An Analysis of Patent Search Systems

Shigeyuki Sakurai[1] and Alfonso F. Cardenas[2]

1. Introduction

The number of patent applications is growing steadily all over the world. Approximately 445,000 and 467,000 patent applications are filed to the USPTO in FY 2006 and FY 2007 respectively[3]. As for international patent applications under PCT, approximately 149,000 and 156,000 applications are filed to the WIPO in 2006 and 2007 respectively[4]. The number of PCT applications is continuously growing by around 5% per year[5].

The enormous number of patent applications becomes a large backlog for patent authorities such as the USPTO, and furthermore, the enormous number of patents granted means much more difficulty in prior-art search during future patent examination.

At the same time, for the applicants and private companies, the unprecedented volume of patent applications makes it much more difficult to research and find contending technologies of other companies.

To analyze the increasing enormous patent documents, a powerful search system for patent documents and the continuous development of it are significantly important. It helps patent authorities improve their efficiency of prior-art search and applicants improve their efficiency of search for contending technologies of other companies.

Although commercial patent search systems don't disclose the inside structure and the algorithms of their systems clearly, in this research we analyze the existing public and commercial patent

1 Visiting Scholar at Computer Science Department, University of California Los Angeles. Also Patent Examiner at 4th Patent Examination Department, Japan Patent Office, Tokyo, Japan

The opinions in this article are merely the current opinions of the authors.

2 Professor at Computer Science Department, University of California Los Angeles.

3 USPTO, *Performance and Accountability Report Fiscal Year 2007*, page 109, 2007.

4 WIPO, 'Unprecedented Number of International Patent Filings in 2007' (PR/2008/536), Feb. 21, 2008. *http://www.wipo.int/pressroom/en/articles/2008/article_0006.html.*

5 WIPO, *WIPO Overview 2007 edition*, page 32, 2007.

search systems. After that, we consider the evaluation of patent search systems. Also, we consider evaluation data sets essential for promoting development of patent search systems.

2. Analysis of Public and Commercial Patent Search Systems

2-1. Patent search systems

In this section we address existing public, i.e. free, and commercial patent search systems. Table 1 shows search algorithm highlights, searchable fields and operators, result ordering, patent corpus, and cost of each existing patent search system.

The most basic patent search system is the full-text patent search system of the USPTO. We can access the system on the Internet at no cost. Within the system there are two web pages corresponding to two databases, issued patents and published applications. Therefore we can conduct full-text search of publication of issued patents and published unexamined applications. The system offers basic operators such as AND, OR, and ANDNOT. The query "A ANDNOT B" retrieves documents that contain A and do not contain B. After retrieval, it shows a list of documents in reverse chronological order, which means newer publications come first.

Many patent authorities are adopting simple word finding full-text search, because it is fail-safe. Word finding full-text search retrieves all documents that contain all words in a user query. However, the result, i.e. a list of documents, contains a lot of documents, so users have to use more qualifying words to narrow the result.

Google Patent Search was launched in December 2006; thus it is a rather new system. This system scores retrieved patent documents and shows higher scored documents first. It says "As with Google Web Search, we rank patent results according to their relevance to a given search query. We use a number of signals to evaluate how relevant each patent is to a user's query, and we determine our results algorithmically"[6]. But the detail of the search algorithm of Google Patent Search is not provided. (Note that Google Web Search scores results based on PageRank algorithm[7].)

Delphion is a commercial patent search system. This system scores results by the frequency and location of search words. Basic search allows searching of the front page for free, while Premier search allows searching for the full-text of the whole document at a price.

There is a study, [Larkey 1999], which also uses search word frequency. This study makes a ranked list of retrieved patents based on tf-idf, i.e. term frequency-inverse document frequency, weighting score. Therefore it calculates how many times the search words appear in each document, normalizes the word counts, and makes a vector of the word counts for each document.

PatentCafe also uses a vector of word counts. Judging from the explanation of PatentCafe, it compares the vector of search words to vectors on the patent database and ranks the result. PatentCafe is a commercial system.

6 Cited from 'About Google Patent Search' *http://www.google.com/googlepatents/about.html.*

7 Lawrence Page, Sergey Brin, Rajeev Motwani, Terry Winograd, 'The PageRank Citation Ranking: Bringing Order to the Web', Stanford Digital Library Technologies Project, 1998.

