

국내 광역 과학 지도 생성 연구*

Making a Science Map of Korea

이재윤(JaeYun Lee)**

초 록

전 학문 분야를 포괄하는 광역 과학 지도는 학문 분야 사이의 구조적인 관계를 시각적으로 분석하는데 사용되고 있다. 이 연구에서는 광역 과학 지도에 대한 선행 연구를 개관한 후 새로운 방법으로 국내의 학술 활동을 반영하는 광역 과학 지도를 생성하였다. 광역 과학 지도에 대한 연구는 ISI사(현재 Thomson Scientific)의 Garfield와 Small에 의해서 촉발되었고 최근에는 스페인 그라나다 대학의 SCImago 연구팀과 미국 인디애나 대학의 Borner 교수팀이 활발히 연구 결과를 발표하고 있다. 이들은 자신들이 만들어 발표하고 있는 지도를 과학 지도 또는 사이언트그램이라고 부르며, 이에 관련된 활동을 과학지도학(scientography)이라고 하였다. 기존의 광역 과학 지도는 대부분 학술 논문 사이의 인용 분석에 근거하여 제작되었으나, 국내 학술 논문에 대한 인용 데이터베이스는 아직 미비한 상태이다. 따라서 이 연구에서는 국내의 광역 과학 지도를 만들기 위해서 학술진흥재단에 신청된 과제 제안서의 텍스트를 활용하였다. 학문 분야 사이의 연결 정보를 네트워크로 표현하는 수단으로 널리 사용되고 있는 패스파인더 네트워크(PFNet) 알고리즘으로 광역 과학 지도를 생성한 후, 이의 대안으로 개발된 클러스터링 기반 네트워크(CBNet) 알고리즘으로 다시 지도를 생성하였다. 최종적으로 두 지도에 나타난 상반된 관점을 통합하도록 CBNet 지도를 수정하여 국내 광역 과학 지도를 제시하였다.

ABSTRACT

Global map of science, which is visualizing large scientific domains, can be used to visually analyze the structural relationships between major areas of science. This paper reviewed previous efforts on global science map, and then tried to making a science map of Korea with some new methods. There are several research groups on making global map of science including Dr. Small and Dr. Garfield of ISI (now Thompson Scientific), SCImago research group at the University of Granada, and Dr. Borner's InfoVis Lab at the Indiana University. They called their maps as science map or scientogram and called the activity of mapping science as scientography. Most of the previous works are based on citations between scientific articles. However citation database for Korean journal articles is still under construction. This research tried to make a Korean science map with the text in the proposals suggested for funding from Korean Research Foundation. Two kinds of method for generating networks of scientific fields are used. One is Pathfinder network (PFNet) algorithm which has been used in several published bibliometric studies. The other is clustering-based network (CBnet) algorithm which was proposed recently as an alternative to PFNet. In order to take into account both views of the two algorithms, the resulting maps are combined to a final science map of Korea.

키워드 : 과학 지도, 패스파인더 네트워크, 클러스터링 기반 네트워크, 과학지도학, 도메인 분석
science map, pathfinder network, clustering-based network, scientography,
domain analysis

* 이 연구의 일부 분석 결과는 제14회 한국정보관리학회 학술 대회에서 발표된 바 있음.

** 경기대학교 문헌정보학 전공 조교수 (memexlee@kgu.ac.kr)

- 논문접수일자 : 2007년 8월 26일
- 게재확정일자 : 2007년 9월 10일

1. 서 론

광역 과학 지도(global map of science), 혹은 광역 학문 지도는 개별 학문 분야 안에서의 지적 구조를 분석하는 일반적인 지적 구조 분석 연구와 달리 전체 학문 분야 사이의 구조적인 관계를 시각화하여 분석하기 위한 도구이다.

광역 과학 지도 연구의 선구자라고 할 수 있으며 여러 차례 관련 논문을 발표한 Small(1999)은, 과학 문헌을 공간상에 표현하는 것이 학문 분야 사이의 개념적인 관계와 전개 양상을 이해하는데 도움이 되며, 동시대 지식의 현황에 대한 통찰을 제공할 수 있다고 하였다. 그에 따르면 과학을 지도화할 수 있으리라는 발상이 예전에 전혀 없었던 것은 아니지만 이를 분명하게 논문에서 적시한 사람은 Price(1966)였다.

Small과 함께 연구한 Garfield는 계량서지적 정보로부터 지도를 만드는 것을 scientography(여기서는 과학지도학[科學地圖學]으로 번역함)라고 칭하였다(Garfield 1994). 최근 광역 과학 지도에 관한 단행본을 출간한 Vargas-Quesada와 Moya-Anegón(2007)은 '도메인 시각화'나 '정보/지식 시각화'와 같은 유사 용어가 널리 사용되고 있지만, 학술 활동의 산물로 도표를 그리는 활동 및 그 효과에 대한 표현으로는 과학지도학이 가장 적절하다고 주장하였다.

이와 같이 Price 이래의 아이디어가 과학지도학으로 현실화 된 것은 인용DB를 이용하는 동시인용분석(Small 1973)이 등장한 덕분이라고 할 수 있다. 동시대 지식의 현황을 살펴 보기 위해서 Small은 1985년에 Garfield와

함께 SCI와 SSCI DB에 나타난 동시인용을 분석해서 1983년의 '과학 지도(maps of science)'를 만들었다(Small and Garfield 1985). 이는 다양한 결합 수준에서 과학의 지형도를 보여주는 것이었다. 여기에는 과학 전체를 포괄하는 광역 수준에서부터 개별 논문 단계까지 포함되었다. 이후 ISI사는 연례적으로 전 분야에 대한 매핑을 시도할 예정이라고 언급하였고, Garfield(1998)도 강연에서 Small의 도움으로 작성한 1996년의 과학 지도를 소개하였다. 또한 Small은 1999년에 발표한 논문에서 1985년에 사용한 기법을 약간 수정하여 1995년의 과학 지도를 보고한 바 있다.

한편 정보시각화 기술과 대용량 데이터 처리 기술이 발달함에 따라서 Small과 Garfield의 뒤를 이어서 개선된 광역 과학 지도를 생성하려는 시도가 나타났다(이재운 2005; Boyack, et al. 2005; Moya-Anegón et al. 2004; Vargas-Quesada & Moya-Anegón 2007). 그러나 국내의 연구 동향과 학문 분야간 관계를 파악할 수 있는 광역 과학 지도가 발표된 사례는 없었다. 국내 연구 동향을 반영하는 계량서지적 분석은 김영진(1986)의 연구 이후 문헌동시인용과 저자동시인용 분석 사례가 여러 건 발표되었지만, 모두가 특정 학문 분야 내부의 지적 구조만 살펴보는 연구였다.

국내 연구 동향에 대한 광역 과학 지도를 생성하는데 있어서 가장 큰 걸림돌은 국내 연구 성과물에 대한 인용 데이터베이스가 완성되어 서비스되고 있지 않은 점이다. 개별 학문 분야에 대한 분석을 위해서는 연구자가 개인적으로 인용 자료를 구축하는 것이 제한적으로는 가능하다. 기존의 국내 연구동향에 대한 인용분석

연구에서는 대부분 연구자가 직접 인용 자료를 수집하여 데이터베이스화하고 분석하였다. 그러나 수십 개 이상의 학문 분야를 다루는 광역 과학 지도를 생성하기 위한 인용 데이터베이스는 개인적 차원에서 구축하기가 거의 불가능하다. 한국학술진흥재단에서 등재학술지와 등재 후보학술지를 대상으로 2001년부터 한국형 인용색인 데이터베이스인 KCI를 준비하고 있으나 아직까지는 다음과 같은 문제점이 해결되지 않은 상태이다(홍중선 2007).

첫째, 구축된 데이터의 수록기간이 짧으며 수록 범위가 제한적이어서 평가를 위한 충분한 기초 데이터를 제공하지 못하고 있다.

둘째, 인용색인 데이터베이스의 기본정보인 인용정보의 생성이 불완전하다.

셋째, 인명정보가 통제되지 않아 정확한 저자 및 기관 관련 평가와 분석이 불가능하다.

넷째, 평가지표 개발에 필요한 관련 항목이 미비하여 안정적이고 개선된 지표 개발을 할 수 없다.

다섯째, 기능적 측면이 미약하여 일반에 공개하여 활성화를 기대하기 어렵다.

따라서 국내 연구 동향을 폭넓게 반영하는 광역 과학 지도를 생성하기 위해서는 현재로서는 인용이 아닌 다른 수단이 필요하다.

