

# Dominant technology identification of wind power generation

Hongying Wang<sup>1,2</sup> and Bing Sun<sup>1,2,\*</sup>

<sup>1</sup>School of Economics and Management, Harbin Engineering University, Harbin, China

<sup>2</sup>Present address: No. 145, Nantong Street, Nangang District, Harbin City, Heilongjiang Province, China

Wind energy, the most commercialized prospect of renewable energy, is being developed and utilized on a large scale. Major institutions and universities have invested a lot of manpower, capital and technology in researching the wind power technology, which has made remarkable progress. In the era of green economy, the leading technology of wind power deserves more attention. Dominant technology represents the development direction of technology area. Identification of the dominant technology is of great significance for the technological choice and strategic layout of enterprises. In contrast to traditional technology identification methods, here we propose a visual analysis model based on patent citation relationship. First, the patent mutual citation data are input into the visualization software Gephi to identify the leading technology in the visual processing stage. The PageRank algorithm is used to cross contrast the technical value to build the leading technology recognition model. Second, we evaluate the technical value of each patent with five indices, including 'number of patent cited', 'patent number of the same family', 'scope of patent coverage', 'claim number of each patent', and 'number of patent litigations' to verify the accuracy of the visual model. Taking the database of wind power generation technology from European Patent Office as an example, we obtain 7421 patents from 1900 to 2017. The results of visual processing, evaluation of the index and PageRank judgment show that the visual model has a significant effect in the identification of the leading technology. The results also explain the applicability of the PageRank algorithm and the five indicators are the most scientific and reasonable for dominant technology identification.

**Keywords:** Leading technology identification, patent citation, search algorithms, visual analysis model, wind power generation.

UNDER the dual pressure of global ecological environment deterioration and depletion of fossil energy, the research and utilization of new energy has become the primary focus at present. According to the distribution and utilization of energy, with the gradual depletion of non-renewable energy and the limited hydroelectric

capacity, there are many unstable factors in power generation, so seeking new forms of power generation will be the first difficult problem of the electric power industry. At present, renewable energy wind power generation has the characteristics of early start, mature development and better industrialization; so it is expected to become the third largest power generation after thermal and hydroelectric<sup>1</sup>. Due to the several advantages of improving the ecological environment, optimizing the energy structure and promoting the sustainable development of society and economy, all countries in the world are developing and researching wind power generation and related technologies. Wind energy depends on air flow. Because the solar radiation causes uneven heating of all parts of the earth's surface, pressure distribution in the atmosphere is unbalanced. Under the action of the horizontal pressure gradient, air moves along the horizontal direction to form winds. Wind energy is an environmental-friendly and green energy source. The leading technology in this field will play a decisive role in the development of wind projects in relevant enterprises and countries.

A leading technology not only represents the direction of social development in a certain time-period, but also causes significant changes in social technology system. With the increasing level of science and technology every year, the strategy of science and technology is progressing continuously. The age of rapid development has brought technological competition among enterprises to a new height. Mastery over the leading technology will be the key factor to improve core competitiveness of enterprises. It also highlights the importance of the leading technology studies. The identification of leading technology could help enterprises make rapid decisions regarding technology selection, scientific layout of related technology patents, consolidating their own core competitiveness, and providing reference for the technology R&D strategy. However, currently there is no uniform definition of leading technology. According to Malone *et al.*<sup>2</sup>, when a technology occupies more than half the market, it holds a dominant position. The aforementioned judgment is not based on a combination of changes in technical competition, the technology which occupies more than half of the market for a long time could not become the dominant technology. The reason is that similar technology may imitate and gradually seize the market. In

\*For correspondence. (e-mail: heusun@hotmail.com)

order to solve the problem of lacking in uniform definition, Suarez<sup>3</sup> gives a more comprehensive evaluation: First, there is an obvious phenomenon that reflects the most important alternative technology withdrawing from the confrontation. Secondly, there is a distinct advantage in market share, and the recent market trend reflects the advantage in continuous strengthening. If there are one or two cases, it can be judged that a technology occupies the dominant position. Some scholars have focused on the key technology of the new energy industry. Sun *et al.*<sup>4</sup> used the Kaya equation to determine the core standards of new energy leading technologies: energy security and emission reduction. They also analysed the external and internal environments of China's new energy technology extension through the technology diffusion model. Johan *et al.*<sup>24</sup> used citation network analysis and patent analysis to analyse knowledge flows between wind firms, and to identify and compare the position and role of each firm in the knowledge network. Huming *et al.*<sup>25</sup> used the Beijing IncoShare Patent Database as well as patent maps for statistical analyses on the evolution, regional distribution, variation in type and International Patent Classification (IPC) of patent applications for offshore wind power in China, to track the development of leading technology in wind energy industry.

