Composing technology roadmapping according to bibliometrics: hybrid model and empirical study

Yi Zhang[1+], Donghua Zhu[1], Xuefeng Wang[1], Ying Guo[1], Hongshu Chen[1], Changli Liu[1],

Jing Xu[2], Xiaofan Heng[1], Xiaoguang Lu[1]

[1]School of Management and Economics, Beijing Institute of Technology, Beijing 100081, China

[2]The Science of Science and the Sci-Tech Management Magazine, Tianjin 300011, China

Abstract. Technology roadmapping is considered as a powerful instrument for technology assessment and forecasting, and the composing model for technology roadmapping is also a hot topic in the field of technology visualization. This paper tries to introduce a bibliometric method involving with terms frequency analysis and association analysis to construct the objects and their relationships in the technology roadmapping, and blend with expert knowledge to locate each object in the appropriate layer and time interval. Technology roadmapping is drawn in a half automatic way, in which terms and IPCs are retrieved and clustered by systematic software, and the classification and adjustment in technology roadmapping are finished by experts. Moreover, an empirical study of EV (electric vehicle) proves the applicability and effectiveness of this hybrid composing model for technology roadmapping.

Keywords: Technology roadmapping, Bibliometrics, Terms frequency analysis, Association analysis, EV,

1. Introduction

The booming of emerging technology is an obvious feature of today's science and technology development. In this way, how to dominate a favourable position in the fierce global competition should be considered seriously by both governments and enterprises. Organizations operating in competitive environments demanding process improvements, new product introductions, or technology-enhanced services must obtain and use information on emerging technologies [1]. Therefore, the research on emerging technology and its dynamic development trend is significant in both theoretical and practical ways.

Technology roadmapping is a flexible technique that is widely used within industry to support strategic and long-ranged planning. The approach provides a structured (and often graphical) means for exploring and communicating the relationships between evolving and developing markets, products and technologies over time [1]. Technology roadmapping emphasizes the combination of technology visualization and qualitative methodology, and identifies the relationship between technology and time. At present, researchers are trying to introduce some quantitative techniques into technology roadmapping. However, quantitative method and qualitative method could not be blended with each other in an appropriate way. For this reason, this paper proposes to build a hybrid composing model for technology roadmapping with both qualitative and quantitative methodologies.

As one of the most important techniques of quantitative methodology, bibliometric method is introduced. Bibliometrics use counts of publications, patents, or citations to measure and interpret scientific and technological advances [2], summarize scientific activities, and indicate research performances over time [3]. Bibliometric method is economical, non-invasive, and simple to implement massive quantitative data in a manner that individuals could not handle [4]. In this consideration, bibliometric analysis has been used as a common research instrument for systematic analysis in many research fields.

The remainder of this study is organized as follows. A brief literature review of technology roadmapping and bibliometric method including terms frequency analysis and association analysis is in Section 2. And then the proposed composing model for technology roadmapping with bibliometrics is described in Section 3.

[+] Corresponding author. Tel.: +86 10 68918560.
  E-mail address: yi.zhang.bit@gmail.com

Section 4 gives a case of electric vehicle to verify the hybrid model. Finally, some conclusions are provided in Section 5.

2. Literature Review

Technology roadmapping represents a powerful technique for supporting technology management and planning, especially for exploring and communicating the dynamic linkages between technological resources, organizational objectives and the changing environment [1]. Generally, technology roadmapping is a kind of visualized method to illustrate the development path of technologies mainly in a qualitative method [5-8]. What is more, some quantitative techniques especially bibliometric method are also used for technology roadmapping independently [9-11].

After the definition of modern bibliometrics by Derek Price [12], Alan Porter defined technology opportunity analysis by identifying related component technologies, developers who are active with the technologies, locations where the developers are, and related strengths between the identified developers [2]. It's a challenge to look for the relationship between different items (individuals, companies and institutions) in a bibliometric way.

There are numbers of papers introducing various forms of analyses involving with bibliometric methods. In global, bibliometric method is usually based on mathematic statistic analysis and data mining technique involving with literature analysis, patent analysis [13-14], citation analysis [3,15-18] and terms analysis[19], etc. Science publications such as papers in Web of Science or Engineering Index database are the foundation of bibliometric method, and now, more and more kinds of literatures are used. Furthermore, the research on the science mapping based on bibliometric method is also becoming an important trend [20].

From the literature review, we find that researchers in the field of technology roadmapping are mostly committed themselves to only qualitative or quantitative method, and few of them combine the two methods together. For this reason, we are trying to use bibliometric method and technology roadmapping technique together with the support of experts, and introduce a composing model for technology roadmapping with bibliometrics, in which terms frequency analysis and association analysis are used.

