Technological Forecasting & Social Change 77 (2010) 1147–1155

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

Extracting the commercialization gap between science and technology — Case study of a solar cell

Naoki Shibata [a,*], Yuya Kajikawa [a], Ichiro Sakata [b]

[a] Innovation Policy Research Center, School of Engineering, The University of Tokyo, #9-201, 2-11-16 Yayoi, Bunkyo Ward, Tokyo 113-8656, Japan
[b] Todai Policy Alternative Research Institute, The University of Tokyo, 7-3-1 Hongo, Bunkyo Ward, Tokyo 113-0033, Japan

ARTICLE INFO

Article history:
Received 30 September 2009
Received in revised form 6 March 2010
Accepted 19 March 2010

Keywords:
R&D management
Technology roadmap
Research front
Bibliometrics
Citation analysis
Patent analysis

ABSTRACT

In this paper, we compared structures of the citation network of scientific publications with those of patents, and discussed the differences between them. A case study was performed in a solar cell to develop a method of detecting gaps between science and technology. Scientific research has tended to be more basic, especially in terms of cell design, whereas patents have focused on more applied technology used in solar cell modules. Of the major citation clusters of scientific publications, only two, namely silicon and compound solar cells, corresponded semantically with patent clusters. Conversely, there were no patent clusters corresponding to the other two scientific research fronts, namely dye-sensitized and polymer solar cells. These research areas could be regarded as opportunities for industrial commercialization because scientific activities exist but not technological applications. Our results could offer an intellectual basis for discovering potential opportunities for industrial commercialization.

1. Introduction

It is widely accepted that basic research in science provides a fundamental basis for technology-oriented innovations. There are three main layers in the technology-oriented innovation processes; science, technology and industry. In this model, scientists create the seeds of innovation, companies take up these seeds, develop technologies, and then industrialize. Although such a linear model is often criticized [1,2], technological inventions tend to be strongly connected to material in scientific outputs [3]. For example, Mansfield estimates 10% of new products and processes would not have been developed in the absence of recent academic research [4]. Moreover, as the innovation cycle shortens, the linkage between technology and basic science is increasing [5,6].

In today's increasingly global and knowledge-based economy, competitiveness and growth depend on the ability to keep pace with the seeds of innovation in science and swiftly develop technological applications as also emphasized by Kessler and Chakrabarti [7]. Watts and Porte illustrate a way for “innovation forecasting” by combining technological trends, interdependencies and competitive intelligence [8]. Walsh indicated the model of roadmapping for disruptive technologies focusing on the fundamental differences between sustaining and disruptive technologies [9]. Daim et al. showed a method to create forecasting by bibliometrics and patent analysis [10]. Kostoff and Schaller also emphasized the significance of not only understanding science and technology respectively, but also constructing high-quality roadmaps to integrate both [15]. Therefore, for R&D managers and policy makers focusing on future technology, understanding the relationship between science and technology has become a key task. This understanding helps not only the interaction between R&D and marketing [11–13] but also the uncertainty reduction [14], which are significant for technological innovations. This relationship can be categorized into four types based on the existence of academic papers and patents as shown in Fig. 1(a). Some technology is in area A, where science and technology coevolve. In those areas, we expect science and technology to have a reciprocal influence while the intensity of

* Corresponding author. Tel./fax: +81 3 5841 7672.
E-mail addresses: shibata@ipr-ctr.t.u-tokyo.ac.jp (N. Shibata), kaji@ipr-ctr.t.u-tokyo.ac.jp (Y. Kajikawa), isakata@ipr-ctr.t.u-tokyo.ac.jp (I. Sakata).

0040-1625/$ – see front matter
doi:10.1016/j.techfore.2010.03.008

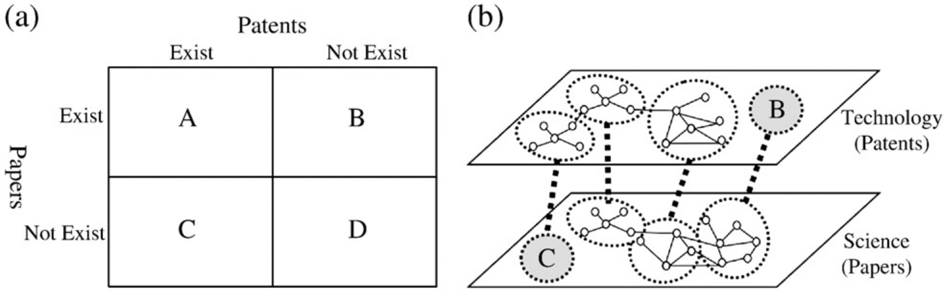

Fig. 1. Relationships between science and technology. (a) Four types of relationship; (b) gap between science and technology.

integration varies depending on the level of science linkage. In B, although academic papers exist, there are hardly any patents, whereas the opposite applies to C and D is an area to which neither scientists nor engineers pay attention. The topics in C could be regarded as overly applied research or reflect a possible lack of science. Conversely, although topics in B might be too basic, they could represent chances or opportunities for industry in the near future. Therefore, investigating the correlative relationship between science and technology is significant in enabling the detection of the gap between both layers, like area B as shown in Fig. 1(b).