이 논문에서는 기존의 광역 과학 지도 생성 사례를 살펴본 후, 국내 학술 연구 동향을 반영하는 광역 과학 지도를 인용분석이 아닌 새로운 기법으로 생성하여 제시해보았다. 이를 통해서 국내 과학지도학 연구가 특정 분야에 대한 분석에 그치지 않고 학문 연구 동향 전체에 대한 분석 차원으로 나아가는데 일조하고자

한다.

2. 광역 과학 지도 생성 사례

광역 과학 지도를 생성한 주요 선행 연구를 정리하면 <표 1>과 같다. Small과 Garfield의 연구 이후 21세기에 들어서 해외에서 발표된 연구는 스페인 그라나다 대학 문헌정보학과 의 Moya-Anegón 교수가 이끄는 연구팀과, 미국 인디애나 대학 문헌정보학과 의 Börner 교수가 이끄는 연구팀에 의해서 각각 수행되었다. 이 장에서는 세 연구팀과 국내에서 시도된 사례 한 건에 대해서 구체적으로 살펴보기로 한다.

2.1 Small과 Garfield의 과학 지도

본격적으로 동시인용분석을 통해 광역 과학 지도를 생성한 첫 사례라고 할 수 있는 Small과 Garfield(1985), 그리고 이후의 Garfield(1998), Small(1999)의 세 연구는 분석 대상 시기만 다르고 분석 기법은 거의 동일하다. Small과 Garfield는 자신들이 만든 지도를 ‘과학 지도(maps of science)’라고 불렀다. Small은 과학 지도가 학문 분야에서부터 개별 문헌이나 저자에 이르는 여러 차원에서 각각이 어떤 관계인지를 공간적인 근접성과 상대적인 위치를 통해 지도처럼 표현하는 것이라고 하였다(Small 1999, 799). 이들이 사용한 방법을 간단히 정리하면 다음과 같다.

- ① 인용문헌 선별 - 인용빈도 기준치를 설정하여 일정한 값 이상의 인용문헌을 선

〈표 1〉 광역 과학 지도 생성을 시도한 주요 연구

연구자	대상 DB	대상 자료	대상 영역	관계 분석 기법	표현 기법
Small & Garfield(1985)	SCI & SSCI 1983	피인용문헌 50,994건	사회과학 및 자연과학	문헌 동시인용	다단계 군집화, MDS
Small(1999)	SCI & SSCI 1995	피인용문헌 164,612건	사회과학 및 자연과학	문헌 동시인용	다단계 군집화, 삼각측량법
Moya-Anegón et al.(2004)	Web of Science (SCI-E, SSCI, A&HCI) 2000년	주소 항목에 스페인이 포함된 문헌 26,062건, 범주 222 항목	JCR 수록 전 영역	범주 동시인용	MST
Vargas-Quesada & Moya-Anegón (2007)	Web of Science (SCI-E, SSCI, A&HCI) 2002년	전 세계 - 범주 219 항목, 문헌 901,493건 스페인 - 범주 219 항목, 문헌 294,778건	JCR 수록 전 영역	범주 동시인용	패스파인더 네트워크
Boyack, et al. (2005)	SCI & SSCI 2000년	인용을 주고 받은 학술지 7,121종	사회과학 및 자연과학	학술지 동시인용, 학술지 상호인용	VxOrd
이재윤(2005)	KISTI 과학기술분야 해외학술지	논문 551,599건, 범주 202 항목	과학기술 영역 (경영, 경제 포함)	분류 항목별 학술지 구성 유사도	패스파인더 네트워크

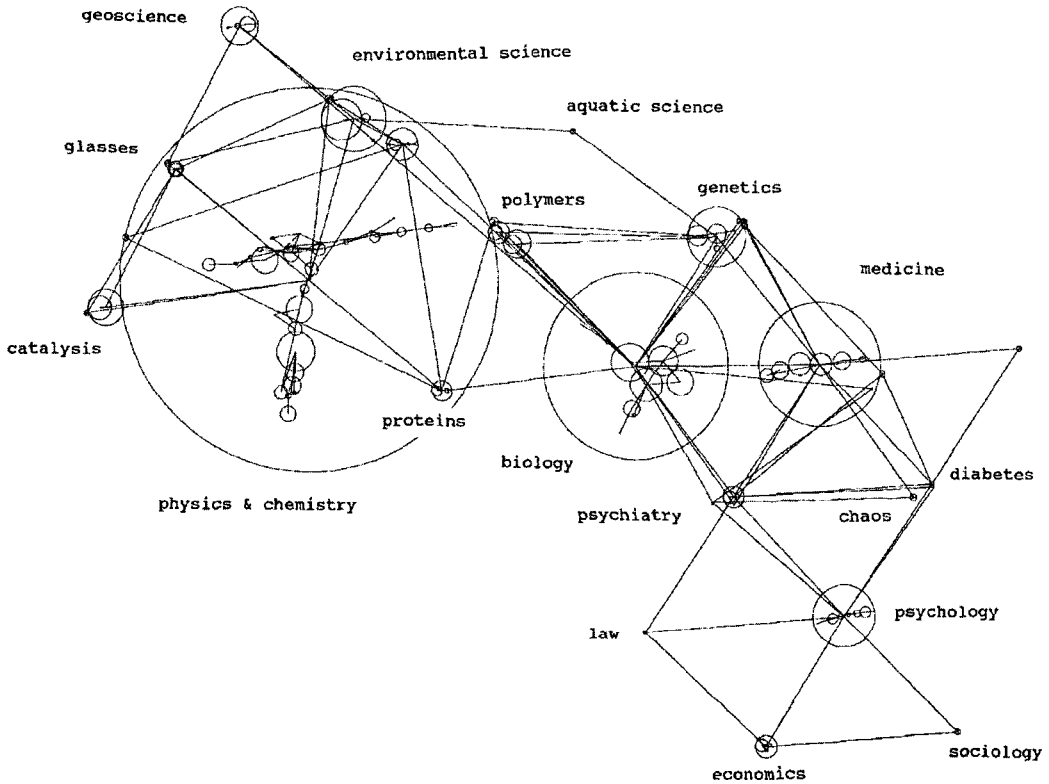
정함. 이때 단순 상수를 적용하면 학문 분야별 차이에 따라 편향된 결과가 얻어 지므로 한 논문으로부터의 인용은 그 논문의 참고문헌 수로 나눈 값으로 셈하였다.

- ② 동시인용값 산출 - 동시인용 빈도를 산출하고 코사인 정규화(각 문헌의 인용빈도를 곱한 것의 제곱근으로 동시인용빈도를 나눔)를 수행하였다.
- ③ 클러스터링 - 속도 문제로 단일연결 클러스터링을 사용하되, 한 군집이 과점하는 문제를 방지하기 위해서 단계적인 클러스터링을 수행하였다. 최대 60개짜리 군집이 생성되면 중단한 후 다시 군집을

단위로 하여 단일연결 클러스터링을 반복하는 방법으로 대형 군집(mega cluster)을 생성하였다.

- ④ 지도 생성 - 단계적 단일연결 클러스터링을 네 다섯 차례 반복한 후 남은 군집들 사이의 동시인용값 행렬을 입력 데이터로 하여 다차원 척도법(multidimensional scaling; MDS)이나 이보다 단순한 삼각측량법(triangulation)을 적용해서 지도를 생성하였다.

Small(1999)은 이와 같은 절차를 거쳐서 1995년에 출판된 문헌에서 인용된 1981-1995



〈그림 1〉 1995년 SCI와 DB의 문헌동시인용을 이용한 학문 영역 지도 (Small 1999)

년 사이 15년간 발표된 문헌 사이의 동시인용 정보를 사용하여 〈그림 1〉과 같은 광역 과학 지도를 생성하였다. 15년간의 문헌 중 164,612건 사이의 동시인용을 분석 대상으로 삼아서 5단계로 결집한 결과 가장 큰 군집에 전체 문헌의 29%인 36,720건이 포함되었다. 이 군집은 처음 36,720건의 문헌이 1단계에서 4,341개의 군집으로 결집되고, 2단계에서 699개, 3단계에서 148개, 4단계에서 35개의 군집으로 결집된 후 5단계에서 하나로 모인 집단이다.

생성된 군집 및 군집간 관계 시각화를 위해 Small은 1985년 논문에서는 MDS를 사용하였고, 1999년에 발표한 논문에서는 이보다 간

단하지만 전역 배치 최적화를 수행하지 않는 삼각측량법을 사용하여 지도를 생성하였다. 삼각측량법은 개별 노드를 하나씩 차례대로 배치하되, 새로 자리 잡을 노드의 위치는 이미 자리 잡은 노드들과의 거리에 따라 결정되는 방법이다. 이는 노드의 배치 순서에 따라서 결과가 달라지는 것이 단점이지만 MDS에 비해서 월등히 속도가 빠르므로 Small은 여러 번 반복해서 좋은 경우를 선택하였다.