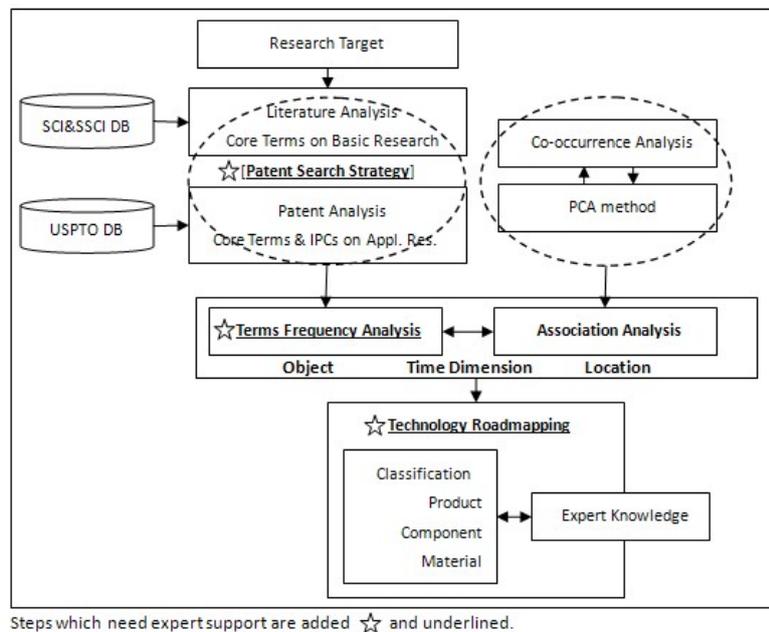## 3. Methodology: the technology roadmapping process

Fig.1 The process of hybrid composing model for technology roadmapping

This study proposes a process integrating expert knowledge and bibliometric methods, which include terms frequency analysis and association analysis, to engage the challenge of technology roadmapping. Terms frequency analysis lays emphases on the technology core terms and IPCs retrieved from publication and patent documents. Among them, literature analysis gathered terms on basic research, while patent

analysis puts an eye on the terms of application research and especially the IPCs. Association analysis is used to find the relationship between terms and IPCs by co-occurrence analysis and PCA method. In the other hand, terms and IPCs are recognized as the objects in the technology roadmapping. Experts should classify the objects in an appropriate layer for technology development, and order them in the time dimension considering the relationships we got from association analysis. In this way, the composing model for technology roadmapping has been finished, the process of which is illustrated in Fig.1. In addition, aims to blend quantitative and qualitative methodologies together, bibliometric method is connected with expert knowledge closely to make sure the authority, credibility and objectivity.

3.1. Objects and Relationships Retrieving for Technology roadmapping

Science publications focus on the hot topics of academic research fields on basic research, which could be induced briefly by the terms. Hence, the set of terms from publications in a long time describes the global characteristics of research results, the internal relationships between research contents and the future directions of research trends. In this condition, technology core terms on basic research are gathered by literature analysis, and then, the patent search strategy is composed of core terms by experts.

Patent connects the inventor and the government [5], represents the origin and features of technology, and has long been recognized as a very fruitful source of data for the study of innovation and technical change and thus for technology management research [22]. Not only technology core terms could be retrieved from patent documents, but also the IPCs (International Patent Classification) which define the research fields of patents and help to identify the relationship between patents and should be an important evaluation index for patent analysis.

Association analysis is an important analysis method of data mining, which combines the terms co-occurrence analysis and PCA (Principal Component Analysis) method. The terms co-occurrence analysis is a kind of quantitative method, which is based on the pattern of terms occurring together in the records. If two terms occur together in the records more frequently than expected, there is a presumption of relationship between them. On the other hand, PCA is a useful technique for extracting the main relationships implicit in a data set. A PCA-based approach called latent semantic indexing generates conceptual indices instead of individual words to improve information retrieval [21]. According to the association analysis, we find the relationships between different technology core terms and IPCs, and cluster the related terms in time dimension preliminarily.

There is no doubt that SCI and SSCI database is one of the most important international literature search tools and databases for science statistic analysis and assessment, and there are also several reasons why we choose USPTO (United States Patent and Trademark Office) as the patent database in this paper. Firstly, USPTO covers the patents from all countries, and has the largest number of patents in the world. Secondly, the U.S. market is an important market for technology transfer and international trade, combined with the territoriality of patent protection, thereby luring inventors to file patent applications in the U.S [11]. Considering of these reasons, the literature and patent databases are designated.

