Structural analysis of R and D division from patent documents

Iino, Yurie
Graduate School of Information Science and Electrical Engineering

Yamada, Yasuhiro
Department of Mathematics and Computer Science, Shimane University

Hirokawa, Sachio
Research Institute for Information Technology, Kyushu University

出版情報：IEEE International Conference on e-Business Engineering, ICEBE'08 - Workshops: AiR'08, EM2I'08, SOAIC'08, SOKM'08, BIMA'08, DKEEE'08, pp.423-428, 2008-12-01
バージョン：
権利関係：

Structural Analysis of R & D Division from Patent Documents

Yurie Iino
Graduate School of Information Science and
Electrical Engineering, akioiino@d6.dion.ne.jp

Yasuhiro Yamada
Department of Mathematics and Computer Science,
Shimane University, yamada@cis.shimane-u.ac.jp

Sachio Hirokawa
Research Institute for Information
Technology, hirokawa@cc.kyushu-u.ac.jp
Kyushu University, Hakozaki 6-10-1,
Fukuoka 812-8581, Japan

Abstract

This paper describes a method to draw the network of inventors and their technology. It is based on the co-occurrence analysis of inventors of a company described in their patent documents. Empirical evaluation, using 16,375 cosmetic related patent documents, shows that the method discloses the structure of research and development activities of companies.

1. Introduction

The increasing amount of information are available on the Web. The technological information and academic research information are not exception, which were used to be published only in scientific journals. The situation concerning to the patent documents is changing as well. The patent documents used to be available mostly in commercial sites e.g., PATOLIS[1]. In March 1999, Japan Patent Office (JPO) opened the site IPDL[2] where a user can search and download the patent documents. In European Patent Office[3], patent documents of 71 countries are provided for free. Google[4] started free patent search service in 2007. The patent have three features that are different from non-patent documents such as scientific articles. The first feature is that they are written in the uniform format that contains title, abstract, publication date, document number, assignee name, application serial number, inventor name, international classification etc. The second feature is the coverage of fields. Any kinds of invention in any technology fields can be applied as patents. On the other hand, most of all non-patent document databases focus some specific technological field. The third feature is that the documents guarantee the activity of enterprises as intellectual property. Intellectual property is very important for the enterprises to maintain their business activities in markets. From the time that Koizumi administration established the Strategic Council on Intellectual Property, Japanese government has promoted enterprise's activities for intellectual property[5]. In April 2007, JPO released the collection of intellectual property strategy examples: "Toward Strategic Intellectual Asset Management: Improving Technology Management Capacity" [6]. This report announced that intellectual property is very important to strengthen technology management, and showed progressive activities in Japanese big companies. By analysing patent documents of a company, we can guess the policy of the company's research and development.

It is necessary and indispensable to maintain stable intellectual property for managing a enterprise. Moreover, the evaluation of the literary property is closed up in M&A and the value of the intellectual property becomes an important judging material in the market. The value and range of the patents should be determined compared to the competitors. In some circumstances, it will be wise to consider licensing the patent to competitors. That is, it is necessary to analyse patent information from the aspect of the accounting. However, it is difficult to understand the correlation of the patent of the huge number of cases at one view, and is a factor to annoy management. Thus, more convenient tools are required for no-expert of patents. This paper proposes a method that enables non-patent experts to analyse relationship of inventors of target company.

The contribution of this paper is as follows:

[1] http://www.patolis.co.jp/
[2] http://www.ipdl.inpit.go.jp/homepg.ipdl
[3] http://ep.espacenet.com/
[4] http://www.google.com/patents
[5] http://www.kantei.go.jp/foreign/policy/titeki/konkyo_e.html

1: We proposed a novel approach for analyzing organizational structure of research & development division of companies. A naive approach would require some information on personnel affairs, which is not available in general. Instead, we use the information of co-inventors described in patent documents. We apply the notion of concept graph [16] which uses co-occurrence frequencies of inventors and keywords.
2: We showed case studies of the proposed method for cosmetic companies. It turned out that there is a clear difference that identifies the structure of concept graph of inventors of Japanese company from that of European company. We observed that raw material companies and consumer product companies are able to distinguished by the same characteristics of the graph.
3: We proposed another method of "matrix map" [14] to extract key inventors and to identify research group. The inventors who listed in the first place are used as x-axis of the matrix and other inventors are used as y-axis.

