Patent citation: A technique for measuring the knowledge flow of information and innovation

P. Sharma [a, *], R.C. Tripathi [b]

[a] GB University-Noida, India
[b] Nehru Gram Bharti University, India

ABSTRACT

Knowledge cannot be bound, restricted or categorized. Knowledge is precisely an intangible strength that has a definite economic importance if well utilized and commercialized. Knowledge spillover is an occurrence, which is imaginable but difficult to have an effective measurement of it. Patents citation is a developing concept and has gained momentum in recent past. Patents citation contains valuable data and if analyzed well, may sometimes reveal concealed mysteries of the information flow between countries, laboratories, companies, and universities. Profuse technical research has been conducted on this topic by many scientists. Through these experiments, scientists have tried to show that the innovative information hidden in patents crosses every barrier and is taken by the research labs for its further expansion. Patents citation reveals the diffusion of information and its applicability into many other technical fields which give birth to a new technology. This paper presents a comprehensive survey of patents citation analysis covering and promoting the landmark research done in the field of patents citation, informing readers to consider this important segment of patent document as a field for analysis. Also, this paper presents an innovative methodology for generating patent citation network with the help of techniques of Information Retrievals.

© 2017 Elsevier Ltd. All rights reserved.

Contents

1. Introduction

Since 1990, theorist and economist have started diverting their attention towards the technical data. Perhaps the world understood the importance of technology and its increasing involvement in

* Corresponding author.
   E-mail address: psh.3005@gmail.com (P. Sharma).

human life. This was one of the main reasons behind blooming and constant use of patent data for the research work. Moreover, the digitalization of patent data increased its value and usability in the research. Economist, corporate-managers and policy makers were moved by the realization that technology is a major crucial determinant which is required to be studied properly for national competitiveness, endogenous growth of macroeconomics and its acceptability in terms of micro impacts. An important and less noticed section of a patent document is the reference or citation segment. Patents citation reflects the relationship between existing and current patent and previous patents. It is similar to a research paper reference later on and has emerged as a newer tool to identify the strong link existing between the technologies. Patents citation study gained momentum after the heroic effort of National Bureau of Economic Research (NBER) scientists involvement in providing the citation information of US patents from 1963 to 2006 (in different volumes). With the improved quality citation information, it became easier for researchers to utilize the patents citation information in identifying the linkage and changing trends in technologies. Patents application goes through a very rigorous prior art search, with least possibilities of infringing any other patent. These patents citation provided by the inventor prove the originality of inventions. For this reason patents citation information is considered as a reliable proxy to measure activities like technology transfer, technology life cycle, base patents and technology diffusion [1,2]. Patents with high number of citations are of higher quality and are considered as a base patent for any technology, *ceteris paribus* [3]. Though being a very important aspect, the research in terms of patents citation has been very limited. Most of the concrete research in patents citation has been performed only after the availability of NBER data and has gained momentum after year 2006—2007. Models related with patents citation relationship are scarce and the earlier research has been only at aggregate level [4].

Patents citation research has eradicated the traditional method of simple patents citation count. Poisson regression model or negative binomial model, deals simply with citation count data. Being too generic in analysis, it cannot identify the complexity in relationship between the technologies. Patents citation holds enough of data to be utilized in research practices. If assessed and analyzed properly with computational models it can reveal various aspects of technology growth and consequential developments. It can provide information like base patents, technology life cycle, technology convergence, technology mergers with other technology, scientific research converting into real time technology, etc. For all this, we need to first understand, *what is a patent citation*?

Patents citation is nothing but prior work for any invention. Citations are of two types —

Citations provided by the applicant are termed as 'prior art' and are known as 'applicant citation'. Citations which are left over and are found by the patent examiner of the country patent office during patent search are known as 'examiner citation' [5]. The cited material has some bearing to the idea(s) being patented and its claims [6]. Patent holds large amount of valuable data which cannot be perceived by the first look of patent. A vigilant analysis is required to find out this indirect linkage of information with the patent. Inventors generally see the claims and inventive part of the patent with no inquisitions towards the most valuable part of harness patent-the citation. This un-surfed citation data can be utilized in number of ways to the value of patents. Not only the citing patent but also the cited patents have enough relevancies for performing a proper analysis of the technology growth. Patent citations count usually has a direct relationship with the market value of the product [7]. The more citation a patent receives the more commercial value it holds. Patents citation is also considered

as tool for assessing the commercial value of any patent. However, they are not the only parameter of assessing the value. Other factors also play considerable role. Narin et al. [8] also conducted a research and concluded that more citation a patent receives the more value it has. But they did not mention any thing about self-citation. Companies may cite their own patent for increasing number of citations of their patents.

Patent citations study is a developing concept in use since past two decades. The literature review of the patents citation theory reveals the fact that the study has been targeted mostly to understand the various technical dimensions of many popping up technologies. It depicts the importance which has been established in the field of research opening various new dimensions and scopes. Delving the research work of several inventors, one may find numerous uses of patents citation in various areas of researches. Majority have been used to understand the technology trajectories *i.e.* the road mapping of technologies and knowledge flow. For instance, Xuan Ting Ye et al. have used patents citation network for measurement of International knowledge flow [9]. Park Y et al. have applied patents citation analysis to find the innovation across industries and other domains for various prospects [10]. Patents are technical data considered as a main indicator for mapping the technical progress in real terms. They provide possible explanation for the birth of technical research problems. Here we have discussed some significant works of researchers which have been cited considerably by other authors. Also it fulfills the need for understanding in brief the pioneer works done in the field of Patent Citation Network (PCN).

2. Patents, citations & innovations: NBER research

United States, National Bureau of Economic Research[1] (NBER) researchers [11] built a comparable framework which neo theorist have widely applied in their research models for obtaining desired results. Kuznets [12] Foster [13] took initiative to introduce the idea of utilizing patents citation information. Their work was further carried away by Griliches [14] to identify those intangible valuables which push forward the economic growth. To achieve the task they utilized the massive US patent data filed at USPTO into a channelized way. They identified the generality and originality in patents, identified parameters, validated these parameters for understanding the technological changes along with the subtle economic issues emerged during the research. Their reckoned research is considered as pioneering in the field of patents citation and has been quoted by more than hundred researchers in their work related with patents citation. Their work is a well comprehended description of US patent database. The aim of their research was to make it widely available for public research use. It contains data of patents filed and granted by USPTO during 1963—2006 in various segments. These indicators show the various changing trends registered in past 44 years of technology. These patents exhibit a range of changes the technology has faced across the diverse categories. The complete dataset which consists of 3 million patents and their citations, totaling to approximately 16 million is open for public research use and is freely downloadable from the official website of NBER.

Their research also discussed some crucial issues which researchers may encounter during the use of these citation data, and suggested measures to handle them efficiently. Two alternate approaches have been used in their methodology —

---

[1] http://www.nber.org/