February 2021

Network Visualization of a Retracted Article: Repeated Proliferation of Error through Citation Networks

Rosy Jan
University of Kashmir, India, hakimrosy@gmail.com

Kirat Nazir
University of Kashmir, India

Kushboo Rani
University of Kashmir, India

Irtiza Peer
University of Kashmir, India

Waseem Farooq
University of Kashmir, India

Network Visualization of a Retracted Article: Repeated Proliferation of Error through Citation Networks

Rosy Jan
Sr. Assistant Professor
Deptt. of Library and Information Science, University of Kashmir
hakimrosy@gmail.com
Kirat Nazir
Deptt. of Library and Information Science, University of Kashmir
Kushboo Rani
Deptt. of Library and Information Science, University of Kashmir
Peer Irtiza
Deptt. of Library and Information Science,
University of Kashmir
Waseem Farooq
Deptt. of Library and InformationScience,
University of Kashmir

Abstract

*Retraction is used as an optimum tool to uphold and safe-guide the integrity of scholarly literature. However, knowingly or unknowingly the authors build the work on these false claims by citing the retracted articles. Such dependencies on retracted articles may become implicit and indirect causing profound and long-lasting threat to the credibility of the literature. Consequently, it is important to detect and analyze such threats. The article aims to demonstrate dependency of citing articles on retracted article with reference to the rest of the literature. A case study of highly cited (as reported by retraction watch) retracted article "**Spontaneous human adult stem cell transformation**" published in **Cancer Research** in 2005 by **Rubio, D** as lead author is visualized in terms of bibliographic coupling of citing journals and network and density visualizations of co-citations of authors. The study concludes that there is high-order citation dependency of scientific literature on retracted article.*

Keywords: *Scholarly communications, Research misconduct, Retractions, Retraction watch, Network Visualization.*

INTRODUCTION

The lack of dedication and honesty of a researchers/scientists leads to loss of ethics, misbehavior or fraud in research. Numerous undesirable practices such as plagiarism, falsification of results, data inconsistency, image duplication and compromised peer review are result of scientific fraud. The identification of research misconduct in a research article leads to its retraction **(Greitemeyer, 2014)**. **Van (2011)** defines retraction as "science's ultimate post-publication punishment: retraction, the official declaration that a paper is so flawed that it must be withdrawn from the literature". Before retracting an article having slight error or incorrect information an alteration message may be send to author or in a more acute case "*expression of concern*" may be issued **(Grieneisen & Zhang, 2012)**. Retractions were least or uncommon in the past, however number of studies suggested that retractions are on the rise "with reference to overall growth in scientific literature" (**Marcus & Oransky, 2014)**. This is an alarming trend. "Any retraction speaks to an enormous misuse of scientific assets and the publication of retracted literature can erode the faith of public in science"**(Fang & Casadevall, 2011)**. Retraction of an article can take many years from the time of its publication till retraction depending on the

reason of retraction. Articles involving misconduct take longer time to be retracted than erroneous papers **(Steen, 2011; Fang, Steen & Casadevall, 2012; Moylan & Kowalczuk, 2016).** However, studies suggest that articles continue to be cited even after their retraction **(da Silva & Dobranszki, 2017; da Silva & Cimenti, 2017)**. Error propagate when retracted literature is continuously cited and such propagation of error can be particularly dangerous in the field of medicine **(Steen, 2011).** Number of problems arises when researchers favorably cite an erroneous article. Citations to erroneous paper make such papers credible. Finally, a researcher prompted by the invalid point may incorporate it in his writings and becomes a means for propagation of an error **(Cor & Sood, 2017)**. Thus it is necessary to study and showcase the problem in more explicit form. It is important to find out the extent to which retracted articles are interwoven with the rest of the scientific literature. More importantly how such flawed literature is firmly entrenched in co-citation networks. The study specifically demonstrates the potential of a visual analytics approach to examine and monitor not only retracted articles, but also articles that might be at risk of contamination. Construction & visualization of bibliometric maps of co-occurrence of data is done in the study by using VosViewer software. The software is developed by ***Nees Jan van Eck and Ludo Waltman,*** researchers at the CWTS Leiden of Leiden University in Leiden. "The software was built for the analysis of scientometric data, but the software has a broader relevance. In particular, VosViewer is particularly good at producing textual maps of any sorts, not just from scientometric datasets but its 2.0 version is capable of handling larger datasets and broadening its focus explicitly targeting non-scientometricians" **(Sangam, S. L., & Mogali, M. S. S, 2012).**

