

기술기회발굴을 위한 문서인용기반 기술진화 경로 생성 모형

Models for Technology Evolution Path Creation Based on Citation Tree
to Investigate Technology Opportunity Discovery

이재민(Jae-Min Lee)*, 이방래(Bangrae Lee)**,
문영호(Yeong-ho Moon)***, 권오진(Oh-Jin Kwon)****

목 차

- | | |
|-------------------|-------------------|
| I. 서론 | III. 사례 연구 |
| II. 인용기반 핵심문서 계보도 | IV. 결론 및 향후 연구 방향 |

국 문 요 약

본 연구에서는 기술문서의 인용 정보를 기반으로 직접인용과 간접인용을 고려하여 핵심 문서를 선별하였고 선별된 문서 간의 인용네트워크 트리(citation network tree)를 생성하여 기술진화 경로를 추출하고자 하였다. 활용 예시로 OLED(유기발광다이오드) 분야 특허데이터를 분석하여 핵심문서를 선별한 후 문서간 인용관계 트리 계보도를 생성하였고, 그 중 OLED 원천기술이 어떤 기술진화 경로를 통해 반도체 관련 기술로 전이되었는지 분석하였다. 다른 한편으로 그래핀(graphene) 분야 논문의 인용관계 트리 분석을 통하여 간접인용을 고려한 가중치 계산방법이 직접인용만을 고려한 피인용 회수 계산법보다 현실을 더 잘 반영함을 고찰하였다.

핵심어 : 인용 네트워크, 간접인용, 기술진화, 유기발광다이오드(OLED), 그래핀(graphene)

※ 논문접수일: 2011.10.25, 1차수정일: 2011.11.11, 게재확정일: 2011.11.28

* 한국과학기술정보연구원 선임연구원, jmlee@kisti.re.kr, 02-3299-6292

** 한국과학기술정보연구원 선임연구원, brilee@kisti.re.kr, 02-3299-6052

*** 한국과학기술정보연구원 책임연구원, yhmoon@kisti.re.kr, 02-3299-6090

**** 한국과학기술정보연구원 선임연구원, dbajin@kisti.re.kr, 02-3299-6097, 교신저자

ABSTRACT

We selected core documents from enormous number of documents by analyzing citation relation of papers or patents including direct citation and indirect citation. Then we tried to creat the technology evolution path from citation network tree. By applying the method to the patent DB of OLED (Organic Light Emitting Diode), we obtained genealogical citation network of core patents of OLED. We analyzed how the one of OLED technology was transferred to the semiconductor related technology and we named the process of transition as 'technology evolution path' of OLED technology. And we also analyzed the genealogical citation network of papers on graphene. From the analysis, we found that the weight count method including indirect citation was better in evaluating the value of technology of a paper than the times cited method.

Key Words : Citation network, Indirect citation, Technology evolution, OLED(Organic Light Emitting Diode), Graphene

I. 서 론

기술개발 단계는 크게 기술기획, 기술개발 그리고 사업화단계로 구분할 수 있다. 우리나라 604개 중소기업을 대상으로 한 설문조사 결과에서 제품개발 단계별 중요도 평가를 보면 기술 기획부분에서 아이디어 검토가 전체에서 7.8% 그리고 사전기술평가가 5.7%의 비중을 갖는 것으로 조사되었다(KISTI, 2011). 아이디어 검토와 사전기술평가의 단계가 중소기업의 R&D 과정에서 차지하는 비중이 작지 않다. 또한 기술기획과 아이디어 검토는 초기 기술개발의 방향을 정하는데 결정적인 영향을 주기 때문에 이후에 진행되는 전체 R&D 과정에 소요되는 경제적 비용과 시간 그리고 기술 사업화의 성패를 좌우하는 단계이다. 따라서 초기에 아이디어 검토와 사전기술평가를 얼마나 성공적으로 수행했는지에 따라 사업화의 성패를 좌우한다고 할 수 있다. 본 논문에서는 아이디어 검토 및 사전기술평가에 도움이 될 수 있는 방법론으로써 기술 문헌간 인용관계를 기반으로 한 기술진화경로생성모형을 소개하고자 한다. 이를 통해 기술의 진화경로, 기술의 전이 방향에 대한 분석 그리고 가중치 계산에 기반한 평가의 효용성을 제시하였다.

아이디어 검토 및 사전기술평가를 위해서는 내·외부로부터의 정보자원과의 접근이 필수적이다. 논문과 특허는 R&D 성과의 대표적 성과물로서 비록 라이선스라는 제약이 있지만, 비교적 일반 사용자가 접근하기 쉬우며 경제적인 비용으로 획득할 수 있는 정보 자원이다. 특히 특허 정보의 경우 산업과의 높은 연관성으로 인하여, 특허 분석을 통해 미래유망기술 아이템을 발굴하는 방법론이 연구되어졌고, 특허정보를 통해 기술의 파급효과를 측정하려는 시도도 있었다(고병열, 2005. 유선희, 2007). Kerr(2006)의 인텔리전스 검색 양식 분류를 적용해 본다면 논문과 특허문헌은 기업 외부의 개방형인 ‘스캔(Scan)’과 통제형인 ‘타깃(Target)’에 해당하며, 본 논문에서 다루고자 하는 정보자원은 탐색의 의미를 갖는 개방형 정보 검색인 ‘스캔’에 좀 더 가깝다고 할 수 있다.

기업이 기술 혁신을 이루기 위해 시도하는 여러 행위들 중에 정보의 수집이 그 출발점이라 할 수 있다. 그런 측면에서 전문적인 과학기술 정보를 담고 있는 논문과 특허는 규격에 맞게 잘 정리되어 있으며 전문가 평가 등의 검증된 평가를 통해 선별된 것으로써 신뢰할 수 있는 정보자원이다. 논문과 특허는 양질의 정보를 포함하고 있지만, 그 양이 방대하고 또한 4년마다 2배씩 기하급수적으로 증가하고 있다(Van Rann, 2005). 논문과 특허를 효율적으로 활용하기 위해서는 해당 분야에 대한 전문적인 지식을 갖는 것도 중요하지만 무엇보다도 막대한 정보 속에서 의미 있는 정보를 찾는 것이 중요하며 그 방법도 효율적이어야 한다.

본 논문에서는 과학기술문서(논문과 특허)의 인용관계를 근거로 의미 있는 기술문서를 선

별하고자 하며, 직접인용과 간접인용을 동시에 고려하였다. 간접인용에 대해서는 간접인용 단계에 반비례하는 값을 곱하여 계산한 후 직접인용의 값에 합산하였으며, 그 값을 바탕으로 검색데이터 집합에서 핵심 문서를 선별하였다. 선별된 핵심 문서들은 연도별로 배열된 후 네트워크를 형성하게 된다. 이때 시각화되어 보여지는 문서 네트워크는 인용관계와 함께 기술의 역사적 관계를 보여준다. 본 논문에서는 유기발광다이오드(OLED, Organic Light Emitting Diode) 기술을 주제로 3890여건의 특허를 검색하였고, 검색 집합 전체에 대한 가중치 계산을 통해 핵심문서 30건을 선별하였다. 이 중 “Multi-color OLED(US-5707745)” 기술 특허를 중심으로 인용 네트워크 계통도를 도시하였고 핵심 문서간의 연결 관계와 역사적 관계를 바탕으로 기술이 어떤 형태로 진화하고 전파되었는 지를 분석하였다. 한편으로는 그래핀(Graphene) 관련 학술 논문 1만여 건에 대해서 간접인용 가중치 계산을 수행하였고, 가중치 계산에 바탕을 둔 순위와 피인용회수 순위를 비교하여 네트워크 가중치 계산의 유효성을 확인하였다.