Several studies have been conducted on the leading technology identification methods. Lee<sup>5</sup> used text mining, network analysis and patent research to identify the role of an enterprise in technology development, to improve core competitiveness and innovation power of the enterprises. Noh *et al.*<sup>6</sup> used the network analysis to analyse the citing and cited patents, with the aim to study the technical distance of patents and the structure of patent reference network, and then identify the future key technology areas. They also developed a patent-based leading technology identification framework using 5G-telecom technology patent data to obtain the core patents published by the main organization in the field, and used literature coupling and text mining to determine its formation trajectory<sup>6</sup>. Edsand<sup>7</sup> utilized the comprehensive framework of system function and landscape factor to analyse the features that restrict the spread of wind energy technology in Columbia, which confirms the value of landscape analysis. The results showed that landscape factors should be considered as endogenous factors to promote the diffusion of energy technology<sup>7</sup>. Liu *et al.*<sup>13</sup> used of the input-output analysis method and Python programming language to present a patent citation matrix for identification of the induction coefficient and influence coefficient in the leading technical field and advanced technology field. The results showed that the frontier technology was the research and development of marine energy technology, while the leading technology was wind power generation energy technology, and that the wind power generation technology has a key role in the development of the whole new energy field<sup>8</sup>. Wang<sup>26</sup>

used Gephi to calculate PageRank value for a comparative study of leading technologies in OLED industry between China and other countries. Dalcin<sup>22</sup> used Gephi for a visualization approach of the existing network available in Brazil to achieve the targets of the Global Strategy for Plant Conservation (GSPC) for 2020 for that country. The approach uses the Brazilian National Biodiversity Strategies and Action Plans and other key documents to identify and map institutions, and roles and actions related to each of the GSPC targets<sup>27</sup>.

Most of the above-mentioned studies focus on the recognition of the core patents and core technologies, which lack the scientific and complete identification model in dominant technology. To solve this problem, visual analysis method is applied to develop a leading technology identification model, and each patent with five indices is evaluated to verify the accuracy of the model, which aims to identify the leading technology at industrial level. The present authors have conducted an empirical study taking the leading technology identification in wind power generation energy field.

### Exposition visualization model construction

Gephi is an open-source complex network analysis software based on Java Virtual Machine. It is mainly used for the interactive visualization of various networks, complex systems, dynamic and layered graphs. Developed modules can import, visualize, spatialize, filter, manipulate and export all types of networks. The visualization module uses a special render engine to render graphs in real time. This technique uses the computer graphic card, similar to video games, and thus the CPU can be used for other computing. It can deal with large networks (over 20,000 nodes) and because it is built on a multi-task model, it takes advantage of multi-core processors. Node design can be personalized; instead of a classical shape it can be a texture, panel or a photograph. Highly configurable layout algorithms can be run in real time on the graph window. For instance, speed, gravity, repulsion, auto-stabilize, inertia or size adjust are real-time settings of Force Atlas, a special force-directed algorithm developed by the present authors. Several algorithms can be run simultaneously, in separate workspaces without blocking the user interface. The text module can show labels on the visualization window from any data attribute associated with nodes<sup>8</sup>. A special algorithm named Label Adjust can be run to avoid label overlapping.

The popular visualization software currently used by researchers include Datawrapper, Tableau Public, Chart.js, Raw, Gephi, Infogram, etc. Gephi is an open-source, interactive, visualization and detection tool built on NetBeans platform for various networks and complex systems, dynamic and hierarchical diagrams. It can be used for: exploratory data analysis, link analysis, social

network analysis, biological network analysis, etc. It not only has the advantages of simple operation process, but is also more convenient for visualized graphics. The software has powerful functions, excellent visualization results and relatively simple operation. Therefore, the patent inter-citation network in this study uses Gephi for visualization processing and analysis. We identify the leading technology through visualization of patent-citation relationship. So the data used in this study come from the global patent statistics database (Worldwide Patent Statistical Database, PATSTAT), which is the main database of patent documents of the European Patent Office (EPO Master Documentation Database, DOCDB). PATSTAT aims to provide researchers with a patent database that can be run on personal computers for statistical analysis. Since its release to the public in 2007, PATSTAT has been widely used in the academic world due to its characteristics of statistical analysis and unified data compliance<sup>9</sup>. We use the retrieval expression ‘TI(Title) = wind energy or TI = wind power or TI = wind \* resource or TI = wind \* resources or TS(Title or Abstract) = wind energy or TS = wind power or TS = wind \* resource or TS = wind \* resources, and the retrieval time is 19000101-20171231’.

