The Application of Proximity Measure in Patent Citation Ranking

Budi Nugroho
Informetric and Socio Informatics Research Group
Research Center for Informatics – Indonesian Institute of Sciences
Indonesia
{budi.nugroho@lipi.go.id}

ABSTRACT *: Citation ranking is the process of ranking authors in a collaborative network, according to the number of references each author has scored. Imagine a graph-based patent citation structure. One of the basic operations on such a network is to search for a new patent and connections. We concentrate on the approaches to patent citation based on measures for analyzing the proximity of vertices in a network. The notion of closeness is defined by a value called proximity measure. In this paper, we analyze the proximity measures named PageRank, Adamic-Adar Index, Common Neighbours, Cosine similarity, and Jaccard coefficient using a citation network dataset of The U.S. patent dataset that is maintained by the National Bureau of Economic Research. The citation graph includes all citations made by patents granted between 1975 and 1999, totaling 16,522,438 citations. We conducted the paired t-test analysis to know whether there is a significant difference between patent citation scores or not. We show the results in a comparison table between each proximity measures. The paired t-test analysis shows that there is no significant difference between patent citation scores.*

Keyword: Proximity Measure, Patent Citation Ranking, Paired T-test Analysis

References

[1] Bollen, J., de Sompel, H., Rodriguez, M. A. (2008). Towards Usage-Based Impact Metrics: First Results from the Mesur Project. *In*: Proceedings of the 8th ACM/IEEE-CS Joint Conference on Digital Libraries, 231–240. https://doi.org/10.1145/1378889.1378928

[2] Bollen, J., Van de Sompel, H., Smith, J. A., Luce, R. (2005). Toward alternative metrics of journal impact: A comparison of download and citation data. Information Processing & Management, 41(6), 1419–1440. https://doi.org/https://doi.org/10.1016/j.ipm.2005.03.024

[3] Cobo, M. J., Martínez, M. A., Gutiérrez-Salcedo, M., Herrera, M., Herrera-Viedma, E. (2014). Identifying citation classics in fuzzy decision making field using the concept of H-classics. *Procedia Computer Science*, 31, 567–576. https://doi.org/10.1016/j.procs.2014.05.303

[4] DELGADO-LÓPEZ-CÓZAR, E., CABEZAS-CLAVIJO, Á. (2013). Ranking journals: could Google Scholar Metrics be an alternative to Journal Citation Reports and Scimago Journal Rank? *Learned Publishing*, 26(2), 101–114. https://doi.org/10.1087/20130206

[5] Érdi, P., Makovi, K., Somogyvári, Z., Strandburg, K., Tobochnik, J., Volf, P., Zalányi, L. (2013). Prediction of emerging technologies based on analysis of the US patent citation network. *Scientometrics*, 95(1), 225–242. https://doi.org/10.1007/s11192-012-0796-4

[6] Eysenbach, G. (2011). Can Tweets Predict Citations? Metrics of Social Impact Based on Twitter and Correlation with Traditional Metrics of Scientific Impact. *J Med Internet Res*, 13(4), e123. https://doi.org/10.2196/jmir.2012

[7] Franceschet, M. (2010). Ten good reasons to use the EigenfactorTM metrics. *Information Processing & Management*, 46(5), 555–558. https://doi.org/https://doi.org/10.1016/j.ipm.2010.01.001

[8] Fu, S.-C., Chan, K.-M. (2014). Ranking of orthopaedic journals: A challenge to the citationbased metrics. *Journal of*

Orthopaedic Translation, 2(3), 131–138. https://doi.org/https://doi.org/10.1016/j.jot.2014.03.005

[9] Giatsidis, C., Nikolentzos, G., Zhang, C., Tang, J., Vazirgiannis, M. (2019). Rooted citation graphs density metrics for research papers influence evaluation. Journal of Informetrics, 13(2), 757–768. https://doi.org/https://doi.org/10.1016/j.joi.2019.03.006

[10] Guan, J. C., Gao, X. (2009). Exploring the h-index at patent level. *Journal of the American Society for Information Science and Technology*, 60(1), 35–40. https://doi.org/10.1002/asi.20954

[11] Haddawy, P., Hassan, S.-U., Asghar, A., Amin, S. (2016). A comprehensive examination of the relation of three citation-based journal metrics to expert judgment of journal quality. *Journal of Informetrics*, 10(1), 162–173. https://doi.org/https://doi.org/10.1016/j.joi.2015.12.005

[12] Kahng, A. B., Luo, M., Nam, G., Nath, S., Pan, D. Z., Robins, G. (2015). Toward metrics of design automation research impact. 2015 *IEEE/ACM International Conference on Computer-Aided Design* (*ICCAD*), 263–270. https://doi.org/10.1109/ICCAD.2015.7372579

[13] Keselman, L. (2019). Venue Analytics: A Simple Alternative to Citation-Based Metrics. 2019 ACM/IEEE Joint Conference on Digital Libraries (JCDL), 315–324. https://doi.org/10.1109/JCDL.2019.00052

[14] Krapivin, M., Marchese, M., Casati, F. (2009). Exploring and Understanding Scientific Metrics in Citation Networks. In J. Zhou (Ed.), *Complex Sciences* (p 1550–1563). Berlin, Heidelberg: Springer Berlin Heidelberg.

[15] Krapivin, M., Marchese, M., Casati, F. (2010). Exploring and Understanding Citation-based Scietific Metrics. *Advances in Complex Systems*, 13(01), 59–81. https://doi.org/10.1142/S0219525910002487