Moving Integrated Product Development to Service Clouds in the Global Economy
J. Cha et al. (Eds.)

doi:10.3233/978-1-61499-440-4-153

Using Patent Co-Citation Approach to Explore Blu-ray Technology Classifications

Yu-Hui WANG[a,1], Pin-Chen KUO[b] and Tzu-Han CHOW[b]
[a] *Department of Information and Finance Management, National Taipei University of Technology, Taiwan, R.O.C*
[b] *Institute of Services and Technology Management, National Taipei University of Technology, Taiwan, R.O.C*

Abstract. Blu-ray Disc (BD), the next generation in optical discs, offers clearer and sharper image, better sound quality than DVD. In 2013, more than 72 million households have BD compatible devices in U.S. This paper aims to explore objectively key technology fields and intellectual intelligence of this emerging BD technology for participants. The patent co-citation approach (PCA), a patent classification system that is adaptive to the characteristics of a specific industry, is applied to draw key technology fields of BD technology. The results show that BD patents can be classified into eight technological categories. Most patents are classified into the two factors - factor 1.1 recording medium, recording and reproducing process, and factor 1.2 information recording medium, defect management. The intelligence can benefit patent management to make technological forecasting, research planning, and technological positioning for BD technology.

Keywords. Blu-ray Disc, patent co-citation approach (PCA), technology classification

Introduction

With the fast-advancing technology, patents play the role of strengthening the competitive advantage of enterprises [1]. When dealing with large amount of patents, an efficient patent classification can further benefit patent analysis. The current studies on patent analysis use the International Patent Code (IPC), developed by the Wide Intellectual Property Office (WIPO), or the United States Patent Code (UPC), developed by the United States Patent and Trademark Office (USPTO), to identify patent classifications. However, both the IPC and the UPC systems are too general to satisfy the needs for a specific industry [2]. Thus, the patent similarity based patent classification system, patent co-citation approach (PCA), was proposed to benefit the understanding of the essential patents for a specific industry, and the relationships among clusters of technology. The result of PCA can offer explicit intelligence for patent management, technological forecasting, research planning, technological positioning and strategy making [3].

Optical storage technology has developed three generations: Compact Disc (CD), Digital Versatile Disc (DVD), and Blu-ray Disc (BD). The technology has been applied

[1] Corresponding Author.

to the market for information storage, automotive electronics, and audio appliances and so on. With the rising need for audio entertainment quality, BD format offers an immense storage capacity (up to 50GB) that is perfect for High Definition video recording and distribution, as well as for storing large amounts of data. Including Dell Inc., Hewlett Packard Company and almost 100 founding members initiated the Blu-ray Disc Association (BDA) to promote BD format in 2005 and the market of BD has rapidly expanded. According to numbers compiled by the Digital Entertainment Group (DEG) with input from retail tracking sources, the number of Blu-ray homes continued to grow. Overall, consumer spending on digital content rose 17 percent in 2013. Blu-ray Disc consumer spending remained consistent, up about five percent for the year. The total household penetration of all Blu-ray compatible devices is more than 72 million in U.S. homes in 2013 [4].

However, when users use "CD, DVD, or Blu-ray" keywords to search for patent classifications of the IPC or the UPC, only general and rough classifications of optical storage technology are found. To sufficiently differentiate techniques among CD, DVD, and BD and catch up with BD technology classifications, this paper adopts the PCA to illustrate BD related patent classification system.

1. Patent Co-citation Approach (PCA)

Information about patent citations offers patent bibliography analyzers the initial judgment to realize the context of technology development and to evaluate the importance of patents [5]. Patent citation documents the course of the accumulation of the technical knowledge, and makes connections among the related patents. These connections demonstrate the correlations between relevant patents [3].

Lai and Wu (2005) proposed the PCA, based on co-citation analysis of bibliometrics, to create a patent classification system. The conception and the application of the PCA are shown in Figure 1. For instance, Q1-Q6 are target patents, and P1-P4 are basic patents selected from target patents. According to the similarity measured by co-cited frequency, basic patents are classified into two groups, representing different technology categories. P1 and P2 are covered by F1 category, while P3 and P4 are assigned to F2 category. The target patent Q1 cites the basic patent P1, so Q1 belongs to the F1 category [3].