In this study, we choose the citing patent publication number as ‘source’ and cited patent publication number as ‘target’, so as to get the patent relationship in different stages. Since the patent data of the PATSTAT website show only 500 patents per page, and the total number of wind power generation technology patents is over 10,000 from 1900 to 2017, the author divides the data in chronological order and downloads no more than 500 patents on each page to ensure the integrity of the data. We then estimate the following indicators: number of patents cited, patent number of the same family, scope of patent coverage, claim number of each patent, and number of patent litigations. We find that some of the indicators have a value of 0. Thus, the possibility of such technologies becoming dominant is small. The patent data of three or more of the five indicators with value 0 are manually cleared to identify the dominant technology more accurately. After reprocessing, 9640 patents are selected for study, and the total number of patent applications in the wind power generation energy field is obtained (Figure 1). The data are cleaned and rearranged according to the CSV format for the Gephi data model.

From the data collected, we find that the number of citations and cited patents are different; some patents are widely cited, but there are some technologies that are recognized only in small areas.

Technologies should not become ‘technical islets’; each technology may have evolved from another, and the emergence of a technology is based on numerous existing patents. Therefore, we take the patent as a node to study the citation correlation between different patents. According to the attribute value of the data, we choose the circle

as the whole frame to arrange the data, and technologies that are more relevant to other technologies in the centre, so as to form a prototype network diagram with hierarchical structure (Figure 2).

Figure 2 shows that patents located in the peripheral positions have less connection with others; on the contrary, nodes in core positions have more connection with other nodes. Figure 2 intuitively shows a patent which in the central position contains more quoted content and has more value to study and meaning of diffusion. Figure 3 shows the citation distribution of 25 core patents.

PageRank (proposed by Sergey Brin and Larry Page) is an algorithm used by Google search to rank web pages, we have applied the same for sorting of the retrieved results. The basic idea of PageRank comes from citation analysis of traditional bibliometrics<sup>10</sup>. An enterprise patent reference network similar to web link network was developed by Shu, and he also designed the evaluation algorithm of enterprise technical influence<sup>11</sup>.

In this study, the number of link-in web pages corresponds to the number of cited patents, and the number of link-out web pages corresponds to the number of citing patents. PageRank is more advantageous in the present study than analysis of the web process; patent references do not involve loops, and there will not appear mutual citing of each other between A and B. If patent A cited patent B, it shows that the publication time of patent A is later than that of patent B, and patent B will not use patent A. Based on the advantages of this method we used PageRank algorithm to construct the model.

The PageRank algorithm can be briefly described as follows: Let a small group be composed of four pages – A, B, C, D; if all the pages are linked to A, then the PR (PageRank) of A is the total value of B, C, and D, i.e.  $PR(A) = PR(B) + PR(C) + PR(D)$ . If B can also be linked to C, and D has linked to A, B, C. Since one page has only one vote, B can only have half a ticket for each page. In the same way, D can only have three points and one vote is cast to the PageRank value of A, so that

$$PR(A) = \frac{PR(B)}{2} + \frac{PR(C)}{1} + \frac{PR(D)}{3}.$$

According to the calculation rules, the PageRank value of A could be calculated by the following formula.  $L(X)$  means the link-out pages’ number of X,

$$PR(A) = \frac{PR(B)}{L(B)} + \frac{PR(C)}{L(C)} + \frac{PR(D)}{L(D)}.$$

We used the plug-in unit of PageRank to get the ‘network rankings’ window. In order to treat all the link-out pages equally, we choose  $q = 0.85$  ( $q$  is the damping factor which is the probability of a user going to a page and continuing to browse).  $1 - q = 0.15$  is the probability that users will no longer click a new URL. The PageRank

