

학제적 분야의 정보서비스를 위한 학술지 인용 분석에 관한 연구: Y대학교 생명공학과를 중심으로*

Journal Citation Analysis for Library Services on Interdisciplinary Domains: A Case Study of Department of Biotechnology, Y University

유소영(So Young Yu)**

이재윤(Jae Yun Lee)***

초 록

이 연구에서는 자관의 학술지 상호인용 및 동시인용 분석을 통하여 단순 피인용빈도 이상의 학술지 인용 패턴 분석을 시도 하였다. 이 연구를 통해 학술지의 중요도 파악에 있어서 자관 인용 네트워크의 구조적 분석이 인용빈도 이상의 자관 인용 패턴에 대한 설명을 하고 있는지와, Web of Science에서 제공하는 JIF 이외의 일반적 인용 지수 서비스들을 고려해야 할 필요성이 있는지를 살펴보았다. Y대학교 생명시스템대학 생명공학과 전·현직 교수진이 2006년과 2007년에 발표한 학술논문의 인용 네트워크 분석 및 Web of Science 이외의 일반적 인용 지수들간의 관계를 분석한 결과는 다음과 같다. 첫째, 자관의 상호인용 네트워크를 통해 자관의 연구 분야를 확인할 수 있었다. 둘째, 자관의 동시인용 네트워크 지수들은 자관 인용 네트워크의 구조적 속성을 반영하는 인용 패턴의 설명이 가능하며 이는 피인용빈도와 유사하면서도 추가적인 설명력을 가지는 것을 확인하였다. 셋째, 일반적 인용지수로는 JIF 외에도 합산지향지수, h-index와 같은 다양한 일반적 인용 지수들의 설명력이 다양하므로 이를 이용하여 다각적으로 고려하는 것이 필요한 것으로 파악되었다. 또한 학술지 평가에서 인용 색인 데이터베이스의 수록범위보다는 지수의 유형에 따른 설명력 차이가 크다는 것을 확인하였다. 이와 같은 자관의 인용 네트워크 분석은 정보서비스의 여러 분야에서 유용하게 사용될 수 있을 것으로 기대된다.

ABSTRACT

In this study, we testify that network structural attributes of a citation network can explain other aspects of journal citation behaviors and the importances of journals. And we also testify various citation impact indicators of journals including JIF and h-index to verify the difference among them especially focused on their ability to explain an institution's local features of citation behaviors. An institutional citation network is derived using the articles published in 2006-2007 by biotechnology faculties of Y University. And various journal citation impact indicators including JIF, SJR, h-index, EigenFactor, JII are gathered from different service sites such as Web of Science, SCImago, EigenFactor.com, Journal-Ranking.com. As a results, we can explain the institution's 5 research domains with inter-citation network. And we find that the co-citation network structural features can show explanations on the patterns of institutional journal citation behavior different from the simple cited frequency of the institution or patterns based on general citation indicators. Also We find that journal ranks with various citation indicators have differences and it implies that total-based indices, average-based indices, and hybrid index(h-index) explain different aspects of journal citation pattern. We also reveal that the coverage of citation DB doesn't be a matter in the journal ranking. Analyzing the citation networks derived from an institution's research outputs can be a useful and effective method in developing several library services.

키워드: 네트워크 분석, 인용분석, 동시인용, 인용지수, JIF, 도서관 서비스

network analysis, citation analysis, cocitation, journal citation indicators, JIF, library services

* 이 논문은 2008 한국문헌정보학회/한국정보관리학회/국립중앙도서관 도서관연구소 추계공동학술대회(2008년 11월 7일, 국립중앙도서관)에서 발표한 것을 수정·보완한 것임.

** 연세대학교 문헌정보학과 박사과정(sweety798@yonsei.ac.kr) (제1저자)

*** 경기대학교 문헌정보학과 조교수(memexlee@kgu.ac.kr) (공동저자)

■ 논문접수일자: 2008년 11월 19일 ■ 최초심사일자: 2008년 11월 20일 ■ 게재확정일자: 2008년 12월 2일
■ 情報管理學會誌, 25(4): 283-308, 2008. [DOI:10.3743/KOSIM.2008.25.4.283]

1. 연구의 배경 및 목적

도서관에서 장서개발 및 주제전문 서비스 개발을 위하여 자관의 학술지 이용 현황을 살펴보고자 할 때 다양한 분석 방법 중의 하나로 인용 분석이 사용된다. 인용 분석을 통해서 사서는 소속 기관 이용자들의 연구동향과 주요 이용 학술지를 파악할 수 있으며, 이렇게 파악된 정보는 주제서비스, 장서개발, 이용자 교육 등의 여러 도서관 서비스를 고도화할 수 있는 단서가 된다. 이와 같이 자관 이용자들의 연구성과물에 나타난 학술지 인용정보를 분석하는 소속 기관 내 인용 분석은 사서의 업무를 전문화할 수 있는 유망한 방안이다.