〈그림 1〉의 원이나 점은 4단계에서 생성된 35개 군집이고 원 안에 포함된 작은 원은 3단계에서 생성된 군집이다. 이처럼 단계적으로 세부 차원의 지도를 만들어서 볼 수 있다. 연

결선은 동시인용값이 상대적으로 높은 군집간 관계를 표현한 것이다. 각 군집의 제목은 소속 문헌의 제목과 수록 학술지의 분야를 감안하여 연구자가 판정하였다. 전체적으로 오른쪽 아래에서 왼쪽 위 방향으로 사회과학, 생명과학, 자연과학의 순으로 자리를 잡았다. 자연과학 영역의 범위가 넓게 나타났으며 생명과학 영역이 중앙에 위치한 것이 눈에 띄는 특징이다.

2.2 Moya-Anegón과 Vargas-Quesada의 사이언토크램

스페인 그라나다 대학 문헌정보학과와 Moya-Anegón 교수와 그가 이끄는 SCImago 연구팀의 일원인 Vargas-Quesada는 Atlas of Science 프로젝트를 수행하면서 스페인 연구자들이 SCI 등재 학술지에 게재한 연구로 구성된 학문 영역 지도를 발표하였다(Moya-Anegón et al. 2004). 이들은 하나의 문헌에서 두 학술지가 동시인용된 경우를 세어서 학술지간 동시인용을 산출한 후 이를 각 학술지가 속한 ISI의 학술지 분류표, 즉 Journal Citation Reports에서 각 학술지에 지정한 학문분야 단위로 합산하여 학문 영역간 관계를 산출하였다.

SCImago 연구팀은 자신들의 광역 과학 지도를 사이언토크램(scientogram)이라고 불렀으며 이후 2007년에 단행본을 출간하면서 방법론을 보완하여 전 세계, 유럽, 스페인과 같은 여러 차원에서 사이언토크램을 생성하여 제시하였고, 프로젝트 홈페이지(www.atlasofscience.net)를 통해서 중남미 국가들과 포르투갈, 스페인의 학문 영역 지도

를 탐색해볼 수 있도록 서비스하고 있다.

네트워크 표현을 위해서 2004년 연구에서는 최소신장트리(minimum spanning tree)를 이용했으나 이후에는 <그림 2>와 같은 패스파인더 네트워크(Schvaneveldt 1990)를 이용하고 있다. <그림 2>는 2002년 Web of Science의 인용 자료 전체를 이용하여 생성한 사이언토크램이다. Small의 연구와 달리 SCI와 SSCI 뿐만 아니라 A&HCI까지 포함하였기 때문에 인문과학 영역도 포함되어 있다. 이 그림에서는 소속 논문이 54,972건으로 가장 많은 생화학 및 분자생물학 범주가 중앙에 위치하며 오른쪽에 자연과학, 왼쪽에 사회과학이 자리잡고, 인문학은 왼쪽 상단의 큰 노드인 역사학으로부터 가지 끝으로 전개되고 있다. 문헌정보학은 오른쪽 위의 전기전자공학에서 컴퓨터과학으로 전개되는 가지의 맨 끝에 위치하고 있다. 전체적인 배치 면에서 생명과학 영역이 중앙에 위치하고 좌우에 여타 자연과학 영역과 사회과학 영역이 자리잡은 것은 Small(1999)의 결과와 유사하다.

SCImago 연구팀은 학술지 사이의 동시인용 빈도에 다음과 같은 표준화 동시인용 계수 SCM(standardized cocitation measure)를 적용하여 연관도를 산출하였다(Vargas-Quesada & Moya-Anegón 2007, 66).

$$SCM(ij) = Cc(ij) + \frac{Cc(ij)}{\sqrt{c(i) \cdot c(j)}}$$

이 공식에서 $Cc(ij)$ 는 학술지 i 와 j 사이의 동시인용빈도이고, $c(i)$ 는 학술지 i 의 인용빈도이다. 풀어보자면 위의 공식은 두 학술지 사이의 동시인용빈도를 Salton 지수 공식으로 정규

화한 다음 이를 원래의 동시인용빈도에 더한 것이다. Salton 지수는 코사인 계수의 일종으로서 0에서 1사이의 값을 가지며, 동시인용빈도는 당연히 정수로 산출된다. 결국 SCM 공식에 의한 연관도의 우열은 동시인용빈도에 전적으로 좌우되며, 동시인용빈도가 같은 경우에만 코사인 계수의 크기가 의미를 가진다. 즉, 동시인용빈도에 의해서 연관도를 산출하되 동시인용빈도가 같은 경우의 우열을 가리는 용도로만 Salton 지수를 사용하고 있다.

2.3 Börner와 Boyack의 과학 지도

미국 인디애나 대학 문헌정보학과의 Börner 교수가 이끄는 정보시각화 연구실(InfoVis Lab)에서는 다양한 정보시각화 연구의 일환으로 학문 영역 시각화를 연구하고 있다. 이들은 2005년 *Scientometrics*에 발표한 논문에서 2000년 SCI와 SSCI DB에서 추출한 7,000여 종 학술지의 상호인용(inter-citation)에 근거한 '과학 지도'를 <그림 3>과 같이 제시하였다(Boyack et al. 2005).

이들은 학술지의 상호인용 빈도와 동시인용 빈도, 그리고 이들을 코사인이나 자카드 계수 공식 등으로 정규화한 지표까지 포함하여 8가지 방식의 학술지간 연관도를 산출해보았으며, 평가 결과 학술지간 상호인용 빈도를 자카드 계수로 정규화하는 방식이 미미한 차이지만 가장 좋았다고 보고하였다. 이들의 평가 기준은 ISI 학술지 분류표에서 동일 범주에 속한 저널이 유사한 주제를 다룬다는 가정에 근거한 것이다. 구체적으로는 각 연관도 공식에 의해 산

출한 학술지간 연관도로 k-평균 클러스터링을 수행한 후, 생성된 학술지 클러스터와 ISI 학술지 분류표의 범주가 유사할수록 더 좋은 방식이라고 판정하였다. 그러나 이와 같은 방식은 생성된 지도가 ISI 분류표와 얼마나 더 부합하는지를 평가한 것이지, 더 품질이 좋은가를 평가했다고 볼 수는 없다. ISI 분류표가 절대적으로 옳은 기준이라면 굳이 개별 학술지별로 인용 빈도를 확인할 이유가 없기 때문이다. 상호인용이나 동시인용에 의해서 파악되는 학술지 사이의 관계는 미리 정해진 주제 분류표에서의 관계와 다를 수 있다는 점을 감안하면, Boyack 등이 내린 판단처럼 상호인용에 의한 지도가 동시인용에 의한 지도보다 반드시 더 적합하다고 볼 수는 없다. 다만 생성된 지도가 ISI 분류표와 부합될수록 결과를 해석하기가 수월해진다는 점에서는 의의가 있다. 이들이 사용한 상호인용 자카드 지수 척도(IC-Jaccard index measure)는 다음과 같다.

$$JAC_{i,j} = \frac{C_{i,j} + C_{j,i}}{\sum_{k=1}^n C_{i,k} + \sum_{k=1}^n C_{j,k} - (C_{i,j} + C_{j,i})}$$

여기서 $C_{i,j}$ 는 학술지 i 가 학술지 j 를 인용한 횟수를 뜻한다. 따라서 이 공식은 결국 두 학술지간 상호인용 빈도를 각 학술지가 다른 학술지를 인용한 총 횟수로 정규화한 것이다. 두 학술지의 참고문헌에서 서로가 등장한 비율이 높아질수록 상호인용 자카드 지수 척도도 높아지게 된다.

<그림 3>은 SCI와 SSCI의 2000년 DB에서 상호인용이 검출된 7,121종 학술지간의 관계를 상호인용 자카드 지수로 산출한 후 Börner

교수팀이 자체 개발한 VxOrd 기법으로 시각화한 것이다. 그림에서 각 점은 학술지 한 종을 나타내며 두 점이 가까이 위치할수록 상호 인용이 빈번함을 뜻한다. 특정 지역의 분야명은 주변에 모인 학술지의 ISI 분류표 범주명 중에서 가장 우세한 것으로 선택하였다. 따라서 앞서 언급한 바와 같이 ISI 분류표와 가급적 부합되는 군집을 도출하는 것이 결과 해석에 유리하다. 특정 지역에 모이는 학술지의 ISI 분류 범주가 가급적 통일되면 주제 판정이 용이하기 때문이다. 그림에서 저널 종수가 많

은 분야는 큰 글씨로, 그렇지 않은 분야는 작은 글씨로 표현하였다. 전체적인 구성은 지도의 오른쪽에 수학, 물리학, 화학, 지구과학 등의 자연과학 영역이 자리잡고, 왼쪽 아래에는 의학을 비롯한 생명과학, 왼쪽 위에는 심리학과 사회학을 비롯한 사회과학 영역이 위치해있다. 문헌정보학(LIS)은 그림의 맨 위쪽에 컴퓨터 과학(Comp Sci) 옆에 나타나서 앞서 살펴본 SCImago 연구팀의 사이언토크램과 유사하지만, 여기서는 여타 사회과학 분야와도 그리 멀지 않은 것으로 표현되었다.