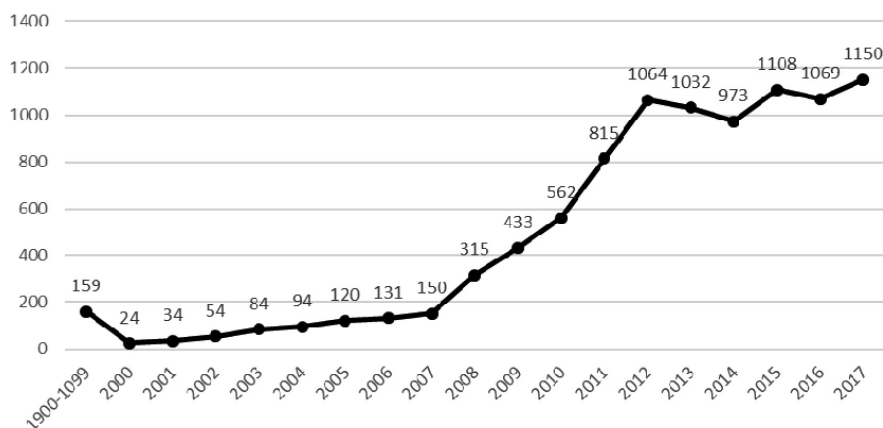


Figure 1. Annual total number of patent applications in wind power generation energy industry.

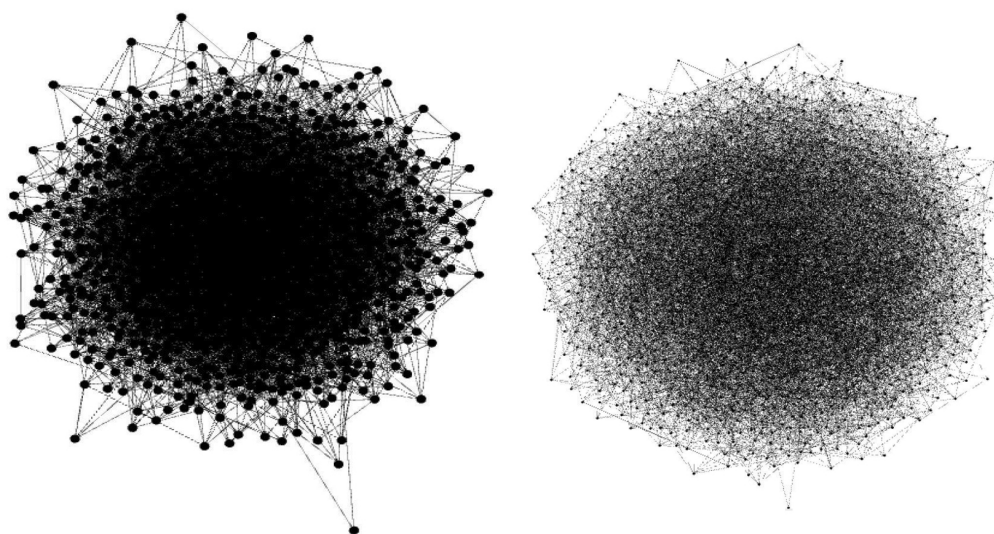


Figure 2. Patent distribution diagram in different layouts.

algorithm applies  $q = 0.85$  to all pages to reflect the probability that a page can be bookmarked on the internet. The PR value of a page without links will be 0; so Google gives the initial PR value of each page based on the following relation

$$\text{PageRank}(p_i) = \frac{1-q}{N} + q \sum_{p_j} \frac{\text{PageRank}(p_j)}{L(p_j)} \quad (1)$$

In eq. (1),  $p_1, p_2, \dots, p_i$  are the studied pages. There is a link in the network from pages  $P_j$  to  $p_i$ .  $L(p_j)$  is the number of link-out pages from  $p_j$ .  $N$  is the number of all pages. The PR value of all pages is a feature vector in a special matrix, which is shown in eq. (2)

$$R = \begin{bmatrix} (1-q)/N \\ (1-q)/N \\ \vdots \\ (1-q)/N \end{bmatrix} + q \begin{bmatrix} l(p_1, p_1)l(p_1, p_2) \cdots l(p_1, p_N) \\ l(p_2, p_1) \\ \vdots \\ l(p_N, p_1) \end{bmatrix} \quad (2)$$

Therefore, the PR value of a page is calculated by the PR values of other pages, and the PageRank algorithm constantly repeats the PR value of each page. If each page is given a random PR value (not 0), then the PR values of these pages tend to be normal and stable after repeated calculations.

Figure 4 shows the results of PR distribution. Due to the large number of nodes, the PageRank score is close to 0. Based on the results, 20 leading patents are obtained through cross-comparison between the top 25 patents, and the visualization results (Table 1).

### Construction of index evaluation system

In order to verify the recognition results of visual models, we used the index system to comprehensively evaluate each patent. After a comprehensive comparison, we chose five indicators, including ‘number of patents cited’, ‘patent number of the same family’, ‘scope of patent coverage’, ‘claim number of each patent’, and ‘number of

patent litigations' for recognition. The specific reasons are as follows.

