Counting the Number of Highly Cited Papers

B. Elango
Library, IFET College of Engineering,
Villupuram, India

Abstract

The aim of this study is to propose a simple method to count the number of highly cited papers. To illustrate the method, literature on the field of tribology covered in Science Citation Index Expanded of Web of Science published during 2001 – 2010 is used. As tribology is highly multidisciplinary in nature, which can be demonstrated by the proposed method, the method can be used for other subject disciplines.

Keywords: Bibliometrics, Scientometrics, Citation Analysis, Highly Cited Papers

Introduction

Citations are important to position an investigators' research, to retrospectively evaluate research performance, to detect scientific trends, and to acknowledge the contribution of others **(Wohlin 2005)**. Citation analysis is widely used to assess the characteristics and performance of various articles, journals, disciplines, institutions, and countries **(Moed 2005; 2009)**. A better measure of an individual paper is the actual number of times the paper is cited **(Madhan et al., 2010)**. Although the citation rate is not a direct measure of the impact or importance of a particular scholarly work, it does provide a marker of its recognition within the scientific community **(Shadgan et al., 2010).**

Highly-cited articles provide an interesting and useful insight into which authors, articles and topics are influencing the scientific research over time **(Smith, 2008)**. According to Zitt, **Ramanana-Rahary, and Bassecoulard (2005)**, "highly-cited articles are among the most commonly used indicators" for measuring "excellence". Further, highly cited papers obtained through bibliometric analysis can be considered in the collection of 'suggested readings' which may provide the outline of a particular research area **(Lee et al. 2005)**.

There is no standard criterion to list out the highly cited papers **(Glanzel and Shubert 1992)** and it is advocated by **Manisha and Mahesh (2013)**. Different authors or researchers counting the highly / most / top cited papers with different citation thresholds (see table 1). It is observed from table 1 that the present counting method of highly cited papers is not affected by either total

papers or total citations or total period. In this study, an innovative model has been proposed to count the number of highly cited papers.

Table 1 – Studies on highly cited papers		
Author (s) and Year	Citation Thresholds	Remarks
Ahmed and Gupta (2013)	50	Not affected by total papers, total citations, and period
Rajendran, Elango, and Manickaraj (2014)	50	
Kademani et al (2013)	400	
Bala and Gupta (2010)	230	
Malarvizhi, Wang and Ho (2010)	300	
Madhan, Chandrasekar and Arunachalam (2010)	100	
Ho Y S (2013)	1000	
Mazhari (2013)	Top 100	
Feijoo, et al (2013)	Top 100	
Levitt and Thelwall (2009)	Top 0.01%	Not affected by total citations and period
Elango, Rajendran and Bornmann (2013)	Top 1%	

Data

The sample data (literature on tribology) for this study was obtained from Science Citation Index – Expanded of WoS and ten year publication data is used for illustration purpose. To retrieve the bibliographic records related to tribology research for the period of 10 years from 2001 to 2010. The following keywords are used in the combined fields of title, abstract, author keywords and keywords plus: *tribolog* OR "tribosyst*" OR "tribo-syst*" OR "tribo-chem*" OR "tribochem*" OR "tribotechn*" OR "tribo-physi*" OR "tribophysi*"* (Elango et al., 2015; Elango & Rajendran 2015). The search is refined to restrict the literature to articles only. The data was retrieved on May 2015.

Proposed Method

An innovative simple model to count the highly cited papers has been proposed. All the properties like total citations, total papers and period of study have been integrated into this model.

$$MC \text{ in } HCP = \frac{TC}{TP} \, X \, \frac{\sum Yn - Yi}{N}$$

Where, MC = Minimum Citations

HCP = Highly Cited Papers

TC = total citations received by the papers

TP = total papers

Yn = Date of database access for citations

Yi = One of the year of publications in a data set

N = Number of years

Justification of variables used

According to Hirsch (2007), total paper and total citations are among the standard indicators. Papers published earlier would have garnered more citations than papers published more recently, given that they have been in circulation longer (Lee et al., 2005). Hence circulation period of papers should be normalized.

Need of the proposed model

The previous studies (table 1) analyzed the highly cited papers which have not been affected either by size of the total citations or total papers or years in a discipline. In the proposed method, all the properties such as, total papers, total citation and period of circulation have been integrated.

Example : Consider the following data set.