The rest of this paper is organized as follows. We review of related work in Section 2 followed by a discussion of co-inventors' analysis. Section 4 begins with the exposition of concept graph applied to patent documents. Section 5 describes the two organizational styles of R&D division observed by the concept graphs of inventors. Section 6 analyses key inventors using matrix maps. Finally, the paper is concluded in Section 7.

2. Related Work

There have been several commercial patent services, but there have not been many academic research in patent retrieval before SIGIR2000 [6] and NTCIR project[7]. Recently, the patent data are attracting many researchers in information retrieval(IR) field [2]. The aim of these research are mainly in technological ones, for example, cross language retrieval, summarization, question and answering and clustering, categorization. One reason of this movement comes from the political change toward "pro-patent" [6, 11]. On the other hand, multidisciplinary research emerged to analyse and evaluate the influence of scientific research to industry. [12] and [18] disclosed "science linkage" by applying citation analysis for references of patent and scientific articles. These researches belong to information science and sociometrics.

The goal of the present paper is not to analyse the relationship between patent documents or scientific articles but to analyse the relationship of inventors of patents. It can be considered as one subject of social networks. Kautz et al. [8] used the search results of a person's name and extracted other names and evaluated their relationship by Jaccard coefficient. Asada et al [1] extracted human relationship from researchers' activities that can be observed in Web pages. [10] showed the human network with keywords that connect people. [4] applied this approach to discovery of interfirm networks. These studies aim to develop new methods of competitive intelligence [7] which analyse relationship of people and organization. The present paper differs from those researches in that we use patent documents and that we analyse the co-inventor relationship.

The "brainmap" of CHI Research Inc. [13] uses the same information and visualises relationship of co-inventors in specified patents of the same company. The relationships are shown in a matrix where the names of inventor are lined in y-axes and the patent id's are lined in x-axes. On the other hand, the "concept graph" of the present paper visualises an inventor as a node and the relationship with co-inventors are displayed as directed edges.

3. Analysis of Co-inventors

Patent search is indispensable to confirm the intellectual property of the firm who are going to apply a patent. Patent search yields valuable information of competitors[7] which is useful for strategy decision of own company as well as for acquisition or merge of the target company. There are several reports on the analysis of the research organization by the official body and the think tank. They are based on the questionnaire to the enterprises and the researchers. However, the questionnaire is not possible for a rival enterprise or for all companies in a specific industry for a company. Analysis tools based on objective material are needed for structural analysis of research and development division of target companies. The brainmap of CHI is another such tool where co-inventors who worked together for some period are connected in lines. They can be used to find weak and strong technology areas of a company. Patent analysis can be useful to consider merger/acquisition target purely from a technological view.

Inventors play a key role in activities of an enterprise. The records of the inventors and their activities, which can be measured through patent documents, can be used to guess the company's R&D direction. The number of researchers reflects the investment of the company for research division. In fact, it is known that there is a high correlation between the average sales per employee and the patent application number per inventor [3]. The main target of key researchers will show the current status of technological development level of the company.

The present paper analyses how inventors are organized. In some companies, very small number of researchers work together for one project. In other companies, there are several people who have different level of contribution for one

[6] http://research.nii.ac.jp/ntcir/sigir2000ws/
[7] http://research.nii.ac.jp/ntcir/outline/prop-ja.html

project and who appear as co-inventors in many patent documents. The concept graph [16] of inventors visualizes the organizational difference of companies. If we augment the keywords into the graph, we can interpret the group of inventors.

This paper demonstrates some case studies based on the patent documents that were applied to Japan Patent Office during 2000 and 2005 and were classified with IPC (international patent classification) "A61K7" that covers cosmetic related technical field. We prepared 16,375 patent documents with "A61K7". We constructed a concept graph search engine and a matrix search engine for the data set and analysed the inventors of each company.

