

# 특허 인용분석을 이용한 영향력 있는 특허를 찾는 방법 및 특허정보 분석 시스템 설계

이재현<sup>o</sup>

인포클루

{amuro@infoclue.co.kr<sup>o</sup>}

The method find powerful patent or patent information analysis system using patent citation analysis

JaeHeon Lee<sup>o</sup>

Infoclue corp

요 약

산업이 발전함에 따라 지적재산권에 대한 가치는 점점 중요해지고 있고, 기업의 흥망성쇠를 결정짓는 가장 큰 요인중 하나로 자리잡고 있다. 엄청난 비용을 투자하여 사업화하고 있던 세계 굴지의 기업들도 경쟁회사의 핵심특허 단 몇 건만으로 순식간에 몰락하는 경우도 발생하고 있다. 이제 지적재산권은 살아남기 위한 유일한 방어수단이자 공격수단이 되었다. 이와 같은 급박한 상황에서 각 기업은 경쟁회사의 지적재산권을 분석하고, 이를 통해 자사의 연구개발 전략을 수립하기 위해 노력하고 있다. 그러나 이와 같은 분석작업은 다수의 전문가 일지라도 상당한 시일이 소요되는 방대한 작업이며, 이를 자동으로 실행하기 위해 많은 소프트웨어 회사들은 특허분석용 툴을 개발하였다. 그러나 이들 종래의 특허분석용 툴들은 대부분 외형적인 정량분석 정보만을 제공할 뿐, 실제 필요로 하는 정성적인 정보는 거의 제공하지 못하였다. 즉, 종래의 특허분석용 툴에서는 자사 및 경쟁사의 특허에 대한 가치가 어느 정도이며, 경쟁회사 특허에 대해 자사 제품의 침해 가능성은 어느 정도이며 회피할 수 있는 전략은 무엇인지, 자사 특허에 대한 경쟁사 제품의 침해 가능성은 어느 정도인지 등을 제공하지 못하였다. 이에 본 논문에서는 특허 출원시 인용했던 선행특허들의 인용 관계를 분석하여 특허간 관계를 규명하고, 각 기술 분야에서 영향력이 높은 특허를 찾는 방법과 시스템을 제시하고자 한다.

## 1. 서 론

기술수준을 측정하기 위해서는 과학기술활동의 투입-과정-성과-영향에 관련된 지표를 체계화해야 할 필요가 있으나 과학기술의 성과와 영향은 매우 측정이 어렵기 때문에 이를 간접적으로 잘 나타내 줄 수 있는 대리변수로서 특허정보가 널리 활용되고 있다. 특허는 신기술을 이해하는데 있어서 빼놓을 수 없는 객관적이고 표준적인 기술정보이며 기술수준과 기술혁신 흐름뿐만 아니라 기술혁신 방향 등을 분석하는데 널리 활용되고 있다 [1].

특허 자료를 이용한 분석은 대개 서지적 사항을 이용하게 된다. 즉, 출원 번호, 출원 일자, 제목, 발명자, 인용 특허 등과 같은 사항을 이용함으로써 발명의 특성 파악과 공백 기술의 발견 등이 가능하다. 이와 같은 특허의 서지적 사항 분석에 대한 연구는 기술 개발 영역의 현 상태와 추세를 도출하는 방법에 대해 활발하게 진행되고 있다 [2]. 이러한 연구들 중 활용의 범위나 정도에 있어 매우 중요하게 인식되고 있는 특허 인용 분석은 특허가 거시적 차원에서 갖는 경제적 가치를 측정하고 미시적으로는 기업이 보유하고 있는 기술의 중요도 평가, 타 기업과의 기술 중복의 발견 등에 이용될 수 있다. 또 한 지식의 흐름의 양상을 인용의 측면에서 조명하여 기술간, 기업간, 산업간 흐름을 규명하는데 기초를 제공하였다. 따라서 특허 인용 분석은 방대한 양의 특허 자료를 분석하는 데에 적당한 방법을 찾지 못하던 기존의 연구들에 유용한 도구로서 자리 잡고 있다고 할 수 있

다 [3].