Table 1: Patent Search Systems

Name	Search Algorithm, Other Features	Searchable Fields, Operators	Result Ordering	Patent Corpus	Cost
USPTO Patent Full-Text and Full-Page Image Databases http://www.uspto.gov/patft/index.html	Full Text Search	-All Fields -Specific Fields: Patent Number Issue date US Classification and so on -Operators: [Quick search]: AND, OR, ANDNOT [Advanced search]: AND, OR, ANDNOT, Unifier parenthesis "()", Phrase operator "<phrase>"	Reverse chronological order: newer comes first	Full text of US patent publication (1976-) (Patents of 1790-1975 are searchable only by Issue Date, Patent Number, and Current US Classification.) Full text of US patent application publication (Mar. 15th 2001-)	$0
Google Patent Search http://www.google.com/patents	Using a number of signals to evaluate how relevant each patent is to a user's query (cited from "About Google Patent Search")	-All Fields -Specific Fields: Patent Number Issue date US Classification and so on -Operators: AND, OR, Without operator "-" (=ANDNOT), Phrase operator "<phrase>"	Scored Ranking: higher scored comes first	US patent publication (1790s-)	$0
Delphion (The Thomson Corporation) http://www.delphion.com/simple	Front page Search (Basic) Full Text Search (Premier) Scoring by the frequency and location of search word(s) (cited from Delphion) Proximity Search, Wild card, Thesaurus (Premier, Unlimited) Citation Link tool (which shows citation graph)	-Front page (Basic) -All Fields (Premier, Unlimited) -Specific Fields: <in> operator -Operators: AND, OR, NOT(=ANDNOT) Wild card: ? (1 character) * (0- characters) Proximity: <near/n> <order> <sentence> <paragraph> Thesaurus: <thesaurus>	Scored Ranking: higher score comes first	US, EP, DE, JP, CH, WO, DWPI, INPADOC US patent publication (full text is 1974-, bibliographic text is 1971-) Basis registration: Quick searching against the US granted bibliographic collection and Patent number searching against worldwide collections	Basic: Free Premier: $124.50/Mo Unlimited: $249/Mo

Table 1: *(continued)*

Name	Search Algorithm, Other Features	Searchable Fields, Operators	Result Ordering	Patent Corpus	Cost
PatentCafe http://www.patentcafe.com/	Linguistics engine converts entered words into a concept, represented by a mathematical vector. ProSearch matches and ranks the vector to similar vectors in PatentCafe's Semantic database (cited from PatentCafe)	PatentCafe allows to use full calims section as a query. Also it allows certain specific paragraphs in specification as a query (cited from PatentCafe)	Scored Ranking: higher score comes first	US, EP, JP, DE, CA, FR, GB, WO(PCT),	$499/Mo
FreePatentsOnline (started by a former patent searcher) http://www.freepatentsonline.com/	Advanced search techniques: -proximity searching -search term weighing -latent semantic search	-All Fields -Specific Fields -Operators: Proximity: "cat dog"~5 ('cat' within 5 words of 'dog') Relevancy Weight: cat^5 OR dog (Cat is 5 times more important to the relevancy of documents than dog) Wildcard: ? *	Users can choose Reverse chronological order or Relevancy order	US, EP, PAJ(Patent Abstracts of Japan), WO(PCT)	$0
MicroPatent Patent Web (The Thomson Corporation) http://www.micropat.com/static/patentweb.htm	Full Text Search (contains Forward and backward citation data)	-Multi-field search -Operators: Boolean, proximity, and truncation	(unknown)	FullText database: US, DE, EP, GB, JP, and WO INPADOC database (EPO examiner's database)	First $50 deposit is needed. Daily or Annual subscription
AUREKA (charged tool of MicroPatent) http://www.micropat.com/static/aureka.htm	citation analysis (this tool shows citation graph)	-	-	-	Charged
LexisNexis TotalPatent http://www.lexisnexis.com/ip/totalpatent/	(unknown)	-Multiple Field -Multiple Operators -Multiple Wildcards	(unknown)	22 full-text databases of the major patent authorities	subscription basis, also available transactionally