Total papers   -   8038 (2001 – 2010)

Total citations -   110461

Database accessed on May 2015

Yn = 2014

Yi = 2001, 2002,…..2010

Substitute the values in the equation.

$$MC\ in\ HCP = \frac{TC}{TP}\ X\frac{\sum Yn - Yi}{N} = \frac{110461}{8038}\ X\frac{85}{10} = \mathbf{116.81}$$

R.off = 116 (papers received 116 or more citations can be considered as highly cited papers)

With this new method, there are 49 articles received 116 or more citations and these 49 articles can be considered as highly cited articles.

Conclusion

An innovative model has been proposed to count the number of highly cited papers. The model has been illustrated with a sample of publications in the field of tribology. Since the tribology is a highly multidisciplinary field, the results of this study can be considered as a proxy.

References

- Glanzel W and Shubert A (1992). Some facts and figures on highly cited papers in the sciences, 1981 – 1985. *Scientometrics*, 25 (3), 373-380.

- Manisha M and Mahesh G (2013). Bibliometric characteristics of champion works of China and India. *Scientometrics*, 98 (2), 1101-1111.

- Ahmed K K M and Gupta B M (2013). India's contribution on antioxidants : a bibliometric analysis, 2001 – 2010. *Scientometrics*, 94 (2), 741-754.

- Rajendran P, Elango B, and Manickaraj J (2014). Publication trends and citation impact of tribology research in India: a scientometric study. *Journal of Information Science Theory and Practice*, 2 (1), 22-34.

- Kademani B S, et al (2013). Publication trends in materials science : a global perspective. *Scientometrics*, 94 (3), 1275-1295.

- Bala A and Gupta B M (2010). Research Activities in Biochemistry, Genetics and Molecular Biology during 1998 – 2007 in India : A Scientometric Analysis. *DESIDOC Journal of Library and Information Technology*, 30 (1), 3-14.

- Malarvizhi R, Wang M H and Ho Y S (2010). Research Trends in Adsorption Technologies for Dye Containing Wastewaters. *World Applied Science Journal*, 8 (8), 930-942.

- Madhan M, Chandrasekar G and Arunachalam S (2010). Highly cited papers from India and China. *Current Science*, 99 (6), 738-749.

- Ho Y S (2013). The top-cited research works in the Science Citation Index Expanded. *Scientometrics*, 94 (3), 1297-1312.
- Mazhari S (2013). The 100 top-cited articles published in psychiatric journals. *Journal of Psychiatric Practice*, 19 (4), 327-338.

- Feijoo, et al (2013). The 100 most cited articles in dentistry. *Clinical Oran Investigations*, 18 (3), 699-706.

- Levitt J M and Thelwall M (2009). The most highly cited Library and Information Science articles: Interdisciplinarity, first authors and citation patterns. *Scientometircs*, 78 (1), 45-67.

- Elango B, Rajendran P and Bornmann L (2013). Global Nanotribology Research Output (1996 – 2010): A Scientometric Analysis. *PLOS ONE*, 8 (12); e81094.

- Elango B, Rajendran P and Bornmann L (2015). A scientometric analysis of international collaboration and growth of literature at the macro level. *Malaysian Journal of Library and Information Science*, 20 (2), 41-50.

- Elango B and Rajendran P (2015). Global tribology research output (1998-2012): a macro level scientimetric study. *Journal of Information Science Theory and Practice*, 3 (4), 35-48.

- Hirsch J E (2007). Does the H-index have predictive power? *Proceedings of the National Academy of Sciences of the United States of America*, 104 (49), 19193-19198.

- Lee J D, Cassano-Pinche A, Vincente K J (2005). Bibliometric analysis of Human Factors (1970 – 2000): A quantitative description of scientific impact. *Human Factors*, 47 (4), 753-766.
- Madhan M, et al. (2010). Highly cited papers from India and China. *Current Science*, 99 (6), 738-749.

- Zitt, M., Ramanana-Rahary, S., & Bassecoulard, E. (2005). Relativity of citation performance
- and excellence measures: From cross-field to cross-scale effects of field-normalisation.
- Scientometrics, 63(2), 373–401.

- Shadgan B, Roig M, HajGhanbari B, Reid WD (2010). Top cited articles in rehabilitation. Arch Phys Med Rehabil, 91, 806-815.

- Moed HF (2005). Citation analysis in research evaluation. ISSI 2005. In: Proceedings of the 10th international conference of the International Society for Scientometrics and Informetrics 1–2:437–441.

- Moed HF (2009). New developments in the use of citation analysis in research evaluation. Arch Immunol Ther Exp 57(1), 13–18.

- Wohlin C (2005). Most cited journal articles in software engineering. *Information and Software Technology*, 47, 955.
Follow us on: IRJLIS, Facebook, Twitter