How can we extract the relationship between the scientific outcomes and the pieces of industrial technology? One straightforward method is to analyze the science linkage, which is a citation from a patent to a non-patent reference such as a scientific paper. Although the number of science linkages supports the hypothesis that public science pushes technology development [5,6], measuring the science linkage in order to track the relationship between academic papers and patents is imperfect. The first problem is the differences among countries in terms of how non-patent references are cited [16]. The United States Patent and Trademark Office (USPTO) requires applicants to indicate the scientific sources related to the technology on the front page. Moreover any lack of such scientific references could be a reason for rejection. In contrast however, the Japan Patent Office (JPO) does not necessarily require the same. These differences could be problematic if we analyze the patent data globally. Secondly, especially in the US, citations from patents to papers might not always be intended by the authors themselves. To avoid rejection caused by a lack of citations to related papers, agents such as attorneys tend to add more citations to related scientific research. The final problem is the semantic similarity between a patent and a paper cited by the same. Meyer pointed out that the contents of patents are not always related to the paper [17].

The other method involves measuring the relationship based on the semantic similarity of topics in both layers. Before comparing the topics in science and technology, we have to extract the representative topics discussed in each layer. To categorize the documents and extract the topics, there are two main methods: text-mining and citation-mining. Kostoff et al. extracted the taxonomic structure of energy research by analyzing multi-word phrase frequencies and phrase proximities [18]. The latter approach, namely citation-based, is widely applied to investigate the structure of knowledge and it is assumed that citing and cited documents have similar topics, even if the document is an academic paper or a patent. For instance, Small investigated the co-citation network of academic publications in order to track the emerging growth of a certain research area [19], while Mowery et al. explored the citation pattern among patents to measure the interfirm knowledge transfers within strategic alliances [20]. In this paper, we applied a citation-based method, which is appropriate to detect the emerging topics [21]. Using the topological clustering method, we categorized the documents (papers/patents) in each layer, compared the topics of each group, and detected the gap between science and technology. The aim of this paper is to develop a methodology to discover research areas where scientific activities exist, but not technological applications, by investigating the corresponding relationships between academic papers and patents. The value of this process is emphasized by illustrating the case of the solar cell. Our results could offer an intellectual basis for detecting possible opportunities for industry.

2. Methodology and case study

2.1. Methodology

Our basic concept to extract the gap between science and technology consists of two phases. The first is to classify the documents (academic papers and patents) into clusters in each layer respectively, as shown in Fig. 2(1) to (6). This method was reported as a means of detecting an emerging research knowledge domain [21]. As the second phase, we compared the topics of all major clusters in the science layer to those in the technology layer, and extracted the gap with experts of the research domain as shown in Fig. 2(7). The clusters of academic papers, which are young and do not correspond to any clusters of patents, could have an industrial potential because these academic research projects could potentially be industrialized. Experts are only required for steps (6) and (7), while the rest could be automatically calculated by computer.

In the first phase, step (1) involves collecting the documents of each layer within a certain domain. We searched databases of papers and patents using the same query in order to determine the latter. The databases of academic papers used were the Science Citation Index Expanded (SCI-EXPANDED), the Social Sciences Citation Index (SSCI), and the Arts & Humanities Citation Index (A&HCI) compiled by the Institute for Scientific Information (ISI), because SCI-EXPANDED, SSCI, and A&HCI are three of the best

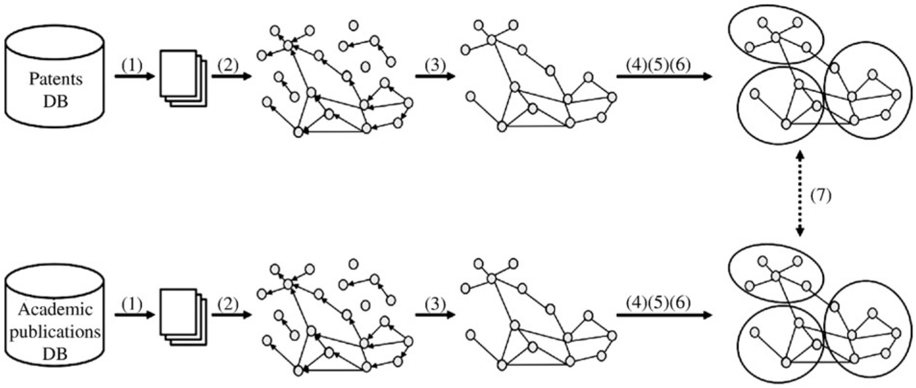

Fig. 2. Proposed method in this paper.

sources for citation data. For the patent data, we collected patents and citation data from the database named "FOCUST-J" provided by Wisdomain Inc., which covers the patents of US Granted, US Applications, European Applications, European Granted, PCT Publications, Chinese Patent Publications, and Chinese Utility Publications. These databases enabled us to obtain both the attribute data of each document such as the year published, title, author(s), abstract and so on, and relational data, i.e. citation data.

Subsequently, as step (2), we constructed citation networks by regarding papers/patents as nodes and intercitations as links. The network created in each year facilitates the time-series analysis of citation networks. According to a previous study, intercitation, which is also sometimes known as direct-citation, is the best way to detect emerging trends [22]. In network analysis, only the data of the largest-graph component was used, because this paper focuses on the relationship among papers, and we should therefore eliminate those not linked with any other papers in step (3).