SCOPE

The scope of study is confined to one of the retracted article "***Spontaneous human adult stem cell transformation".***

METHODOLOGY

List of highly cited retracted articles were retrieved using "***Retraction Watch",*** devoted to the examination of retracted articles as "a window into scientific process". One the highly cited retracted article list on retraction watch was selected for analysis and examination in terms of networks visualization of citations using VOSviewer. The retracted article was searched in ***Web of Science*** (WoS) and a total number of 650 citations as on October 2018 were retrieved. The results obtained were exported to VOSviewer for Constructing and visualizing bibliometric networks of data.

REVIEW OF LITERATURE

Redman, Yarandi and Merz (2008) analyzed 315 retracted articles in Pub-Med from 1995-2004 and found that these articles were cited 3942 times before retraction and 4501 times post retraction. **Da Silva and Cimenti (2017)** studied the problem of post retracted citations and traced various works that have observed that articles continue to be cited post retractions almost similarly as they were cited before retraction **(Budd, Sievert & Scoville, 1999; Unger & Couzin, 2006; Neale, Northup, Dailey & Abrams, 2007; Van Der Vet,. & Nijveen (2016).** However, the recent studies are interested in Bibliometric mapping. It has become an important research topic in the field of bibliometrics **(Börner.,Chen., & Boyack, 2003).** Construction of bibliometric maps and the graphical representation of such maps are the two aspects of current research in bibliometric mapping. However, there seems to be a trend

towards larger maps **(Boyack et al., 2005; Ioannidis., Klavan., & Boyack, 2018;. Leydesdorff, 2004; Van Eck et al., 2006,),** and for such maps simple graphical representations are inadequate. The graphical representation of large bibliometric maps can be much enhanced by means of zoom functionality, special labeling algorithms, and density metaphors. However, such kind of functionality is not integrated into the computer programs, frequently used by bibliometric researchers. The requirement was fulfilled by the software introduced by (**Van & Waltman, 2009**), the program is used for bibliometric mapping. This program pays special attention to the graphical representation of bibliometric maps. VOSviewer, where VOS stands for *visualization of similarities* is a program developed for constructing and viewing bibliometric maps. The software is used in various study to study the bibliometric mapping and citation clustering **(Chen., Hu., Milbank., & Schultz, 2013).; Leydesdorff., Carley., & Rafols, 2013; Derrick, Meijer., & Van , 2014; Waltman, 2017)**

Data Analysis and Interpretation

The article ***"Spontaneous Human Adult Stem Cell Transformation,"*** published on 15 April, 2005 in *Cancer Research*. The article was retracted since the authors have been unable to reproduce some of the reported spontaneous transformation events and suspect the phenomenon is due to a cross-contamination artifact. However, the retracted article is cited continuously in the literature.

Table 1: Context of Citations Received by the Retracted article

Retracted Article	Authors	Citing Articles before retraction	Citing Articles after retraction	Total cites in Web of Science
Spontaneous Human Adult Cell Transformation	***Rubio D, Garcia-Castro J, Martín MC, de la Fuente R, Cigudosa JC, Lloyd AC, Bernad A.***	293	357	650*

*Citations received by article as on October 2018

Table 1 lists citation to retracted article. It was observed that out 650 citations, 293 citations are received before the article is retracted and 357 citations are received by the article after retraction. Thus it shows that a majority of articles are using the retracted works after retraction. It implies that it could have a direct implication on the citing literature.