Table 1: *(continued)*

Name	Search Algorithm, Other Features	Searchable Fields, Operators	Result Ordering	Patent Corpus	Cost
DEPATISnet (German Patent and Trade Mark Office) http://www.dpma.de/service/depatisnet.html	Full Text Search	-Multiple Field -Operators: AND, OR, NOT(=ANDNOT) Wild card: ! (1 character) ? (0- characters) # (0-1 characters) Proximity: (P) terms within same paragraph (L) terms within same field (A) finds search terms in any order, with no words between them. (#A) # defines the maximum number of words between the search terms. (W) finds search terms in order with no words between them. (NOTW) search terms in order, but not immediately next to each other (#W) # defines the maximum number of words between the search terms.	User can select how to order the result. Chronological, Reverse-Chronological, Application date, title, and so on.	DE (Full Text) WO, GB, US, JP (Title & Abstract only)	$0
[Larkey 1999][8] (University of Massachusetts Center for Intelligent Information Retrieval)	Search using tf-idf algorithm	All fields	ranked list with tf-idf score	US (1980-1996)	(Used at USPTO)
[Li et al. 2007][9] (University of Arizona Department of Management Information Systems)	Classification algorithm based on "citation network" on patents	Citation field	ranked list of similarity of citation network labeled with subclass	US (451,853 patents, 1/1/1990-)	
[Fall et al. 2003][10] (WIPO, and so on)	IPC classification using Support Vector Machine algorithm	Claims, First 300 words, Abstracts	Assigned IPC class and subclass	WIPO-alpha (1998-2002)	

8 Leah S. Larkey, 'A Patent Search and Classification System', Proc. DL-99, 4th ACM Conference on Digital Libraries, pages 179-187, 1999.

9 Xin Li et al., 'Automatic Patent Classification using Citation Network Information: An Experimental Study in Nanotechnology', Proc. JCDL-07, ACM 2007 conference on Digital Libraries, June 2007.

10 C.J. Fall, A. Törcsvári, K. Benzineb, G. Karetka, 'Automated categorization in the international patent classification', ACM SIGIR Forum archive, Volume 37, Issue 1, Pages 10–25, Spring 2003.

FreePatentsOnline is a free patent search system. Its web page mentions that FreePatentsOnline is using advanced search techniques such as search term weighing and "latent semantic search". However, what "latent semantic search" means is unclear. FreePatentsOnline is interesting because it offers weighting operator for search words to modify scored ranking.

MicroPatent has an interesting tool named AUREKA, which contains the function of citation analysis. (FIGURE 1) The function presents a citation graph, in which cited patents are graphically connected to a patent citing it using a line. Delphion also has a similar tool named Citation Link. (FIGURE 2)

FIGURE 1: AUREKA Advanced Patent Analysis of MicroPatent[11]

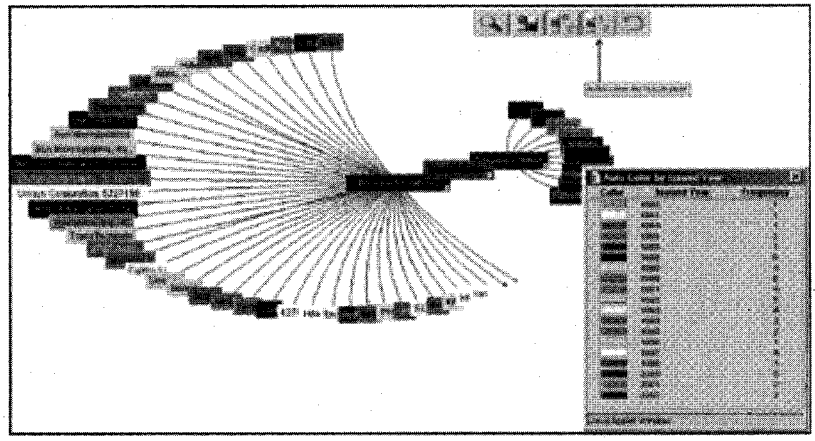
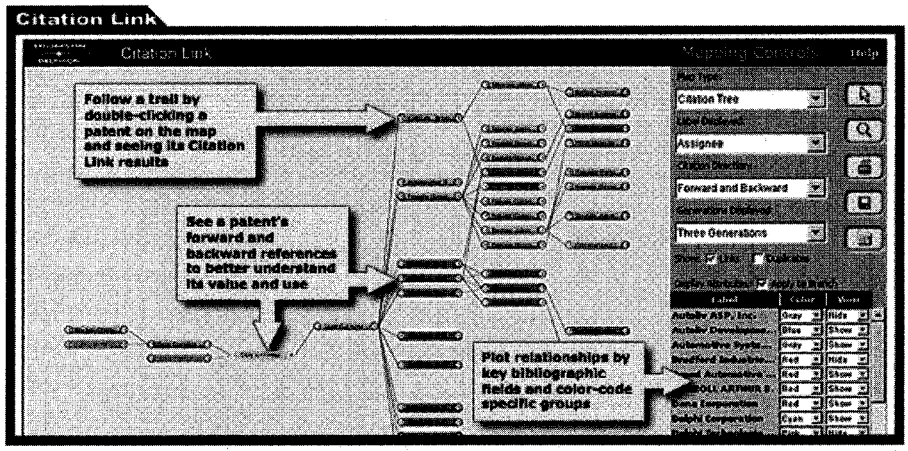