After extracting the largest connected component, in step (4), the network was divided into clusters using the topological clustering method [23], which is not fuzzy. Newman's algorithm discovers tightly knit clusters with a high density of within-cluster edges, which enables the creation of a non-weighted graph consisting of many nodes. After clustering, we visualized the citation networks and named the major clusters of emerging topics as steps (5) and (6), respectively. In step (5), in order to visualize citation maps, we used a large graph layout (LGL), an algorithm developed by Adai et al. [24], capable of dynamically visualizing large networks in the order of hundreds of thousands of nodes and millions of edges, and which applies a force-directed iterative layout guided by a minimal spanning tree of the network to generate two- or three-dimensional coordinates for the nodes. We visualized the citation network by expressing inter-cluster links in the same color, based on which the clusters are intuitively understood. In step (6), experts of the research domain assigned the name to each cluster manually after the titles and abstracts of the papers/patents in each cluster. After the documents in each layer were independently classified into clusters, experts figured out the corresponding relationships between the clusters in the science and technology layer. Given the bibliographic information contained in all major clusters, such as the titles and abstracts of highly-cited papers, experts were requested to determine the paper cluster(s) which dealt with similar topics to a given patent cluster semantically.

In the second phase, experts compared the topics of all major academic clusters to those in the technology layer, as step (7). Where no patent cluster corresponds to a certain academic cluster, the research topic of the latter could be considered a seed of innovation for the near future. Conversely, if a certain patent cluster does not correspond to any paper clusters, the technology of the former may also be applicable.

In sum, the first phase, corresponding to steps (1) to (6), classifies the documents (academic papers and patents) into clusters in each layer respectively. The second phase is the comparison the topics of all major clusters in the science layer to those in the technology layer as of step (7). Throughout these steps, the scientific areas, which are emerging and not commercialized yet, could be retrieved.

2.2. A case: solar cell

Sustainable and renewable energies have been widely accepted as a key concept for our common future [25]. A solar cell or photovoltaic cell, which is a device that converts solar energy into electricity via the photovoltaic effect, represents a promising research front for our future sustainable ecosystem. For instance, previous research investigated the citation networks of 152,514 academic publications published in the 68 journals classified by the Science Citation Index (SCI) under the category Energy and Fuels and found that one of the emerging and large research domains was solar cells [26].

3. Results

In this section, the structures of knowledge in both science and technology layers are shown in Sections 3.1 and 3.2, respectively. Through this process, the major clusters and their names could be extracted. Subsequently, in Section 4, these structures are compared in order to detect the gap between science and technology.

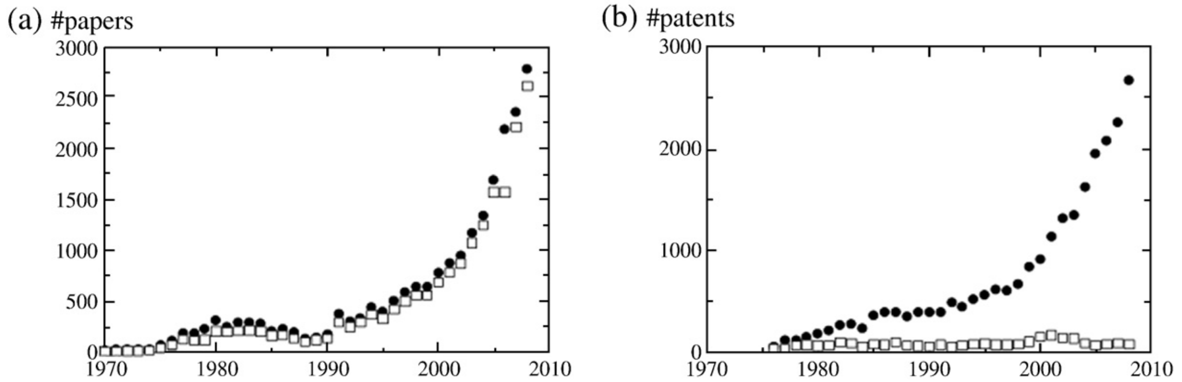

Fig. 3. Annual number of (a) papers and (b) patents including "solar cell" in the title or abstract. White rectangles indicate the number included in the largest connected component in 2008 and black circles the total.

Historically, there has been abundant academic research and technological knowledge about solar cells, hence for this paper, we searched the papers and patents using the term "solar cell*" as the query. Consequently, we obtained the data of 19,063 papers and 10,694 patents published up to 2008 and the number of annual publications is shown in Fig. 3, which corresponded to steps (1), (2) and (3). In Fig. 3, white rectangles indicate the number included in the largest connected component in 2008 and black circles the total.

3.1. Result of analysis in the science layer

After the largest connected component was constructed, step (4) involved academic papers being divided into specific clusters using the topological clustering method. Fig. 4 is a plot of the sizes and average ages of each cluster, while Fig. 5 shows the chronological evolution of each cluster from 1998 to 2008. For instance, in 2008, there were four main clusters, two of which (S4 and S5) were typically young and the other two older. As shown in Fig. 4, an emerging cluster S1, which included 1415 papers and which was 3.3 years old, appeared in 2003. Fig. 5 shows that S1 evolved and was divided into S2 (2247 papers and 2.7 years old) and S3 (1404 papers and 3.7 years old) in 2006 and that S4 (3604 papers and 2.9 years old) and S5 (2876 papers and 3.0 years old) were the outcome of the evolution of S2 and S3, respectively. However, the semantic similarities of papers included in the S4 and S5 clusters were modest, despite originating from the same cluster S1. In other words, the majority of papers in S4 (S5) in a certain year will also reappear in S4 (S5) in subsequent years from 2006 to 2008. This result of parent–child relationship of clusters shown in Fig. 5 also indicates that there appeared an emerging research area, which was S2 and became S4.