Table 2: Top Ten Source/Journals Citing Retracted Article

S.NO	NAME OF CITING JOURNAL	TIMES CITING RETRACTED ARTICLE
1	STEM CELL AND DEVELOPMENT	24
2	STEM CELLS	22
3	CYTOTHERAPY	18
4	PLOS ONE	17
5	CELL TRANSPLANTATION	15
6	CANCER RESEARCH	13
7	STEM CELLS INTERNATIONAL	13
8	EXPERIMENTAL CELL RESEARCH	11
9	EXPERT OPINION ON BIOLOGICAL THERAPY	10

Out of the total number of 650 citations received by retracted article. Table 2 lists top ten journals, which cite the retracted articles highest no. of times. The reputed journals like "***Stem Cell and Development, Stem Cells, Cytotherapy, Plos One*** etc cite retracted articles often in their articles.

Construction & Visualization of Bibliometric Maps of data

Fig.1: Bibliographic coupling of Sources/Journals

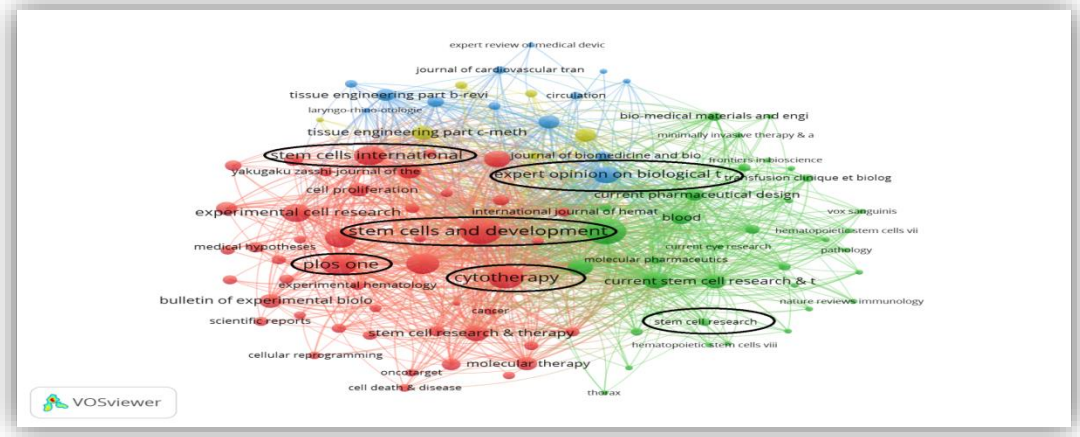
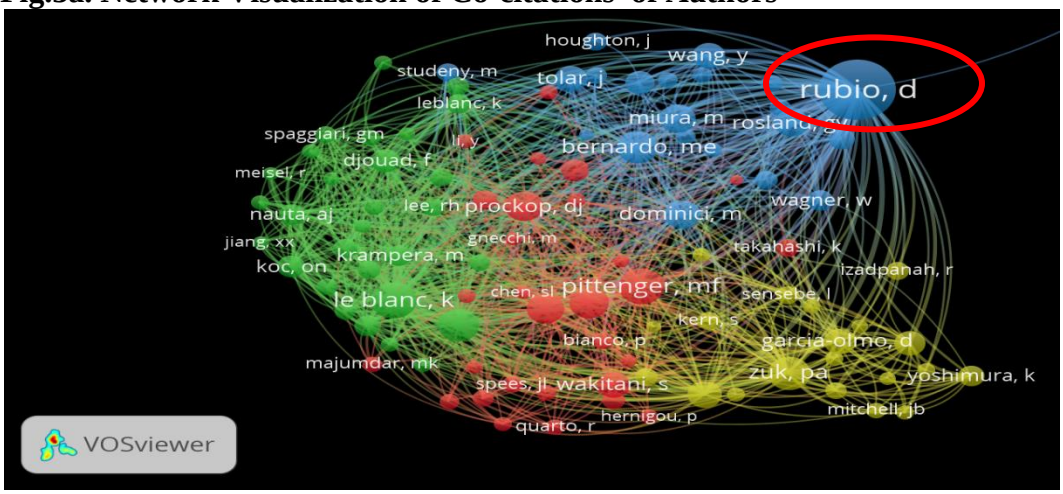