그러나 기존의 특허정보 시스템 및 특허분석 프로그램 등에서는 검색결과 얻어진 특허들의 정량적 분석을 제공하지만 정성적인 부분을 제시하지 못하고 있으며, 인용/피인용의 횟수와 특허들만 제시하고 있다. 또한 특허 인용을 그 기술에 대한전문가나 변리사가 판단함에 따라 시간과 비용이 발생하고 판단자의 주관이 개입됨으로써 정확성이 떨어질 수 있다. 따라서 본 논문에서는 특허 출원시 인용했던 선행특허들의 인용 관계를 분석하여 특허간 관계를 규명하고, 각 기술 분야에서 영향력이 높은 특허를 빠르고 효과적으로 찾는 특허정보 분석시스템을 제안한다.

## 2.배경이론 및 기존의 검색시스템

### 2.1 특허인용분석(Patent Citation Analysis)

특허 인용은 한 특허가 다른 특허나 비 특허 분야에서 인용(citation) 된 횟수로 정의된다 [4]. 특허 개발자나 기업이 기존 특허들을 인용하는 의도에 대한 세부적 증거의 부족에도 불구하고 타 특허에 인용된 횟수가 많은 특허가 높은 기술적 가치를 지닌다는 것은 다양한 연구를 통해 증명된 바 있다.특허 인용 분석은 특허들 간의 인용 관계에 대한 분석을 통하여 상호 연관 관계 뿐만 아니라, 상대적인 중요도 등에 대한 분석을 하는 것을 의미한다 [5]. 많은 연구자들은 이를 통해 기술의 질과 영향력, 기술 정보의 확산 등을 분석하였다. Pavitt은 특

허의 공통 인용 분석 특허 인용 분석이 활용되는 실용적, 학문적 분야는 다양하였으며 [6], Narin과 Noma는 생명공학에서 과학과 기술 사이의 지식 흐름의 구조를 특허 인용 분석을 통해 보았다 [7].

대표적으로는 기술의 영향력이나 기술의 질을 평가하는 기본적인 지표로서 인용된 횟수를 고려하는 인용 빈도(highly cited patents)가 있다. 이외에도 NPL(number of Non-Patent Link)과 CII(Current Impact Index) 등 여러 종류의 지표들이 존재한다. NPL은 특허화된 기술의 과학에 대한 의존 정도를 나타내는 지표이고 CII는 일정한 기간동안 특정한 특허가 인용된 빈도수로 특허의 현재 영향력 정도를 평가하는 지표이다.

2.2 데이터 마이닝을 이용한 특허 인용분석

논문에서는 특허 문서에서 특허의 특성을 설명할 수 있는 주제어를 텍스트 마이닝을 통해 추출하고, 이를 토대로 각 특허간의 관련성 정도를 파악하여 연관성을 분석하였다.

2.3 상관관계 [8]

두 변량 사이에 한쪽이 증가하면 다른 쪽도 증가(또는 감소)하는 경향이 있을 때, 이 두 변량 사이에는 상관관계가 있다고 한다.

상관관계는 상관계수를 구하여 알 수 있는데, 상관계수가 +1이면 하나의 변수가 양으로 증가함에 따라 다른 변수가 같이 양으로 증가하는 양의 상관관계를 의미한다. 그러나 하나의 변수가 양으로 증가함에 따라 다른 변수가 역으로 감소하는 음의 상관관계인 경우에는 상관계수가 -1이 되며, 상관계수가 0이면 두 변수사이에 상관관계가 없다고 본다.

2.3 기존의 검색 시스템의 인용관계

0000 0000 0000

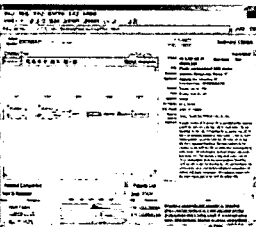
B 5	0	7
B 4	0	
B 3	0	
B 2	0	
B 1	7	
Root	1	1
F 1	135	135
F 2	0	
F 3	0	
F 4	0	
F 5	0	
		143

[표1] wips

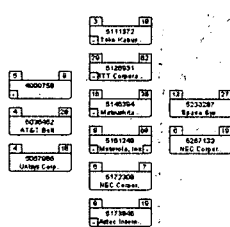
Forward references (135) Backward references (7)

Patent	Pub.Date	Assignee
US4654867	1987-03	Motorola, Inc.
US4754453	1988-06	U.S. Philips Corporation
US4763322	1988-08	U.S. Philips Corp.
US4766753	1988-08	U.S. Philips Corporation
US4876740	1989-10	Motorola, Inc.
US4910510	1990-03	Motorola, Inc.
US4914649	1990-04	Motorola, Inc.