After clustering, as step (5), citation networks were visualized as Fig. 6. Fig. 6(a) is the citation network in 2008, where S4 was colored yellow and S5 white. Fig. 6(b) shows the timeline evolution of citation networks from 1999 to 2008. The color and position were calculated based on the data of 2008 and then fixed in other years. S4 (yellow) and S5 (white) were typically emerging clusters, while the presence of traditional clusters colored orange, sky blue and blue was also noted.

Focusing on the visualization in 2008, four major clusters emerged, each containing more than 1000 papers. Clusters #1, #2, #3, and #4 contained 6937, 3853, 3604, and 2876 papers, which were published in 1996.9, 1999.6, 2005.1, and 2005.0 on average, respectively. Of the four, two clusters, #3(S4) and #4(S5), were so young that they could be regarded as emerging research fronts at that time. The final step here, namely step (6), involved solar cell experts deciding the topics of each cluster. Semantic information such as the titles and abstracts of highly-cited academic papers in each cluster were given to experts, who subsequently decided the name of each cluster as shown in Fig. 7. Traditionally existing clusters #1 and #2 were silicon- and compound-based solar cells respectively, while emerging clusters #3(S4) and #4(S5) dealt with dye-sensitized and polymer

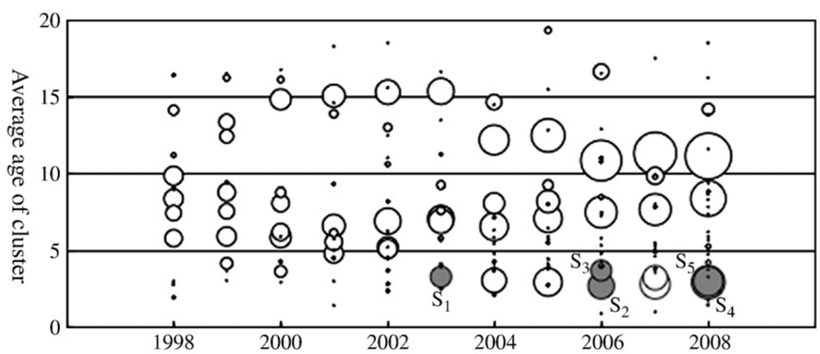

Fig. 4. Cluster size and average age of academic papers related to solar cell.

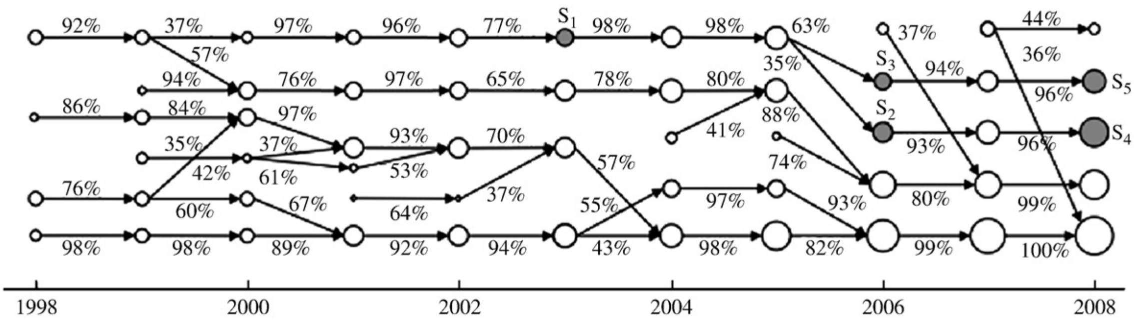

Fig. 5. Timeline visualization of the development of each cluster of academic papers related to solar cell.

related topics, respectively. To sum up, as of 2008, there were four major clusters, two of which, silicon and compounds, had a long research history, while the remaining two, dye-sensitized and polymer, could be regarded as emerging topics.

3.2. Result of analysis in the technology layer

The same procedure was applied to the citation network of patents. There were seven major clusters, containing more than 100 patents, as shown in Table 1 and Fig. 8, while the topics shown in both the latter were also named by experts. Clusters #1, #2, #3, #4, #5, #6, and #7 represented the cell design of silicon-based solar cells (517 papers, published in 1991.6 on average), metal contacts (438 papers, published in 1991.7), battery and converter (436 papers, published in 1994.9), panel (368 papers, published in 1994.9), diode (223 papers, published in 1994.4), CIS (149 papers, published in 1994.6), and solar concentrator (104 papers, published in 1988.7), respectively.

3.3. Correspondence relationships between the clusters in the science and technology layer

The correspondence relationships are drawn in Fig. 9. Of seven major clusters, two were significantly similar to academic clusters: #1 in the technology layer "cell design of silicon" corresponded to #1 "silicon" in science and #6 "CIS" did to #2 "compounds". There was no semantic relationship among the other clusters. In the technology layer, the topics discussed in the other five clusters (#2, #3, #4, #5, and #7) were applied technologies rather than scientific outcomes, whereas the other two (#3 and #4) in the science layer pursed a scientific outcome to design the cells for future solar cells.