4. Concept Graph of Patent Documents

The concept graph engine [16] is a search engine that displays the search result as a directed graph of keywords that appear in the documents. The system extracts the characteristic words that appear almost only in the search results and the hypernym/hyponym relationship of the keywords are estimated using their frequencies and shown as a diagram. In [16], a word u is said to be more general to another word v with a threshold α, denoted $u >_\alpha v$, if

$$df(u, R) > df(v, R) \quad (1)$$

$$df(u * v, R)/df(v, R) > \alpha \quad (2)$$

where $df(w, R)$ is the number of documents in the search result R that contain the keyword w, $df(u * v, R)$ is the number of search results that contain both u and v, and $0 \leq \alpha \leq 1.0$ is a threshold. When we use the threshold $\alpha = 0.5$, $u >_\alpha v$ means that u is used much often than v and that most documents which concerns v have something to do with u.

If we consider inventors as well as keywords that appear in patent documents, the concept graph draws the relationship of co-inventors and their research subjects at tha same time. Fig.1 is such a graph for 584 patent documents applied by a Japanese company "A" in 2000-2002 which contains 35 inventors and 41 keywords(nodes colored pink). We can see 4 large groups (connected nodes) of researchers and 9 small groups. We also can interpret the activity of the groups by the keywords.

5. Concept Graph of Inventors

We study concept graphs to analyse the organizational structure of R&D division of cosmetic companies. The aim of the analysis is to extract key researchers and research groups. In this section, we do not consider the order of the inventors in a patent document and use the number of

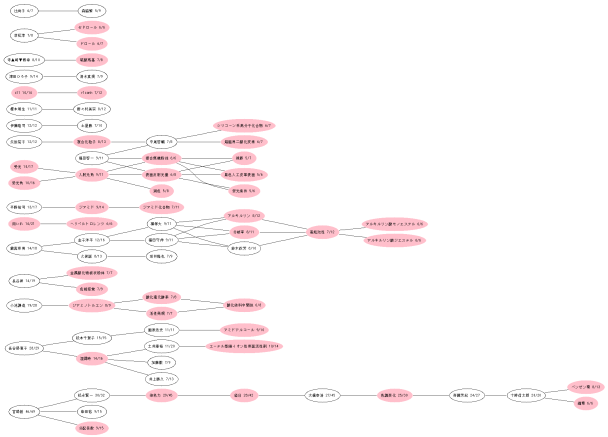

Figure 1. Concept Graph of Inventors and Keywords

patents by a researcher as his/her contribution. The first authorship is considered in the next section.

The lessons we learned are that there are two kinds of organizational style of R&D division which consists of either

Type A	a small number of large groups which contain hierarchy of subgroups, or
Type B	many separated groups of inventors.

5.1. Cluster of Subgroups of Inventors (Type A)

Fig.2 displays the concept graph of inventors of 584 patent documents applied by a Japanese cosmetic company "A" that develops consumer products. The graph is drawn with the threshold $\alpha = 0.5$, which implies that more than half jobs of an inventor v, who is linked to the right side of another inventor u, have something to do with u's job and that u and v belong to the same group where u plays a more important role. Thus, it seems that the company "A" has 18 independent research groups. But it is too naive to understand the situation. If we lower the threshold α, we see an organizational characteristics of R&D division of the company. Fig.3 is the graph with $\alpha = 0.1$ where separated groups in Fig.2 are linked and form a large group. The implication is as follows. This company has several distinct research groups with key person, but they have flexible collaboration. We observed the similar situation in other Japanese cosmetic companies who produce consumer products.

Fig.4 shows the concept graph for 79 patent documents applied by a Japanese company "B" in 2000-2002, who provides raw material to other cosmetic companies. Even though the threshold $\alpha = 0.5$ is tight, the graph is much more complex compared to the graph Fig.3 of the company