Fig.2: Bibliographic coupling of Sources in cluster View with left hand side and bottom panel providing details about clusters and link strength

The above network visualization map shows bibliographic coupling patterns of the 100 citing journals of selected retracted article. Bibliographic coupling network includes the journals with the largest number of bibliographic coupling links. The distance between two journals in the visualization approximately indicates the relatedness of the journals in terms of bibliographic coupling. Smaller the distance between two nodes, the higher is their relatedness. The color of an item is determined by the cluster to which the item belongs. Lines between items represent links. As shown in fig.1 there are four clusters represented Red, Green, Blue and yellow. Journals like *Stem cells and development*, *Plos One*, *Cytotherapy* and other top ten citing journal fall in red zone or in cluster 1 consists of 46 items, cluster 2 consists of 35 items, cluster consists of 13 items and cluster 4 consists of 6 items accounting to 100 items out of 650 with a total of 4950 links and total link strength of 222835.

Fig.3a. Network Visualization of Co-citations of Authors



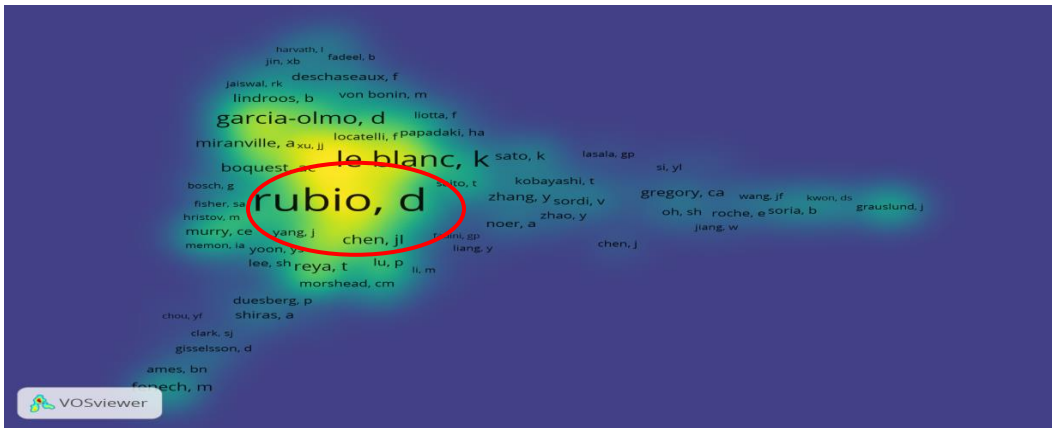
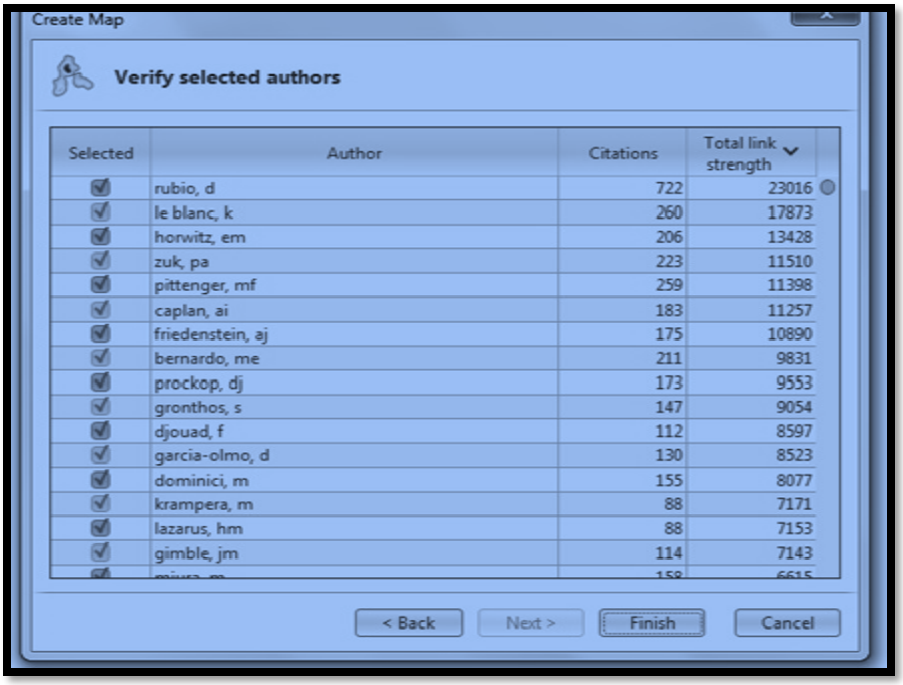


