

## TOPICS and Communities in the FIELD of patent analysis with SNA

Li Jiajia<sup>1,a</sup>, Ma Tiejun<sup>2,b</sup>

<sup>1,2</sup>Business School, East China University Of Science and Technology, Meilong Road 130, Xuhui District, Shanghai, 200237, China

<sup>a</sup>jubby\_li@163.com, <sup>b</sup>tjma@ecust.edu.cn

**Keywords:** social network analysis; bibliometric analysis; network; patent.

**Abstract:** SNA (Social Network Analysis) has been frequently used as a tool for patent analysis. This study analyzes the popular topics and communities in the field of patent analysis with SNA with publications included in Web of Science. This study reveals that (1) the research in the field of patent analysis has kept increasing in the last fifteen years; (2) researchers in this field are mainly from South Korea, America, Taiwan province of China, and some European countries. This research also identifies researcher communities and institution communities in the field of patent analysis with SNA.

### Introduction

Patent information become one of the most important intelligence resources with the fierce international competition. Patents carry the various information such as IPC Index, title, abstract, applicant, and institution, so the researchers can get some meaningful and useful findings though the disintegration, transformation, integration, and statistics of these patent information<sup>[1]</sup>. Patent analysis can help the organizations to improve outdated technologies, develop new technologies and identify the future key technologies by keeping recent developments of technologies under surveillance. In academia, recent years have witnessed a significant increase in attempts to devise appropriate models, method and tool for patent analysis<sup>[2]</sup>.

Social Network Analysis (SNA) provides a perspective of visualized analysis and has been frequently used as a tool for patent analysis. Wasserman proposed social network in 1994 for the first time. He treated community structure as the social relation network where nodes represent people and lines represent relationship among people<sup>[3]</sup>. Then Watts and Strogatz(1998) put forward the famous model of ‘small-world’ network and Barabasi and Albert (1999) proposed the model of scale-free network<sup>[4][5]</sup>. These two pioneering literature can be treated as the symbol of the beginning of complex networks research. Newman and Girvan used the Modularity to measure the partition of the network, and Newman gave some summary and comments to the present community detection algorithms in 2004<sup>[6][7]</sup>.

SNA can transform the complicated information of literature on patent analysis into visualized network graphics, which provide the structured information and relationship among nodes. The paper drew the collinear network of keywords, authors and institutions, and the results will present the general topics and communities in the field of patent analysis and give some instructional information to the researchers who want to have the further study on this field.

### Date Source and Preprocessing

The paper chooses Web of Science as the date source. The search condition is: theme= patent analysis or patent citation (when the title, abstract or keywords cover the theme, this literature will appear in the search result list). There are 2696 pieces of literature from 2000 to 2015 and 2364 pieces of literature from 2005 to 2015.

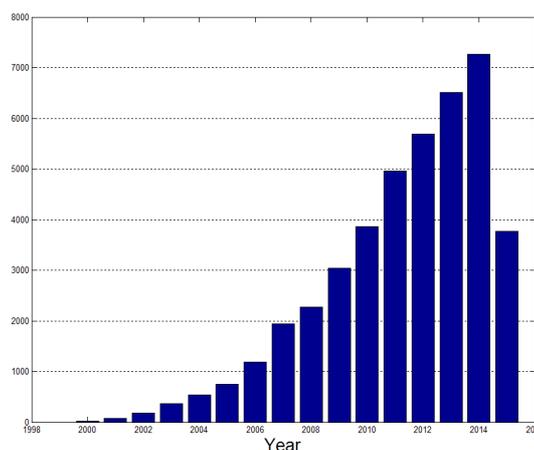
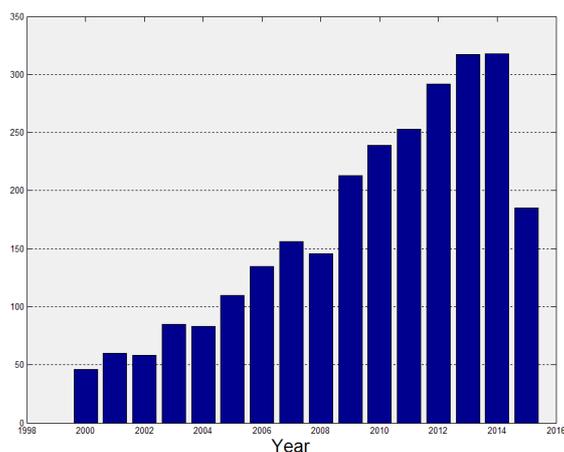


Fig.1. Number of literature from 2000to 2015      Fig.2. Number of citations from 2000to 2015

Above figures show the number of literature and the number of citations from 2000 to 2015. We can see that the research in the field of patent analysis has kept increasing in the last fifteen years. An obvious rise occurs in 2005 both in the number of literature and the number of citations, which means the literature after 2005 have the stronger influence and are more meaningful for the patent analysis. Hence, this research chooses the literature from 2005 to 2015 as the initial data sources. However, some irrelevant literature exists in these date sources (covers patent protection, patent policy or other else) which will make a distraction on the analysis result, so it's necessary to remove the irrelevant information from the date sources. 1002 pieces of literature left after the elimination.