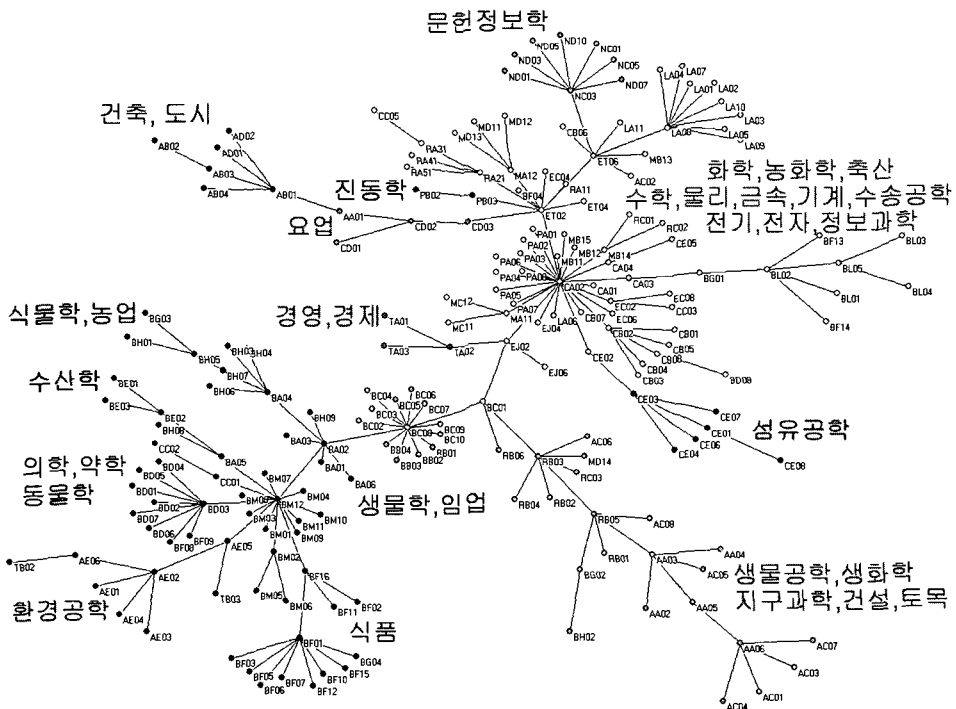
〈그림 3〉 2000년 SCI, SSCI 수록 7,121개 학술지 상호인용 분석 결과 지도 (Boyack et al. 2005)

Börner 교수팀은 이 지도를 가공하여 색상을 입히고 재미있는 이미지를 추가한 후 초등학교 수업에서 학문 세계에 대해서 가르치는 부교재로 활용하는 색다른 응용도 시도하였다.

2.4 KISTI 해외학술지 논문 분류를 이용한 학문범주 네트워크

이재운(2005)은 KISTI의 해외 과학기술분야 학술지 논문 데이터베이스를 자료로 하여, 인용 정보가 아닌 KISTI 표준분류표의 분류항목간 유사성에 기초해서 202개 학문범주의 패스파인더 네트워크를 <그림 4>와 같이 생성하였다. 이 그림에서 분류항목간 유사성은 학술

지 공유성, 즉 분류항목에 논문이 할당된 학술지가 서로 겹치는 정도를 반영하여 산출하였다. 연구자들이 논문을 발표하고 읽는 핵심 무대인 학술지가 상당수 겹치는 두 분야가 있다면 둘 사이는 가깝다고 판단한 것이다. 이는 학문 분야 사이의 관계를 게재 학술지에 근거해서 판단하는 방식이다. 구체적으로는 한 분야에 논문이 할당된 학술지의 논문 건수로 그 분야의 프로파일을 표현한 다음, 분야 프로파일간의 내적유사도를 산출하여 분야간 유사성을 구하였다. <그림 4>를 보면, 일반적으로 인접학문이라고 생각하는 학문분야가 분리된 경우가 드러난다. 건축/도시와 건설/토목의 이격, 식물학과 동물학의 이격, 농업과 임업의



<그림 4> KISTI 해외학술지 논문 분류 정보를 이용한 KISTI 표준 분류표 항목간 유사도 네트워크 (이재운 2005)

이격 등이 그 사례이다. 반면에 건축/도시는 요업과 인접하고, 건설/토목은 지구과학과 인접한다. 또한 동물학은 의학과, 식물학은 농업과, 생물학은 임업과 각각 인접하게 위치한다. 이와 같이 현재 분류체계 상에서는 떨어져 있는 분야이지만 논문이 수록되는 학술지를 살펴보면 가까운 경우가 존재한다. 이는 기존 열거식 계층형 분류체계가 가지는 자연스러운 한계이므로, 이 연구와 같이 학술지 공유 관계 분석을 통해서 주제분류에 대한 재검토와 학문 분야 구조의 새로운 해석이 가능함을 알 수 있다.

이 지도는 국내에서 최초로 시도된 광역 과학 지도이지만, 해외학술데이터베이스를 대상으로 분석한 것이므로 국내의 연구동향이 아닌 KISTI에서 수집하고 있는 해외학술지에 반영된 연구동향을 나타내는 지도라고 할 수 있다. KISTI의 학술지 구독 범위가 매우 넓기는 하지만, 특정 기관의 구독 학술지만 대상으로 하였다는 점에서 생성된 지도의 가치는 다소 제한된다. 다만 인용분석이 아닌 다른 방법으로도 광역 과학 지도의 생성이 가능함을 보였다는 점에서 의의를 가진다.

3. 국내 연구 동향을 반영하는 광역 과학 지도 작성

3.1 자료 수집 및 처리

앞에서 살펴본 사례들은 모두 전 세계적인 연구 동향이나 다른 나라의 연구 동향을 반영한 과학 지도라고 할 수 있다. 이 연구에서는 국내 학문 연구 동향을 반영하는 광역 과학 지

도의 제작을 시도하였다.

여러 학문 영역을 망라하여 구축되어 서비스가 이루어지고 있는 국내 인용색인 데이터베이스가 없으며, 분석을 위한 인용 정보를 자체적으로 데이터베이스화하기는 매우 어려운 일이다. 따라서 이 연구에서는 인용이 아닌 다른 수단을 모색한 결과, 한국학술진흥재단(KRF)에 연구비 지원을 신청한 연구과제 정보를 기초 자료로 사용하였다. KRF에 신청되는 과제는 연구자들이 국내외 기존 연구 동향을 감안하여 연구의 필요성과 의의가 타당하다고 생각한 주제를 다루기 마련이므로 어느 정도는 국내의 연구 전선(research front)을 반영한다고 볼 수 있다.

개별 신청 과제는 학술진흥재단 연구분야 분류표(이하 KRF 분류표로 표현함)에 따라 분야가 지정되어 있으므로, 각 학문 분야에 지정된 신청 과제가 해당 분야를 반영하는 것으로 가정하였다. 학문 분야 사이의 연관성은 소속된 신청과제 정보로부터 자동추출한 색인어가 얼마나 유사한가를 반영하여 산출하였다. 즉 특정 학문 분야로 신청된 과제 지원서에 사용된 용어로 해당 학문 분야를 나타내는 프로파일들을 비교하여 분야간 연관성을 산출한 것이다.

2001년부터 2005년까지 5년간 신청된 과제 중에서 학생 연구 지원 사업과 같이 주제성이 약한 사업을 제외한 모든 사업에 신청된 80,502건을 대상으로 과제의 제목과 한글 키워드 필드로부터 지능형형태소분석기를 사용하여 명사 색인어를 추출하였다. 추출된 색인어에 대해서 KRF 분류표의 소분류 범주 항목을 기준으로 역범주빈도 가중치를 산출하고,

여기에 중분류 수준에서의 출현빈도를 곱하여 각 중분류 범주에 대한 개별 색인어의 가중치를 산출하였다.