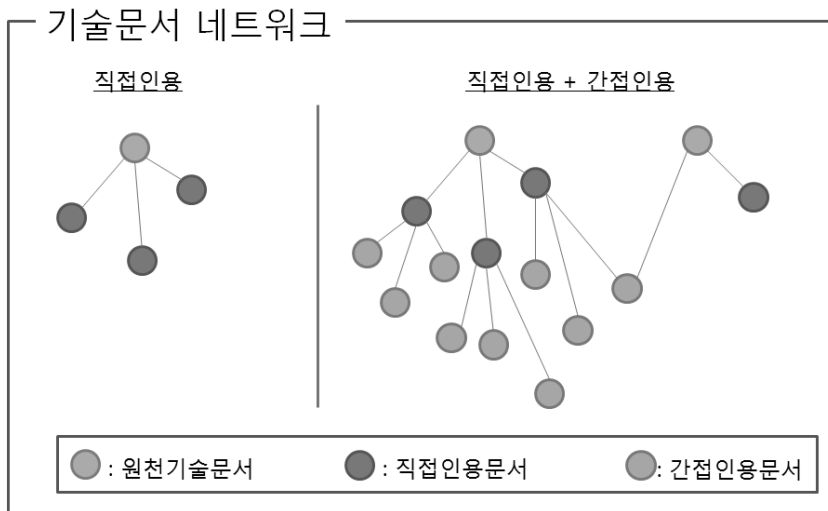
II. 인용기반 핵심문서 계보도

1. 인용분석의 활용

과학기술 문서의 수준을 평가하는 보편적인 지표로써 피인용 회수를 들 수 있다. 인용은 선행 연구자의 연구 업적을 차용함에 따른 반대급부이자 평가이며, 표절 의혹을 피하기 위한 안전장치이기도 하다. 1964년 Eugene Garfield는 “The Use of citation data in writing the history of science”를 저술하면서 인용 분석(Citation Analysis)의 기본적인 개념을 소개하였다. 문서간 인용에 기반한 ranking, Science Citation Index 등의 개념이 연구되었다(Garfield, 1964). 또한 Garner는 문서간 인용 관계를 하나의 거대 network 로 인지하여 분석하였다(Garner, 1967). 특허에서의 인용 관계에 대한 연구도 다음과 같이 보고된 바 있다. 특허 인용분석의 경우 특허인용은 한 특허가 다른 특허에 의해 인용된 정보에 기반하고 있으며 그 회수를 비교하여 평가할 수 있다(Karki, 1997). 그리고 특허에 대한 인용은 단순한 의도에서 출발하지 않으며 여러 요인에 의한 것이지만 결과적으로 높은 피인용 회수를 갖는 특허가 그렇지 않은 특허보다 높은 기술적 가치를 갖는 것이다(Albert, M. 1991). 따라서 논문과 특허 모두에서 인용 관계는 지식의 흐름 그리고 평가의 척도로서 기능할 수 있다.

2003년 Garfield는 SCI(Science Citation Index), SSCI(Social Science Citation Index) 그리고 AHCI(Arts and Humanities Citation Index) 논문들의 서지 정보의 인용관계를 이용하

여 인용네트워크를 연대순으로 배치해주는 “HistCite”라는 시스템을 발표하였다. HistCite는 WoS(Web of Science)에서 데이터 집합을 생성하여 사용하도록 설계되었으며, 검색된 데이터 집합 내에서의 피인용 회수를 기반으로 한 LCS(Local Citation Score)와 전체 논문 집합에서의 피인용 회수에 기반한 GCS(Global Citation Score)를 모두 지원한다(Garfield, 2003). 하지만 HistCite는 오로지 논문에 대한 인용 정보만을 다루며 특허정보는 활용할 수 없다. 따라서 응용성이 높은 산업 기술 탐색에는 제약이 있으며, 무엇보다 간접인용을 배제하고 직접 인용관계만을 반영하여 영향력이 높은 문서에 의한 인용과 낮은 문서에 의한 인용을 동일하게 계량하였다.



(그림 1) 기술문서 네트워크에서(직접인용 네트워크) Vs (직접인용 + 간접인용 네트워크)

본 연구진은 문서간 간접인용을 반영하고 논문은 물론 특허문서에 대한 인용네트워크 연대표를 생성하는 시스템을 구축하고 지속적으로 연구 개발 중이다. 간접인용은 ‘A’라는 대상을 인용하는 개별 문서의 질이 다르다는 가정에서 출발한다. Chen et al(2007)은 구글의 PageRank 알고리즘을 이용하여 Physical Review 논문에 대해 순위(rank)를 매기고 이를 피인용 회수 순위와 비교하였다. PageRank 알고리즘을 바탕으로 하면 많은 링크를 갖고 있는 page(Paper)에 의해 인용을 받을 때, 인용받은 page(Paper)의 rank 점수가 높아진다. 그 결과 피인용 회수는 낮더라도 PageRank 점수는 높게 나온 논문들은 물리학 전문가들에게 높은 평가를 받는 것임을 확인하였다. 결과적으로 간접인용을 반영한 인용분석이 논문 순위에 대한 현실적인 평가를 가능케 하였다(Chen, 2007 & Maslov, 2008). 위의 (그림 1)은 기술문서 네트워크에서 직접인용만 고려한 경우(왼편)와 직접인용과 간접인용을 동시에 고려한 경우(오른편)를 도식으로 표현했다. 간접인용까지 고려하는 경우 네트워크의 더 많은 부분을 반영하고 동시에 근접한 원천 기술에


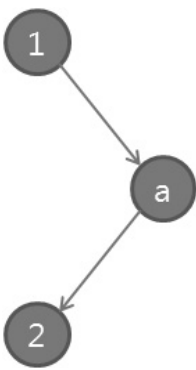
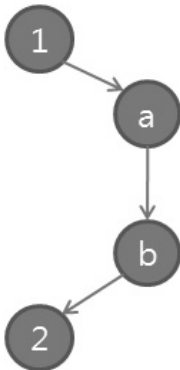
의한 영향(인용)도 반영될 수 있다. 본문에서는 본연구진이 연구 개발 중인 ‘핵심문서 계보도(영문명: Genealogical Core Document graph, 이하 GCD graph)’ 를 소개하고 이론적 배경을 설명한다. 또한 GCDgraph를 활용한 기술진화경로생성 과정을 설명하고 분석하고자 한다.

2. Genealogical Core Documents(GCD) 기본 원리

인용-피인용 네트워크 분석은 인용(Citation)을 매개로 연결된 문서간의 연결 관계를 하나의 방향을 가진 네트워크로 인지하여 분석하는 방법이다. 본 논문에서 기술하고 있는 Genealogical Core Documents(GCD) graph는 인용-피인용 관계를 간접인용(인용의 인용)까지 포함하여 분석하였으며, 이를 연도별 트리(tree)구조의 그림으로 도시하였다(Kwon, 2008).