학술지 인용 분석은 오랫동안 연구되어온 주제이지만 핵심적인 방법은 영향력 지수(impact factor)로 대표되는 인용빈도 위주의 측정이었다. 최근에 들어서는 학술지 순위화에 있어서 인용의 빈도 뿐 아니라 인용의 질(quality)을 측정하여 반영해야 한다는 인식이 증가하고 있다. 이는 학술지 평가에 각 학술지의 위세성을 반영하는 방식에 대한 관심으로 이어지고 있다 (Ball 2006; Hascall & Hanson 2008).

한편 생명공학과 같은 학제적 학문 분야에는 다양한 연구 영역이 공존하고 있고 연구 기관마다 세부 연구 주제의 구성이 전형적이지 않은 경우가 많으므로 소속 기관내 연구 동향 분석의 필요성이 더 높다. 더군다나 학제적 분야에서는 한 학술지가 얼마나 자주 이용되었는지 뿐만 아니라, 어떤 주제 분야들의 학술지들 간에 동시에 자주 이용되었는지도 고려하여야 한다. 왜냐하면 이는 학제적 분야의 학문적 특성을 반영할 수 있을 뿐만 아니라 그 분야 연구자

들의 연구 경향 및 연구 전선을 발견하여 도서관 서비스 및 장서 개발에 이를 반영하는 것을 용이하게 할 수 있기 때문이다. 따라서 학제적 분야의 인용 네트워크 분석은 기존의 인용빈도만을 고려하던 학술지 인용 분석 및 평가를 네트워크의 구조적 특성을 반영함으로써 보완하는 수단이 될 수 있을 것이다.

생명공학은 대표적인 학제적 분야로 의학, 약학, 식품학, 미생물학 뿐 아니라 화학 및 물리학까지 매우 다양한 연구 분야가 공존하고 있으며 이러한 분야의 학술지들이 주로 이용될 것으로 생각해 볼 수 있다. 또한 대부분의 학술 데이터베이스마다 생명공학 분야 해당 학술지를 정의하는 것이 다양할 뿐만 아니라 각 대학별/연구기관별로도 연구 특성화된 세부 주제 분야가 다양할 수 있다. 따라서 학제적 분야의 학술지는 자관의 이용특성(locality)을 파악하여 자관의 장서 개발 및 참고 봉사 서비스에 반영하는 것이 꼭 필요하다고 할 수 있다.

특히 학술지구독 선정을 위한 후보 리스트를 각 과에 제시할 때, 해당 학과 전공분야의 학술지 이용을 기반으로 하여 구독 학술지 후보 리스트를 작성할 수 있다면 좋을 것이다. 학제적인 분야의 경우, 여러 학문 분야가 혼재해 있으므로 앞서 언급한 바와 같이 특정한 색인 데이터베이스에만 의존하여 학술지 후보 리스트를 작성할 경우 자관 내 특정 과에서 이용을 자주 하는 학술지가 누락될 가능성이 크다.

따라서 자관의 학술지 인용 패턴을 파악하여 일반적인 학술지 순위와는 다른 자관 특유의 인용 패턴을 파악하는 것이 필요하다. 또한 자관에서 발표한 학술논문 내에 인용된 학술지를 파악함으로써 자주 이용되는 학술지를 유추해

볼 수 있고 이를 이용자 교육 및 주제 전문 참고 봉사의 주요 참고 자료로 활용하는 것도 가능하다.

이 연구를 통해 살펴보고자 하는 연구 문제는 다음과 같다.

첫째, 학술지의 중요도 파악에 있어서 자관 인용 네트워크의 구조적 분석을 통해 인용빈도 이상의 자관 인용 패턴에 대한 설명을 얻을 수 있는가?

둘째, 학술지의 중요도 파악에 있어서 Web of Science에서 제공하는 JIF 이외의 일반적 인용 지수 서비스들을 고려해야 할 필요성이 있는가?

위의 연구 문제들을 확인하고자 Y대학교 생명시스템대학 생명공학과 전·현직 교수진이 2006년과 2007년에 발표한 학술논문의 인용 분석을 통한 상호인용 네트워크를 시각화하여 연구 분야 및 연구경향을 살펴보았다. 또한 동시 인용 네트워크를 이용하여 자관 내 학술지 이용 특성을 다양한 네트워크 지수들을 통해 분석해 보고 이를 일반적인 학술지 이용 데이터라고 할 수 있는 다양한 학술지 인용 지수들과 비교하여 그 관계를 살펴보았다. 덧붙여, JIF를 포함한 일반적 학술지 인용 지수들 사이의 관계를 살펴봄으로써 JIF 이외의 일반적 인용 지수 서비스들을 고려해야 할 필요성도 확인해 보고자 하였다.