4. Discussion

First of all, let us discuss the potential to discover opportunities for industry. In the case of the solar cell, scientific research has tended to focus on more basic research, such as cell design, while patents focus on more applied technology, including the components used in solar cell modules. The fundamentals are shown in Fig. 10. As shown in the above, all major clusters in the science layer focused on the cell design of the solar cell. There were four main types classified in terms of the materials used in the cell: silicon, compounds, dye-sensitized, and polymer. Of these, not only have silicon and compound solar cells already been commercialized, but they also correspond to some patents, which focus on the cell design of these two types of solar cell. Conversely, the other two topics, dye-sensitized and polymer related solar cells, have yet to be commercialized. Since no

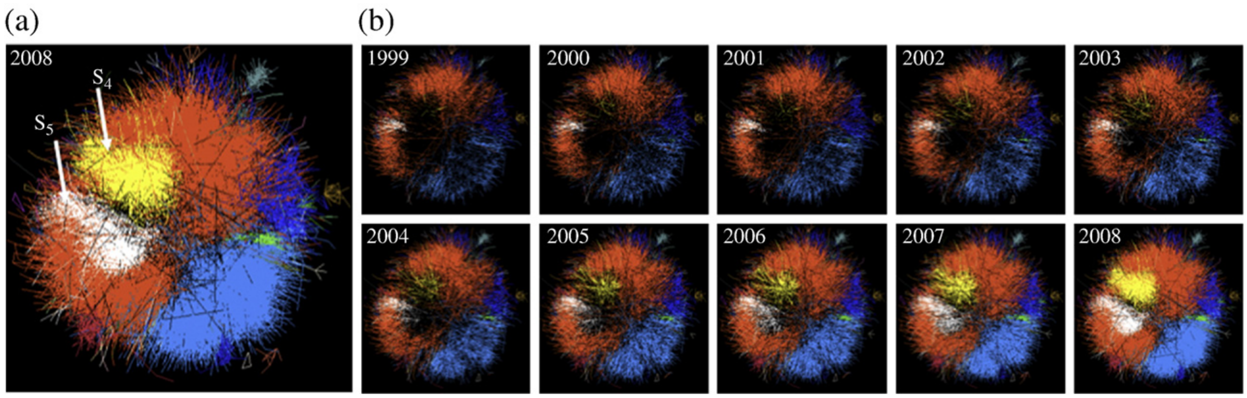

Fig. 6. Visualization of the evolution of a citation network of academic papers related to solar cell.

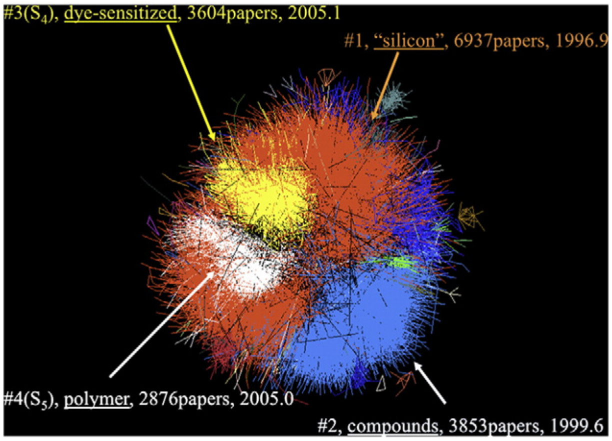

Fig. 7. Visualization of a citation network of academic papers related to solar cell with topics in 2008.

significant patent clusters corresponded to these two, huge opportunities still remain to industrialize these two materials for more efficient solar cells in future. Our approach could, as a minimum, provide a new viewpoint to develop a technology management strategy (MOT) to determine the domain in which to focus and invest.

Our results are partially supported by a technology roadmap, the Photovoltaics (PV) Roadmap toward 2030 (PV2030), established by the New Energy and Industrial Technology Development Organization (NEDO). NEDO is responsible for R&D project planning and formation, project management, and post-project technology evaluation in Japan and established the PV2030 as a long-term strategy for PV R&D in 2004 [27]. In PV2030, although the target is the crystalline Si solar cell, which has the highest PV market share and also high-efficiency compound-based solar cells, as GaAs-based or CuInSe2 (CIS), dye-sensitized cells were the subjects of discussion as post-2010 seed research. It was also pointed out that research on dye-sensitized solar cells is focusing on the improvement in cell performance to a conversion efficiency of 15% through the development of new dyes and advanced cell structures, as well as that of production technology for large-area modules with integrated circuits on various substrates. This suggests that room exists for dye-sensitized cells to be used in the future solar cell market, while the number of patents issued remains low.

Although previous research revealed that our citation-based approach is appropriate to overview academic research fronts [21,22], someone might point out the negative aspects of our method, especially the procedure of step (3), constructing the largest connected component of citation networks. As shown in Fig. 3(b), many patents were eliminated from the largest component and could not be included in the analyzed data. However, the hypothesis that our method might overlook significant patents could be dismissed because the more significant and valuable a patent, the more it tends to be cited [28]. This meant that our method could, as a minimum, determine patents that are significant in terms of value. Of course, some patents dealing with dye-sensitized and polymer related solar cells also existed. Fig. 11 shows a plot of the number of patents containing "dye-sensi*" or "polymer* OR organ*" in the title or abstract. In Fig. 11, white rectangles indicate the total in 2008 and black circles indicate the number included in the largest connected component. In 2008, there were 143 and 150 patents related to dye-sensitized and polymers respectively, most of which were not included in the largest connected component because they had no citation linkage to existing patents.

Table 1
Major clusters in the citation network of patents related to solar cell with topics in 2008.