Fig.3b. Density Visualization of Co-citations of Authors



The co-citation network Visualization (**Fig. 3a**) and Density Visualization (**Fig. 3b**) shows the groups of authors with the greatest co-citation is in the centre, while the authors who have relatively least connected in terms of co-citations are situated on the periphery in fig.3b. It is revealed by the data in VOSviewer that a total strength of co-citation links of authors ***Rubio, d*** has the greatest total link strength of 23016 in the data set.

Discussion and Conclusion:

Our study aims to raise the awareness of the increasing prevalence of citations to retracted article by showcasing how retracted article is cited hundreds of times in the scientific literature. Visualizations of co-citation networks of the selected retracted article demonstrate that it is deeply interwove with the rest of literature. We have demonstrated with visualization and science mapping techniques that many retracted articles are highly cited as part of vibrant lines of research. In other words, these retracted articles are potentially more dangerous than are retracted articles in less active areas of research, especially when no effective tools are readily available to track down closely related articles. We recommend that the study of scientific literature should be done routinely such that retracted articles and closely related articles can be identified in a timely manner. We have demonstrated how a visual analytics approach can be used to facilitate the study of the role played by retracted articles. Article citing retracted works are not methodically reexamined and there are no set guidelines to stop citation to retracted articles. Hence, new articles may unknowingly cite a chain of such articles. More important, verifying the validity of articles on citation chains becomes increasingly challenging as new publications are added to the literature, and their validity may be taken for granted because they are not directly involved in any retractions. New mechanisms for checking plagiarism, duplication, and indirect citations to retracted articles in new manuscripts should be considered as an integral part of a manuscript-management workflow.

References:

Börner, K., Chen, C., & Boyack, K. W. (2003). Visualizing knowledge domains. *Annual Review of Information Science and Technology*, *37*(1), 179–255.

Boyack, K.W., Klavans, R., & Börner, K. (2005). Mapping the backbone of science. Scientometrics, 64(3), 351–374.

Budd, J. M., Sievert, M., & Schultz, T. R. (1999). Phenomena of retraction: reasons for retraction and citations to the publications. *Jama*, *280*(3), 296-297.

Chen, C., Hu, Z., Milbank, J., & Schultz, T. (2013). A visual analytic study of retracted articles in scientific literature. *Journal of the American Society for Information Science and Technology*, *64*(2), 234-253.

Cor, K., & Sood, G. Propagation of Error: Approving Citations to Problematic Research.

Da Silva, J. A. T., & Bornemann-Cimenti, H. (2017). Why do some retracted papers continue to be cited?. *Scientometrics*, *110*(1), 365-370.

Da Silva, J. A. T., & Dobránszki, J. (2017). Highly cited retracted papers. *Scientometrics*, *110*(3), 1653-1661.

Derrick, G. E., Meijer, I., & Van Wijk, E. (2014). Unwrapping "impact" for evaluation: A co-word analysis of the UK REF2014 policy documents using VOSviewer. In *Proceedings of the science and technology indicators conference* (pp. 145-154).

Derrick, G. E., Meijer, I., & Van Wijk, E. (2014). Unwrapping "impact" for evaluation: A co-word analysis of the UK REF2014 policy documents using VOSviewer. In *Proceedings of the science and technology indicators conference* (pp. 145-154).