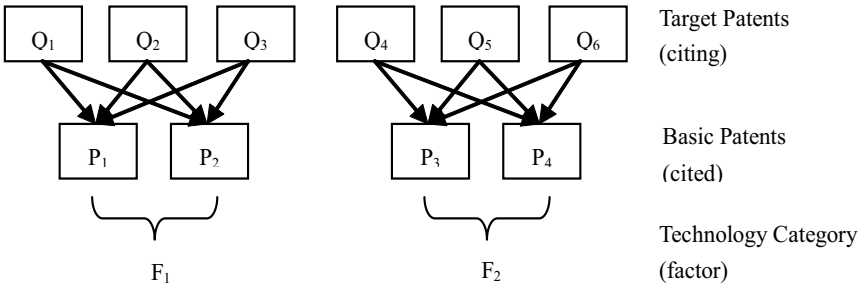

Figure 1. The conception and the application of the PCA [3]

The analysis of this approach is divided into three phases. Phase I selects appropriate databases to conduct patent searches. Phase II uses the co-cited frequency

of the basic patent pairs to assess their similarity. Phase III uses factor analysis to establish a classification system and assess the efficiency of the proposed approach [3].

1.1. Phase 1 Searching and identifying Basic Patents

According to the purpose of the research, researcher chooses proper database and search target patents and candidates of basic patents. Target patents are citing patents to be classified. Candidates of basic patents are those patents that are cited by target patents [3]. We denote Qi as target patent i and CPj as the candidate for basic patent j, respectively. The referential relationship between target patents Qi and the candidate for basic patents CPj is shown as the matrix where M is the amount of target patents, and N is the amount of candidates for basic patents.

$$\left[\alpha_{ij}\right]_{M\times N}, \alpha_{ij} = \begin{cases} 1 & Q_i \ cites \ CPj \\ 0 & otherwise \end{cases} \tag{1}$$

The elder patent has more opportunity to be cited. Hence, besides considering the cited frequency, the time being cited should be considered as well. In order to eliminate the bias from patent age, and select basic patents, the equation is adjusted and shown as follows:

$$ST_j = \sum_{i=1}^{M} \alpha_{ij} \times wt_i \ , \ 1 \le j \le N \tag{2}$$

Where wt_i is the weight of the target patent i, which is obtained by subtracting the standard year. Define the candidate CPj with $ST_j \ge c$ as basic patent [3]. The value of c is the threshold in selecting basic patents, and it will influence the comprehensiveness of the classification system and the complexity during the process of analyzing. The citation relationship between target patents Qi and basic patents CPj is shown as the new matrix$\left[\varepsilon_{ij}\right]_{m\times n}$where

$$\varepsilon_{ij} = \begin{cases} 1 & Q_i \ cites \ P_j \\ 0 & otherwise \end{cases} \tag{3}$$

Where m is the amount of target patents which can be classified by basic patents , and n is the amount of basic patents [3].

1.2. Phase 2 Evaluating the Similarities in Basic Patent pairs

The PCA employs the Pearson correlation coefficient to access the similarity for basic patent pair. There are three steps to calculate the similarity of each basic patent pair:

- Step1: Calculate the co-cited frequency of each basic patent pair
  Given patent j and j', the co-cited frequency of this patent pair is

$$\omega_{jj'} = \begin{cases} \sum_{i=1}^{m} \varepsilon_{ij}\varepsilon_{ij'} & if \ j \ne j' \\ 0 & if \ j = j' \end{cases} \quad 1 \le j \le n, 1 \le j' \le n \tag{4}$$

A symmetrical matrix $\left[\omega_{jj'}\right]_{m\times n}$ can be obtained after calculating all of the co-cited frequency of n basic patent pairs.