ID	# Patents	Ave. year	Topics	Applicants (top 5)
#1	517	1991.6	Silicon (cell design)	Canon Kabushiki Kaisha (51), Siemens Aktiengesellschaft (33), RCA Corporation (18), Sharp Kabushiki Kaisha (17), Atlantic Richfield Company (11)
#2	438	1991.7	Metal contacts	Mobil Solar Energy Corporation (24), The Boeing Company (24), Massachusetts Institute of Technology (17), Evergreen Solar, Inc. (13), The United States of America as represented by the Administrator of the National Aeronautics and Space Administration (12)
#3	436	1994.9	Battery and converter	Canon Kabushiki Kaisha (49), Sanyo Electric Co., Ltd. (11), Sharp Kabushiki Kaisha (8), The Boeing Company (5), RCA Corporation (5)
#4	368	1994.9	Panel	Canon Kabushiki Kaisha (94), Hughes Electronics Corporation (10), TRW Inc. (10), RCA Corporation (9), Sanyo Electric Co., Ltd. (8)
#5	223	1994.4	Diode	Sharp Kabushiki Kaisha (13), Texas Instruments Incorporated (12), Mitsubishi Denki Kabushiki Kaisha (11), Emcore Corporation (7), The United States of America as represented by the United States Department of Energy (6)
#6	149	1994.6	CIS	Matsushita Electric Industrial Co., Ltd. (15), Midwest Research Institute (6), Atlantic Richfield Company (6), The Boeing Company (5), Siemens Aktiengesellschaft (5)
#7	104	1988.7	Solar concentrator	RCA Corporation (5), Hughes Aircraft Company (4), The United States of America as represented by the Administrator of the National Aeronautics and Space Administration (3), Hughes Electronics Corporation (2), Mobil Tyco Solar Energy Corporation (2)

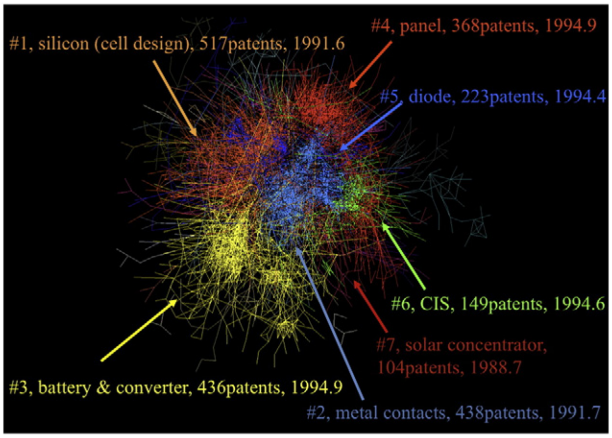

Fig. 8. Visualization of a citation network of patents related to solar cell with topics in 2008.

Some limitations do apply to our approach. The first is the existence of time lag. Up to two years are required until a paper/patent receives citations from other papers/patents. It also takes some months from the academic/technological outcome to the publications of these findings on academic papers or patents. We should not ignore these disadvantages and indeed should balance them out with other information, such as expert opinions and insights, because they can sometimes access unpublished information via a network of their research community or internal community among firms, which is sometimes called invisible. Experts' insights might be able to detect other opportunities for industry, which could not be expressed in the citation networks. Investigating the time lag needed to commercialize a scientific discovery is also significant for our future study.

Secondly, we should use the outcomes as an intellectual basis for constructing an R&D strategy or policy, rather than as a strategy itself. Although the outputs of our method suggest opportunities for future industrialization, there is no guarantee of success in any case. Although the citation network approach is a powerful tool to visualize and compare the overall structure of academic research and technology in a manner that an expert cannot perform and to support experts in constructing an R&D strategy or policy, the strategy should be decided by human beings.

Thirdly, it is not necessarily the case that all potential academic knowledge, which does not exist as patents, could represent opportunities for industry. As mentioned above, there is a possibility that some of these, namely the technologies of area B in Fig. 2, are so basic that they could not be applied to industry. Moreover, even when the academic outcomes applicable for industry are promising, someone might hesitate to write patents if it is obvious that it takes more than twenty years to become a huge market in terms of business.

The forth one is the broadness of the query we used. Our intention was to select academic papers with less noise. If we use *solar cell* OR photovoltaic** as the query, the number of retrieved documents increase. However, this document set includes applications relating photovoltaic effects such as photoreaction and photocatalysis, which are not strongly related to solar cell. Although we limited the query to solar cell based on the reasons above, there is still a room to investigate how to choose the query.

The final point relates to our proposed procedure. As specified in Section 3.2, the proportion of patents involved in the largest connected component is smaller than that of academic papers. Since patents tended to have fewer citations than academic

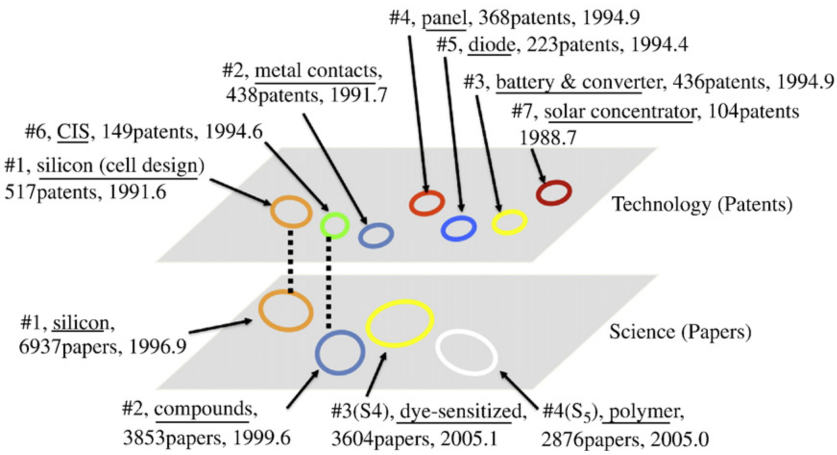

Fig. 9. Comparison of clusters of academic papers and those of patents related to solar cell.