GCDgraph는 핵심문서를 선별하고 추출하여 그들의 관계를 네트워크 형태로 도시해 준다. 핵심 문서의 선별은 각 문서에 대한 직접 인용과 간접 인용을 모두 고려하여 가중치를 계산하여 점수화한 후, 서열을 정하여 얻게 된다. 우선 직접 인용이 간접 인용보다 높은 지표값을 주도록 설계하였으며 간접인용은 어떤 알고리즘을 선택하느냐에 따라 달라진다. 여기서 소개하는 간접인용에 대한 지표 계산은 다음과 같다. 직접 인용에 의한 점수를 1점이라고 할 때, 특정 노드(문서)에 도달하기 위해 n개의 경로를 거쳐야 하는 문서가 있다면 지표값은 1/n 배 해주어 계산한다. 좀 더 구체적인 이해를 돕기 위해 아래 표에 그림과 함께 정리하였다.

〈표 1〉 간접인용을 고려한 가중치 계산

연결된 노드의 수	1	2	3
인용 유형	직접인용	직접 + 간접인용	직접 + 간접인용
network graph (from 1 to 2)			
1번 문헌의 가중치 계산	1	$1 + 1 \times \frac{1}{2} = \frac{3}{2}$	$1 + 1 \times \frac{1}{2} + 1 \times \frac{1}{3} = 1\frac{5}{6}$

지표 = (인용 경로의 개수)*(경로별 가중치 계산값)

위 지표 계산에서 중요한 것은 어떠한 경로(직접 인용 혹은 간접 인용)를 통해 인용네트워크가 그려지는가? 그리고 각 경로의 수가 몇 개인가? 이다. 전체 노드(문서)와 링크(citation)가 복잡하게 얽혀있는 네트워크에 대해서 하나하나 개별 경로를 세고 계산하는 일은 쉽지 않다. 따라서 복잡한 네트워크 구조에 대한 인용관계를 계산하기 위해 인접행렬을 도입한다. 여기서 인접행렬은 노드간의 직접 인용 관계를 행렬로 표현한 것으로, n 개의 노드가 있다면 $n \times n$ 인접행렬이 생성된다. A 라는 인접행렬을 가정하면 A 는 직접 인용의 정보를 담고 있으며, 간접 인용에 대한 정보는 A^m 에서 찾을 수 있다. A^m 에는 $(m-1)$ 개의 노드를 거치는 간접인용의 개수에 대한 정보를 갖고 있다.

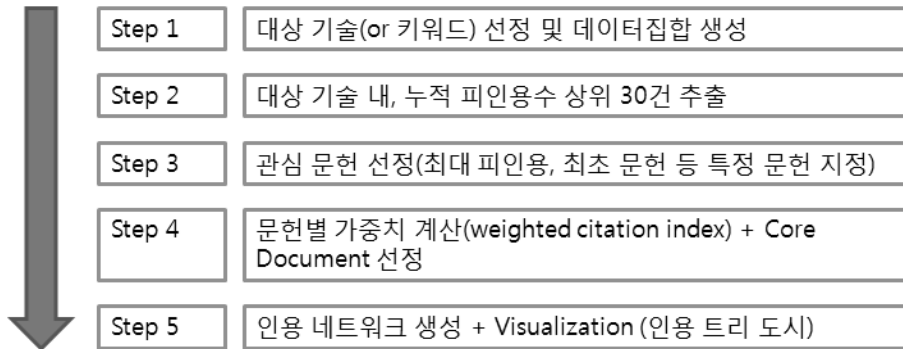
GCDgraph는 기본적으로 좌측에 발행년도(혹은 등록년도)를 표시하고 이를 기준으로 각 노드와 링크를 배열한다. 노드의 크기는 앞에 계산한 가중치를 포함한 지표에 비례하여 커지게 되나, 링크의 굵기 등은 동일하게 표현하였다. 각 노드에 대한 정보, 예를 들면 발명자(저자), 기관 등에 대한 정보는 노드를 클릭하거나 하단에 표시하여 전체 인용트리(citation tree)를 따라가며 비교하며 분석할 수 있도록 배치하였다.

3. GCD Graph 활용 데이터

GCDgraph가 대상으로 삼는 것은 정형화된 형태를 갖는 문서로써 현재는 미국등록특허(US Grant Patent)와 WoS(Web of Science)의 논문 데이터 DB를 활용할 수 있도록 설계되어 있다. 차후 필요에 따라 SCOPUS, EU Patent, 일본등록특허 데이터의 활용 가능여부를 검토할 수 있다. 특허DB의 경우 관심이 되는 특허의 등록번호만 알고 있으면 그 등록특허를 근간으로 하여 특허계보를 그릴 수 있는데, 이는 GCDgraph 내부에 저장된 전체 미국 등록특허에 대한 인용관계 데이터를 불러들여 계산할 수 있기 때문이다. 이런 개방된 플랫폼은 몇 가지 이점이 있다. 우선 인용트리(Citation tree) 안에서 기대하지 않던 핵심문헌(Core document)를 발견할 수 있는데, 이렇게 발견된 핵심문헌은 검색의 부주의함으로 놓친 것이거나 아니면 다른 기술분야와의 융합으로 창발된 것임을 알 수 있다. 그리고 무엇보다도 별도의 입력파일(input file)을 만들기 위해 소요되는 수고와 시간을 절약할 수 있다. 또한 전체 특허집합에 대해서 특정 특허의 위치, citation tree를 얻을 수 있다는 장점이 있다.

문헌 인용 트리를 생성하기 위해서는 몇 단계의 과정을 거치게 된다. 여기서는 특허 문서를 기준으로 기술하고자 한다. 먼저 관심을 갖는 대상 기술 분야를 선정하고 초기 데이터 집합을 구성하여야 한다. 이때 준비하는 데이터 집합은 검색을 통해 얻어진 해당 기술 분야의 특허번호로 이루어진 텍스트 파일이다(논문에 대한 데이터 집합은 WoS(Web of Science)의 데이터

형태를 준비하여 입력한다). 초기 데이터 집합은 인용필드(reference field)를 기반으로 노드(문서)간의 인용-피인용관계를 계산하여 피인용 회수 기준으로 재배열되는데, 이때 피인용회수는 직접인용까지 포함한 지표점수이다. 나열된 특허문서들은 해당 기술 분야에서 자주 인용되어 영향력이 높은 문서라고 할 수 있다.