(1) Number of patents cited (NPC): In the current patent evaluation system, the number of times cited is the highest degree of identification. In general, the more number of times a patent is cited, the more important it is in the technical field<sup>12</sup>. In addition, Li<sup>13</sup> points out that the number of times a patent cited is more important than the number of patents; it is a symbol of the technological competitiveness of enterprises<sup>13</sup>. Albert *et al.*<sup>14</sup> reports that the number of patents can be a good reflection of its importance, because the patent uses core technology when it is invented; so the number of patents can be used to measure whether it is a core patent or not.

(2) Country number of patented family (CNPF): This refers to the number of applications submitted by a patent to different countries. Existing research indicates that more the family number, greater is the cost of patent input, commercial strategic significance and market value. Dietmar *et al.*<sup>15</sup> found that the number of family patents is effective when evaluating patent indicators. Markus *et al.*<sup>16</sup> showed that the size of the patent family could effectively reflect the technical quality of the patent. Schettino *et al.*<sup>17</sup> concluded that the size of the patent family can reflect the importance of a technology while studying the European patent data.

(3) Scope of patent coverage (SPC): Generally, patent coverage refers to content of a technology and the technical scope covered. Petruzzelli *et al.*<sup>18</sup>, found that larger the patent coverage, higher is the number of patent citations from the follow-up technology; it presents an inverted U-type relationship. Lerner<sup>19</sup> used the first four different IPC classification numbers to indicate the scope

of patent technology. The results showed that wider the scope of a technology, the more advanced the innovation, and better quality of the patent.

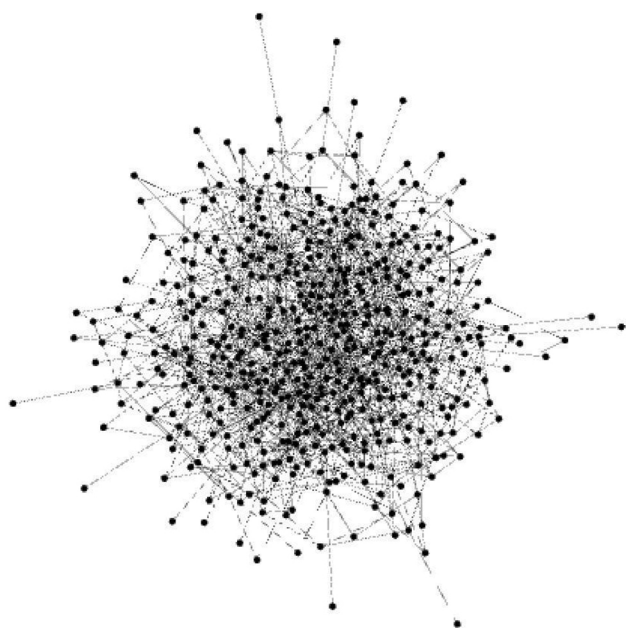
(4) Number of patent claims (PN): This index represents the technical content and technical level of a patent. More the number of patent claims, wider is the scope of protection of a technology and better is the quality of the patent. After studying the relationship between technological creativity and the number of claims, Tong and Davidson<sup>20</sup> found that the number of claims has an important impact on innovation ability. The authors randomly selected 7531 patents from the USPTO patent database and found that more the number of rights in the patent literature, better is the technological innovation of the patent<sup>20</sup>.

(5) Number of patent litigations (PL): Patent litigations refers to all litigations concerning patent disputes, and patents involved in dispute cases have a more important position in the technology field. Elettra and Rossella<sup>21</sup> used patent litigation index in patent value evaluation which could effectively identify core patents. According to Cremers<sup>22</sup> more valuable the patents, the more likely that they will encounter lawsuits.

From different identification methods, single-index recognition method could quickly identify important patents; it is easy to operate, but the recall and accuracy of patents are poor. In comparison, the combination-index identification method is a more scientific and comprehensive assessment of each patent and the value of the index can be evaluated comprehensively, so as to identify the dominant patent. The five indicators mentioned above play an important role in measuring patent value. First, they cover the four aspects of patent quality influence, area protection range, technical protection range and technical level, which measure the patent value in a more comprehensive way. Second, the indicators are the most frequently used in the index system of OECD, CHI Research and the National Intellectual Property Office of China. Third, the patent statistics database (PATSTAT) could provide citation index data of the global patent, which include the five indices selected in this study; so it is feasible to use these indicators in the evaluation system. Therefore, we have chosen these indicators for a more comprehensive and scientific operational measurement of patent value.