[표2] Delphion



[그림1] wisdomain



[그림2] wips

[표1], [표2]는 wips[9]와 delphion[10]에서는 선후 인용관계를 표로 제공 한다. 표에 있는 숫자를 클릭하면 해당 특허가 인용한 특허(Backward Reference, B/R)와 해당 특허를 인용한 특허(Forward Reference, F/R)들의 특허번호를 보여준다.

[그림1][11]과 [그림2]에서는 인용관계 특허들의 연결을 그림으로 제공하고 있다. 그림에서는 해당 특허들이 인용되고 있는 모습을 선으로 나타내고 있다.

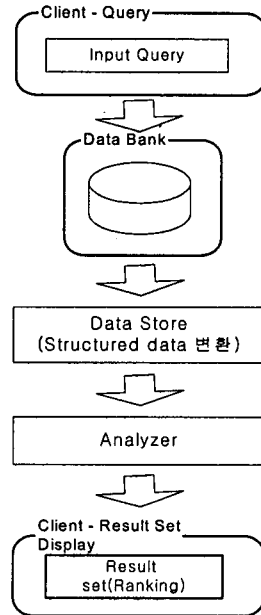
기존의 시스템에서는 선후 인용관계를 표와 도면으로 사실 정보만 제공한다. 따라서 해당 특허가 관련 기술분야 미치고 있는 영향력의 정도를 제시 하지 못한다.

3. 시스템의 구성

3.1 전체 시스템의 구성

[그림3]은 본 논문에서 제안하는 특허정보 분석시스템(Pinfos: Patent information Analysis System)의 전체 구조를 나타내고 있다.

Pinfos는 4단계의 절차를 거친 후 최종 결과를 제공한다. 가중치를 부여하기 전에 인용 특허간 상관계수 값에 따라, 가중치 부여를 차등으로 한다.



[그림3] Pinfos의 전체 구조

[표3]은 시스템에서 제공하는 인용관계 특허리스트를 나타낸 것이다.

특허번호	Forward Ref 수	Forward Ref 특허	Backward Ref 수	Backward Ref 특허	Family 특허수	가중치	순위
US5042082	135	..	7	..	29		
US5067147	112	..	11	..	16		
..	..	..	..	..	..		
..	..	..	..	..	..		
..	..	..	..	..	..		

[표3] 인용관계 특허리스트

첫째, 사용자 질의에서 사용자는 자연어로 또는 불리언

질의를 작성하여 검색을 수행한다.

둘째, 검색이 완료된 후, 특허간의 관계를 규명하기 위해 특허 문서를 구조화된 자료로 변환한다.

셋째, Analyzer에서는 특허 간 연관관계를 분석하기 위해 4단계 프로세스를 수행하며, 인용관계리스트를 생성하여 데이터들을 관리한다.

- ① 'Forward Reference'가 높은 순서로 대상 특허리스트의 가중치 필드에서 가중치를 부여한다.
- ② Family 특허수를 조사하여 해당 필드에 특허수를 추가한다.
- ③ 'Backward Reference'가 낮은 순서로 대상 특허리스트의 가중치 필드에서 가중치를 부여한다.
- ④ 'Backward Reference'로 언급된 특허들을 분석한후, 대상 특허리스트와 'Backward Reference'로 언급된 특허들에서 중복 여부를 확인한다. 대상 특허리스트에 특허가 없는 경우 기존의 대상특허리스트에 특허를 추가한 후, ① ② ③절차를 수행한다.

넷째, 분석 결과는 순위를 부여한 후 사용자에게 제공한다. 순위가 낮을수록 관련 기술 분야에서 영향력이 큰 특허들이다.

#### 4. 연구 결과

본 연구는 미국에서 2005년까지 3월까지 등록된 특허 중 GSM 기술 분야 Handover 기술에 관련하여 분석을 수행 하였다. 초기 검색 결과 2000건의 특허에서 F/R가 있는 1018건의 특허를 갖고 분석을 시작하였다.

##### 4.1 상관관계 분석

분석 자료의 상관관계를 분석한 결과는 아래와 같다.