FIGURE 2: Citation Link of Delphion[12]



11 Cited from *http://www.micropat.com/static/advanced.htm.*

12 Cited from *http://www.delphion.com/products/research/products-citelink.*

There is an interesting study using the citation relation among patents. [Li et al. 2007][13] leverages citation relation for automatic classification. It compares random walk paths of a patent to be classified and prior patents. The random walk paths are labeled with classification categories. (FIGURE 3) They found that the algorithm of the labeled graph kernel (K_Gra) has 86.67% accuracy and 88.04% class averaged F-measure.

DEPATISnet managed by the German Patent and Trade Mark Office has some unique features, although its main corpus is Germany patents. One feature is the number and variety of useful proximity operators, e.g. an operator that finds search words within the same paragraph or the same field, and an operator that finds search words in any order (or in order) within the maximum number of words between them. Delphion also has the similar proximity operators. Another feature of DEPATISnet is result ordering. Users can select result ordering from among several options such as reverse-chronological, application date, title, and so on.

2-2. Summary of search algorithms

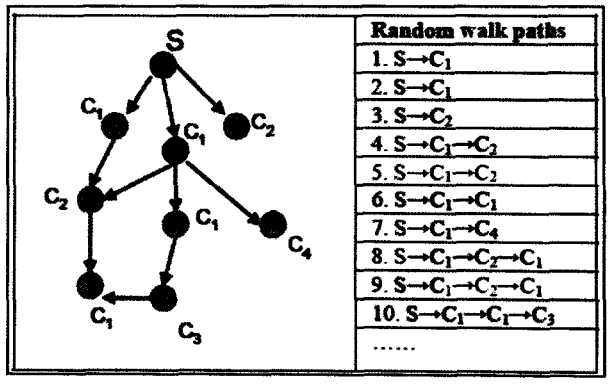
We made a survey of the public and commercial patent search systems. We can summarize that there are three types of patent search algorithms:

(1) Word finding full-text search with index (e.g. USPTO, DEPATISnet)
(2) Semantically scoring search (e.g. tf-idf, Delphion, PatentCafe)
(3) Search using metadata (e.g. citation relation)

Of course it is possible to combine two of these three types in order to implement a patent search engine.

Almost all free patent search systems are adopting word finding full-text search using index. Some non-free search sys-

FIGURE 3: Random walk paths on a patent citation network[14]



S: Start (patent to be classified or prior patent)
C: Cited patent node labeled with classification category

13 Xin Li et al., 'Automatic Patent Classification using Citation Network Information: An Experimental Study in Nanotechnology', Proc. JCDL-07, ACM 2007 conference on Digital Libraries, June 2007.

14 Cited from [Li et al. 2007].

tems implement so-called semantic search algorithms, which are not clearly explained in the web pages of the systems. Some vendors, e.g. Delphion and AUREKA, offer the visualization tool of citation relation among patent documents.

2-3. Summary of searchable Fields

We can conduct field specific search, e.g. patent number search or inventor name search, with almost all patent search vendor software. This is essential for retrieving a specific patent document.