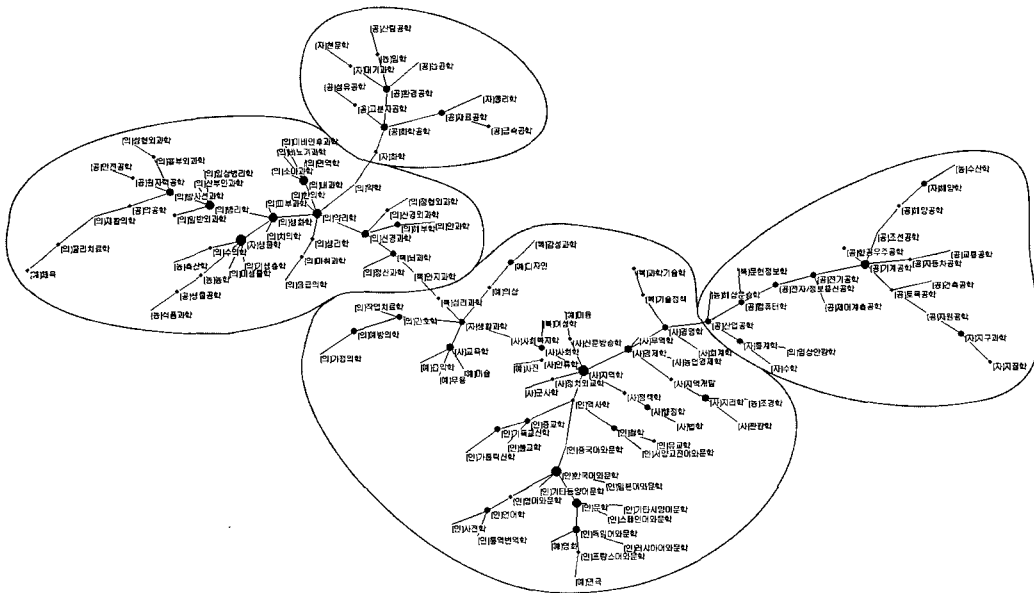
KRF 분류표 상의 중분류 범주에는 '일반'이나 '기타'와 같은 표현이 포함된 범주가 일부 있는데 여기에 범주에 포함된 연구 과제는 주제가 일관되지 않으므로 이들을 제외하고 남은 140개 학문 범주를 분석 단위로 삼고 벡터코사인 계수를 적용하여 연관성을 측정하였다.

140개 학문 범주 사이의 연관성으로 구성된 크기가 140×140인 행렬로부터 국내 광역 과학 지도를 시각화하고 군집을 분석하기 위해서 우선 패스파인더 네트워크(PFNet) 알고리즘과 병렬 최근접이웃연결 클러스터링(PNNC) 기법(이재운 2006b)을 사용하였다. 분석 결과 PFNet으로 표현된 지도에서 일부 단점이 파

악되었으므로 이의 대안으로 클러스터링 기반 네트워크(Clustering-based Network; CBNet) 알고리즘(이재운 2007)과 집단내 평균연결 클러스터링 알고리즘을 이용한 지도를 생성한 후, 두 지도를 참조하여 보완한 최종 국내 광역 과학 지도를 제시하였다. 네트워크의 시각적 표현을 위해서는 공개용 네트워크 분석 소프트웨어인 Pajek(de Nooy et al. 2005)을 사용하였다.

3.2 PFNet으로 표현한 국내 광역 과학 지도

학술진흥재단에 2001년부터 2005년까지 신청된 연구과제 정보를 활용하여 KRF 분류표 상의 중분류 범주 사이의 연관성을 산출하여



(그림 5) PFNet과 PNNC 기법으로 표현한 국내 광역 과학 지도

생성한 PFNet은 <그림 5>와 같다. 그림에서 학문분야명 앞에 대괄호로 둘러싼 글자는 KRF 분류표 상의 대분류 영역 구분 약자이다([공]-공학, [복]-복합학, [사]-사회과학, [인]-인문학, [예]-예술체육, [의]-의약학, [자]-자연과학).

<그림 5>에서 각 노드의 크기는 최근접이웃 중심성(이재운 2006a)에 비례하도록 설정하였다. 최근접이웃중심성은 해당 노드를 가장 가까운 노드로 간주하는 이웃 노드의 수로 측정되며 지역 중심성을 나타낸다. 대체적으로 표현된 네트워크 상에서 직접 인접하는 학문 분야가 많을수록 크게 표현된다. 따라서 수학과 같은 기초학문 분야보다는 지역학과 같이 다면적인 학문 분야의 크기가 크게 표시되고 연결선도 많이 보유한 것으로 나타난다. 140개 분야 중에서 최근접이웃중심성이 매우 높은 분야를 제시하면 <표 2>와 같다. 이 표에서 상대적 최근접이웃중심성은 최근접이웃중심성을 이론상 가능한 최대값으로 나눈 것이다.

<표 2> 최근접이웃중심성 상위 분야

순위	분야	최근접이웃중심성	상대적 최근접이웃중심성
1	[자]생물학	5	0.036
1	[사]지역학	5	0.036
1	[인]한국어와문학	5	0.036
4	[의]약리학	4	0.029
4	[의]생화학	4	0.029
4	[의]내과학	4	0.029
4	[의]병리학	4	0.029
4	[공]기계공학	4	0.029
4	[인]문학	4	0.029

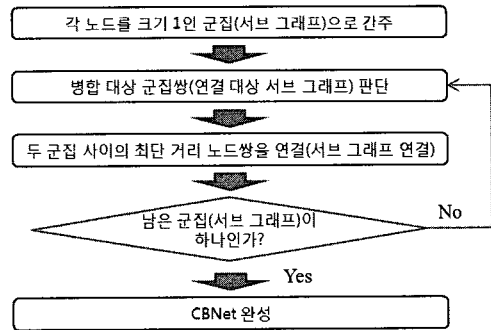
<그림 5>는 PNNC 기법에 의해서 크게 네 개의 영역으로 분할되는데, 가운데에는 인문·사회과학이, 왼쪽에는 생명과학이 각각 위치한다. 그런데 자연과학/공학은 뒤섞여서 왼쪽 끝과 오른쪽 끝으로 분리되어 있으므로 부자연스러운 결과이며 일반적인 학문 영역 구분과 다른 구조로 나타났다. 또한 이 그림에서 문헌정보학을 찾아보면 오른쪽 중간의 컴퓨터학에 연결되어서 자연과학/공학 계열 군집에 속해있다. 이는 SCImago 연구팀이 만든 <그림 2>의 사이언토크램과 유사한 결과이다. 그러나 Web of Science에 나타난 국제적인 연구동향에 비하면 국내 문헌정보학 분야 연구가 자연과학/공학 영역에 자리잡을 만큼 컴퓨터학과 가까워졌는지는 다소 의문이 든다.

지적 구조 표현에 널리 사용되고 있는 PFNet으로 생성한 <그림 5>의 국내 광역 과학 지도는 이처럼 전체적인 영역 배치와 국지적인 학문 분야의 소속 면에서 일부 문제점이 드러나 보인다. SCImago 연구팀은 이와 같이 PFNet으로 생성한 지도에서 이상한 부분이 있더라도 수정하지 말고 생성된 그대로를 제시하되, 이용자에게는 문제점이 있는 것을 감안하고 볼 것을 주지시켜야 한다고 주장하였다(Vargas-Quesada & Moya-Anegón 2007, 76).

3.3 CBNet으로 표현한 국내 광역 과학 지도

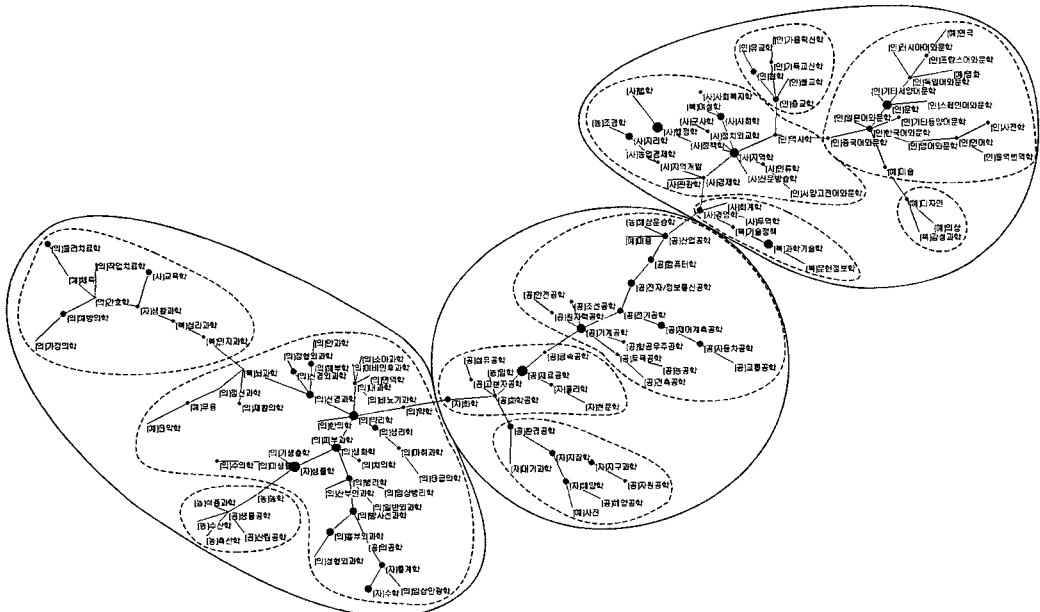
앞 절에서 PFNet으로 표현한 국내 광역 과학 지도가 전체적인 영역 배치와 국지적인 학문 분야의 소속 면에서 문제점을 나타냈으므로, 이 절에서는 PFNet과 유사하지만 다소