### Network Construction and Analysis

SNA is an effective tool to make the abundant and complicated information visualized. The paper drew the collinear network of keywords, authors and institutions using the related software (SATI, UCINET) <sup>[8][9]</sup> with the 1002 pieces of literature selected above.

#### Collinear Network of Keywords

After the extraction of the keywords of the 1002 pieces of literature with SATI, we can get the frequency of each keyword.

Table 1 High-frequency keywords

No	Keywords	Frequency	No	Keywords	Frequency	No	Keywords	Frequency
1	Patents	129	17	Text mining	14	33	knowledge	9
2	innovation	87	18	Knowledge spillovers	13	34	scientometrics	9
3	Patent analysis	83	19	Intellectual property rights	13	35	Japan	9
4	Patent	82	20	Patent value	12	36	Bibliometric analysis	9
5	Patent citation	36	21	Bibliometrics	12	37	Pharmaceuticals	9
6	patent citations	30	22	Citation analysis	12	38	Korea	8
7	analysis	29	23	Patent citation network	12	39	Academic patenting	8
8	Nanotechnology	22	24	collaboration	12	40	University	8
9	China	22	25	Technological	11	41	Research	8
10	Technology	19	26	O31	11	42	networks	8
11	network analysis	18	27	social network analysis	11	43	network	8
12	Biotechnology	18	28	Citation	11	44	licensing	8
13	R&D	18	29	Invention	10	45	Clustering	8
14	Intellectual property	17	30	Knowledge diffusion	10	46	patent mining	8
15	technology transfer	16	31	Citations	10	47	India	8
16	knowledge flows	15	32	Technological innovation	9	48		

For the top 47 high-frequency keywords, a 47\*47 collinear relation matrix was constructed through SATI, and then we can draw a collinear network of keywords using UCINET. The blue nodes represent the keywords and the more relationship the node has with other nodes around it, the bigger the blue nodes will be. Similarly, the lines represent the collinear relationship between two keywords and the more frequently the two keywords appear in the same literature, the bolder the lines will be <sup>[10][11]</sup>.

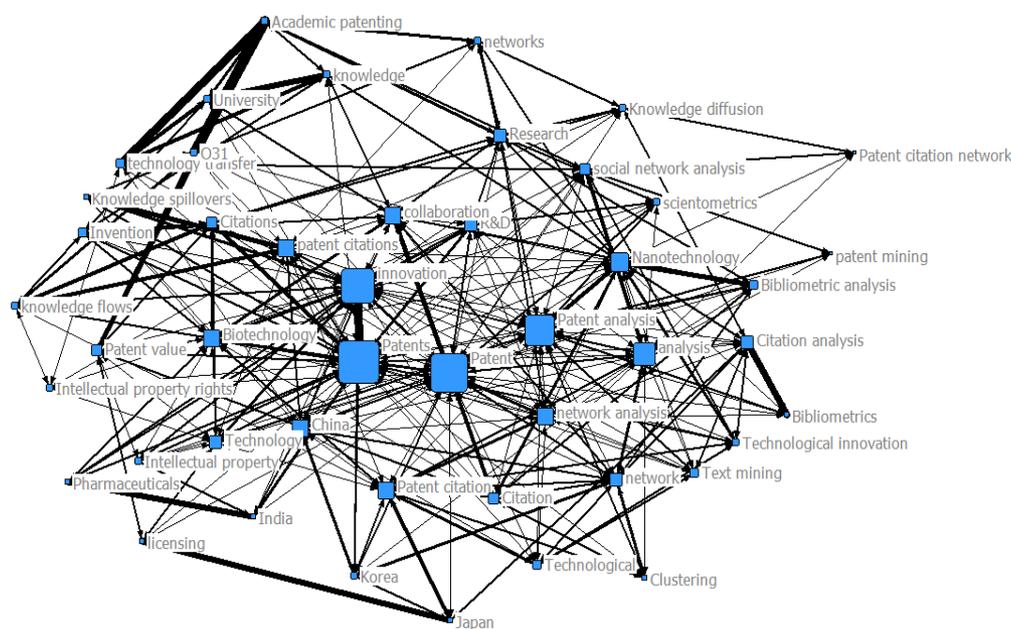


Fig.3. Collinear network of keywords

The findings through Fig.3 are as follows:

- The top four big blue nodes represent patents, innovation, patent and patent analysis, which are accordant to table 1 and the field we intend to study.
- The two biggest nodes (patents and innovation) also have the bolder line between them, which means most of the present researches focus on the point to utilize the patents information to forecast the innovative technology or analyze the evolution of innovative technology.
- Bolder lines exist among nodes representing Academic patenting, technology transfer and O31 (a kind of bacillus), showing there are strong relationship between Academic patenting and technology transfer, Academic patenting and O31. A majority of researchers tend to analysis the technology transfer and O31 though academic patents.