2-4. Summary of operators

The common and popular operators among patent search systems are:

- Boolean: AND, OR, ANDNOT, unifier ()
- Phrase: "<phrase>"
- Wild card: exact 1 character, 0 character or more, 1 character or more
- Proximity: within the same paragraph, within the same field, within n words in order, within n words in any order

2-5. Summary of result ordering

The most basic result ordering is reverse-chronological (newer publication date comes first). Some search systems score the results according to their algorithm and show the results in scored ranking order. Moreover, some systems such as FreePatentsOnline offer a weighting operator for search words to modify scored ranking. DEPATISnet allows to order the result, e.g. reverse-chronological, application date, or patent number order.

3. Evaluation of Patent Search Systems

Let us compare the patent search system above. We selected three search software products, USPTO, Google Patent Search, and FreePatentsOnline, because they are totally free.

We retrieved one sample patent, Patent No. 7,163,290. This patent claims baby sunglasses having a band comprising a mesh piece. With the use of Public PAIR[15] that shows file wrapper of patent applications, we can see the office actions written by the patent examiner in charge of the sample patent. During prosecution, the examiner cites three prior-art patents, No. 5,926,855, 5,481,763, and 5,042,094. So we investigated whether these patent search systems can retrieve the cited prior-art patents.

We made a query suitable for the sample patent referring to the examiner's search strategy in Public PAIR: "((eyeglass OR spectacle OR eyewear) AND band AND mesh) AND 1/1/1976 <=application_date<=12/28/2004". The meaning of the query is to retrieve patent documents that contain the words, "eyeglass OR spectacle OR eyewear" AND band AND mesh, and the patent application date between 1/1/1976 and 12/28/2004. We limited the query with the range of application date because the sample patent was filed on Dec. 28, 2004 and the database of the full-text patent search system of USPTO has data from 1976. The result of retrieval by the three search software products, USPTO, Google Patent Search, and FreePatentsOnline is shown on Table 2. The table shows the results for the query formally indicated in

15 *http://portal.uspto.gov/external/portal/pair.*

the column and in plain English. We retrieved six more sample patents and conducted the same research work showing the results also in Table 2. The numbers of the six patents are 7,239,802, 6,909,889, 7,274,290, 7,232,976, 7,152,184, and 7,095,829.

TABLE 2: The result of retrieval on search systems[16]

Case No.	Patent No. & Issue Date	Query	USPTO		Google Patent Search		FreePatentsOnline	
	(Number of patents cited by examiner)	* The meanings of these queries are displayed in the table below.	number of patents retrieved	cited patents found	number of patents retrieved	cited patents found	number of patents retrieved	cited patents found
#1	7,163,290 Jan. 16, 2007 (3)	((eyeglass OR spectacle OR eyewear) AND band AND mesh) AND Application_date/1/1/1976-> 12/28/2004	69	1	8	1	192	1
#2	7,239,802 Jul. 3, 2007 (3)	(USClass/715/805 OR USClass/707/1) AND select AND button AND Application_date/1/1/1976-> 11/30/2001	383	2	204	1	486	2
#3	6,909,889 Jun. 21, 2005 (3)	(mobile OR wireless OR cell$) AND digital AND (photo OR image) AND "print condition" AND Application_ date/1/1/1976-> 10/5/2001	36	1	1	0	39 (Word Stemming is off because using wildcard)	1
#4	7,274,290 Sep. 25, 2007 (5)	medical AND tray AND wireless AND mobile AND Application_ date/1/1/1976->5/11/2006	317	1	111	1	423	1
#5	7,232,976 Jun. 19, 2007 (4)	"mouse pad" AND therapeutic AND Application_date/1/1/1976-> 1/18/2006	31	1	9	1	51	1
#6	7,152,184 Dec. 19, 2006 (5)	(USClass/707/204 OR USClass/711/162 OR USClass/714/5) AND input AND update AND second AND third AND Application_ date/1/1/1976->11/22/2001	348	3	32	1	522	4
#7	7,095,829 Aug. 22, 2006 (6)	(USClass/707/200 OR USClass/707/204 OR USClass/709/206 OR USClass/709/225 OR USClass/379/88.23) AND (mail OR message) AND archive AND recipient AND Application_date/1/1/1976-> 9/7/2004	148	3	65	2	285	3