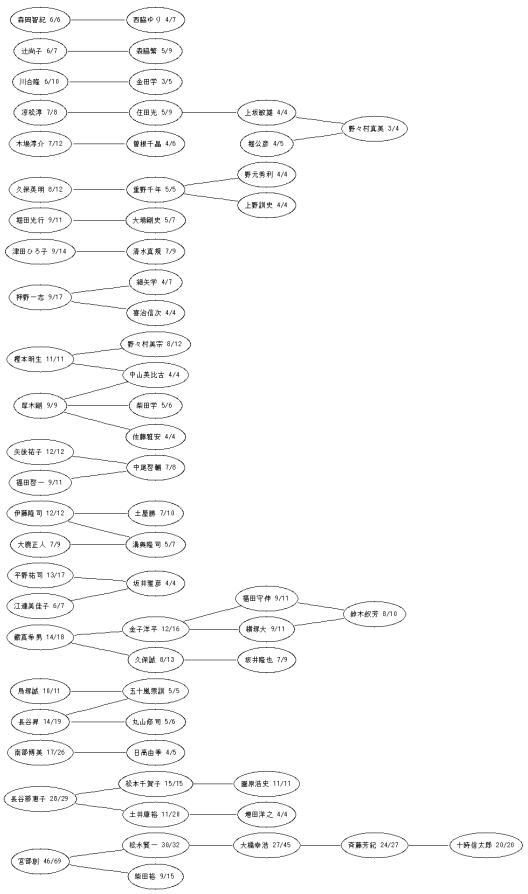

Figure 2. Concept Graph of Inventors ($\alpha = 0.5$)

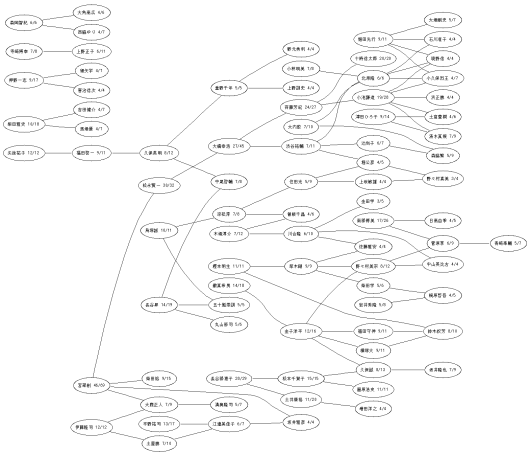

Figure 3. Concept Graph of Inventors ($\alpha = 0.1$)

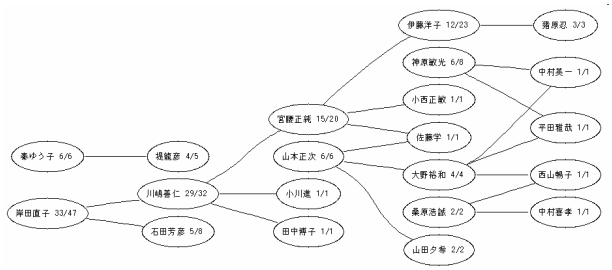

Figure 4. Concept Graph of Inventors of a Raw Material Company($\alpha = 0.5$)

"A" where there are several separated groups. In Fig.4, we see only one large group of inventors where several inventors belong to multiple sub-groups. One possible reason would be that a company who produces raw material tries to keep the intellectual property before they provide to other company who produces consumer product. Thus, the company requires several groups of researchers who consider applicable fields of the material. However, the groups are related to the material and thus the subgroups form a large group.

5.2. Independent Groups of Inventors (Type B)

We analysed another company "C" whose headquarter and research division is settled in a European country. Even though we changed the threshold α from 0.5 to 0.1, there are few change in the concept graph Fig.5. This implies that the company has independent groups whose member are fixed and each group does not work collaboratively. An interpretation of this result would be that an individual inventor has a large responsibility and authority in a specific field, and the relations with other sections are few. In a research report [15], it is pointed out that there are three different points of view in management style of R&D division in Japanese companies and in European companies.

(1) Japanese enterprises adapt lifetime employment system, while European enterprises employ a researcher by a contract that specifies the individual ability.

(2) According to lifetime employment system, most people are expected to promote in the same company in Japanese enterprises, while they prefer to move out of a company instead of moving to other section of the same company in European enterprises.

(3) The most closely related division to R&D division is the sales and marketing section in European enterprises, while it is the manufacturing section in Japanese enterprises.