Fang, F. C., & Casadevall, A. (2011). Retracted science and the retraction index. *Infection and immunity*, IAI-05661.

Fang, F. C., Steen, R. G., & Casadevall, A. (2012). Misconduct accounts for the majority of retracted scientific publications. *Proceedings of the National Academy of Sciences*, *109*(42), 17028-17033.

Greitemeyer, T. (2014). Article retracted, but the message lives on. *Psychonomic bulletin & review*, *21*(2), 557-561.

Grieneisen, M. L., & Zhang, M. (2012). A comprehensive survey of retracted articles from the scholarly literature. *PLoS ONE*, *7*, e44118.

Ioannidis, J. P., Klavans, R., & Boyack, K. W. (2018). Thousands of scientists publish a paper every five days.

Leydesdorff, L. (2004). Clusters and maps of science journals based on bi-connected graphs in Journal Citation Reports. Journal of Documentation, 60(4), 371–427

Leydesdorff, L., Carley, S., & Rafols, I. (2013). Global maps of science based on the new Web-of-Science categories. Scientometrics, 94(2), 589–593.

Leydesdorff, L., Carley, S., & Rafols, I. (2013). Global maps of science based on the new Web-of-Science categories. *Scientometrics*, *94*(2), 589-593.

Marcus, A., & Oransky, I. (2014). Retraction Watch [Internet]. *Available at h ttp://retractionwatch. com.*

Moylan, E. C., & Kowalczuk, M. K. (2016). Why articles are retracted: A retrospective cross-sectional study of retraction notices at BioMed Central. *British Medical Journal Open*, *6*(11), e012047.

Neale AV, Northrup J, Dailey R, Marks E, Abrams J. (2007).Correction and use of biomedical literature affected by scientific misconduct. Sci Eng Ethics. Mar; 13(1): p. 5-24

Redman, B. K., Yarandi, H. N., & Merz, J. F. (2008). Empirical developments in retraction. *Journal of Medical Ethics*, *34*(11), 807-809.

Sangam, S. L., & Mogali, M. S. S.(2012). Mapping and Visualization Softwares tools: a review.

Steen, R. G. (2011). Retractions in the scientific literature: is the incidence of research fraud increasing?. *Journal of medical ethics*, *37*(4), 249-253.

Steen, R. G. (2011). Retractions in the scientific literature: is the incidence of research fraud increasing?. *Journal of medical ethics*, *37*(4), 249-253.

Steen, R. G. (2011). Retractions in the scientific literature: is the incidence of research fraud increasing?. *Journal of medical ethics*, *37*(4), 249-253.

Unger, K., & Couzin, J. (2006). Even retracted papers endure.

Van Der Vet, P. E., & Nijveen, H. (2016). Propagation of errors in citation networks: a study involving the entire citation network of a widely cited paper published in, and later retracted from, the journal Nature. *Research integrity and peer review*, *1*(1), 3.

Van Der Vet, P. E., & Nijveen, H. (2016). Propagation of errors in citation networks: a study involving the entire citation network of a widely cited paper published in, and later retracted from, the journal Nature. *Research integrity and peer review*, *1*(1), 3.

Van Eck, N. J., & Waltman, L. (2009). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, *84*(2), 523-538.

Van Eck, N. J., & Waltman, L. (2011). Text mining and visualization using VOSviewer. *arXiv preprint arXiv:1109.2058*.

Van Eck, N. J., Waltman, L., Van den Berg, J., & Kaymak, U. (2006). Visualizing the computational intelligence field. IEEE Computational Intelligence Magazine, 1(4), 6–10.

Van Noorden, R. (2011). The trouble with retractions. *Nature*, *478*(7367), 26.

Van Noorden, R. (2011). The trouble with retractions. *Nature*, *478*(7367), 26.

Waltman, L. (2017). Citation-based clustering of publications using CitNetExplorer and VOSviewer. *Scientometrics*, *111*(2), 1053-1070.

Waltman, L. (2017). Citation-based clustering of publications using CitNetExplorer and VOSviewer. *Scientometrics*, *111*(2), 1053-1070.