2. 인용 분석과 정보 관리

인용 분석을 장서 개발이나 정보서비스를 위한 기초 자료로 이용하고자 한 연구들은 대부분

일정기간동안 특정 주제 분야의 대표 학술지에 수록된 학술논문에서 인용하고 있는 문헌들을 계량적으로 분석하였다. 이러한 연구들에서는 인용문헌 수, 사용 언어, 주제 분야, 인용문헌의 종류에 따른 분석, 인용문헌이 수록된 학술지의 양적 분석 등이 주로 이루어졌다.

일반적으로 학술논문의 인용 행태를 통하여 장서 개발 및 이용자 참고서비스에 기초 자료로 이용하고자 하는 연구들이 이루어졌다(김홍렬 2005; 유재옥 1993). 이러한 연구들은 주로 특정 주제 분야의 대표 학술지를 선정하여 일정 기간 동안의 인용빈도를 중심으로 인용 행태를 분석하는 연구들이다. 따라서 이러한 연구들은 특정 주제 분야의 인용 행태에 대한 일반적인 양적 데이터는 제공하고 있지만 특정 자관의 인용 행태에서 발생하는 자관특성(locality)을 기술하고 있는 것은 아니라고 할 수 있다. 또 특정 자관의 이용행태나 특정 이용자 집단의 인용 행태를 기술하는 연구들도(고성순, 최상기 2005; 노경란, 한상완 2006) 인용 행태를 양적 데이터로 기술하면서 현상을 설명하는데 중점을 두고 있다고 할 수 있다.

인용 분석에서 대부분의 연구가 사용한 방법은 단순 인용빈도를 측정하는 것이다. 그러나 이는 한 학술지가 인용된 빈도만을 알려줄 뿐 인용된 문헌들간의 관계 정보는 전달해 주지 못한다. 그러나 인용 분석은 인용을 주고받는 개체들에 의한 네트워크를 통해 이들간의 관계를 파악하는 것이 가능한 방법이다. 인용 네트워크의 구조적 속성과 그 네트워크 내에서 학술지가 가지는 역할 및 위치에 의한 중요성은 단순 피인용빈도 이상의 관계에서 나오는 정보를 전달해 줄 수 있다.

인용된 문헌들간의 관계를 반영하는 방법으로는 동시인용 네트워크를 작성하여 다양한 네트워크 지수들을 통해 구조적 속성을 파악하는 방법이 있다.

인용 네트워크 생성 시 자주 사용되는 동시인용은 두 편의 인용 문헌이 후에 출판된 제 3의 문헌 속에 동시에 인용되었을 때 이 두 편의 문헌이 주제적으로 관계가 있다고 보는 방법이다(정영미 2005). 따라서 동시인용빈도는 단순 인용빈도보다 인용된 문헌들 간의 관계를 고려한 인용빈도라고 할 수 있다. 특히 학제적 분야의 경우 다양한 학문 분야가 공존하며 연구되기 때문에 특정 학술지가 자주 인용되는 것뿐만 아니라 어떤 주제 분야의 학술지와 함께 인용되는지도 고려되어야 한다고 볼 수 있다. 이러한 측면에서 학제적 분야의 인용 행태를 분석하는데 있어서 단순인용빈도에 반영되지 않는 인용 학술지간 주제적 관계를 동시인용을 통해 반영할 수 있을 것으로 생각된다.

인용 네트워크 분석에 인용의 빈도뿐만 아니라 위세성(prestige)을 반영하고자 하는 연구들은 대부분 페이지랭크의 개념을 도입하여 학술지를 순위화하고자 한다.

Bollen 등(2006)은 Y-factor라는 지수를 제안하면서 학술지의 인지도(popularity)를 반영하는 ISI의 IF(Impact Factor)와 위세를 반영하는 가중치페이지랭크를 곱하는 방식을 취하였다. 이들은 JCR2003 데이터를 이용하여 인용 네트워크를 작성하고 이 안에서의 IF, 가중치페이지랭크, Y-factor에 의한 상위에 위치한 학술지 순위의 변화를 상관관계 분석을 통해 살펴보았다. 분석 결과 IF와 가중치페이지랭크간의 피어슨 상관계수가 0.5 이하로 낮게 나타

났다. 즉, 인지도지향적인 학술지들과 위세지향적인 학술지들이 각기 다름을 보여주었으며 이는 물리학, 컴퓨터공학, 의학 분야에서 확인되었다. 이 연구의 결과는 학술지 인용 지수를 산출함에 있어서 단순 인용빈도뿐만 아니라 네트워크상의 위세성을 반영해야 한다는 것을 보여주는 사례라고 할 수 있다. Y-factor에 대한 긍정적인 평가(Ball 2006; Hascall & Hanson 2008)와 함께 피부의학 분야에 대한 후속 연구도 발표되었다(Robert et al. 2007). 그러나 Y-factor는 빈도지향지수인 JIF와 위세지향지수인 PageRank를 단순곱으로 처리함으로써 어느 한쪽이 가지는 우세성을 상쇄시켜버리는 단점을 가지고 있다.