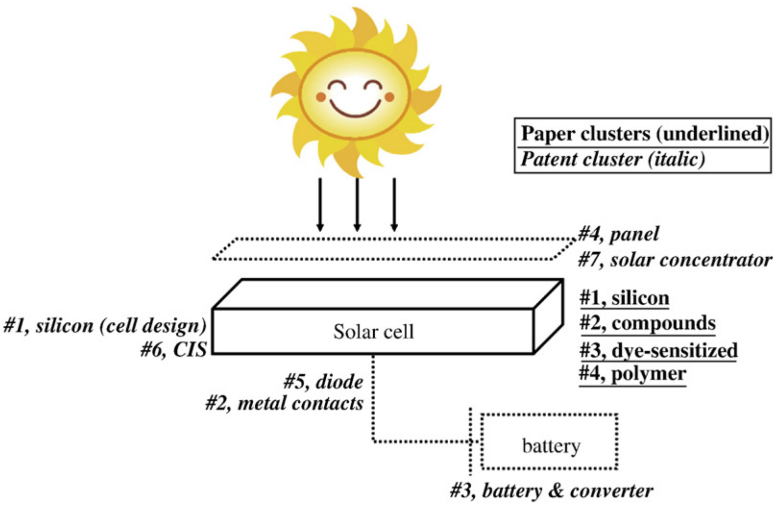

Fig. 10. Knowledge of science and technology related to solar cell.

publications, citation analysis itself could not capture all the characteristics of the former. Text-mining is required for our future study, in particular to track emerging patents which tend to obtain fewer citations.

5. Conclusion

In summary, we performed a case study of a solar cell to develop a method of detecting gaps between science and technology via a combination of citation- and expert-based approaches. Scientific research tended to be more basic, especially the cell design, whereas patents focused on more applied technology around the basic research. Of four science research fronts and seven pieces of technology, only two of each, namely silicon and compound solar cells, corresponded semantically. Conversely, there were no patent clusters corresponding to the other two scientific research fronts, namely dye-sensitized and polymer solar cells. These research areas could be regarded as opportunities for industrial commercialization because scientific activities exist, although there are no technological applications. Our results could offer an intellectual basis for detecting possible opportunities for industrial commercialization.

It is revealed, at least with the case of solar cells, that the commercialization gap between science and technology can be retrieved by comparing the structure of citation networks of scientific publications and patents. Our future study would be 1) the time lag needed to commercialize a scientific discovery, 2) query selection problem, and 3) applying text-mining in order to track emerging technological applications.

Acknowledgment

This research was partially supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Young Scientists (Start-up), 21800011, 2009.

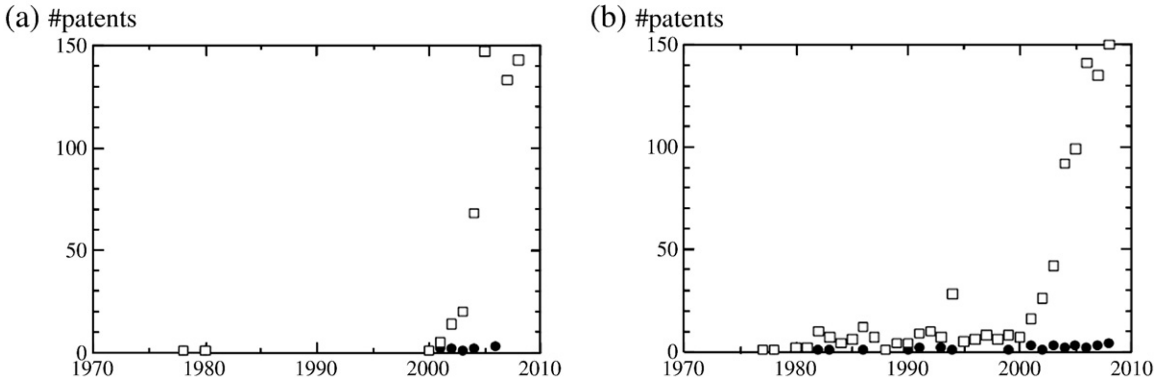

Fig. 11. Annual number of collected patents matched by (a) "dye-sensi*" and (b) "polymer* OR organ*" in the title. White rectangles indicate the total in 2008 and black circles indicate the number included in the largest connected component.