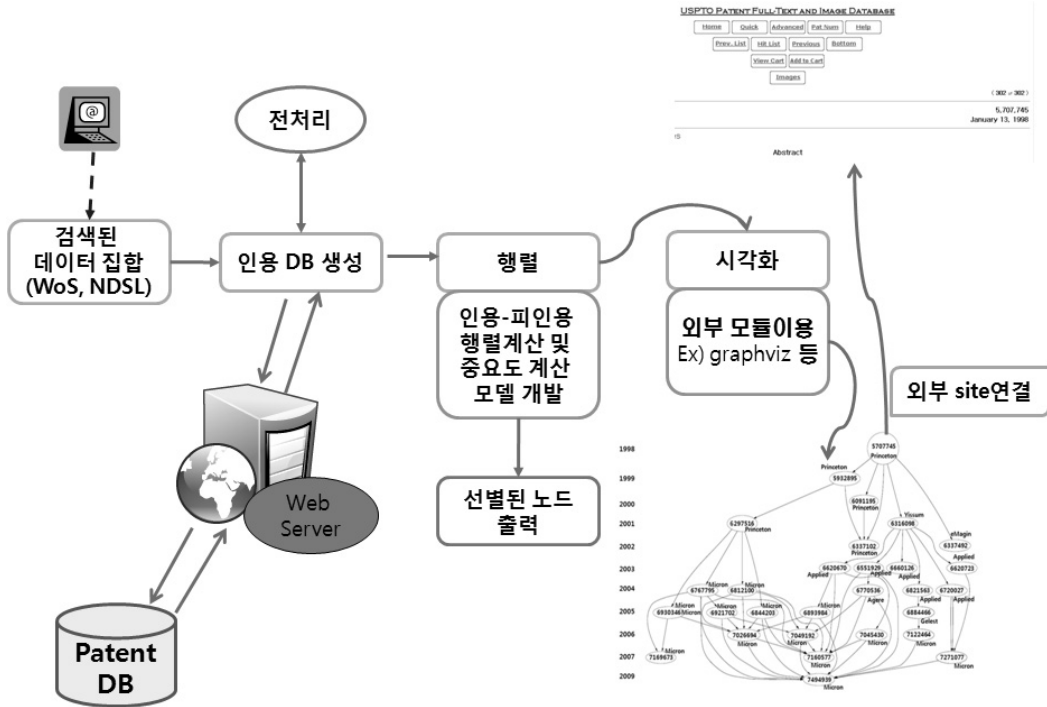
(그림 2) 인용네트워크 트리 생성 순서도

이렇게 나열된 고평인용 회수의 특허문서 중에 하나를 선택하게 되면 그 문서의 인용·피인용 관계를 간접인용에 대한 가중치를 포함하여 계산하며, 그 계산값에 기초하여 인용 네트워크 형태의 그래프로 도시된다. 선택되는 문서는 가장 높은 피인용 점수를 갖는 것이거나 가장 빠른 특허일 수 있으며 별도로 관심있는 특정 문서를 택하면 된다. 일반적으로 원천특허의 경우 그 분야 문서 중에서 년도가 가장 빠르면서 높은 피인용 점수를 갖게 된다.

논문 데이터의 경우, 라이선스 등의 문제 때문에 위와 같이 개방된 인터페이스를 사용할 수 없으며, 검색을 통해 원하는 DB 파일을 별도로 생성하여 사용해야 한다. 특허의 경우에도 원하는 특허를 한정하여 특허번호를 1열(1-column)형태로 저장하여 입력 파일로 사용할 수 있다. 이처럼 통제된 플랫폼의 경우, 검색된 데이터집합(Data set) 안에서 어느 것이 가장 높은 지표점수를 갖는 지 바로 알 수 있으며, 해당 분야의 기술이 누구 혹은 어느 기관(혹은 회사)에 의해 개발되고 발전되었는지 그 연결 고리를 한눈에 조망할 수 있다.

GCDgraph 프로그램의 결과물인 GCDgraph는 해당 분야의 핵심 기술의 시간에 따른 진화 상황을 보여준다. 이는 또한 기술 변천도이기도 하다. 특허 분석의 경우 서로 다른 출원기관의 분쟁을 예측할 수 있으며, 자기인용(self citation)을 제거하고 분석할 경우 기술의 경쟁관계를 파악할 수 있다. 또한 앞의 문단에 기술한 대로 검색집합에 안에서 핵심문헌을 확인하고 요약할 수 있다.

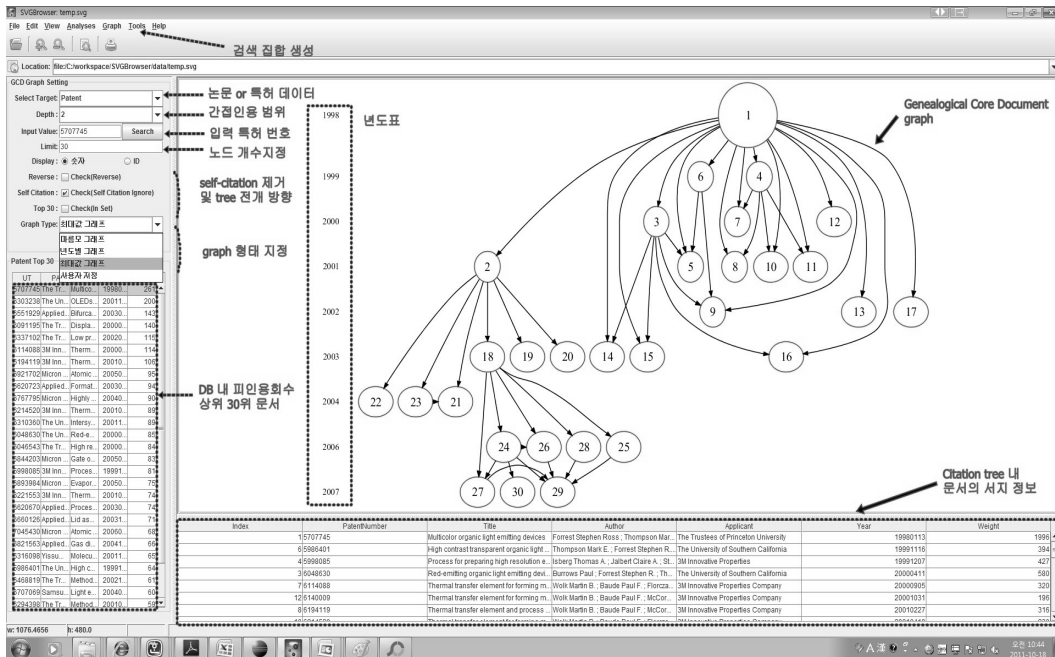
4. GCDgraph 시스템 구현



(그림 3) GCDgraph 프로그램 구동 메카니즘

GCDgraph는 현재 1단계 개발이 완료되어 활용범위를 넓히고 있다. (그림 3)에는 이해를 돕기 위해 GCDgraph 프로그램의 구동 메카니즘을 도식과 그림으로 정리하였다. 사용자는 WoS 혹은 NDSL 을 통해 논문 혹은 특허 데이터 집합을 준비하여 시스템에 입력한다. 입력된 데이터는 웹서버에 저장된 특허 DB(Patent DB)에 접속하여 인용DB를 생성한 후 인용DB를 바탕으로 인용-피인용 행렬계산과 가중치계산 모델을 거쳐 인용의 가중치를 계산한다. 다음으로 이를 시각화하기 위해 외부 모듈인 graphviz를 활용하여 연대별 핵심문서 인용네트워크를 화면에 도식하게 된다. 인용네트워크 화면에서 개별 노드(특허 문서의 경우)를 클릭하면, 입력 데이터가 특허문서인 경우 미국특허상표청(USPTO)의 해당 특허정보 페이지로 넘어가게 되고 논문인 경우는 입력데이터에서 해당정보를 불러들여 해당 논문의 서지정보창이 열리게 된다.