- Step2: Calculate the linkage strength of each basic patent pair
  The equation to calculate the linkage strength of each basic patent pair as follows:

$$\pi_{jj'} = \begin{cases} \frac{\omega_{jj'}}{S_j + S_{j'} - \omega_{jj'}} & if\ j \neq j' \\ 0 & if\ j = j' \end{cases} \quad 1 \leq j \leq n, 1 \leq j' \leq n \tag{5}$$

  Where $S_j = \sum_{i=1}^{m} \varepsilon_{ij}$ represents the cited frequency of the basic patent Pj [3].
  A symmetrical matrix $\left[\pi_{ij}\right]_{n\times n}$ can be obtained after calculating all of the linkage strengths of n basic patent pairs.
- Step3: Calculate the Pearson correlation coefficient of each basic patent pair
  Calculate the Pearson correlation coefficient of each basic patent pair by the symmetrical matrix $\left[\pi_{ij}\right]_{n\times n}$ to obtain the matrix of Pearson correlation coefficient $\left[\gamma_{jj'}\right]_{n\times n}$ [3].

1.3. Phase 3 Creation of the Patent Classification System

In this phase, the result of Pearson correlation coefficient of the basic patents is used to employ factor analysis to classify basic patents. After the factor analysis, the loading of the variables (patent) on the factor (technical category) indicates the degree of importance for the basic patent to the technical category, and it can help naming the technical categories as well [3]. Besides, the correlation coefficient between factors indicates the degree of correlation of technical categories. The performance of classification system can be evaluated by three indicators, cover index, weight cover index, and consistency index [6]. Three indicators are described as follows:

- Indicator 1: cover index
  The definition of cover index is as follows:

$$\text{cover index} = \frac{m}{M} \tag{6}$$

  Where M is the amount of target patents, and m is the amount of patents which can be classified [6].
- Indicator 2: weight cover index
  The definition of weight cover index is shown as follows:

$$\text{weight cover index} = \frac{\sum_{i=1}^{m} W_i}{\sum_{i=1}^{M} W_i} \tag{7}$$

  Where W_i is the frequency that patent i is cited by other target patents [6].
  This indicator is the revised version of the cover index weighted by the importance of each patent, which is measured by the cited frequency.
- Indicator 3: consistency index
  The definition of weight cover index is shown as follows:

$$\text{consistency index} = \frac{m-x}{m} \quad (8)$$

Where x is the amount of target patents that are multiply classified [6].
All high value of the above three kinds of indicators indicate good performance of the patent classification system [6].

2. Empirical case - Blu-ray disc technology

The PCA is a dynamic and self-constructing methodology to create a patent classification system, which can reflect the existing status of the technology [6]. Compared to static classification systems such as the IPC or the UPC, the system created by the PCA has a better ability to reflect the characteristics and existing status of anyone specific technology [6]. In order to offer explicit intelligence for BD patent management, this paper conducts the PCA to sufficiently illustrate BD emerging technology classifications. The processes are described as below.

2.1. Data Collection

Utility patent abstracts, claims, and titles containing keywords and phrases of "blue-ray"[2] or "blu-ray" or "blu ray" were collected in USPTO database. 403 patents are collected and identified as target patents. These patents cite 4,212 patents totally, which were candidates for basic patents. We calculate ST_j by using Eq. (2) with standard year 1977, which is set to let each wt_j remain positive. The results are shown in Table 1. We use STj >= 137 as the criteria to select basic patents, and 192 basic patents are identified from the 4,212 candidates of basic patents, approx top 5% in the 4212 patents.

2.2. Evaluating the Similarities in Basic Patent pairs

The matrix of co-cited frequency of each basic patent pair $[\omega_{jj'}]_{m\times n}$ and linkage strength of each basic patent pair $[\pi_{ij}]_{n\times n}$ were also constructed with Eqs. (4) and (5) respectively. Then we calculate the Pearson correlation coefficient of each basic patent pair, and we obtain the matrix of Pearson correlation coefficient $[\gamma_{jj'}]_{n\times n}$ to process next step.

2.3. Factor analysis

The result of Pearson correlation coefficient of basic patents is employed to factor analysis. This study uses principal component analysis with the promax rotation to extract factors. Based on eigenvalue more than 1 criterion, 7 factors are obtained, which account for 99.88% of the variance. Eigenvalues and variances explained by factors are shown in Table 2.

