, T. (2005). *What is Web 2.0: Design patterns and business models for the next generation of software*. Retrieved March 12, 2013 from http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html?page=1.

Priem, J. & Costello, K (2010). *How and why scholars cite on twittter?*. Retrieved March 14, 2013 from https://docs.google.com/presentation/d/15J41q9EzK3CiMMoIPKY545ACndORpa6Wtj75v1YjrzQ/present?pli=1&ueb=true#slide=id.i0.

Priem, J. & Hemminger, B. (2010). Scientometrics 2.0: Toward new metrics of scholarly impact on the social Web. *First Monday* 15. Retrieved March 14, 2013 from http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2874/2570.

Procter, R., Williams, R., Stewart, J., Poschen, M., Snee, H., Voss, A. & Asgari-Targhi, M. (2010). Adoption and use of Web 2.0 in scholarly communications. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 368, 4039-4056. doi:10.1098/rsta.2010.0155. Retrieved March 14, 2013 from http://rsta.royalsocietypublishing.org/content/368/1926/4039.short#cited-by

Schriger, D., Chehrazi, A., Merchant, R., & Altman, D. (2011). Use of the Internet by print medical journals in 2003 to 2009: A longitudinal observational study. *Annals of Emergency Medicine,* Retrieved March 14, 2013 from http://www.sciencedirect.com.libproxy.lib.unc.edu/science/article/B6WB0-51S6FN7-1/2/2edccd79138b204deee17536d13a2eae

Shirky, C. (2002). *Weblogs and the mass amateurization of publishing*. Retrieved March 12, 2013 from http://www.shirky.com/writings/herecomeseverybody/weblogs_publishing.html

Shirky, C. (2008). *Here comes everybody, the power of organizing without organizations*. New York, NY: Penguin Group USA.

Surowiecki, J. (2004). *The wisdom of crowds: Why the many are smarter than the few*. Retrieved March 14, 2013 from http://www.diplomacy.edu/resources/books/reviews/wisdom-crowds-why-many-are-smarter-few

Wardle, D. (2010). Do "Faculty of 1000" (F1000) ratings of ecological publications serve as reasonable predictors of their future impact? *Ideas in Ecology and Evolution,* 3, 11-15. Retrieved March 14, 2013 from http://library.queensu.ca/ojs/index.php/IEE/article/view/2379