이번 문단에서는 실제 PC에서 구동되는 GCDgraph 화면을 캡처하여 간략히 활용 방법을 소개하도록 한다(그림 4). 우선 검색 데이터를 입력하기 위해서 'Tools' 메뉴에서 입력데이터의 위치와 생성될 인용DB파일이 저장될 위치를 지정해 준다. 또한 'Analysis' 메뉴는 데이터



(그림 4) GCD graph 프로그램 구동모습

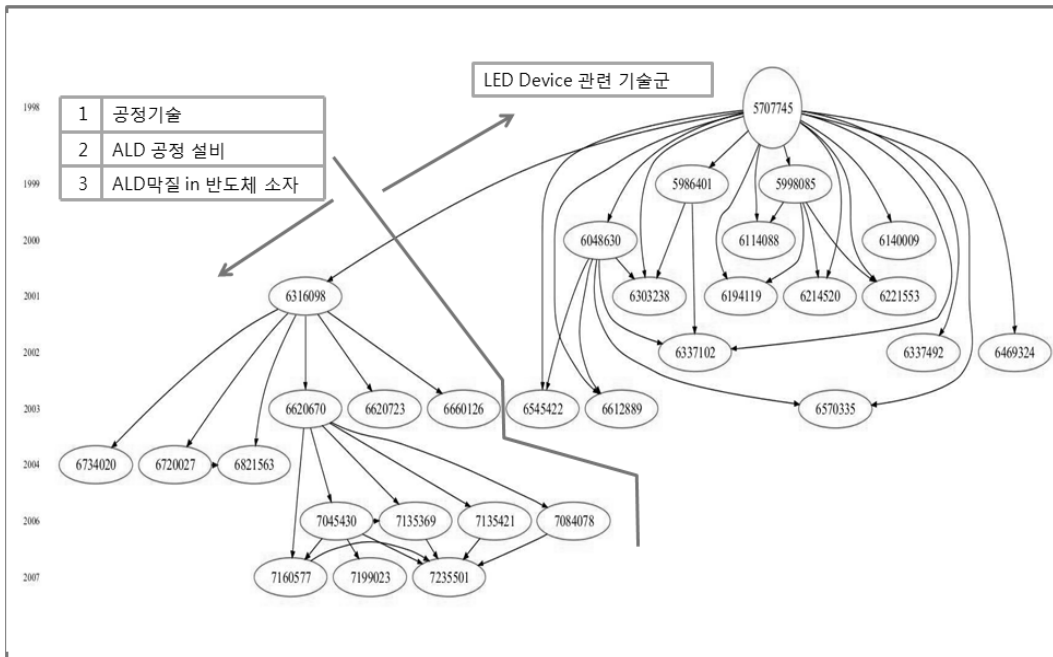
집합에 대한 기본적인 통계를 적용할 수 있도록 설계하였다. 메뉴바 아래의 전체 화면을 크게 세부분으로 나누면, 첫째 좌측 상단의 입력과 모드선택 부분, 둘째 좌측하단과 중앙하단의 문서 정보 표시부분, 그리고 셋째 ‘문서네트워크 계보도’ 부분으로 나눌 수 있다. 입력 부분에는 입력 데이터의 종류(논문 or 특허), 간접인용 범위, 입력 문서 번호, 그리고 그래프 형태 및 자기인용(self-citation) 제거 유무를 선택할 수 있도록 구성하였다. 좌측 하단에 표시되는 정보는 입력데이터 집합 안에서 피인용 회수를 기준으로 상위 30위 건의 문서들을 나열하고 정보를 표시하게 하였다. 사용자는 이들 문서 중에서 관심이 가는 문서를 선택하여 그 문서를 근간(root)으로 하는 새로운 GCDgraph를 얻을 수 있다. 하단에 표시되는 문서정보는 GCDgraph에 표시되는 문서(노드)에 대한 정보로 이 문서들은 피인용 회수 기준이 아니라 가중치 기준으로 재선별된 것이다. 마지막으로 ‘문서네트워크 계보도’는 노드와 인용링크로 구성된 그래프(graph)이며, 분석하고자하는 입력값과 모드 선택에 따라 그래프를 얻을 수 있다. 부가적으로 각 노드를 클릭하면 서지정보가 표시되어 분석을 용이하게 해준다.

III. 사례 연구

1. 기술 문헌 인용기반을 통한 기술진화 경로 분석 사례 - OLED 특허

유기발광다이오드(OLED)는 최근 스마트폰의 디스플레이로 채용되면서 관련 시장이 급속히 성장하고 있는 분야이다. 특히 한국의 경우, 삼성 SMD가 2011년 OLED 세계시장 점유율의 99%를 차지하고 있으며(이투데이, 2011), 더불어 국내 검사장비 업체들이 삼성 SMD에 고정 장비를 공급하면서 관련 장비시장도 급속히 성장하고 있어 OLED 관련 기술에 대한 관심이 지속되고 있다.

분석을 위해 검색을 통하여 OLED 관련 특허 3892건을 수집하였으며 이를 기초로 핵심 특허 문서를 선별하고 이들 간의 인용 네트워크를 분석함으로써 OLED 분야 특허의 인용 네트워크에서 기술진화 경로를 분석하고자 하였다. 우선 3892건의 특허 중에서 피인용 회수 상위 30건을 선별하였고 그 중 최고 피인용회수를 보이며 데이터 집합 중에서 가장 빠른 특허인 'US-5707745'를 기준으로 인용네트워크를 구성한다.



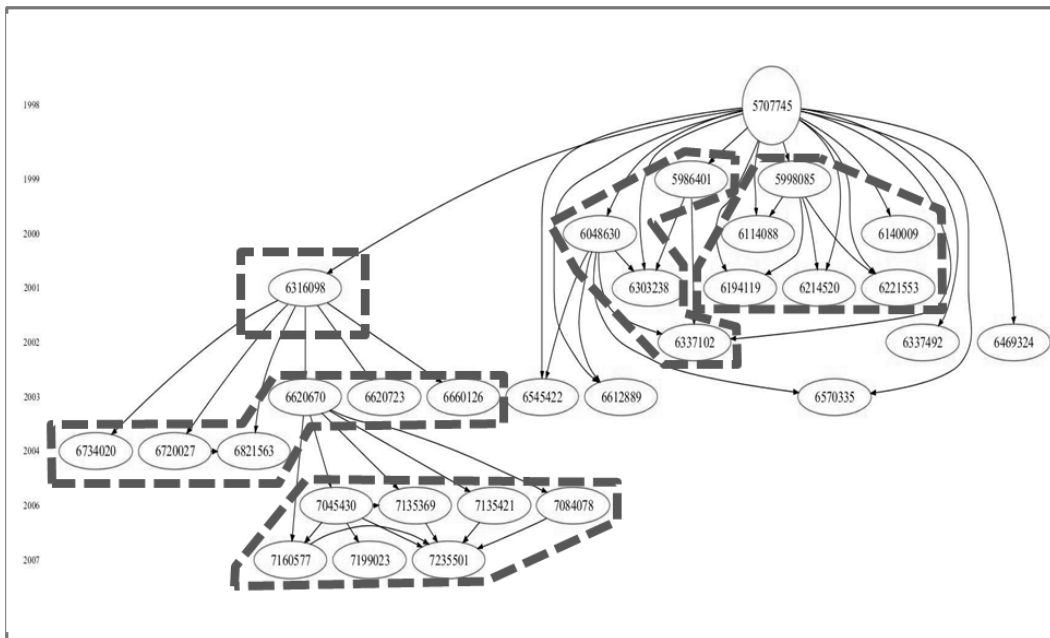
(그림 5) OLED 특허 문서의 인용 네트워크 역사 연대 그래프

〈표 2〉 OLED 특허 인용 네트워크 상위 30.

