Rejecting Moderation: An Entropy-based Indicator System for Measuring Patent Technological Innovation Potential

Yi Zhang[1], **Ying Huang**[2], **Yue Qian**[2], **Ying Guo**[2,*], **Jie Lu**[1]

[1]Decision Systems & e-Service Intelligence research Lab, Centre for Quantum Computation & Intelligent Systems, Faculty of Engineering and Information Technology, University of Technology Sydney, Australia

[2]School of Management and Economics, Beijing Institute of Technology, Beijing, P. R. China

Keywords Patent analysis; Innovative capability; Entropy; Indicator system; Bibliometrics.

Extended Abstract

Following Schumpeter's observations on *Business Cycles* (Schumpeter 1939), invention is considered as a process of recombination (Fleming 2001), and theoretical and systematic explanations of technological innovation have become an important scholarly topic for innovation management. The suitability of patents for indicating technological innovation has been discussed considerably since the 1990s and even before (Basberg 1987). A number of patentometric indicators have been applied to measure technological innovation from diverse econometric perspectives -- e.g., patent citations, cross-collaboration of assignees, number of claims, and other bibliographic information (Archibugi and Planta 1996; Fleming 2001; Qian 2007; Funk and Owen-Smith 2016). Based on statistics and empirical studies, determinants of patent value regarding economic potential were observed, where both quantitative indicators [e.g., backward patent citations, non-patent citations, the number of inventors, and the number of co-assignees (Sapsalis et al. 2006)] and qualitative ones [e.g., technical importance, inventing difficulty, and learning value for competitors (Reitzig 2003)] are involved.

How to evaluate patent value quantitatively and systematically is also a challenge for bibliometrics. As a pioneering study, Pavitt (1985) pursued de Solla Price's argument on the practical needs to explain new empirical data provided by measurement systems (de Solla Price 1983), and foresaw positively on using patent statistics in analyzing technological activities for policy making. Patent indicators were then widely introduced to measure patent value, which could constitute technological value, or direct and indirect economic value (Lee 2009). Such indicators involve not only patent statistics but also legal status information. For a wide range of science, technology, & innovation (STI) studies, such indicators are selected to evaluate a corpus of patents that represents a given technological area or entities like country, organization, and individual (Narin and Hamilton 1996; Meyer and Tang 2007; Zhang et al. 2014).

Indicator systems are not unfamiliar for econometrics, which usually apply regression-based statistic models to measure the relationships between economic outcomes and bibliometric indicators, but a bibliometric indicator system to automatically identify meaningful patents and patent portfolios remains elusive. On the one hand, currently, blending patent citation/co-citation analysis and social network analysis to seek patents at traffic hubs is one mainstream approach to identify "key" patents (Choi and Park 2009; Funk and Owen-Smith 2016), for which it becomes critical to consider citation information. On the other hand, engagement of multiple indicators also introduces issues – e.g., how to weigh those indicators? Delphi-based or Analytic Hierarchy Process-based qualitative approaches could be promising in some sense. However, even if we ignore the biases possibly resulting from subjective opinions of experts, such traditional weighting approaches mostly could lead to "moderation" results – i.e., patents ranked to a top list could well be neither those with highest backward citations nor those with the most active cross-national collaborations, but they would have good-looking values on all indicators. In Chinese philosophy, such a phenomenon is summarized as the Doctrine of the Mean, but it is definitely not good for indicating innovation.

Aiming to address the above concerns, this paper attempts to construct an entropy-based indicator system to measure the technological innovative capability of patents. One basic target is to identify significant patents with high technological innovation rather than those multi-dimensional moderate ones. This paper first proposes a patent indicator system that contains three macro-level perspectives: technological perspective, legal perspective, and market perspective. Each perspective is constituted by a number of patent indicators; we calculate the correlation of these indicators to make sure they are

suitably independent variables. We, then, based on a small training set, apply a learning-based collaborative filtering technique to remove noise and reduce the scale of the target patent corpus. Shannon's entropy (Shannon 1948), well-known as a coefficient for measuring complexity and uncertainty, is introduced to quantitatively weigh indicators. Its basic weighting criterion is that the more common an indicator is, the less weight it would have. In other words, patents with irregular indicator values would be ranked higher. We identify the entropy-weighted value as indicating technological innovative capability. The output of our method is a set of entropy-weighted patents. Aided by expert knowledge, it could be used to seek patents with technological values and innovative potential. Furthermore we consider how the entropy measures could serve to forecast possible technological recombination in the near future. We apply our method to all 26,982 patents with Chinese assignees in the United States Patent and Trademark Office (USPTO) database, covering the period from 1976 to 2014. The results demonstrate the feasibility and efficiency of our method, and also provide interesting insights for related Research & Development (R&D) planning and strategic management.

References

Archibugi, D., & Planta, M. (1996). Measuring technological change through patents and innovation surveys. *Technovation, 16*(9), 451-519.

Basberg, B. L. (1987). Patents and the measurement of technological change: a survey of the literature. *Research Policy, 16*(2), 131-141.

Choi, C., & Park, Y. (2009). Monitoring the organic structure of technology based on the patent development paths. *Technological Forecasting and Social Change, 76*(6), 754-768.

de Solla Price, D. (1983). *Sealing Wax and String: A philosophy of the experimenter's craft and its role in the genesis of high technology.* Paper presented at the Sarton Lecture, AAAS Meeting (May 1983).

Fleming, L. (2001). Recombinant uncertainty in technological search. *Management Science, 47*(1), 117-132.

Funk, R. J., & Owen-Smith, J. (2016). A Dynamic Network Measure of Technological Change. *Management Science*. doi: 10.1287/mnsc.2015.2366

Lee, Y.-G. (2009). What affects a patent's value? An analysis of variables that affect technological, direct economic, and indirect economic value: An exploratory conceptual approach. *Scientometrics, 79*(3), 623-633.

Meyer, M., & Tang, P. (2007). Exploring the "value" of academic patents: IP management practices in UK universities and their implications for Third-Stream indicators. *Scientometrics, 70*(2), 415-440.

Narin, F., & Hamilton, K. S. (1996). Bibliometric performance measures. *Scientometrics, 36*(3), 293-310.

Pavitt, K. (1985). Patent statistics as indicators of innovative activities: possibilities and problems. *Scientometrics, 7*(1-2), 77-99.

Qian, Y. (2007). Do national patent laws stimulate domestic innovation in a global patenting environment? A cross-country analysis of pharmaceutical patent protection, 1978-2002. *The Review of Economics and Statistics, 89*(3), 436-453.

Reitzig, M. (2003). What determines patent value?: Insights from the semiconductor industry. *Research Policy, 32*(1), 13-26.

Sapsalis, E., de la Potterie, B. v. P., & Navon, R. (2006). Academic versus industry patenting: An in-depth analysis of what determines patent value. *Research Policy, 35*(10), 1631-1645.

Schumpeter, J. A. (1939). *Business cycles* (Vol. 1): Cambridge Univ Press.

Shannon, C. (1948). A mathematical theory of communication. *The Bell Systems Technical Journal, 27*(3), 379-423.

Zhang, Y., Zhou, X., Porter, A. L., Gomila, J. M. V., & Yan, A. (2014). Triple Helix innovation in China's dye-sensitized solar cell industry: hybrid methods with semantic TRIZ and technology roadmapping. *Scientometrics, 99*(1), 55-75.