3. 네트워크의 구조적 분석

네트워크의 연결 구조를 반영하는 여러 네트워크 지수는 주로 위세성이나 중심성과 같은 성질을 측정하는 것을 목표로 한다.

네트워크 지수가 반영하는 가장 대표적인 개념으로는 중심성(centrality, 또는 중앙성)이 있다. 중심성 척도는 어떤 노드가 그 네트워크 내에서 집중적인 연결을 받으며 중심에 위치하는 정도를 반영하는 것으로서 다양한 공식들이 제안되어 왔다. Freeman(1979)에 의해 정리된 중심성 척도는 인용 네트워크 내에서의 학술지의 상대적 위치를 나타내주는 척도라고 할 수 있다. 이는 크게 근접중심성(closeness centrality, 인접 중앙성으로도 번역)과 매개중심성(betweenness centrality, 사이중앙성으로도 번역)으로 나뉘어 살펴볼 수 있다.

근접중심성은 한 노드가 다른 노드들과 평균적으로 얼마나 가까이 있는지를 측정해준다. 즉, 여러 학술지들과 인용빈도나 동시인용빈도가 높을수록 근접중심성은 높은 값을 갖게 되므로, 학제적 분야 내에서 자주 이용되는 학술지의 상대적 위치를 살펴보기에 적절한 척도라고 할 수 있다.

근접중심성은 한 노드가 다른 노드들과 갖는 최단 경로 거리의 합의 역수이다. 즉, 한 노드가 다른 노드들과 갖는 최단 경로 거리의 합이 작을수록 근접중심성값은 커지게 되고, 이 노드는 네트워크 전체에서의 중앙에 위치한다고 할 수 있다(Wasserman & Faust 1994).

매개중심성은 한 노드와 다른 노드들 사이에 최단경로 상에 위치하는 정도를 측정하는 중심성 척도이다. 즉, 상이한 집단 간을 연결하는 노드일수록 매개중심성이 높게 나타난다. 따라서 학제적 분야와 같이 여러 주제 분야가 공존하는 학문 분야들 간에서 자주 함께 이용되는 특정 학술지의 상대적 위치를 살펴보기에 적절한 척도라고 할 수 있다.

근접중심성 및 매개중심성이 네트워크 전체 내에서의 역할을 나타내는 전역중심성(global centrality) 지수라고 한다면 연결정도중심성(degree centrality)은 지역중심성(local centrality) 지수라고 할 수 있다. 연결정도중심성은 외향 연결정도와 내향연결정도로 방향성도 나타낼 수 있다. 일반적으로 연결정도중심성은 한 노드의 실제 연결수를 모든 가능한 연결의 수로 나누어 계산해준다.

그러나 중심성 척도는 분석 대상의 수에 비교적 크게 영향을 받으며 Nature지와 같이 여러 학문 분야에서 사용되는 대규모 학술지의

경우 매개중심성이 비정상적으로 높게 나타날 수도 있다는 약점이 있다(Leydesdorff 2007).

근접중심성과 매개중심성, 연결정도중심성은 중심성 척도 중에서도 연결/비연결 여부만을 고려하는 지수들로 연결 정도로 표현되는 가중치를 반영해주지는 못한다. 따라서 이를 보완하기 위해 위세중심성을 사용할 수 있다. 위세중심성은 연결된 노드의 중요성을 고려하며 에이겐벡터와 보나시치 지수(Bonacich Power)가 많이 사용된다.

보나시치 지수는 한 노드의 중심성을 측정하는데 있어서 연결정도에 따라 발생하는 자신의 영향력과 자신에게 연결된 다른 노드들의 영향력을 모두 고려하는 방식을 취한다. 보나시치 지수는 보나시치가 제안한 위세중심성 지수군(Bonacich 1987)의 하나로써 이는 다음의 공식으로 표현된다.

$$c_i(\alpha, \beta) = \sum_j (\alpha + \beta c_j) R_{ij} \quad -1 < \beta < 1$$

여기서 α 는 자기 자신의 영향력에 대한 가중치이며, β 는 자신과 연결된 다른 노드들의 영향력을 말한다. 여기서 $\beta=0$ 이면 단순 연결정도중심성이 된다.