Patent Number	Title	Applicant	Weight
5707745	Multicolor organic light emitting devices	The Trustees of Princeton University	1996
6316098	Molecular layer epitaxy method and compositions	Yissum Research Development Company of the Hebrew Univ. of Jerusalem, ILX	654
6048630	Red-emitting organic light emitting devices (OLED's)	The Univ. of Southern California	580
5998085	Process for preparing high resolution emissive arrays and corresponding articles	3M Innovative Properties	427
6303238	OLEDs doped with phosphorescent compounds	The Univ. of Southern California	409
5986401	High contrast transparent organic light emitting device display	The Univ. of Southern California	394
6114088	Thermal transfer element for forming multilayer devices	3M Innovative Properties Company	320
6194119	Thermal transfer element and process for forming organic electroluminescent devices	3M Innovative Properties Company	316
6337102	Low pressure vapor phase deposition of organic thin films	The Trustees of Princeton University	276
6214520	Thermal transfer element for forming multilayer devices	3M Innovative Properties Company	230
6221553	Thermal transfer element for forming multilayer devices	3M Innovative Properties Company	213
6140009	Thermal transfer element for forming multilayer devices	3M Innovative Properties Company	196
6337492	Serially-connected organic light emitting diode stack having conductor sandwiching each light emitting layer	eMagin Corporation	189
6545422	Socket for use with a micro-component in a light-emitting panel	Science Applications International Corp.	176
6612889	Method for making a light-emitting panel	Science Applications International Corp.	141
6570335	Method and system for energizing a micro-component in a light-emitting panel	Science Applications International Corp.	128
6620670	Process conditions and precursors for atomic layer deposition (ALD) of Al_2O_3	Applied Materials, Inc.	447
6620723	Formation of boride barrier layers using chemisorption techniques	Applied Materials, Inc.	406
6660126	Lid assembly for a processing system to facilitate sequential deposition techniques	Applied Materials, Inc.	356
6821563	Gas distribution system for cyclical layer deposition	Applied Materials, Inc.	148
6734020	Valve control system for atomic layer deposition chamber	Applied Materials, Inc.	135
6720027	Cyclical deposition of a variable content titanium silicon nitride layer	Applied Materials, Inc.	127
7045430	Atomic layer-deposited $LaAlO_3$ films for gate dielectrics	Micron Technology Inc.	340
7084078	Atomic layer deposited lanthanide doped TiO_x dielectric films	Micron Technology, Inc.	218
7135369	Atomic layer deposited $ZrAlO$ dielectric layers including $ZrAlO$	Micron Technology, Inc.	193
7160577	Methods for atomic-layer deposition of aluminum oxides in integrated circuits	Micron Technology, Inc.	191
7135421	Atomic layer-deposited hafnium aluminum oxide	Micron Technology, Inc.	189
7235501	Lanthanum hafnium oxide dielectrics	Micron Technology, Inc.	136
7199023	Atomic layer deposited $HfSiON$ dielectric films where each precursor is independently pulsed	Micron Technology, Inc.	126

〈표 2〉에는 ‘US-5707745’특허와 직·간접적인 인용관계에 있는 특허 중에 가중치(weight) 값이 가장 높은 30개의 특허를 선별하여 정리하였고, 이들 간의 관계를 (그림 5)에 인용 네트워크로 표현하였다. (그림 5)의 네트워크는 년도별로 배치되어(좌측에 년도 표시) 아래로 갈수록 최근의 특허 문서이다. 인용관계와 시간의 흐름에 대해 지식(기술)이 어떤 형태로 변화하는지 알려준다. 따라서 (그림 5)는 단순한 인용 네트워크에 더하여 기술의 역사연대표 역할을 동시에 수행하고 있다.

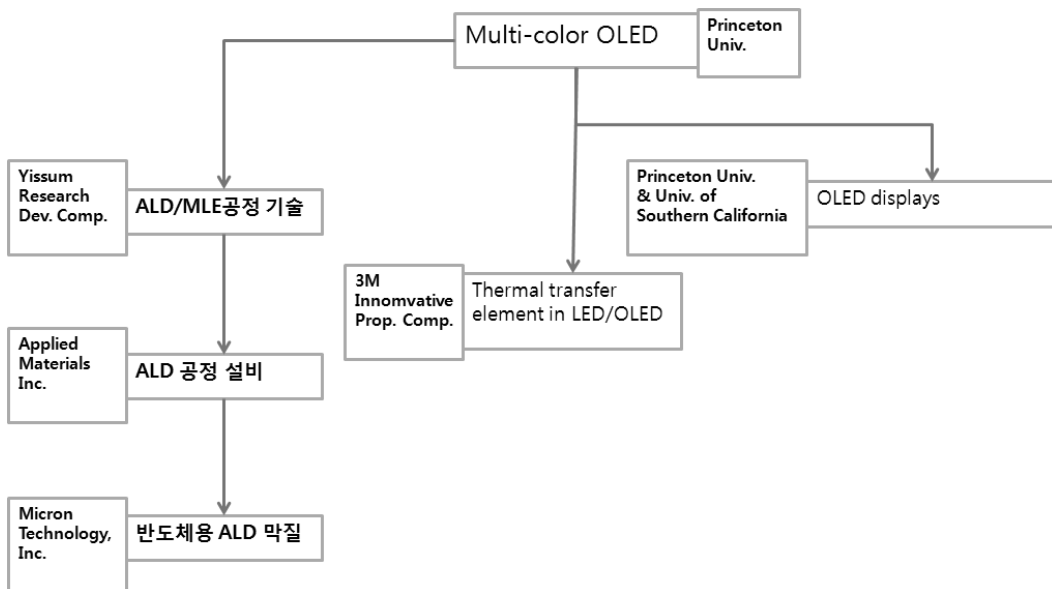
(그림 5)의 인용네트워크 연대표를 보면 형태 상 크게 두 개의 cluster로 나누어 볼 수 있다. 하나는 우측 상단 US-5707745 아래의 문서들 그리고 다른 하나는 좌측 하단 US-6316098 아래의 문서군으로 나눌 수 있다. 각 Cluster의 기술 진화 형태를 파악하기 위해 각 노드의 특허 번호 정보를 출원인(Applicant) 정보 및 제목(title) 또는 초록(abstract) 정보로 치환한 후 분석하여 다시 클러스터(cluster)를 세분화하였다(그림 6).