•

### Collinear Network of Author

36 nodes left when we chose the high-frequency authors (frequency greater than or equal to 5) after the extraction of authors of the 1002 pieces of literature with SATI. The detailed information of the authors and their frequencies were not showed in the paper. Similar to the collinear network of keywords above, the blue nodes represent the authors and the lines represent the collinear relationship between two authors.

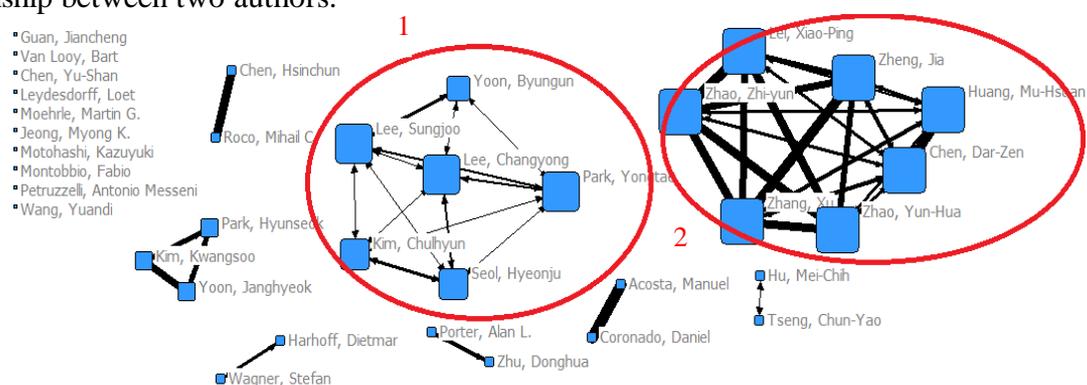


Fig.4. Collinear network of authors

The findings through Fig.4 are as follows:

- We can get two obvious communities through the collinear network of authors, showing there are intimate connections and cooperation among these authors. Upper left corner of Fig.4 are isolated points which means these authors don't have relationship with others.
- Community 1 is composed of the researchers from Korea (Lee, Changyong; Park, Yongtae; Yoon, Byungun; Lee, Sungjoo; Kim, Chulhyun; Seol, Hyeonju). They mainly use the patent text mining to analyze patent information.
- Community 2 is composed of the researchers from Taiwan province of China (Chen, Dar-Zen; Huang, Mu-Hsuan; Zhang, Xu; Lei, Xiao-Ping; Zheng, Jia; Zhao, Yun-Hua; Zhao, Zhi-yun). Their research fields and methods are various, mainly focusing on the country ranking, technical cooperation and technical evolution of various industry through patent information.

### Collinear Network of Institutions

Similar to above, 33 nodes left when we chose the institutions whose frequencies were greater than or equal to 10 after the extraction. The detailed information of the institutions and their frequencies were not showed in the paper and the cooperation network among institutions was developed as shown in Fig.5.