일반적으로 에이겐벡터와 보나시치 지수는 가중치 대칭 행렬에 사용가능하다.

이와 같은 사회네트워크 지수들을 학술지 인용 네트워크에 적용하고자 하는 연구들은 최근 들어 주로 페이지랭크로 대변되는 위세중심성을 이용하고자 하고 있다. 이는 단순 인용빈도가 아닌 인용의 위세성을 반영하고자 하는 것이다. 즉, 영향력 있는 학술지를 인용하는 것이 그렇지 않은 학술지를 인용하는 것보다 더 가

중치를 받도록 하는 것이다.

가중치 네트워크를 위한 지수들로 이재윤(2006b)은 최근접중심성, 평균연관성, 평균프로파일연관성, 삼각매개중심성을 제안하였다. 최근접중심성은 지역중심성을 나타내주는 지수이며 평균 프로파일연관성과 삼각매개중심성은 전역중심성을 나타내주는 지수이다.

최근접중심성(nearest centrality)은 가중 네트워크를 최근접이웃 그래프(이재윤 2006a)로 변환한 다음 내향연결정도중심성을 산출하는 방식을 사용하며 가중 네트워크에서 연결정도중심성의 대안 개념으로 제시되었다.

평균연관성(mean associations)은 한 노드가 다른 노드들과 전체적으로 얼마나 근접한 위치에 있는가를 측정하는 것이므로 전역 중심성을 나타내는 지수로 가중 네트워크에서 근접중심성의 대안 척도로 제안되고 있다.

평균프로파일연관성(mean profile associations)은 가중 네트워크의 전역 중심성 척도로 제안된 것으로 다른 영역 노드들과의 연관성이 높고 낮은 패턴을 이용한다.

마지막으로 삼각매개중심성(triangle betweenness centrality)은 3자간 관계(triad)에 의한 지수로 한 노드가 다른 노드들 사이를 결속시켜주는 능력을 측정하는 지수이다. 이는 다음에 설명될 제약성(constraint) 지수와 유사한 개념이다.

중심성과 함께 네트워크의 구조적 속성을 파악하는 또 다른 개념은 네트워크 내의 '자본의 분배'라는 개념에 의한 설명이다. 이 중 제약성은 연결망의 구조적 위치를 측정하는 개념 중 하나인 '구조적 공백(structural hole)'이라는 개념을 계량적으로 표시하는 하나의 지수이다.

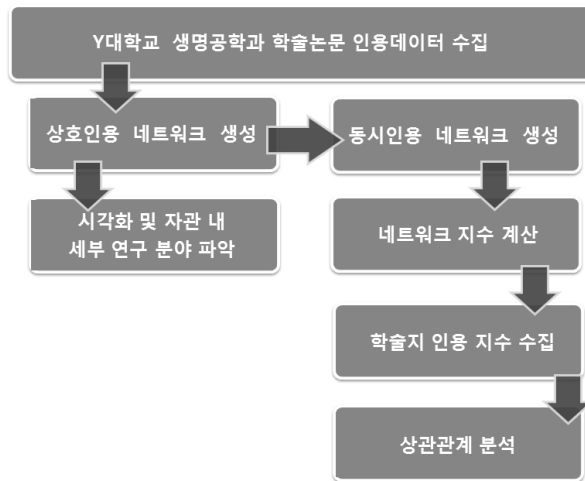
사회 네트워크 분석에서 얘기하는 구조적 공백은 사회적 자본의 개념과 연결되어 있는 것으로, 조직 내 네트워크에서 구조적 공백에 위치할수록 일반적으로 좋은 성과를 내고 양질의 정보를 획득할 수 있다고 한다(Burt 1992; 김용학 2007).

그러나 구조적 공백이라는 네트워크 지수는 존재하지 않으므로 흔히 비 중복성과 낮은 제약성의 개념으로 측정하고 있다(Burt 1992). 일반적으로 중복된 관계에 놓이지 않을 때에 제약성이 낮으며 이러한 낮은 제약성은 구조적 공백을 연결한 경우에 생겨난다고 보고 있다(김용학 2007).

제약성은 한 노드와 다른 노드 그리고 이들 간의 제 3자간의 관계를 고려하는 지수로서, 나(i)와 상대(j) 뿐 아니라 제3자(q)에 대한 나와 상대의 관계를 모두 고려하고 있다(표 2 참조). 일반적으로 제약성 값이 낮으면 구조적 공백이라고 보며 좋은 성과를 낸다고 하므로 이를 인용네트워크에 적용할 경우 인용이 많이 되는 학술지가 좋은 성과를 낸다고 볼 수 있으며 이들의 제약성 값이 낮을 것으로 예상된다.