3.2. Classification and Construction of Technology Roadmapping

Technology roadmapping illustrates the dynamic development path of technology graphically, and should be considered as an important technique for the research of technology assessment and technology forecasting. At present, qualitative analysis which is mostly based on the experiences of experts is the main method to approach the technology roadmapping, and is panned because of the lack of data supporting. This paper prospects to construct the framework of technology roadmapping, which is based on both bibliometric method and expert knowledge. In this framework, the technology core terms and IPCs retrieved from publication and patent documents are used to be the objects in the mapping, and their relationships are substantial evidences for experts' decision.

Generally, there are many phases in the development path of technology. Considering of the general model of emerging technology, this paper designs three phases for the ordinate axis which are material, component/technology and product, while time dimension is used to the horizontal axis. In this way, based on the results of bibliometric method which defined the objects and their relationships, experts classify each

object in an appropriate layer, locate them in the right time interval and adjust locations in order to ensure the cluster. In this way, the composing model for technology roadmapping should be finished.

Specially, not similar with the phase of bibliometric method, qualitative methodology is the main technique during the phase of drawing the technology roadmapping. Therefore, expert knowledge is the foundation while the results of bibliometric method provide effective support. The bibliometric method with massive objective data improves the objectivity of normal composing model for technology roadmapping, at the same time expert knowledge is the assurance. Accordingly, this combination with quantitative and qualitative analysis makes the composing model more scientificalness.

4. The Case for Electric Vehicle

With the increasing energy needs and environmental pressures, electric vehicle which is environment-friendly has become the best strategic choice for the sustainable development of auto industry. After the research on global auto industry and R&D programs, we find that the technology of electric vehicle is becoming the domination of global competition in auto industry. For this reason, this paper proposes an empirical study on EV (electric vehicle), and gives a draft technology roadmapping in this field.

Fig. 2 Time-terms association map of electric vehicle

Fig. 3 Technology roadmapping of electric vehicle

This paper selects SCI and SSCI database and USPTO database as the publication and patent database. Based on the patent search strategy we constructed before and the knowledge of experts, we download all the 3944 publications from 2001 to Feb. 21st, 2010 and 4144 patents. After the process of VantagePoint, which applies a combination of machine learning, statistics and natural language processing to yield what bibliometricians call a mix of "one-dimensional" descriptions(lists) and "two-dimensional"

relationships(matrices) [21], technology core terms in the field of EV have been retrieved. In the other hand, technology core IPCs are retrieved by an automatic patent designed by Dong-hua Zhu who is also a designer of VantagePoint. Finally, some modifications are given by experts. The list is given in Table 1.

Table 1 Core terms and IPCs of EV

Core Terms	Core IPCs
electric vehicle(s)/EV(s)/applications electric vehicles, hybrid electric vehicle(s) /HEV(s)/hybrid vehicle(s), battery/batteries, lead-acid battery/lead-acid batteries/ lead/acid batteries, super capacitor(s), lithium-ion battery/lithium ion battery, energy storage, regenerative braking, control strategy, induction motor, electric propulsion, vehicle simulation, cycle life, power electronics, valve-regulated, motor drives, polymer electrolytes, alternative fuels	B60K6, B60K41, B60W20, B60K, B60L, B60T, B60W, F02D, F16H, H01M, H02K, G06F

These core terms and IPCs represent sub technologies in the field of EV, and should be the objects in the technology roadmapping. In the other hand, based on the association analysis involving with co-occurrence analysis and PCA method, the relationships of objects could be easily found and ordered in the time dimension. The time-terms association map is given in Fig. 2. Core terms and IPCs in the same year are packaged as a single object, and the relationships retrieved from publication and patent documents by text data mining technique are indicated by the links between objects. Therefore, based on the knowledge of experts in the field of EV, we hand pick useful core terms and IPCs, classify them in the correct phase and locate in the time dimension. Finally, with the location of objects, a simple clustering should be used. In this way, technology roadmapping of EV is given in Fig. 3.

In Fig. 3, the three clusters attract both us and experts. The first one is the research on materials before 1990, the second one is the research of propulsion technology from 1991 to 1995, and the third one is a booming research of battery technology starting from 1997 when a hybrid propulsion technology with both batteries and fuels appeared and the EV research was elevated into a new step. These conclusions are not only similar with the clustering we got in Fig.2, but also as the same as conclusions from some reports given by governments and businesses.

5. Conclusions and Suggestions

This paper has made some active attempts to describe the dynamic paths of technology development in a visual mapping, to realize a half automatic process to generate technology roadmapping which should be experienced in the development of automatic technology visualization, and to blend both quantitative and qualitative methodologies together. We suggested a hybrid composing model involving with bibliometric method and experts knowledge for technology roadmapping. In this model, we creatively introduced terms frequency analysis and association analysis to gather technology core terms, IPCs and their relationships, and indentify the objects and their locations in the technology roadmapping. In another hand, we also designed the three technology development layers and put them into the ordinate axis of the mapping, while the time is used to be the horizontal axis. In this way, we ensure the authority and credulity by expert knowledge, and emphasize the objectivity by bibliometric method. Moreover, an empirical study of electric vehicle is given, the comparison with some conclusions from government reports improves that the hybrid composing model for technology roadmapping is a useful tool to identify the historical development of technology.