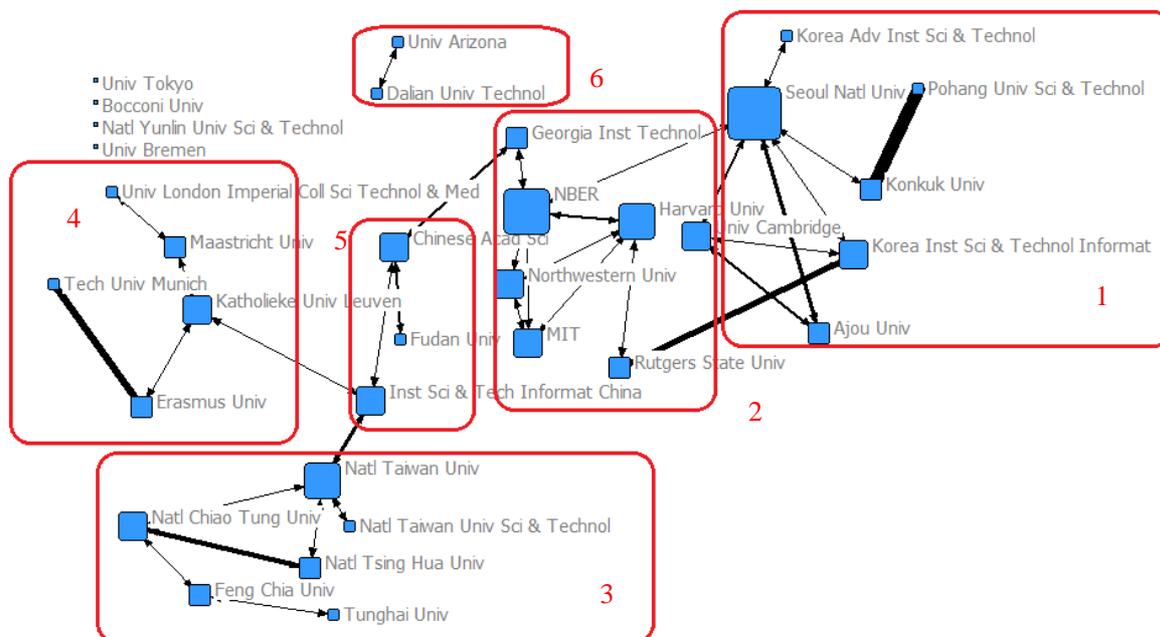


Fig.5. Collinear network of institutions

The findings through Fig.5 are as follows:

- A total of six communities appeared in the collinear network of institutions and four institutions didn't have cooperation with others.
- Communities 1 to 5 represent the cooperation among Korea, America, Taiwan province of China, Europe (mainly in Netherlands) and mainland China. Community 6 told that there were relationship between Dalian University of Technology and University of Arizona, but they didn't have cooperation with other institutions.
- The bigger blue nodes and bolder lines appeared mostly in community 1, showing that researchers in Korea have more research and more achievements in the field of patent analysis than other countries.
- Researchers in mainland China still have much space to improve in the field of patent analysis since there are only four institutions (Dalian Univ Technol, Chinese AcadSci, InstSci& Tech Informat China, Fudan Univ) belong to mainland China and they seldom have relationship with other countries' researchers.

## Conclusion

The paper made the visualized analysis on patent analysis and got the collinear network of keywords, authors and institutions based on the information from Web of Science. We can get that:

- The research in this field increased quickly from 2005 and the researchers can be divided into four communities, namely Korea, America, Taiwan province of China and Europe (especially Korea and Taiwan province of China).
- We can also identify four institution communities similar to the above researcher communities, meaning there are close connections and cooperation within the communities.
- Researchers from Korea made further study on the analysis method, mainly using patent text mining to extract the core patent or core technology.
- Researchers from Taiwan province of China focused on the empirical study (making policy analysis through patent information).

From the analysis on communities we can see the amount of the research institutions in this field in mainland China was small and they have weak cooperation with institutions in Taiwan Province and other countries.

## Acknowledgments

This research was sponsored by NSFC (No. 71125002).

## References

- [1] Jie Zhang, Jing Zhai. The research of SNA in patent analysis [J]. Science and Technology Management Research. 07(2011). (in Chinese)
- [2] Changyong Lee, Bokyoung Kang, Juneseuk Shin. Novelty-focused patent mapping for technology opportunity analysis [J]. Technological Forecasting & Social Change. 90(2015).
- [3] Stanley Wasserman. Social network analysis: Methods and applications [M]. Cambridge: Cambridge University Press. (1994).
- [4] Watts D J, Strogatz S H. Collective dynamics of 'small-world' networks [J]. Nature. 393(1998).
- [5] Barabbas A L, Albert R. Emergence of scaling in random networks [J]. Science. 286(1999).
- [6] Newman M E J, Girvan M. Finding and evaluating community structure in networks [J]. Physical review E. 69(2004).
- [7] Newman M E J. Detecting community structure in networks [J]. The European Physical Journal B-Condensed Matter and Complex Systems. 38(2004).
- [8] Xiaoying Liu. Bibliometric Analysis on Research Papers of Patent Analysis in China [J]. SCI-TECH INFORMATION DEVELOPMENT & ECONOMY. 25(2015). (in Chinese)
- [9] Jiansheng Du, Qian Fang. Bibliometric Study on the Community Detection under Social Networking Context [J]. SCI-TECH INFORMATION DEVELOPMENT & ECONOMY. 24(2014). (in Chinese)
- [10] Ruibin Wei. An Empirical Study of Keywords Network Analysis Using Social Network Analysis [J]. JOURNAL OF INTELLIGENCE. 28(2009). (in Chinese)
- [11] Hong Miao, Haili Liu. Research on the Science and Technology Collaboration of Beijing and Taiwan Based on Patent and Literature [J]. SCIENCE OF SCIENCE AND MANAGEMENT OF S. & T. 35(2014). (in Chinese)