The implication of this report is that European enterprises employ necessary personnel to specific technological

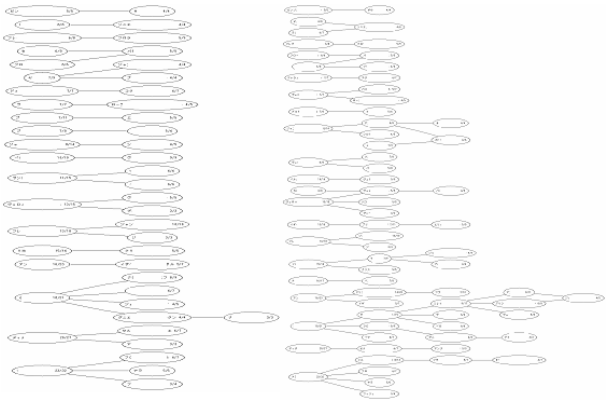

Figure 5. Concept Graph of Inventors of an EU company ($\alpha = 0.5$(Left), $\alpha = 0.1$(Right))

development. The employed person takes the responsibility and the authority for the development and has little in-house reshuffle and relations with other section of the R&D division or with the manufacturing section and has small number of co-inventors. Our analysis of Fig.5 matches this claim.

6. Discovery of Key Inventors

In the concept graphs of inventors, we used the number of the patents of an inventor to evaluate his/her contribution to the company. Group structure is estimated by hypernym/hyponym relationship. As a result, separated groups of researchers form a directed graph where key inventors can be easily recognized as roots of groups. However, there may be another way to describe researchers' contribution by their position in the list of inventors, where the first inventor is assumed to the key person of the patent.

In this section, we analyse groups of inventors by "matrix" engine [14] which displays the search results in two dimensional matrix. There are many ways to display patents as patent maps[8] and there are many patent map softwares [5]. In [14], a user can choose two aspects, e.g., X and Y, of the data to analyse. Each data d_i is displayed in a cell of the matrix that is specified by $(X(d_i), Y(d_i))$. Fig.6 shows a matrix with the IPC and the publication year as the two aspects. The number written in each cell shows the number of inventors whose patent matches the two aspects.

In this section, we use the first inventor as Y-axis and other inventors as X-axis. Fig.7 is such a matrix map for a Japanese cosmetic company "C". The patent documents were searched by specifying the top 3 researchers of the company who made many patents as the first inventors.

[8] http://www.jpo.go.jp/shiryou/s_sonota/tokumap.htm

Figure 6. Matrix Engine

Note that inventors are written in the cells instead of the number of patents. We can see that A,D and F form a group where A and D are key inventors and F is a supporting researcher. We can also see that another group of B, E, I and J where B and E are key inventors and I and J are supporting members.

7. Conclusions

We proposed a method to analyse the relationship of inventors of a company by the concept graph engine. Case studies are shown for cosmetic companies. A clear difference was found in the structure of the graphs between Japanese companies/European companies and raw material companies/consumer product companies.

The present paper is still in an early stage of the analysis. The number of case studies of the present paper is small and much large number of samples will be necessary to confirm the effectiveness of the proposed method. There may be several factors to evaluate key inventors, as mentioned in Section 6. Comparison and combination of these approach is another further work.

References

[1] Asada, Y., Matsuo, Y., Ishizuka, M., Increasing Scalability of Researcher Network Extraction from the web (in Japanese), Trans. Jap. Soc. Art. Intel., Vol..20, No.6, pp.370–378, 2005.

[2] Fujii, A., eds, Special Issue on Patent Processing, Information Processing & Management, Vol.43, Iss.5, pp.1119–1440, 2007.