Based on the mentioned mark-up principles and cited number, 20 dominant technologies were chosen, to compare with the results of visual model recognition. The results show that 17 patents are identified as the dominant technology, the accuracy rate reaching 85%. Compared to other recognition models, the visualization model achieves higher accuracy<sup>28-30</sup>.

'Misjudgement' and 'missed judgement' during identification of the leading technology are possible. When analysing the effect of recognition, the authors not only guarantee the accuracy of identification, but also consider the rate of misjudgement and missed judgement.



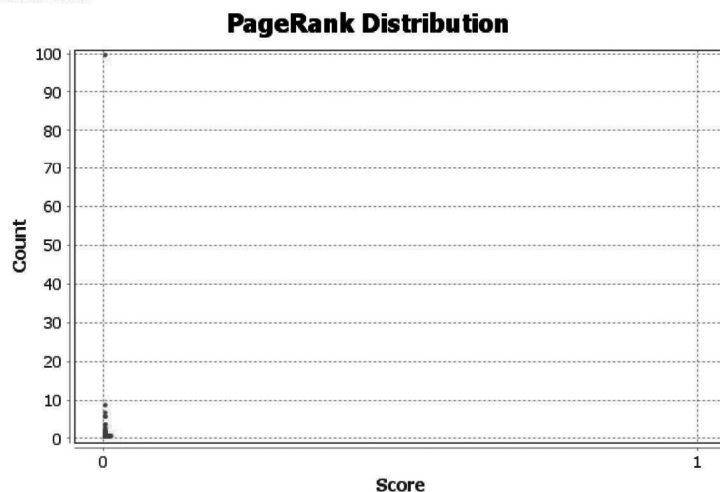
**Figure 3.** Citation distribution diagram of 25 core patents.

## PageRank Report

### Parameters:

Epsilon = 0.001  
Probability = 0.85

### Results:



### Algorithm:

Sergey Brin, Lawrence Page, *The Anatomy of a Large-Scale Hypertextual Web Search Engine*, in Proceedings of the seventh International Conference on the World Wide Web (WWW1998):107-117

Figure 4. PageRank distribution.

Table 1. Visual model results of leading technology recognition

Publication number	Core position ranking	Visual recognition results	PageRank ranking
US7215037 (B2)	1	0.00756378	1
US4832569 (A)	2	0.007026909	2
US5272378 (A)	3	0.006971327	3
US2004183307 (A1)	4	0.006867114	4
US4140433 (A)	5	0.006434224	5
US4303835 (A)	6	0.006140502	6
US5225712 (A)	7	0.00611604	7
US4720640 (A)	8	0.00606281	8
US5182458 (A)	9	0.006047503	9
US6984899 (B1)	10	0.005971123	10
CN101793235 (A)	11	0.005945417	11
CN101943127 (A)	12	0.005793182	12
JP2006183755 (A)	13	0.005732004	13
JP2008063960 (A)	14	0.0055961	14
JPS56150999 (A)	15	0.005328607	15
US2004086373 (A1)	16	0.005313577	16
US2006028025 (A1)	17	0.00521143	17
US2007228838 (A1)	18	0.005195687	18
US2009085354 (A1)	19	0.005139513	19
US2012201679 (A1)	20	0.004722212	20

There are misjudgements in the model proposed here, mainly reflected in the high number of cited numbers but lower number of subordinates. For example, the total cited number of ‘CN101280942(A)’ is ranking tenth, but

it is in the periphery of the visual circular section and the PR value ranks behind. CN101280942(A) is not identified as the dominant technology in the model. There are three similar cases in the data sample. The reason is that

the visual model only tracks the technical reference and does not take into account the coverage of a single patent, the number of litigations, family size and so on; so the attribute of the visual model determines the unavoidable misjudgment. Table 2 shows that the recognition effect of the model is worthy of recognition, which could recognize 17 dominant technologies. Hence the correct rate is far greater than the missed rate. This also confirms the results obtained by Ruimin's research; it has obtained 77.27% accuracy rate in the process of recognition<sup>23</sup>. The PR algorithm model proposed could achieve the recognition result of 85% accuracy rate, thus the scientific character of the model can be judged.