16 Using database of issued patents only, not using database of published applications.

Case No.	Meaning of the query
#1	To retrieve patent documents that contain the words, "eyeglass OR spectacle OR eyewear" AND band AND mesh, and the patent application date between 1/1/1976 and 12/28/2004.
#2	To retrieve patent documents that are classified to 715/805 OR 707/1 under U.S. Patent Classification, and that contain the words, select AND button, and the patent application date between 1/1/1976 and 11/30/2001.
#3	To retrieve patent documents that contain the words, "mobile OR wireless OR cell$ ($ is indicating truncation.)" AND digital AND "photo OR image" AND "print condition" (phrase), and the patent application date between 1/1/1976 and 10/5/2001.
#4	To retrieve patent documents that contain the words, medical AND tray AND wireless AND mobile, and the patent application date between 1/1/1976 and 5/11/2006.
#5	To retrieve patent documents that contain the words, "mouse pad" (phrase) AND therapeutic, and the patent application date between /1/1/1976 and 1/18/2006.
#6	To retrieve patent documents that are classified to 707/204 OR 711/162 OR 714/5 under U.S. Patent Classification, and that contain the words, input AND update AND second AND third, and the patent application date between 1/1/1976 and 11/22/2001.
#7	To retrieve patent documents that are classified to 707/200 OR 707/204 OR 709/206 OR 709/225 OR 379/88.23 under U.S. Patent Classification, and that contain the words, "mail OR message" AND archive AND recipient, and the patent application date between 1/1/1976 and 9/7/2004.

The result shows that the three search systems found different numbers of patents satisfying the query, and found only a subset of the number of prior-art patents cited by the patent examiner in charge of each patent (the number of prior-art patents cited by the examiner is indicated in parenthesis under the patent number in the table).

We found the fact that the three systems retrieved the closest prior-art documents cited by the patent examiner. We call the closest prior-art "the main prior-art". For example, for case #1 (Patent no. 7,163,290), Patent 5,926,855 is cited as a prior-art in the actions written by the examiner and the prior-art patent is the main prior-art document. All three systems retrieved the main prior-art document for case #1. Whether or not a patent search system can retrieve the main prior-art documents that are the closest as a whole to the application at issue is important.

In addition, we can say that after retrieval of the main prior-art patent, examiners conduct other searches using other queries to find the rest of the prior-art documents. These can be combined with the main prior-art document to bridge the difference between the main prior-art document and the patent application at issue (e.g. case #1). So we need other queries to retrieve the rest of the prior-art documents. We call the rest of the prior-art "the sub prior-art".

Our research shows, as illustrated in Table 2, that Google Patent Search retrieves fewer documents when we add more search terms in a query. This is because the aim of Google Patent Search is different from other patent search systems. It places more emphasis on focus, i.e. narrowing down documents retrieved, and scored relevancy ranking of documents retrieved. The search system of the USPTO uses a simple word finding algorithm in order to be fail-safe, and retrieves more documents than Google Patent Search. FreePatentsOnline retrieves the most number of documents that are thought to be

relevant to a user query. One of the reasons for this is the Word Stemming function of FreePatentsOnline, which expands each word of a user query to other forms of the word. Note that within FreePatentsOnline we can choose result ordering: reverse chronological order or relevancy order. This is a useful function.

One possible usage of patent search systems is to combine use of different types of the systems. First we can use Google Patent Search to retrieve a narrowed set of prior-art documents in relevancy order. After that we can use a fail-safe search system such as the USPTO system or FreePatentsOnline to retrieve a more comprehensive set of prior-art documents.

The discussion in this section addresses prior-art patent search. However, we recognize that patent examiners also conduct non-patent literature searches as well as patent searches. Patent examiners often also cite documents of non-patent literature such as proceedings of conference or articles in magazines. Thus, there is a potential need for a system to search non-patent literature.

4. Test Set and Training Set of Patent Publications

Some projects such as WIPO-alpha managed by WIPO[17] and NTCIR in Japan[18] have provided test sets and training sets of patent publications for researchers of patent search algorithms. However, the data sets are available on a registration basis and for a limited time.