4. 연구방법

이 연구의 개요는 <그림 1>과 같다. Y대학교 생명공학과에서 2006년부터 2007년에 발표된 학술논문 및 이들의 인용데이터를 ISI Web of Science에서 검색하였다. 검색 결과 내 학술논문에서 피인용되고 있는 학술지를 확인한 후 이를 이용하여 상호인용 네트워크 및 동시인용 네트워크를 작성하였다.



〈그림 1〉 연구 개요도

상호인용 네트워크를 이용하여 자관의 학술지 이용 및 연구 경향을 살펴보고, 동시인용 네트워크를 이용하여 네트워크 지수를 산출하여 구조적 속성을 살펴보고자 하였다.

자관 동시인용 네트워크를 이용하여 네트워크 지수를 산출한 후, 이 네트워크를 구성하는 학술지의 일반적 인용 지수를 수집하였다. 수집 대상으로는 JCR2006과 JCR2007, EigenFactor.org, Journal-Ranking.com, SCImagojr.com을 이용하였다.

이렇게 수집한 일반적 인용 지수들과 자관 동시인용 네트워크 지수들 각각의 관계를 살펴보았다.

4.1 분석 대상 데이터 수집 및 인용 행렬 생성

자관의 학술지 이용을 파악하기 위하여 Web of Science를 검색하여 2006년과 2007년 Y대학교 생명공학과 소속 교수들이 발표한 학술논문 122건의 인용 데이터를 수집하였다. 122개의 논문 중 55개는 2006년, 67개는 2007년에 발표된 것이다. 분석 대상 데이터의 평균 피인용 학술지 수는 기간별로 〈표 1〉과 같다.

수집된 데이터의 모든 학술지 약어명을 완전 명으로 변환한 후에 피인용빈도 4이상인 172개 학술지만을 분석대상으로 하여 상호인용 네트워크 및 동시인용 네트워크를 작성하였다.

〈표 1〉 분석 대상 데이터

	전 체	2006	2007
발표 논문 수	122	55	67
논문당 평균 피인용 학술지 수	20.83	19.14	22.22

학술지의 상호인용 네트워크는 시각화를 통해 자관의 연구경향 및 세부주제 분야를 살펴 보는데 이용하였고 동시인용 네트워크는 네트워크지수 계산을 통해 자관의 이용특성을 계량화하고 이를 일반적인 학술지 영향력 지수들과 비교하고자 하였다.

4.2 네트워크 지수 산출

자관의 이용특성을 알아보고자 자관의 동시인용 네트워크를 이용하여 단순피인용빈도, 단순피인용학술지수, 최근집중심성, 평균연관성, 평균프로파일연관성, 삼각매개중심성, 가중치 페이지랭크(2가지), 연결정도중심성, 근접중심성, 매개중심성, 보나시치 위세중심성, 에이겐벡터, 제약성, 결속계수(2가지)의 16가지 지수를 산출하였다.

연결정도중심성, 근접중심성, 매개중심성, 결속계수는 네트워크 분석 프로그램인 Pajek (De Nooy et al 2005)을 사용하여 계산하였으며, 위세중심성, 에이겐벡터, 제약성은 UCINET (Borgatti, et al. 2002)을 사용하여 계산하였다.

자관의 인용 네트워크를 이용한 각 지수별 공식 및 표시기호는 <표 2>와 같다.

4.3 일반적 학술지 영향력 지수 수집

산출된 자관 동시인용 네트워크 지수값들을 일반적 학술지의 영향력 지수들과 비교하였다. 비교에는 JCR2006과 JCR2007의 평균, EigenFactor.org, Journal-Ranking.com, SCImago Journal & Country Rank(SCImagojr.com)의 데이터를 이용하였다. JCR, EigenFactor.org, Journal

-Ranking.com은 모두 Web of Science의 데이터를 이용해 학술지의 영향력 지수들을 제시하고 있는 반면에 SCImagojr.com은 SCOPUS의 데이터를 이용하여 학술지 영향력 지수들을 제시하고 있다.

<표 3>은 이 연구에서 비교분석에 사용한 일반적 학술지 영향력 지수들이다.

1) JCR(Journal Citation Report)

자관 동시인용 네트워크를 구성하는 총 172개 학술지를 대상으로 Web of Science에서 제공하는 JCR2006과 JCR2007의 총인용빈도(cites), IF(JIF), 즉시성 지수(Immediacy index: IM), 논문수(articles)를 검색해 내었다. 검색결과 172개 학술지 중 163개 학술지의 데이터를 확보할 수 있었으며 9개의 학술지는 JCR에 포함되어 있지 않았다.