Pearson correlation of B/R and F/R = 0.169  
P-Value = 0.014

correlation 값이 0 보다 크므로 양의 상관관계를 이루고 있다. 따라서 B/R and F/R는 서로 상관관계가 있다.

##### 4.2 분석 결과

Analyzer 단계의 4단계의 프로세스를 진행하면서 ④의 단계에서 기존의 대상 특허리스트에 없는 특허를 60여건 발견을 하여 추가하였다. 이 특허들은 초기 검색식으로 특허를 검색하지 못한 특허들이다.

[표4]은 추가된 특허들 중에서 GSM Handover 기술 분야에서 매우 영향력이 높은 특허들이다.

특허번호	F/R 수	특허번호	F/R 수
US4765753	215	US4759051	83
US4670899	181	US5243598	78
US5303286	163	US5164958	76
US4654867	158	US5109528	75
US4775999	134	US4932049	71
US4475010	113	US5161248	68
US4737978	111	US4696027	63
US4704734	106		

[표4] 추가된 특허 중 중요 특허리스트

'Backward Reference'의 가중치를 부여하기위해, 'Forward Reference'의 값의 범위에 가중치를 차등 부여하였다.

분석결과 현재 GSM Handover 기술 분야에서 가장 영향력이 높은 200여건의 특허를 순위를 부여 선별하였다. 순위가 낮을수록 해당 기술 분야에서 높은 영향력을 발휘 하고 있는 특허들이다.

#### 5. 결론 및 향후 연구

최근들어 자주 등장하는 외국 경쟁사와의 특허분쟁 사례에서도 볼 수 있듯이 원천기술을 확보한 기업과 상용화 기술을 보유한 기업 간 공통기술 분야에서의 특허분쟁은 필연적으로 발생할 수밖에 없는 상황에 이르렀다.

다 특허에 인용된 횟수가 많은 특허가 높은 기술적 가치를 지닌다는 것은 다양한 연구를 통해 증명된 바 있다. [12], 본 논문은 특허 간 인용관계를 분석하여 특허전략의 첫 출발점인 해당 기술 분야의 원천 특허 또는 영향력이 높은 특허를 빠르고 쉽게 찾을 수 있는 방법 및 특허정보 분석시스템을 제시하였다.

향후 연구로는 해당 특허가 실질적으로 영향력을 미치고 있는지의 검증 작업이 필요하며, 미래에 영향력을 발휘할 특허들을 구분해 내는 작업이 향후 연구과제로 진행할 예정이다. 이러한 연구 결과는 관련 기술 분야의 전체 특허를 분석하지 않고 빠르게 중요 특허를 제시함으로써 분쟁 상황에 빠르게 대처 할 수 있도록 한다.

#### 6. 참고논문

1. IT 및 BT 분야의 기술수준 평가 및 정책적 시사점: 미국특허의 인용도 분석, 윤문섭, 이우형 SETPI, 2003
2. Karki M., 'Bibliometric Analysis of Patents: Implications for R&D and Industry', Emerging Trend in Scientometrics(edit by Nagpaul P. et al.), 1999
3. 데이터 마이닝을 이용한 특허 인용분석, 윤병윤, 백재호, 박용태
4. Freeman, L., 'Turning a Profit from Mathematics : The Case of Social Networks', *Journal of Mathematical Sociology*, Vol. 10, pp. 343-360, 1984
5. Karki M., 'Patent Citation Analysis: A Policy Analysis Tool', *World Patent Information*, Vol. 19(4), pp. 269-272, 1997
6. Pavitt, K., 'Uses and Abuses of Patent Statistics', Handbook of Quantitative Studies of Science and Technology(edit by Van Raan), pp. 239-256, 1998
7. Narin, F. and Noma E., 'Is Technology Becoming Science?', *Scientometrics*, Vol. 7, pp. 259-273, 1985
8. <http://www.jeonju.ac.kr/%7Egina/spss/corr.htm>
9. www.wips.co.kr
10. www.delphion.com
11. 한국 공개특허(KR2003-0012516), 특허데이터베이스로부터 특허간의 상화관계를 분석하여 제공하는 방법
12. Albert, M., Avery, D., Narin F. and McAllister, P., 'Direct Validation of Citation Counts as Indicators of Industrially Import Patents', *Research Policy*, Vol. 20, pp. 123-134, 1991