### Conclusion and countermeasures

This study uses visual analysis to develop a leading technology identification model with wind power energy generation as an example. We can make the following conclusions. First, it is scientific and reasonable to construct the leading technology identification model by five indicators mentioned earlier in the text. The reason is that these indicators may affect the number of cited patents, they are independent of one another and will be able to measure the influence of a patent in a more comprehensive way. Second, the visual model is more intuitive than the other discriminant methods in the identification of dominant technology. The PageRank algorithm can be applied to the patent mutual citation problem to calculate the patent value more accurately. We consider the accuracy and omission rate comprehensively through empirical analysis. The recognition effect of the visual model in

the leading technology identification is notable. So the model can effectively identify the dominant technology.

In view of the above leading technology results identified, we provide some countermeasures and suggestions for the development of the wind industry.

First, from the perspective of researchers, it should strengthen patent-related work to improve the quality of patent writing and actively carry out research on the new generation of wind technology. The patent data show that the contents and writing quality of patent applications need to be further improved in some developing countries. The technicians could open thoughts on the basis of existing technology and develop new fields in the wind industry.

Second, from the perspective of relevant enterprises, it should build an interdisciplinary R&D team and achieve technological breakthroughs through school–enterprise cooperation. Wind power technology covers many fields, such as architecture, materials, physics, and so on. Therefore, in the initial stage of each project, the related enterprises must pay attention to the problem of cross-discipline and recruit personnel from different scientific fields, for a cross-disciplinary R&D team.

Third, from the perspective of the Governments, it should promote industrialization of technology and strengthen the construction of industry–university research alliance. The governments could support research and development of the follow-up technology, common technology and future technology of wind power through projects and innovation systems. At the same time, they should promote the development and large-scale industrialization of wind industry by supporting internationalization of talents and the all-round integration of enterprises.

Fourth, from the perspective of patent protection, researchers should avoid intellectual property risks while raising the awareness of patent protection. Both research institutions and enterprise developers should strengthen the protection of new achievements, such as the establishment of rights defense institutions in organizations, and establishment of intellectual property system, including patents and standards. In addition, in the process of product development, production and sales, intellectual property rights (IPRs) must be considered. Enterprises in developed countries attach great importance to the protection of IPRs, while developing country enterprises neglect IPR risks for various reasons.

The limitations of this study are as follows: (1) Identification of dominant technology only depends on patent data and has not been verified by the actual wind energy industry. (2) The dominant technology identification model only uses one technology for verification, and thus lacks reliability. In future, we need to analyse the wind energy industry and further verify the accuracy of leading technology identification. We can also apply this method to other industries, with suitable modifications.

**Table 2.** Comparison of index evaluation results and visual results

Publication number	Cited numbers	Index evaluation results ranking	Visual results
US4832569 (A)	72	1	√
US4303835 (A)	70	3	√
US7215037 (B2)	69	2	√
US2004183307 (A1)	66	6	√
US4140433 (A)	63	4	√
US2006028025 (A1)	52	7	√
US5225712 (A)	51	5	√
US4720640 (A)	51	11	√
US5272378 (A)	49	9	√
CN101280942 (A)	47	8	
US6984899 (B1)	45	12	√
US2004086373 (A1)	40	13	√
US7015595 (B2)	40	10	
CN101943127 (A)	40	14	√
JPS56150999 (A)	37	18	√
JP2008063960 (A)	34	17	√
JP2006183755 (A)	32	15	√
WO2004047284 (A1)	31	16	
CN101793235 (A)	31	19	√
US5182458 (A)	25	20	√