[2] In February 2002, the introduction of the Blu-ray Disc (BD) format was announced. However, the phase of "blue-ray" is still commonly used in patent applications before and after its official name.

Table 1. Weighted frequency of candidate for basic patents

STj	Frequency	Accumulated frequency	STj	Frequency	Accumulated frequency	STj	Frequency	Accumulated frequency
376	1	1	187	3	44	132	7	319
333	1	2	185	13	57	131	1	320
327	1	3	177	1	58	127	1	321
318	1	4	175	33	91	125	1	322
314	1	5	174	1	92	123	1	323
270	1	6	166	32	124	122	3	326
252	1	7	165	2	126	121	7	333
229	1	8	164	1	127	120	1	334
222	7	15	163	3	130	119	1	335
220	5	20	162	3	133	118	12	347
215	1	21	161	11	144	117	1	348
214	2	23	159	12	156	116	1	349
213	1	24	155	2	158	115	1	350
208	1	25	152	2	160	114	5	355
204	1	26	150	1	161	112	1	356
202	2	28	149	1	162	111	1	357
201	3	31	147	1	163	110	3	360
200	1	32	146	1	164	109	1	361
195	1	33	145	17	181	105	7	368
193	2	35	142	1	182	101	84	452
192	1	36	141	1	183	100	1	453
190	1	37	140	2	185	1~99	3759	4212
189	1	38	137	7	192			
188	3	41	133	120	312			

Factor 1 includes 104 patents in which 48 patents show with negative value of loading and 56 patents show positive value. The great difference between positive and negative value of loading impairs naming factors. So the second-round factor analysis is conducted to separate 104 patents of factor 1, and then two sub-factors were obtained, which accounted for 99% of the variance. Eigenvalues and variances explained by these two factors are shown in Table 3.

After the basic patents are classified, the targets patents are classified in the specific factor in which most of its citation belong to. There are 90 target patents which can be classified, and 4 patents are classified duplicately with the same frequency in multiple factors. The result of the basic and target patents classification is shown in Table 4. There are 56 basic patents and 19 target patents in factor 1.1 and then 48 basic

patents and 27 target patents in factor 1.2 respectively. Each factor is named after the titles of basic patents as shown in Table 4.

Table 2. Eigenvalues and variances explained by factors

Factor	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings
	Total	% of Variance	Cumulative%	Total	% of Variance	Cumulative%	Total
1	82.906	43.180	43.180	82.906	43.180	43.180	80.758
2	56.528	29.441	72.621	56.528	29.441	72.621	63.928
3	23.209	12.088	84.709	23.209	12.088	84.709	39.978
4	16.117	8.394	93.104	16.117	8.394	93.104	38.7
5	6.104	3.179	96.283	6.104	3.179	96.283	44.858
6	4.207	2.191	98.474	4.207	2.191	98.474	16.889
7	2.906	1.514	99.988	2.906	1.514	99.988	30.302

Table 3. Eigenvalues and variances explained by two sub-factors of factor 1

Factor	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of squared Loadings
	Total	% of Variance	Cumulative%	Total	% of Variance	Cumulative%	Total
1	79.299	76.249	76.249	79.299	76.249	76.249	68.770
2	23.738	22.825	99.074	23.738	22.825	99.074	63.225

2.4. Evaluating the performance of classification

We use Eq. (6), Eqs. (7) and (8) in section 1.3 to calculate the value of three quantitative indicator and the result is shown in Table 5. The first indicator cover index is 22.3% and it indicates 90 out of 403 target patents are classified by PCA. However, target patents with higher cited frequency (more than 3) are totally classified in this study. The other target patents which are less cited by target patents indicates less important and can be ignored.

The weight cover index is weighed by the frequency that each target patent was cited by other target patents. The value of weight cover index is 67.3%, which is better than prior studies conducted by Chen (2005) [7], Lai and Wu (2005) [3], and Yeh (2005) [8]. Thus, it can be concluded that this classification has a better performance to classify critical patents, of which the importance is measured by the frequency that was cited by target patents.