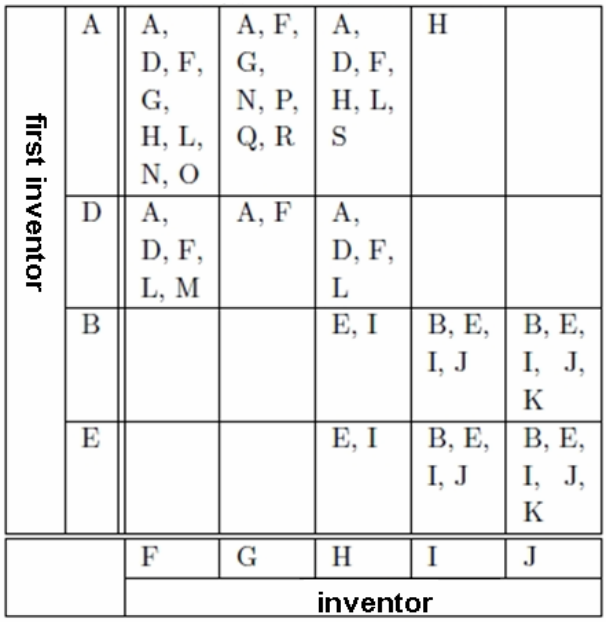

Figure 7. Matrix of first-inventor x inventors

[3] Ichikawa, T., Kasahara, K., Takagi, T., Evaluation Method for the Productivity of Research in Enterprises (in Japanese), Journal of the Japan Society for Management Information Vol. 9, No. 4, pp.19–32, 2001.

[4] Jin, Y., Matsuo, Y., Ishizuka, M., Extracting inter-firm networks from the World Wide Web using a general-purpose search engine Online Information Review Vol.32, Iss.2, pp.196–210, 2008.

[5] The third subcommittee, Intellectual Property Information Committee, Functional Comparison of Commercial Patent Map Softwares (in Japanese), Intellectual Property Management, Vol.50, No.9, pp.1427–1440, 2000.

[6] Japan Patent Office, The collection of intellectual property strategy examples: Toward Strategic Intellectual Asset Management: Improving Technology Management Capacity, 2007.

[7] Kahaner, L., Competitive Intelligence: How to Gather, Analyse, and Use Information to Move Your Business to the Top, Simon & Schuster 1997.

[8] Kautz, H., Selman, B., and Shah, M., The Hidden Web, AI magazine, Vol. 18, No.2, pp.27–35, 1997.

[9] Mase, H., Ohnishi, N., An Analysis of Term Usage in Patent Documents by Applicant and an Evaluation of its Effect to Patent Retrieval Accuracy (in Japanese) IPSJ SIG Notes Vol.2006, No.22, pp.77–84, 2006

[10] Mori, J., Matsuo, Y., Ishizuka, M., Faltings, B., Keyword Extraction from the Web for Creation of Person Metadata, 3rd Int'l Semantic Web Conference (ISWC2004), pp.45–46, 2004.

[11] Motohashi, K., Japan's Patent System and Business Innovation:Reassessing Pro-patent Policies, RIETI Discussion Paper Series 03-E-020,2003 http://www.rieti.go.jp/jp/publications/dp/03e020.pdf

[12] Narin, F., Hamilton, K., Olivastro, D., The increasing linkage between U.S. technology and public science. Research Policy Vol. 26, pp.317–330, 1997.

[13] Narin, F., From science papers to technology patents to company financial performance: The essential role of quality in R&D, WIPO-OECD Workshop on Statistics in the Patent Field, 2003

[14] Seki, T., Wada, T., Yamada, Y., Ytow, N., Hirokawa, S., Multiple Viewed Search Engine for an e-Journal - A Case Study on Zoological Science, Springer LNCS 4553, pp. 989–998, 2007

[15] Setani, M., Arico, F., Lambert, G., Llerena, P., Comparative Study on Research and Development Management of Japanese Enterprise and France Enterprise (in Japanese), National Institute of Science and Technology Policy, Research Material, No.5, 1997. http://www.nistep.go.jp/achiev/ftx/jpn/mat049j/pdf/mat049j.pdf

[16] Shimoji, Y., Wada, T., Hirokawa, S., Dynamic Thesaurus Construction from English-Japanese Dictionary, in:Proc. The Second International Conference on Complex, Intelligent and Software Intensive Systems, pp.918–923, 2008.

[17] Shinmori, A., Okumura, M., Marukawa, Y., Iwayama, M., Structure Analysis of Japanese Patent Claims Using Cue Phrases (Natural-Language Processing) (in Japanese), Transactions of Information Processing Society of Japan Vol.45, No.3, pp. 891–905, 2004.

[18] Tamada, S., Kodama, F., Gemba, K., A Study on Science Linkage of Japanese Patents : An analysis on patents in the field of genetic technology by constructing a citation database, The journal of science policy and research management Vol.17, No.3/4, pp.222–230, 2004.