(그림 6) 출원기관별, 유사 기술별 세부 cluster

세부 클러스터 간의 기술 및 출원인 관계를 다시 (그림 7)에 도식으로 정리하여 기술의 진화와 흐름을 파악할 수 있도록 하였다. 핵심특허인 US-5707745를 인용한 문서 중에서 US-6316098 이하의 인용문서들은 다시 출원기관별, 그리고 기술별로 구별된다. 우선 US-6316098는 ALD (Atomic Layer Deposition) 그리고 MLE(Molecular Layer epitaxy) 성장 기술을 포함하고

있는 특허로서 Hebrew Univ. of Jerusalem의 Yissum Research Development Company 사 소유의 특허이다. 그 아래의 US-6620670, US-6620723, US-6660126, US-6734020, US-6720027 그리고 US-6821563 특허는 모두 US-6316098을 인용한 문서로써 동시에 모두 글로벌 반도체 장비 메이커인 Applied Materials, Inc. 가 출원한 특허이다. 이 6 개의 특허는 반도체 최신 공정인 ALD 공정 장비와 관련된 기술 특허들이다. 다시 이들 특허 중에 US-6620670 와 직접인용 및 간접 인용 관계에 있는 US-7045430, US-7135369, US-7135421, US-7084078, US-7160577, US-7199023 그리고 US-7235501 특허는 미국 메모리 반도체 회사인 Micron Technology, Inc. 가 출원한 특허로서 주로 메모리 반도체의 transistor 나 capacitor 에 사용하는 유전물질막(Dielectric layers) 기술을 포함하고 있다. 정리하면 전체 인용 트리 중에서 US-6316098 아래의 인용 문서들은 단계별로 ALD/MLE 공정기술 - ALD 공정 설비 기술 -반도체용 ALD 막질 기술로 이어지는 진화단계를 밟아 발전했으며, 또한 US-6316098 이하 인용 트리는 US-6316098특허를 매개로 상위 특허인 US-5707745(Multi-color OLED)의 기술에 영향을 받았으나 결과적으로 다른 제품영역에서 비롯된 것이다.



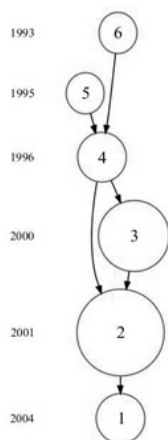
(그림 7) OLED 문서 인용 네트워크의 도식화/기술 진화 흐름도

또 다른 cluster는 (그림 6)에서 US-5707745 바로 아래에 전개된 특허들이다. (그림 7)의 흐름도에서 각각 OLED display 소자 기술과 LED/OLED 에 사용하는 thermal transfer element 생성 및 사용에 관한 특허이다. 이 두 그룹은 각각 Princeton Univ. 와 Univ. of

Southern California 그리고 3M Innovative Prop. Comp.에서 출원한 특허들이다. 각기 OLED 제조와 관련된 기술로 원천기술인 US-5707745와 동일한 제품 영역 안에 한정되어 기술 진화가 이루어졌다.

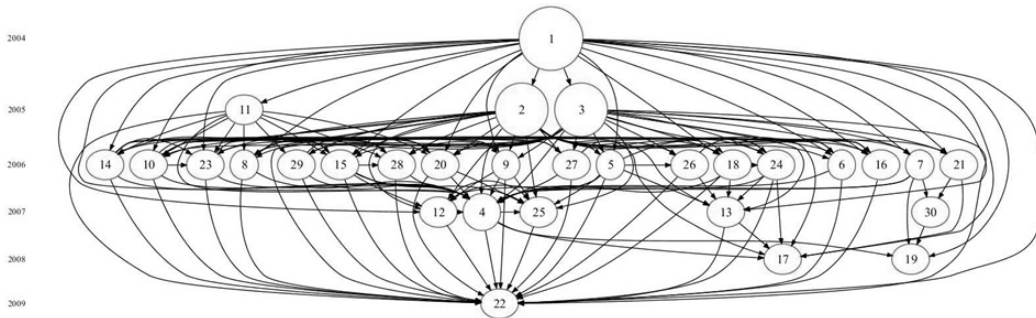
2. 기술 문헌 인용기반을 통한 기술진화 경로 분석 사례: graphene논문

그래핀(graphene)은 2010년도 노벨물리학상의 주제가 되면서 대중적으로도 많이 알려진 분야이다. 탄소 단일층으로 구성된 그래핀은 기존의 실리콘 반도체 대비 높은 전도성과 안정성으로 실리콘을 대체할 물질로 주목받고 있으며, 또한 플렉서블 디스플레이의 소재로서도 연구되고 있다. 본 단락에서는 2004년에 러시아와 영국의 두 과학자인 노보셀로프(Novoselov)와 가임(Geim)의 연구 논문을 시작점으로 해서 그래핀 분야에 대한 논문 인용 네트워크 계통도를 분석하고자 한다.



(그림 8) Novoselov(2004) 논문의 Backward citation tree

(그림 8)은 Novoselov의 2004년 논문에 기재된 인용논문들의 과거 인용을 추적하기 위해 후방인용트리(Backward citation tree)를 보여주고 있다. 또한 (그림 9)는 Novoselov(2004) 논문의 일반적인 인용 트리이며, 이를 전자와 구별하기 위해 전방인용트리(Forward citation tree)로 명명하였다. 특정 논문 A를 인용한 논문들과 다시 그 논문들을 인용한 논문들의 인용 관계를 연도별로 표현한 것을 전방인용트리라고 하였으며, A가 인용한 논문들과 다시 그 논문들이 인용한 논문들을 거슬러 올라가며 인용관계를 연도별로 표현한 것을 후방인용트리라고



(그림 9) Novoselov(2004) 논문의 Forward citation tree.

명명하였다. 두 그림의 인용관계는 연도 상 2004년 그리고 Novoselov의 논문(index '1')을 기점으로 명확한 대비를 보여준다. 2004년 이전의 경우, 그래핀 관련한 연구는 제한적인 연구자들에 의해서만 진행되는 소수자 연구영역이었다. 하지만 2004년 이후 폭발적으로 그래핀 연구 결과가 증가하였다. 다만 연도별 인용 트리에서 다른 연구분야와 연결되는 핵심 문서가 발견되지 않은 것으로 미루어 보면 아직 다른 영역으로의 확장은 상대적으로 미미한 수준이라고 해석된다.

(그림 9)에 표시된 논문들 중에는 가중치(weight)값 순위와 피인용회수 순위 사이의 역전이 발견되기도 하는데, 어느 값을 중시하는지에 따라 기술 및 과학의 가치에 대한 평가가 바뀔 수도 있다. 아래에는 피인용 회수 상위 5위권 논문의 피인용 회수와 가중치를 비교한 표이다.

〈표 3〉 그래핀(graphene) 분야 가중치(weight) 점수 상위 5위권 논문

Index	Author	교신저자	Year	Journal	Times cited	Weight
1	Novoselov	Geim	2004	Science	4174	38428
2	Novoselov	Geim	2005	Nature	2953	32178
4	Geim	Geim	2007	Nat. Mat.	2905	22616
3	Zhang	Kim, P	2005	Nature	2582	32151
22	Castro	Castro	2009	Rev. Mod. Phys.	1426	4792

우선 index 3번인 Zhang & Kim, P(김필립 교수)의 논문의 피인용회수(Times cited) 점수는 4위이지만 간접 인용을 포함한 가중치를 고려한 점수는 2위권에 해당한다. '3'번 논문은 직접인용과 함께 간접인용을 통해 그래핀 연구 그룹에 많은 영향을 미쳤음을 추론할 수 있다. 이는 김필립 교수가 2010년 노벨상의 세번째 수상자 자격이 있다고 주장하는 일련의 견해와 일치하는 부분이다(경향신문, 2010. 11. 29). 그리고 index 22번, Castro의 논문은 Times

cited 점수가 5위이지만 weight 점수는 22위에 불과하다. 원인은 Castro의 논문은 리뷰(review) 성격의 논문으로 이 분야에 새로 진입한 신진과학연구자에게는 많이 인용되지만 이 분야의 대가들 혹은 전문가들에게는 인용되지 않았기 때문이다. 참고로 그 분야 전문가들의 논문은 높은 피인용회수(times cited)값을 가질 가능성이 크다. 따라서 전문가의 인용을 받는 논문에 대해 간접인용을 고려한 가중치 계산을 할 경우 그 값이 직접인용만 고려한 경우보다 크게 나오게 된다.