Actually, there are still some limitations in this hybrid model for technology roadmapping. Firstly, the only way we gather technology core terms and IPCs is the terms frequency analysis with publication and patent documents, but some emerging technologies, especially the latest ones which have just appeared for a short time, could not be found in the list of high frequency terms. So low-occurrence analysis also should be noted. Secondly, the drawing of technology roadmapping is half-automatic. The bibliometric methods we used in the model are all implemented by software, and the locations of objects are hand-worked with the help of experts. In the future, the design of the software should be more intelligent.

6. Acknowledgements

7. Reference

[1] R. Phaal, Technology roadmapping—A planning framework for evolution and revolution, Technol. Forecast. Soc. Change 71 (2004) 5-26.

[2] A.L. Porter, M.J. Detampel, Technology opportunity analysis, Technol. Forecast. Soc. Change 49 (1995) 237-255.

[3] Y. Kajikawa, Y. Takeda, Citation network analysis of organic LEDs, Technol. Forecast. Soc. Change 76 (2009) 1115-1123.

[4] G. Abramo, C.A. D'Angelo, A. Caprasecca, Allocative efficiency in public research funding: Can bibliometrics help? Research Policy 38 (2009) 206-215.

[5] S.T. Walsh, Roadmapping a disruptive technology: A case study-The emerging microsystems and top-down nanosystems industry, Technol. Forecast. Soc. Change 71 (2004) 161–185.

[6] M.L. Garcia, O.H. Bray, Fundamentals of Technology Roadmapping, NM: Sandia National Laboratories Report SAND97-0665 (1998) 3-34 .

[7] K. Cuhls, A. Beyer-Kutzner, W. Ganz, The methodology combination of a national foresight process in Germany, Technol. Forecast. Soc. Change 76 (2009) 1187-1197.

[8] R.N. Kostoff, R. Boylanb, G.R. Simons, Disruptive technology roadmaps, Technol. Forecast. Soc. Change 71 (2004) 141-159.

[9] Z. M'Chirgui, Dynamics of R&D networked relationships and mergers and acquisitions in the smart card field, Research Policy 38 (2009) 1453-1467.

[10] S. Lee, Y. Park, Customization of technology roadmaps according to roadmapping purposes: Overall process and detailed modules, Technol. Forecast. Soc. Change 72 (2005) 567–583.

[11] S. Lee, B. Yoon, C. Lee, Business planning based on technological capabilities: Patent analysis fortechnology-driven roadmapping, Technol. Forecast. Soc. Change 76 (2009) 769–786.

[12] D.J. deSolla Price, Little Science, Big Science, Yale University Press, New Haven, CT, 1963.

[13] T.U. Daim, G. Rueda, H. Martin, Forecasting emerging technologies: use of bibliometrics and patent analysis, Technol. Forecast. Soc. Change 73 (8)(2006) 981-1012.

[14] F. Liu, Y. Sun, A comparison of the spatial distribution of innovative activities in China and the U.S., Technol. Forecast. Soc. Change 76 (2009) 797-805.

[15] Y. Shen, S. Chang, G.T.R. Lin, A hybrid selection model for emerging technology, Technol. Forecast. Soc. Change 77 (2010) 151-166.

[16] R.N. Kostoff, Normalization for Citation Analysis, Cortex 37 (2001) 604-606.

[17] M. Meyer, K. Debackere, W. Gla¨nzel, Can applied science be 'good science'? Exploring the relationship between patent citations and citation impact in nanoscience, Scientometrics 85 (2010) 527-539.

[18] A.H. Goodall, Highly cited leaders and the performance of research universities, Research Policy 38 (2009) 1079-1092.

[19] H. Fu, Y. Ho, Y. Sui, A bibliometric analysis of solid waste research during the period 1993–2008, Waste Management 30 (2010) 2410-2417.

[20] L. Waltman, N.J. van Eck, Ed C.M. Noyons, A unified approach to mapping and clustering of bibliometric networks, Journal of Informetrics 4 (2010) 629-635.

[21] D. Zhu, A.L. Porter, Automated extraction and visualization of information for technological intelligence and forecasting, Technol. Forecast. Soc. Change 69 (2002) 495-506.

[22] C. Choi, Y. Park, Monitoring the organic structure of technology based on the patent development paths, Technol. Forecast. Soc. Change 76 (2009) 754–768.