To promote research for automatic patent search algorithms, patent authorities such as the USPTO should provide an experimental set of patent documents. We suggest a possible experimental set of patent documents that contains:

Training collection of patent publications
Target test collection of published patent applications
- Prior-art publications as an answer to target patent applications

We suggest providing the patent publications cited in the office actions written by patent examiners, i.e. Non-Final Rejections and List of references cited by examiner, as an answer to target patent applications. Patent examiners try to find the closest prior-art during examination. Therefore, we can use the patent publications cited by patent examiners as an answer to target patent applications. The patent publications cited by patent examiners are also shown in the field of "References Cited" on the cover page of patent publications with asterisk.

5. Conclusion

In this research, we analyzed existing public and commercial patent search systems, investigated the search algorithms and search results for sample patents, and considered the common operators and the result ordering. Continuous research and development of patent search systems are

17 C.J. Fall, A. Törcsvári, K. Benzineb, G. Karetka, 'Automated categorization in the international patent classification', ACM SIGIR Forum archive, Volume 37, Issue 1, Pages 10–25, Spring 2003.

18 Sumio Fujita, 'Technology survey and invalidity search: A comparative study of different tasks for Japanese patent document retrieval', Information Processing and Management: an International Journal archive, Volume 43, Issue 5, pages 1154-1172 September 2007.

significantly important to accelerate patent examination and streamline search for contending technologies. To promote the research, providing testing and training sets of patent documents is essential. We hope for a more active and open discussion of established patent search systems and algorithms.

REFERENCES

USPTO, *Performance and Accountability Report Fiscal Year 2007*, page 109, 2007.

WIPO, 'Unprecedented Number of International Patent Filings in 2007' (PR/2008/536), Feb. 21, 2008. *http://www.wipo.int/pressroom/en/articles/2008/article_0006.html.*

WIPO, *WIPO Overview 2007 edition*, page 32, 2007.

'About Google Patent Search' *http://www.google.com/googlepatents/about.html.*

Lawrence Page, Sergey Brin, Rajeev Motwani, Terry Winograd, 'The PageRank Citation Ranking: Bringing Order to the Web', Stanford Digital Library Technologies Project, 1998.

USPTO Patent Full-Text and Full-Page Image Databases http://www.uspto.gov/patft/index.html.

Google Patent Search http://www.google.com/patents.

Delphion http://www.delphion.com/simple. *http://www.delphion.com/products/research/products-citelink.*

PatentCafe http://www.patentcafe.com/.

FreePatentsOnline http://www.freepatentsonline.com/.

MicroPatent Patent Web http://www.micropat.com/static/patentweb.htm. *http://www.micropat.com/static/advanced.htm.*

AUREKA http://www.micropat.com/static/aureka.htm.

LexisNexis TotalPatent http://www.lexisnexis.com/ip/totalpatent/.

DEPATISnet http://www.dpma.de/service/depatisnet.html.

[Larkey 1999] Leah S. Larkey, 'A Patent Search and Classification System', Proc. DL-99, 4th ACM Conference on Digital Libraries, pages 179-187, 1999.

[Li et al. 2007] Xin Li et al., 'Automatic Patent Classification using Citation Network Information: An Experimental Study in Nanotechnology', Proc. JCDL-07, ACM 2007 conference on Digital Libraries, June 2007.

[Fall et al. 2003] C.J. Fall, A. Törcsvári, K. Benzineb, G. Karetka, 'Automated categorization in the international patent classification', ACM SIGIR Forum archive, Volume 37, Issue 1, Pages 10–25, Spring 2003.

USPTO Public PAIR
http://portal.uspto.gov/external/portal/pair.

[Fujita 2007] Sumio Fujita, 'Technology survey and invalidity search: A comparative study of different tasks for Japanese patent document retrieval', Information Processing and Management: an International Journal archive, Volume 43, Issue 5, pages 1154-1172, September 2007.

[Lamirel et al. 2003] Jean-Charles Lamirel, Shadi Al Shehabi, Martial Hoffmann, Claire François, 'Intelligent patent analysis through the use of a neural network: experiment of multi-viewpoint analysis with the MultiSOM model', Proceedings of the ACL-2003 workshop on Patent corpus processing, Volume 20, Pages 7-23, 2003

[Chen et al. 2003] Liang Chen, Naoyuki Tokuda, Hisahiro Adachi, 'A patent document retrieval system addressing both semantic and syntactic properties', Proceedings of the ACL-2003 workshop on Patent corpus processing, Volume 20, Pages 1-6, 2003