위의 사실은 흥미로운 결과로써 단순 피인용회수(직접인용)에 기반한 평가보다는 전체 문헌 네트워크에서 차지하는 위치를 함께 고려한 점수(간접 인용까지 고려한 가중치)가 좀 더 의미있는 결과를 줄 수 있음을 의미하며, 이는 본 논문에서 제시하는 기술 인용관계 트리(가)가 논문과 특허를 좀 더 현실적으로 평가할 수 있는 방법론이며, 이를 통해 학계와 산업에 의미 있는 정보를 보여준다.

IV. 결론 및 향후 연구 방향

본 연구는 특허와 논문의 인용 정보를 활용하여 직접인용뿐만 아니라 간접 인용까지 고려한 인용관계를 연도별 그래프로 표현하였다. 활용 예시로써 특허의 경우 OLED 분야 특허 3892건에 대해서 인용관계 계통도를 분석하였고 이를 통해 OLED 분야의 핵심문헌(Core document)의 기술 진화 과정을 분석하였다. OLED에서 인용관계를 통해 분화된 기술은 반도체 분야로 전이되어 반도체 공정설비기술과 반도체 신물질 기술로 진화되어 갔으며, 일부 기술은 OLED 소자 관련 기술로 지속적으로 진화되었다. 또한 graphene 관련 논문 10,293건에 대한 인용관계 계통도 분석을 통하여 간접인용까지 고려한 방법이 단순 피인용회수를 계산하는 방법보다 좀 더 현실에 부합하는 평가 방법임을 확인하였다. 본 연구진은 향후 각 문서의 제목 및 초록에 대한 자연어 처리와 분석을 통하여 인용 관계로 연결된 두 문서의 기술 관계를 비교하여 기술의 흐름 및 진화경로를 객관적으로 기술할 수 있는 방법론을 연구할 계획이다.

참고문헌

한국과학기술정보연구원(KISTI) (2011), 「시맨틱 기술 기반 기술 기회 탐색 지원 체제 연구개발 사업 기획 보고서」, 서울: 한국과학기술정보연구원.

- 고병열 (2005), “기술-산업 연계구조 및 특허 분석을 통한 미래유망 아이템 발굴”, 『기술혁신학회지』, 8(2) : 860-885.
- 유선희 (2007), “특허정보분석을 이용한 기술과급효과 측정에 관한 연구”, 『기술혁신학회지』, 10(4) : 687-705.
- Garfield E, Sher IH, Irving H. Sher, Torpie RJ. (1964), “The Use of Citation Data in Writing the History of Science”, *Institute for Scientific Information Inc.*
- Garner R., Lunin L., and Baker L. (1967), “Three Drexel Information science research studies”, *Drexel University Press.*
- Karki, M. (1997), “Patent Citation Analysis: A Policy Analysis Tool”, *World Patent Information*, 19(4), pp. 269-272.
- Kerr, C.I.V., Mortara, L., Phaal, R. and Probert, D.R. (2006), “A conceptual model for technology intelligence”, *Int. J. Technology Intelligence and Planning*, Vol. 2, No. 1, pp. 73-93.
- Alber, M., Avery, D., Narin, F. and McAllister, P. (1991), “Direct Validation of Citation counts as Indicators of Industrially Import Patents”, *Research Policy*, 20, pp. 123-134.
- Garfield E., Pudovkin A.I., and Istomin V.S. (2003), “Mapping Output of Topical Searched in the Science Citation Index, Social Sciences Citation Index, Arts and Humanities Citation Index” *Special Libraries Association (SLA) Meeting in New York, NY.*
- Van Raan (1988), “Handbook of Quantitative Studies of Science & Technology”, *Dordrecht, North Holland.*
- Chen P., Xie. H., Maslov S., Redner S. (2007), “Finding scientific gems with Google’s PageRank algorithm”, *Journal of Informetrics* 1, pp. 8-15.
- Maslov S., Redner S. (2008), “Promise and Pitfalls of Extending Google’s PageRank Algorithm to Citation Networks”, *The Journal of Neuroscience*, 28(44), 11103-11105.
- Kwon O., Lee B. Seo J., Noh K., Lee J., Kim J. (2009), “A method to make the genealogical graph of core documents from the directed citation network”, *Information-An International Interdisciplinary Journal*, Vol. 12, Iss. 4, 875-888.
- 이투데이 (2011), 반도체·LCD 새 승부수, “차세대 기술로 불황돌파”, (2011.09.26).
- 경향신문 (2010), “김필립 교수, 위원회 실수로 노벨물리학상서 제외됐다”, (2010.11.29).

이재민

서울대학교에서 "Study on Fe and Fe₃O₄ ultra-thin film growth on GaAs(100) at low temperature and its application as efficient spin-injection device."로 이학박사학위를 취득하였으며, 삼성전자 반도체 연구소 책임연구원을 거쳐 현재 한국과학기술정보연구원 선임연구원으로 재직 중이다. 주요 연구 분야는 과학계량학(Scientometric), 기술지능(Technology Intelligence), 지식과학(Knowledge Science), 사회네트워크 등이며, 현재 기술기회발굴(Technology Opportunity Discovery)모델 연구를 수행 중에 있다.

문영호

한국과학기술원에서 "양생온도와 양생시점에 따른 콘크리트 강도예측에 관한 연구" 논문으로 공학박사학위를 취득하였으며, 산업연구원, 산업기술정보원, 한국과학기술정보연구원 책임연구원을 거쳐 현재 한국과학기술정보연구원 정보분석본부장으로 재직 중이다. 주요 연구 분야는 과학계량학(Scientometric), 미래유망기술, 정보분석시스템, 지식과학(Knowledge Science), 정보 구조화 등이다.

이방래

서울시립대학교에서 박사 수료하였으며, 현재 한국과학기술정보연구원 선임연구원으로 재직 중이다. 주요 연구 분야는 데이터마이닝, 지식계량학, 사회네트워크분석 등이다.

권오진

서울시립대학교에서 "과학계량학 기반 과학기술 지식 네트워크 구조 분석 모델 개발" 논문으로 이학박사학위를 취득하였으며, 산업기술정보원 책임연구원을 거쳐 한국과학기술정보연구원 선임연구원으로 재직 중이다. 주요 연구 분야는 과학계량학(Scientometric), 정보분석시스템, 지식과학(Knowledge Science), 정보 구조화 등이며, 현재 기술기회발굴(Technology Opportunity Discovery) 모델 연구를 수행 중에 있다.