2) EigenFactor.org

EigenFactor는 학술지의 영향력을 지수화할 때 인용 네트워크의 전체 구조를 반영하는 Google의 PageRank나 위세중심성과 유사한 EigenFactor 알고리즘을 사용하고 있다. 즉 영향력이 높은 학술지와 연결된 학술지의 영향력이 높게 계산된다. 또한 이 알고리즘에서는 학술적 인용 네트워크 내에서 어떤 학술지를 이용한 시간이 그 학술지의 중요성을 나타내준다고 보고 있다.

EigenFactor 지수는 정규화된 상호인용 행렬에 학술지의 영향력 가중치를 이용한 에이겐벡터값을 결합하는 위세중심성 방식의 인용 지수이다. 정규화된 상호인용 행렬에서 어떤 학술지와도 인용을 주고받지 않는 학술지(damping

〈표 2〉 자관 인용네트워크 분석 지수

유형	지수	기호	공식
피인용빈도 지수	피인용빈도	cFrq	-
	피인용학술지수	cJnum	-
가중치 페이지랭크 (Xing & Ghorbani 2004)	wPR(자기인용 포함)	wPR(s)	$(1-d) + d \sum_{v \in B(u)} PR(v) W_{(v,u)}^{in} W_{(v,u)}^{out}$ $W_{(v,u)}^{in} = \frac{I_u}{\sum_{p \in R(v)} I_p} : \text{inlink popularity}$ $W_{(v,u)}^{out} = \frac{O_u}{\sum_{p \in R(v)} O_p} : \text{outlink popularity}$
	wPR(자기인용 미포함)	wPR(x)	
중심성 지수	최근접중심성	Cn	설명으로 대체
	평균연관성	Cm	설명으로 대체
	평균프로파일연관성	Cmp	설명으로 대체
	삼각매개중심성	Ctb	설명으로 대체
	근접중심성	Close	$[\sum_{j=1}^g d(n_i, n_j)]^{-1}$ $d(n_i, n_j): \text{노드 } i, j \text{ 간의 최단 경로 거리}$
	매개중심성	Between	$\frac{\sum_i \sum_j g_{ij}(m)}{\frac{g_{ij}}{(N^2 - 3N + 2)/2}}, i \neq j, i < j$ $g_{ij}(m): \text{노드 } i, j \text{ 간의 최단 경로 내 노드 } m \text{ 이 위치하는 경우의 수}$
위세 중심성 지수	에이겐벡터	Eigen	$P_i = \alpha \sum_{j=1}^{N-1} p_j Z_{ji}$ $Z_{ji}: \text{모든 노드 } j \text{로부터 노드 } i \text{로 가는 관계의 수}$ $P_i: \text{노드 } i \text{의 power/중심성}$
	보나시치 지수	Bona	$P_i(\alpha, \beta) = \alpha \sum_j p_j (\alpha + \beta p_i) Z_{ji} \quad (\alpha + \beta = 1)$ $Z_{ji}: \text{모든 노드 } j \text{로부터 노드 } i \text{로 가는 관계의 수}$ $P_i: \text{노드 } i \text{의 power/중심성}$
구조적 공백 지수	제약성	Constr	$C_{ij} = (p_{ij} + \sum_{q=1}^n p_{iq} p_{qj})^2, q \neq i, j$ $p_{i,j}: i \text{가 } j \text{에게 투자하는 자본의 양}$
결속 계수	결속계수 1	CC1	$\frac{2 \times E(G1(v)) }{k(v) \times (k(v) - 1)}$ $K(v): v \text{의 이웃노드 수}$ $ E(G1(v)) : \text{노드 } v \text{의 1단계 이웃노드들간의 연결선 수}$
	결속계수 2	CC2	$\frac{ E(G1(v)) }{ E(G2(v)) }$ $ E(G1(v)) : \text{노드 } v \text{의 1단계 이웃노드들간의 연결선 수}$ $ E(G2(v)) : \text{노드 } v \text{의 1, 2단계 이웃노드들간의 연결선 수}$

〈표 3〉 일반적 학술지 영향력 지수

유형	지수	기호	공식
EigenFactor.org	EigenFactor	EF	$100 \times \frac{H\pi^*}{\sum_i [H\pi^*]_i}$ H: 정규화된 상호인용 행렬 π^* : 에이겐벡터, 학술지 영향력 가중치
	Article Influence	AI	$AI_i = 0.01 \frac{EF_i}{\alpha_i}$ EF_i : 학술지 i의 eigenfactor 값 α_i : 정규화된 아티클 벡터의 i 번째 값
Journal-Ranking.com	Journal Influence Index	JII	비공개
	Paper Influence Index	PII	$1000 \times \frac{JII_i}{Articles_i}$ $Articles_i$: 학술지 i의 아티클 수
Scimago Journal & Country Rank	SJR	SJR	$SJR_i = \frac{SR_i}{Art_i}$ $SR_i = \frac{(1-d-e)}{N} + e \cdot \frac{Art_i}{\sum_{j=1}^N Art_j} + d \cdot \sum_{j=1}^N \frac{C_{ij} \cdot SR_j}{C_j} \cdot \frac{1 - (\sum_{k \in D_n} SR_k)}{\sum_{h=1}^N \sum_{k=1}^N \frac{C_{hk} \cdot SR_k}{C_k}} + d \cdot [\sum_{k \in D_n} SR_k] \cdot \frac{Art_i}{\sum_{j=1}^N Art_j}$
	h-index	hindex	설명으로 대체