1. Li, J. *et al.*, Current situation and prospect of wind power generation in China. *China Sci. Technol. Invest.*, 2007, **11**(1), 25–28.
2. Malone, E., Hultman, N. E., Anderson, K. L. and Romeiro, V., Stories about ourselves: how national narratives influence the diffusion of large-scale energy technologies. *Energ. Res. Soc. Sci.*, 2017, **31**(9), 70–76.
3. Suarez, F. F., Battles for technological dominance: an integrative framework. *Res. Policy*, 2004, **39**(7), 271–286.
4. Sun, M., Gao, C., Jia, C., Ni, F. D. F. and Zhang, J., The selection and promotion of core technology to China's energy goals. *Energ. Proc.*, 2016, **104**(3), 233–238.
5. Lee, H., Analysis and citation analysis of patents by social network analysis. *Telecommun. Policy*, 2007, **27**(2), 293–312.
6. Noh, H., Song, Y. and Lee, S., Identifying emerging core technologies for the future: case study of patents published by leading telecommunication organizations. *Telecommun. Policy*, 2016, **40**(10), 956–970.
7. Edsand, H., Identifying barriers to wind energy diffusion in Colombia: a function analysis of the technological innovation system and the wider context. *Technol. Soc.*, 2017, **49**(5), 1–15.
8. Bastian, M., Heymann, S. and Jacomy, M., Gephi: an open source software for exploring and manipulating networks. *Int. AAAI Conf. Weblogs Soc. Media*, 2009, **10**(2), 361–362.
9. Zhang, J., Yang, G. C. and Liu, H. J., Review of the global patent statistics database (PATSTAT). *Digit. Lib. Forum*, 2015, **12**(5), 63–65.
10. Narin, F., Patent bibliometrics. *Scientometrics*, 1994, **30**(12), 147–155.
11. Wang, S. and Wang, J., Evaluation of technological influence based on enterprise citation network. *Sci. Res.*, 2011, **29**(11), 396–402.
12. Carpenter, M. P., Narin, F. and Woolf, P., Citation rates to technologically important patents. *World Patent Inf.*, 1981, **4**(2), 160–163.
13. Li, Y. A., Borders and distance in knowledge spillovers: dying over time or dying with age? – evidence from patent citations. *Eur. Econ. Rev.*, 2014, **71**(10), 152–172.
14. Albert, M. B., Avery, D., Narin, F. and Mcallister, P., Direct validation of citation counts as indicators of industrially important patents. *Res. Policy*, 1991, **20**(3), 251–259.
15. Dietmar, H., Frederic, M. S. and Katrin, V., Citations, family size, opposition and the value of patent rights. *Res. Policy*, 2003, **32**(8), 1343–1363.
16. Markus, R., Joachim, H. and Christopher, H., On sharks, trolls, and their patent prey – unrealistic damage awards and firms' strategies of 'being infringed'. *Res. Policy*, 2007, **36**(1), 134–154.
17. Schettino, F., Sterlacchini, A. and Venturini, F., Inventive productivity and patent quality: evidence from Italian inventors. *J. Policy Model.*, 2008, **35**(6), 1043–1056.
18. Petruzzelli, A. M., Rotolo, D. and Albino, V., Determinants of patent citations in biotechnology: an analysis of patent influence across the industrial and organizational boundaries. *Technol. Forecast. Soc. Change*, 2015, **91**(2), 208–221.
19. Lerner, J., The importance of patent scope: an empirical analysis. *RAND J. Econ.*, 1994, **25**(2), 319–333.
20. Tong, X. S. and Davidson, F. J., Measuring national technological performance with patent claims data. *Res. Policy*, 1994, **23**(2), 133–141.
21. Elettra, A. and Rossella, A., An application of fuzzy methods to evaluate a patent under the chance of litigation. *Exp. Syst. Appl.*, 2011, **38**(10), 13143–13143.
22. Cremers, K., Settlement during patent litigation trials. An empirical analysis for Germany. *J. Technol. Trans.*, 2009, **34**(7), 182–195.
23. Ma, R. and Wei, X., Research on core patent prediction from the perspective of technology subdivision. *J. Infor. Sci.*, 2017, **12**(5), 1279–1289.
24. Johan, N., Yuan, Z. and Xiao, Z., Innovation core, innovation semi-periphery and technology transfer: the case of wind energy patents. *Energy Policy*, 2018, **120**(9), 213–227.
25. Huiming, Z., Yu, Z. and Dequn, Z., Selection of key technology policies for Chinese offshore wind power: a perspective on patent maps. *Mar. Policy*, 2018, **93**(7), 47–53.
26. Bing, S. and Hongying, W., Comparative study on Chinese and global OLED industry based on patent data. *IEEE Access*, 2018, **12**(6), 72381–72391.
27. Eduardo, D. and Wyse, J., A network-wide visualization of the implementation of the global strategy for plant conservation in Brazil. *Rodriguésia*, 2018, **4**(10), 1613–16391.
28. Abdolreza, M. and Katja, R., Identification and monitoring of possible disruptive technologies by patent-development paths and topic modeling. *Technol. Forecast. Soc. Change*, 2016, **104**(3), 16–29.
29. Chulhyun, K., Hakyeon, L. and Hyeonju, S., Identifying core technologies based on technological cross-impacts: an association rule mining (ARM) and analytic network process (ANP) approach. *Exp. Syst. Appl.*, 2011, **38**(10), 12559–12564.
30. Hyojeong, L. and Yongtae, P., Identification of technological knowledge intermediaries. *Scientometrics*, 2010, **84**(3), 543–561.

ACKNOWLEDGEMENTS. The work was supported by the National Natural Science Foundation of China under Grant 71774035 and Grant 71372020.

Received 14 December 2018; revised accepted 1 February 2019

doi: 10.18520/cs/v116/i9/1525-1532