The value of consistency index is 95.6%, as shown in Table 5, which shows that seldom patents were classified into multiple factors. Therefore, this study has high consistency in classification system.

Table 4. Names of Factors

Factor	Sub-factor	Name	Amount of basic patents	Amount of target patents*
Factor 1		Information recording medium, recording and reproducing process, method and manufacture, information management and processing.	104	45
	1.1	Recording medium, recording and reproducing process, method and manufacture, rewritable compact disk, erase content, recovering information, protecting copyright, identifying code, optical disc and apparatus, discriminating system.	56	19
	1.2	Information recording medium, defect management, maintaining data, defective area processing, information management, recording, reproducing processing, replacement process, spare area management.	48	27
Factor 2		Image processing, graphics display, video processing apparatus, image information combine, encoding/decoding, subtitle processing, reproducing data.	40	7
Factor 3		Electrochromic printing medium, piracy-protected recording, theft deterrent coating, RFID security for optical disc.	18	12
Factor 4		Controlling Interactive media, plurality of data streams, multiple sources, controlling timing signal, text subtitle data synchronized, organizing data, configuration functions.	16	8
Factor 5		Optical information medium, optical recording medium production method, production apparatus, program, and recordable optical disc.	7	12
Factor 6		Decoding information, reproduction method and apparatus, recording apparatus and playing apparatus.	4	3
Factor 7		Bitmap data encoding, display format, video compression method and system.	3	6

*4 patents are classified duplicately with the same frequency in multiple factors: one patent was classified in factor 1.1 and 1.2.; one is in factor 5 and factor 6; two patents are in factor 1.2 and factor 5.

Table 5. Quantitative indicators

Indicators	Value
Cover index	22.3%
Weight cover index	67.3%
Consistency index	95.6%

3. Conclusion

BD, the next generation in optical discs, offers three major improvements over DVD: a clearer, sharper image; better sound quality; and more special features. The total household penetration of all Blu-ray compatible devices is more than 72 million in U.S. homes in 2013. To sufficiently differentiate techniques among Compact Disc (CD), Digital Versatile Disc (DVD), and BD and catch up with BD technology classifications, this paper adopts the PCA classification system to extracts 8 BD patent categories. The result shows the consistency index is 95.6% and the weight cover index is 67.3% and indicates great classification performance. The result also indicates most basic and target patents are classified in factor 1.1 recording medium, recording and reproducing process, and factor 1.2 information recording medium, defect management. The result can offer explicit intelligence for patent management, technological forecasting, research planning, technological positioning and strategy making for BD technology.

References

[1] Y. Wang, and B. Liu, Innovation Effect on Patent Pool Formation: Empirical Case of Philips' Patents in Digital Versatile Disc (DVD) 3C, *International Journal of Automation and Smart Technology* **3** (2013), 155-167.
[2] D. Archibugi, and M. Planta, Measuring technological change through patents and innovation survey, *Technovation* **16(9)** (1996), 451–468.
[3] K. Lai, and S. Wu, Using the patent co-citation approach to establish a new patent classification system, *Information Processing and Management* **41** (2005), 313-330.
[4] The Digital Entertainment Group (DEG), DEG year-end 2013 Home Entertainment Report, January 7, 2014.
[5] J. O. Lanjouw, and M. Schankerman, Patent quality and research productivity: measuring innovation with multiple indicators, *Economic journal* **114 (495)**, 2004, 441-465.
[6] S. Wu, *Using Patent Co-citation Analysis to Establish Patent Classification Support System- Illustrated with Foundry Industry*, National Yunlin University of Science and Technology, Yunlin, Taiwan, 2003.
[7] K. Chen, *Using Highly Cited and Patent Co-citation to Foresight PDA Technology Trajectory*, National Yunlin University of Science and Technology, Yunlin, Taiwan, 2005.
[8] S. Yeh, *A Study of Patent Portfolio Based on PCA - Case of TFT-LCD Panel Manufacturer*, National Yunlin University of Science and Technology, Yunlin, Taiwan, 2005.