References

[1] R. Williams, D. Edge, The social shaping of technology, Res. Policy 25 (865) (1996) 899.
[2] J. Niosi, Fourth-generation R&D: from linear models to flexible innovation, J. Bus. Res. 45 (111) (1999) 117.
[3] L. Fleming, O. Sorenson, Science as a map in technological search, Strateg. Manage. J. 25 (909) (2004) 928.
[4] E. Mansfield, Academic research and industrial innovation, Res. Policy 20 (1) (1991) 12.
[5] F. Narin, K. Hamilton, D. Olivastro, The increasing linkage between U.S. technology and public science, Res. Policy 26 (1997) 317–330.
[6] F. Narin, D. Olivastro, Status report: linkage between technology and science, Res. Policy 21 (237) (1992) 249.
[7] E.H. Kessler, A.K. Chakrabarti, Innovation speed: a conceptual model of context, antecedents, and outcomes, Acad. Manag. Rev. 21 (4) (1996) 1143–1191.
[8] R.J. Watts, A.L. Porte, Innovation forecasting, Technol. Forecast. Soc. Change 56 (1) (1997) 25–47.
[9] S.T. Walsh, Roadmapping a disruptive technology: a case study — the emerging microsystems and top-down nanosystems industry, Technol. Forecast. Soc. Change 71 (1–2) (2004) 161–185.
[10] T.U. Daim, G. Rueda, H. Martin, P. Gerdsri, Forecasting emerging technologies: use of bibliometrics and patent analysis, Technol. Forecast. Soc. Change 73 (8) (2006) 981–1012.
[11] W.E. Souder, Managing relations between R&D and marketing in new product development projects, J. Prod. Innov. Manage. 5 (1) (2003) 6–19.
[12] R.K. Moenaert, W.E. Souder, An information transfer model for integrating marketing and R&D personnel in new product development projects, J. Prod. Innov. Manage. 7 (2) (2003) 91–107.
[13] R.K. Moenaert, W.E. Souder, A. De Meyer, D. Deschoolmeester, R&D–marketing integration mechanisms, communication flows, and innovation success, J. Prod. Innov. Manage. 11 (1) (2003) 31–45.
[14] W.E. Souder, R.K. Moenaert, Integrating marketing and R&D project personnel within innovation projects: an information uncertainty model, J. Manage. Stud. 29 (4) (2007) 485–512.
[15] R.N. Kostoff, R.R. Schaller, Science and technology roadmaps, IEEE Trans. Eng. Manage. 48 (132) (2001) 143.
[16] J. Michel, B. Bettels, Patent citation analysis, Scientometrics 51 (1) (2001) 185–201.
[17] M. Meyer, Does science push technology? Patents citing scientific literature, Res. Policy 29 (3) (2000) 409–434.
[18] R.N. Kostoff, R. Tshiteya, K.M. Pfeil, J.A. Humenik, G. Karypis, Power source roadmaps using bibliometrics and database tomography, Energy 30 (709) (2005) 730.
[19] H. Small, Tracking and predicting growth areas in science, Scientometrics 68 (595) (2006) 610.
[20] D.C. Mowery, J.E. Oxley, B.S. Silverman, Strategic alliances and interfirm knowledge transfer, Strateg. Manage. J. 17 (77) (1996) 91.
[21] N. Shibata, Y. Kajikawa, Y. Takeda, K. Matsushima, Detecting emerging research fronts based on topological measures in citation networks of scientific publications, Technovation 28 (11) (2008) 758–775.
[22] N. Shibata, Y. Kajikawa, Y. Takeda, K. Matsushima, Comparative study on methods of detecting research fronts using different types of citation, J. Am. Soc. Inf. Sci. Technol. 60 (3) (2009) 571–580.
[23] M.E.J. Newman, Fast algorithm for detecting community structure in networks, Phys. Rev. E 69 (2004) 066133.
[24] A.T. Adai, S.V. Date, S. Wieland, E.M. Marcotte, LGL: creating a map of protein function with an algorithm for visualizing very large biological networks, J. Mol. Biol. 340 (1) (2004) 179–190.
[25] P. Hennick, M. Fischedick, Towards sustainable energy systems: the related role of hydrogen, Energy Policy 34 (1260) (2006) 1270.
[26] Y. Kajikawa, J. Yoshikawa, Y. Takeda, K. Matsushima, Tracking emerging technologies in energy research: toward a roadmap for sustainable energy, Technol. Forecast. Soc. Change 75 (771) (2008) 782.
[27] F. Aratani, The present status and future direction of technology development for photovoltaic power generation in Japan, Prog. Photovolt. Res. Appl. 13 (463) (2005) 470.
[28] D. Harhoff, F. Narin, F.M. Scherer, K. Vopel, Citation frequency and the value of patented inventions, Rev. Econ. Stat. 81 (3) (1999) 511–515.

Dr. Naoki Shibata is an assistant professor at The University of Tokyo and also a visiting scholar at CSLI, Stanford University. His interests are Management of Technology and Information Science focusing on detecting emerging research fronts by citation analysis. He received his PhD degree from The University of Tokyo in 2009. He has an industrial experience as a product manager at CEO office of Rakuten Inc., the largest internet company in Japan. He is a member of The American Society for Information Science and Technology (ASIS&T), Information Processing Society of Japan, and The Japanese Society for Artificial Intelligence.

Yuya Kajikawa is an assistant professor at the Institute of Engineering Innovation at the School of Engineering in the University of Tokyo. He is also a visiting scholar at research institute for sports industry in Waseda University. He has a PhD in chemical system engineering from the University of Tokyo. He is a multidisciplinary scholar having a wide coverage of knowledge in materials processing, information processing, and technology management. His research background is chemical engineering, and has a number of academic publications in chemical engineering, applied physics, and also materials science. He is also a professional of information science and management science, and has a number of referred journal papers in these disciplines. His current research interest includes R&D management and policy (incl. technology management, technology roadmapping, national innovation system, regional economy), and energy and related technologies (incl. renewable energy, energy policy, thin film processing), and information processing (incl. network analysis, natural language processing, ontology).

Dr. Ichiro Sakata is a Professor at Todai Policy Alternatives Research Institute, The University of Tokyo. His interests include technology management, technology roadmap and innovation network focusing on rapidly growing sectors including healthcare, solar cell and battery. He received his master's degree from Brandeis University in 1997 and PhD from The University of Tokyo in 2003. He has a working experience as a senior policy analyst and policy maker at the Japanese Ministry of Economy, Trade and Industry.