다른 네트워크를 생성하는 기법인 CBNet으로 국내 광역 과학 지도를 생성해보기로 한다. CBNet은 클러스터링 기반 네트워크 (Clustering-based Network; CBNet)의 출입말로서, Ward 클러스터링 알고리즘이나 평균연결 클러스터링 알고리즘과 같은 기존의 계층적 클러스터링 알고리즘을 이용하여 <그림 6>과 같은 절차로 네트워크를 구축하는 기법이다(이재윤 2007). CBNet은 연결 대상 서브그래프를 선정할 때 이용하는 계층적 클러스터링 알고리즘에 따라서 이름이 구분되는데, 이 연구에서는 여러 방식을 모두 시도해본 결과 학문 영역의 구분이 학문분야 분류표와 가장 잘 어울린다고 판단된 집단내 평균연결 알고리즘을 이용하는 CBNet-Wave 방식의 결과를 채택하였다.



<그림 6> CBNet 알고리즘

140개 KRF 분류 범주 사이의 연관성 행렬을 입력자료로 하여 생성한 네트워크인 CBNet-Wave는 <그림 7>과 같다. 이 그림에서 네트워크를 분할한 클러스터의 설정은 CBNet-Wave를 형성하는데 이용된 집단내 평균연결 클러스터링 기법으로 이루어졌다. 클러



<그림 7> 2001-2005년 학술진흥재단 분야별 신청과제 용어 유사성에 근거한 CBNet-Wave

스터링 결과 작은 군집 11개가 생성되었으며 이들이 결집된 큰 군집은 3개로 설정하였다.

〈그림 7〉의 국내 광역 과학 지도는 전체적으로 오른쪽의 인문·사회과학, 가운데의 자연과학/공학, 그리고 왼쪽의 생명과학으로 구분된다. 이는 앞서 PFNet으로 생성한 지도에서 자연과학/공학 영역이 네트워크의 양쪽 끝으로 나뉘었던 것에 비하면 일반적인 학문 영역 구분에 매우 근접한 결과이다.

CBNet-Wave에서 문헌정보학의 위치를 살펴보면 학술진흥재단 분류표에서 함께 복합학으로 구분된 과학기술학과 연결되어서 오른쪽의 인문·사회과학 영역에 위치하였다. 앞서 PFNet에서는 문헌정보학이 컴퓨터학에 연결되어서 자연과학/공학 영역에 위치하였던 것에 비하면 소속 영역 면에서 더 적절한 것으로 판단된다. 이처럼 두 네트워크에서 문헌정보학의 위치가 달라진 이유는, PFNet에서는 문헌정보학과 가장 가까운 한 분야만 고려한 반면에 CBNet-Wave에서는 가장 가까운 학문 소속한 군집(서브 그래프)의 다른 학문 분야도 함께 고려해서 연결 대상을 결정했기 때문이다. 문헌정보학이 컴퓨터학과 매우 가까울 수는 있어도 컴퓨터학과 같이 공학에 속할 수는 없음이 CBNet-Wave에서 확인되었다고 볼 수 있다.

3.4 PFNet과 CBNet의 비교를 통한 광역 과학 지도의 개선

앞 절에서처럼 CBNet-Wave 기법으로 생성한 지도가 PFNet으로 생성된 지도에 비해서 영역 배치가 더 적절한 것으로 판단되지만,

구체적으로 살펴보면 PFNet으로 생성된 지도가 더 나은 점도 일부 발견된다.

〈그림 7〉의 CBNet-Wave에서는 심리과학이나 교육학이 생활과학 분야와 연결되어서 왼쪽 끝의 생명과학 영역에 위치하고 있으나, 〈그림 5〉의 PFNet에서는 이들이 중앙의 사회과학 영역에 자리잡은 것을 볼 수 있다. CBNet-Wave에서는 심리과학과 교육학이 여타 사회과학과 거리가 매우 먼 것으로 표현되었으므로 부적절한 결과라고 볼 수 있다.

PFNet과 CBNet은 각각 나름대로의 기준에 따라서 학문 분야를 배치하였지만, 어느 한 방식에 의한 지도가 모든 면에서 우월하다고 말하기는 어려운 결과가 나타났다. 이는 두 가지 네트워크 생성 알고리즘이 모두 나름대로의 기준에 따르고 있으므로 각자가 지적 구조의 어느 한 측면만을 반영하고 있기 때문이다.

두 그림을 구체적으로 비교해보았을 때 가장 다른 부분은 기계공학 분야와 금속공학 분야 사이의 관계, 그리고 생활과학 분야와 사회복지학 사이의 관계이다. 〈그림 5〉의 PFNet에서는 기계공학 분야와 금속공학 분야가 거의 왼쪽과 오른쪽 끝으로 떨어져 있는데 반해서 〈그림 7〉의 CBNet-Wave에서는 두 분야가 네트워크의 중앙 영역에서 직접 연결되어 있다. 반대로 CBNet-Wave에서 네트워크의 왼쪽과 오른쪽으로 멀리 떨어져 있는 생활과학 분야와 사회복지학 분야는 PFNet의 중앙 영역에서 바로 이웃하여 연결되어 있다.

이처럼 두 분야쌍의 연결 여부가 두 가지 네트워크에서 확연히 다른 결과를 가져오고 있다는 사실은 〈표 3〉의 근접중심성 지표로도 확인할 수 있다. 근접중심성(closeness

centrality)은 한 노드가 다른 노드들과 얼마나 평균적으로 가까이 있는지를 측정하는 전역 중심성 척도이다(Freeman 1979). 여기에서의 거리란 한 노드에서 다른 노드로 건너가기 위해서 거쳐야 하는 링크의 수로 산출되는 경로거리를 뜻한다. 따라서 근접중심성이 높다는 것은 다른 노드들과 평균적으로 가까운 위치, 즉 네트워크의 중앙에 자리잡았음을 뜻한다. PFNet과 CBNet-Wave에서 140개 학문 분야의 근접중심성을 각각 측정해본 결과는 <표 3>과 같다.

<표 3>에서 증가폭은 CBNet에서 측정한 각 분야의 근접중심성 값에서 PFNet에서 측정한 근접중심성 값을 뺀 것이다. 증가폭이 매우 크거나 매우 작을수록 두 네트워크에서 해당 분야의 전역적인 입지가 크게 다르다는 것을 뜻하게 된다. 측정결과 증가폭이 큰 1위와

2위는 금속공학과 기계공학 분야로 나타났다. PFNet에서 멀리 떨어져 있던 두 분야가 CBNet-Wave에서 이어짐으로써 각자의 구조적인 입지가 가장 크게 개선된 것이다. 반대로 증가폭이 가장 작은(감소폭이 가장 큰) 1위와 2위는 생활과학과 사회복지학으로 나타났다. 이번에는 PFNet에서 이어져 있던 두 분야가 CBNet-Wave에서 멀리 떨어짐으로써 각자의 구조적인 입지가 가장 심하게 저하된 것이다.

이와 같이 금속공학과 기계공학, 그리고 생활과학과 사회복지학의 연결 여부가 네트워크 구조에 매우 심각한 영향을 끼치는 점을 감안하여 이들 두 쌍을 모두 연결시킨다면 현재의 PFNet과 CBnet-Wave이 나타내고 있는 상반된 관점을 통합하는 지도가 생성되는지 여부를 시험해보았다. 이를 위해서는 CBNet-Wave의 관점을 반영하여 원래 PFNet에서는 분리되었던 금속공학과 기계공학을 연결하거나, PFNet의 관점을 반영하여 원래 CBNet-Wave에서는 분리되었던 생활과학과 사회복지학을 연결하는 두 가지 방법이 모두 가능하다. 여기서는 상대적으로 더 나은 영역 배치를 보였던 CBNet-Wave에다가 PFNet의 관점을 반영하는 방법을 채택하여 <그림 8>을 생성하였다. 지면 관계상 제시하지는 않았지만 반대의 방법으로 PFNet을 수정한 경우도 <그림 8>과 매우 유사한 형태를 보였다.

CBNet-Wave를 수정한 <그림 8>에서는 중앙에 큰 고리가 형성된 것을 볼 수 있다. 고리의 모양은 원형보다는 6각형에 가까운데, 꼭지점을 이루는 여섯 개의 분야는 12시 방향의 약리학에서부터 시계 방향으로 돌아가면서 생

<표 3> 두 네트워크에서 측정된 분야별 근접중심성의 비교

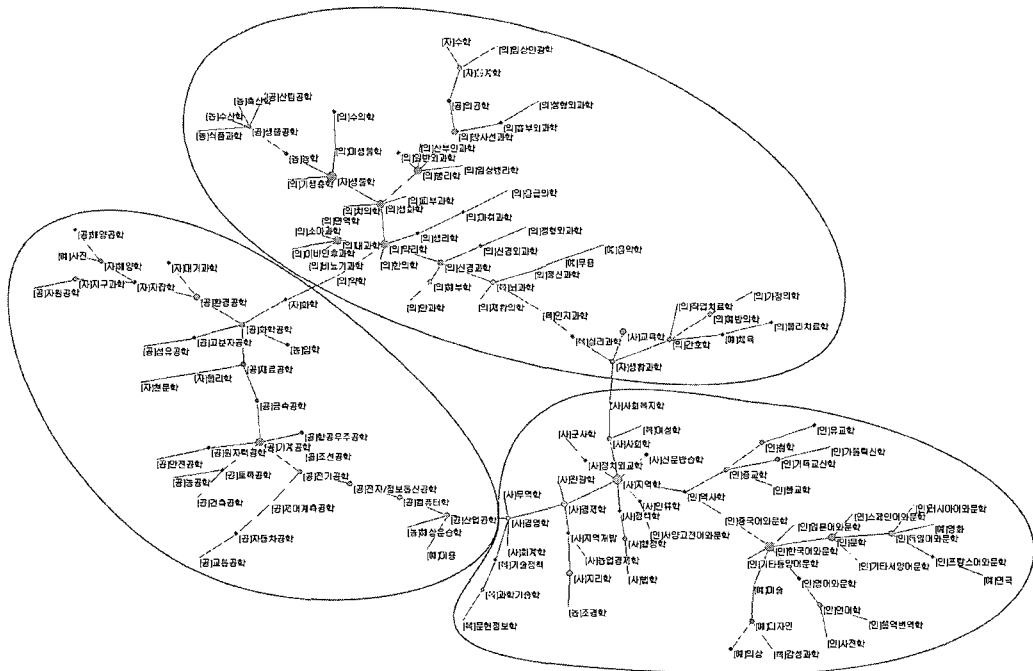
	분야명	PFNet에서의 근접중심성	CBNet에서의 근접중심성	증가폭
1	[공]금속공학	0.079	0.123	0.044
2	[공]기계공학	0.083	0.123	0.040
3	[자]지질학	0.063	0.101	0.038
4	[공]재료공학	0.086	0.124	0.038
5	[공]토목공학	0.077	0.110	0.033
:	:	:	:	:
:	:	:	:	:
136	[사]교육학	0.125	0.072	-0.053
137	[복]심리과학	0.138	0.083	-0.055
138	[예]의상	0.125	0.067	-0.058
139	[사]사회복지학	0.142	0.084	-0.059
140	[자]생활과학	0.142	0.077	-0.065

활과학, 지역학, 경영학, 기계공학, 화학공학 이라고 할 수 있다. 6개 분야 중에서도 앞의 <표 2>에 제시된 바와 같이 최근접이웃중심성 기준으로 공동 4위 이내에 포함된 분야는 하나씩 건너뛰면서 약리학, 지역학, 기계공학의 세 분야이다. 이들 세 분야는 집단내 평균연결 클러스터링에 의해 구분된 세 영역(생명과학, 인문·사회과학, 자연과학/공학)에 각각 하나씩 위치하고 있다. 따라서 생명과학 영역에서는 약리학이, 인문·사회과학 영역에서는 지역학이, 그리고 자연과학/공학 영역에서는 기계공학이 각각 다방면으로 연결되는 다리에 해당한다고 볼 수 있다.

PFNet을 참조하여 CBNet-Wave를 일부 수정한 결과, 오른쪽 아래의 인문·사회과학

영역에서 왼쪽의 화학공학 분야로 연결되는 경로가 두 가지가 되는 효과를 얻었다. 이 중에서 아래쪽에 위치한 지역학-경영학-기계공학-화학공학의 순서로 이어지는 경로는 수정하기 전의 CBNet-Wave에 존재한 것이고, 위쪽에 위치한 지역학-생활과학-약리학-화학공학의 순서로 이어지는 경로는 PFNet에 있던 경로를 수용한 것이다. 이처럼 PFNet과 CBNet-Wave의 두 가지 네트워크에서 가장 상반되게 표현되었던 연결고리(학문분야간 링크) 하나를 수용한 결과 두 네트워크의 다른 관점이 하나의 지도에 수용되는 효과를 얻게 되었다.

이처럼 CBNet-Wave를 보완한 광역 과학 지도는 영역간의 관계가 선형적인 스펙트럼으로 나타나지 않고 순환적인 고리 구조로 나타



<그림 8> PFNet의 관점을 반영하여 수정된 CBNet-Wave

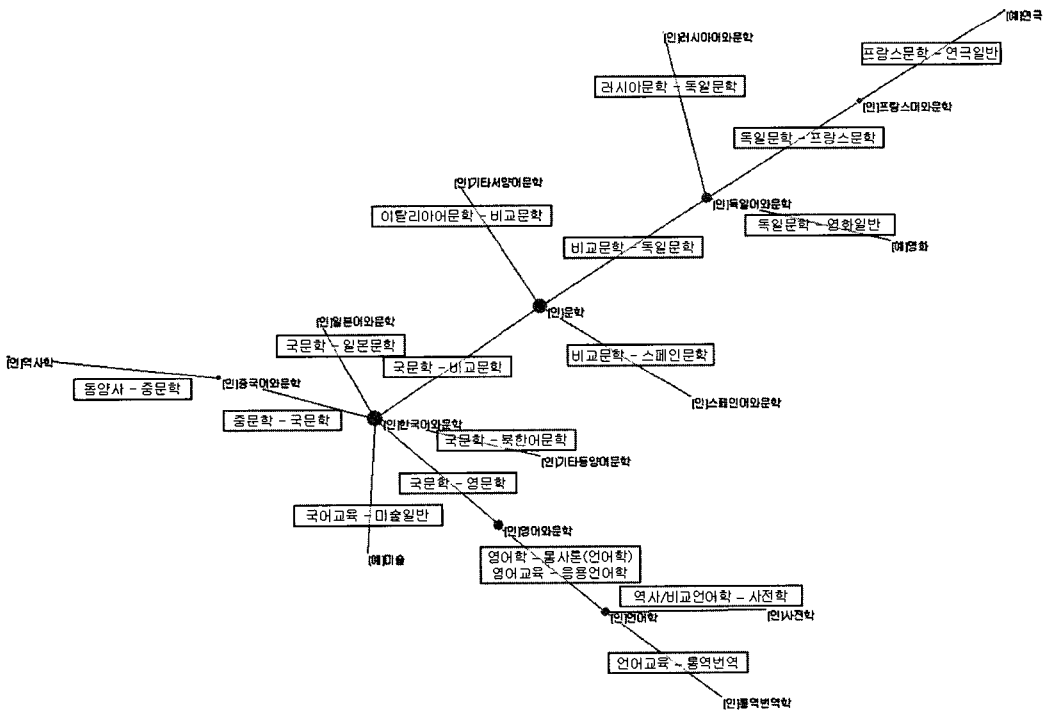
남으로써 학문 영역의 구조를 다면적으로 해석할 수 있는 여지를 남겼다. 예를 들어 생명과학 영역을 살펴보면 인간이 주요 연구 대상인 신경과학, 뇌과학 등의 분과 학문은 인지과학과 심리과학을 통해서 사회과학으로 이어지게 되며, 물질이 주요 연구 대상인 약학은 화학을 통해서 자연과학 및 공학으로 이어진다는 양면성이 지도에 나타난 것이다.

3.5 국지적인 연결 구조의 해석

앞 절에서 제시된 광역 과학 지도를 이루는 140개의 분야는 KRF 분류표 상의 중분류 범주에 해당하므로 각각이 매우 넓은 범위를 포

괄하고 있다. 따라서 어느 두 학문 분야가 가깝다고 판정되어 네트워크 상에서 연결되는 이유를 더욱 구체적으로 파악하려면, 현재의 중분류 수준에서 더 깊이 들어가서 소분류 단계나 세분류 단계에서의 관계를 파악할 필요가 있다. 두 학문 분야가 밀접한 것으로 판정되었다고 해서 각각에 속한 모든 분과 학문도 서로 밀접하지는 않을 것이기 때문이다. 소분류 수준에서 살펴보면 두 학문 분야가 중분류 수준에서 밀접하다고 판정을 받는데 가장 크게 기여한 소분류 항목간의 관계를 찾아낼 수 있다.

〈그림 8〉에서 주로 어문학 분야 학문들로 구성된 오른쪽 아래의 가지만 떼어낸 후, 가지로 직접 연결된 두 학문 분야에 속한 분과 학



〈그림 9〉 광역 과학 지도 중에서 어문학 위주의 영역에 대한 소분류 수준의 연결 관계 해석

문들끼리의 관계 중에서 가장 가까운 조합쌍을 파악하여 제시하면 <그림 9>와 같다. 예를 들어 한국어와문학 분야의 여러 분과 학문과 중국어와문학 분야의 여러 분과 학문 중에서 가장 연관성이 높은 조합쌍은 국문학과 중문학의 조합임이 표시되어 있다.

<그림 9>를 보면 영어와문학 분야를 제외한 대부분의 어문학 계열 분야들은 어학보다는 문학이라는 공통점 때문에 연결되고 있음이 드러난다. 특히 중앙에 위치한 중분류 수준 항목인 문학 분야는 그에 속한 여러 소분류 수준 분과 학문 중에서 비교문학 때문에 다수의 어문학 분야와 연결되고 있음이 드러난다. 한편 영어와문학 분야는 예외적으로 문학이외에 영어학이나 영어교육이라는 분과 학문의 역할도 두드러져서 언어학과 연결되는 것으로 나타났다.

이와 같이 KRF 분류표 상의 소분류 범주 사이의 관계를 활용함으로써 중분류 범주를 기준으로 생성한 광역 과학 지도를 더욱 구체적으로 해석할 수 있다. 다만 분석 대상 범주의 단위가 중분류에서 소분류로, 그리고 세분류로 더욱 더 깊어질수록 각 분류항목을 나타내는 단서 정보가 줄어드는 문제점은 유의해야 한다. 이 연구에서 채택한 단서 정보가 신청과제 정보이므로 특정 분야의 신청 과제가 단지 몇 건에 불과하다면 한 두 건의 과제 정보에 의해서 결과가 좌우될 위험이 있기 때문이다.

4. 결 론

전체 학문 분야 사이의 구조적인 관계를 시각화하여 분석하는 광역 과학 지도에 대한 연

구가 최근에 활발히 전개되고 있다. Small과 Garfield의 연구 이후 최근에는 스페인 그라나다 대학 문헌정보학과의 Moya-Anegón 교수가 이끄는 SCImago 연구팀과, 미국 인디애나 대학 문헌정보학과의 Börner 교수가 이끄는 연구팀의 활동이 두드러지고 있다. 이들은 모두 Web of Science의 인용 정보를 이용하는 광역 과학 지도를 연구하였다.

이 연구에서는 외국과 달리 인용 데이터베이스가 미비한 국내 현실을 감안하여 새로운 방법으로 국내 연구 활동 동향에 대한 광역 과학 지도를 생성해보았다. 이를 위해서 KCI 분류표의 각 분류 범주를 분석 대상인 학문 분야로 간주하고, 2001년부터 2005년까지 학술진흥재단에 지원 신청된 연구과제 신청서의 텍스트를 단서로 하여 학문 분야간 관계를 분석하였다.

PFNet과 CBNNet-Wave 기법으로 각각 생성한 광역 과학 지도 중에서 평균연결 클러스터링 알고리즘을 활용하는 CBNNet-Wave 기법으로 생성한 지도가 인문·사회과학 영역, 자연과학/공학 영역, 생명과학 영역으로 적절히 구분되므로 더 적합한 것으로 판단되었다. 그러나 두 기법에 의한 지도가 학문 분야 사이의 연계 관계를 반영하는 시각이 상호보완적이므로 이를 결합하여 CBNNet-Wave의 결과를 보완하는 시도를 하였다. 그 결과로 생성된 광역 과학 지도는 학문 영역의 전역 배치가 적절히 구분되면서 학문 영역 사이의 연계가 선형적인 스펙트럼이 아닌 순환적인 고리 구조로 나타났다. 따라서 학문 영역의 구조를 다면적으로 해석할 수 있는 가능성을 보였다.

또한 KRF 분류표 상의 소분류 범주 사이의 관계를 활용함으로써 광역 과학 지도에 나타난

중분류 범주 사이의 관계를 더욱 구체적으로 해석할 수 있다.

이 연구에서 사용한 광역 과학 지도 생성 방식은 텍스트가 단서라서 인용DB를 필요로 하지 않으며, 예술·체육 영역을 비롯해서 인문·사회·자연의 전 영역을 망라할 수 있으므로 향후 국내 연구 추세 분석과 연구 지원 정책 수립에 도움이 될 것으로 기대된다.

이 연구는 텍스트를 단서로 하여 광역 과학 지도를 생성하였지만, 향후에는 여러 가지 단서에 의한 광역 과학 지도 생성 연구가 각각

이루어져야 한다. 우선 KCI와 같은 국내 연구 문헌의 인용 DB가 구축되면 인용 정보를 이용한 광역 과학 지도를 생성해야 할 것이며, 학술 연구자의 협력 활동이나 논문 투고 활동과 같은 사회적 연결 고리를 통한 광역 과학 지도의 생성도 가능할 것이다. 이처럼 다각적인 측면에서 광역 과학 지도를 생성하고 이들을 비교하는 연구가 이루어진다면, 현재 인용 분석에 의존하고 있는 과학지도학을 한 단계 발전시킬 것으로 전망된다.

참 고 문 헌

- 김영진. 1986. 『논문의 동시인용을 통한 지적(知的)구조의 규명에 관한 연구』, 석사학위논문, 연세대학교 대학원, 문헌정보학과.
- 이재운. 2005. KISTI 과학기술문헌 서비스에의 자동분류 도입 방안 검토. KISTI 해외정보실 내부 세미나 발표자료.
- 이재운. 2006a. 계량서지적 네트워크 분석을 위한 중심성 척도에 관한 연구. 『한국문헌정보학회지』, 40(3): 191-214.
- 이재운. 2006b. 지적 구조 분석을 위한 새로운 클러스터링 기법에 관한 연구. 『정보관리학회지』, 20(4): 215-231.
- 이재운. 2007. 클러스터링 기반 네트워크 생성 알고리즘. 『제14회 한국정보관리학회 학술대회 논문집』, 147-154.
- 홍종선. 2007. 『KCI 기반 Kor-Factor (Korea Factor) 모형 개발 기초 연구』, 정책연구-2006-지식확산 02, 한국학술진흥재단.
- Atlas of Science Project Homepage. [online] [cited 2007.7.10] <<http://www.atlasofscience.net/>>.
- Boyack, K. W., R. Klavans, and K. Börner. 2005. "Mapping the backbone of science." *Scientometrics*, 64(3): 351-374.
- de Nooy, W., A. Mrvar, and V. Batagelj. 2005. *Exploratory Social Network Analysis with Pajek*. New York: Cambridge University

- Press.
- Freeman, L. C. 1979. "Centrality in social networks conceptual clarification." *Social Networks*, 1: 215-239.
- Garfield, E. 1998. *Mapping the World of Science*. Presented at the 150 Anniversary Meeting of the AAAS, Philadelphia, PA, February 14, 1998. [online] [cited 2007.7.15] <<http://www.garfield.library.upenn.edu/papers/mapsciworld.html>>.
- Garfield, E. 1994. "Scientography: mapping the tracks of science." *Current Contents: Social & Behavioral Sciences*, 7: 5-10.
- Moya-Anegón, F., B. Vargas-Quesada, V. Herrero-Solana, Z. Chinchilla-Rodríguez, E. Corera-Álvarez, and F. J. Munoz-Fernández. 2004. "A new technique for building maps of large scientific domains based on the cocitation of classes and categories." *Scientometrics*, 61(1): 129-145.
- Price, D. J. DeSolla. 1966. "The science of scientists." *Medical Opinion and Review*, 1(10): 88-97. Quoted in Small & Garfield(1985).
- Schvaneveldt, R. W. (ed). 1990. *Pathfinder Associative Networks: Studies in Knowledge Organization*. Norwood, NJ: Ablex.
- Small, H., and E. Garfield. 1985. "The geography of science: disciplinary and national mappings." *Journal of Information Science*, 11: 147-159.
- Small, H. 1973. "Co-citation in the scientific literature: a new measure of the relationship between two documents." *Journal of the American Society for Information Science*, 24(4): 265-269.
- Small, H. 1999. "Visualizing science by citation mapping." *Journal of the American Society for Information Science*, 50(9): 799-813.
- Vargas-Quesada, B., and F. Moya-Anegón. 2007. *Visualizing the Structure of Science*. Berlin: Springer-Verlag.