vector)의 경우에는 그 학술지가 5년 동안 출간한 총 논문수를 데이터베이스내의 5년 동안 출간된 모든 논문수로 나눈 '논문 벡터(article vector)'를 이용하여 보정해준다. 가중치벡터인 에이겐벡터값은 학술지영향력 가중치로써 상호인용 행렬내에서 각 학술지에 투자한 시간에 대응되는 개념이다.

Article Influence 지수는 위에서 계산된 EigenFactor 지수값을 정규화된 학술논문 값으로 나누어준다. 이 사이트에서는 Eigenfactor 지수와 Article Influence 지수를 제시하고 있다. 인용학술지는 2006년을 기준으로 하고 있으며 피인용 기간은 2001년부터 2005년의 최근 5년간의 데이터를 이용한다. 자세한 공식은 위의 〈표 3〉과 같다.

3) Journal-Ranking.com

Red Jasper 사에서 제공하는 학술지 순위 목록으로 현재 2005년까지의 인용 분석에 의한 학술지 순위 목록을 제공하고 있다. 2008년 9월 현재 전체 학문영역을 대상으로 했을 때 총 8,023개 학술지의 순위를 제공하고 있다. 여기에서 제공하는 지수는 JII, PII, B2, B4, B6의 총 5가지이며, PII에 따른 순위도 제공한다.

JII(Journal Influence Index)는 학술지의 영향력 지수를 위세중심성인 에이겐벡터를 적용하여 계산한 것으로 이 값이 클수록 영향력이 높은 것이다. 자세한 공식은 현재 공개되어 있지 않다.

PII(Paper Influence Index)는 JII를 학술논문 수로 나눈 후에 1000을 곱한 값이다.

B2는 JCR의 영향력지수(IF)와 동일한 것이

며, B4와 B6는 이의 변형 형태들로 B4는 2005년 이전 4년간의 인용빈도를, B6는 2005년 이전 6년간의 인용빈도를 고려하여 IF를 계산한다. B2, B4, B6는 모두 2005년을 기준으로 계산되어 있다.

4) Scimago Journal & Country Rank(SCImagojr.com)

Scimago는 스페인 University of Granada의 'Consejo Superior de Investigaciones Cientificas(CSIC)'내의 연구 집단으로 정보분석 및 시각화를 연구하고 있다. 이 연구 집단은 SCOPUS의 데이터베이스를 이용하여 SCImagojr.com 서비스를 운영하면서 학술지 및 국가별 연구 지수(scientific indicator)값을 제공하고 있다. 여기서 제안한 SJR 지수는 페이지랭크에 기반을 둔 방식을 취하면서 동시에 이를 학술논문수로 나눈 공식이다. SJR의 분자가 되는 페이지랭크기반의 영향 지수 계산 공식은 다음과 같다.

$$SR_i = \frac{(1-d-e)}{N} + e \cdot \frac{Art_i}{\sum_{j=1}^N Art_j} + d \cdot \sum_{j=1}^N \frac{C_{ji} \cdot SR_j}{C_j} \cdot \frac{1 - \left(\sum_{k \in D_n} SR_k \right)}{\sum_{h=1}^N \sum_{k=1}^N \frac{C_{kh} \cdot SR_k}{C_k}} + d \cdot \left[\sum_{k \in D_n} SR_k \right] \cdot \frac{Art_i}{\sum_{j=1}^N Art_j}$$

여기서 C_{ji} 는 학술지 j에서 학술지 i로의 인용 수를, C_j 는 학술지 j의 인용수를 말하며, d와 e는 보통 상수로 d=0.85, e=0.10로 설정해준다. N는 학술지 수를 D_n 은 dangling-nodes를 의미하며 Art_j 는 학술지 j의 총논문수를 의미한다.

SCImagojr.com에서는 또한 h-index 지수값

도 제공해 주고 있다. h-index는 Hirsch(2005)가 제안한 지수로 한 연구자의 연구 성과와 연구 영향력을 함께 나타내 줄 수 있는 지수이다. Hirsch(2005)의 정의에 따르면 h-index는 “어떤 연구자의 h-index가 h라는 것은 그가 출간한 N_p 개의 연구논문 중에서 최소한 h회 이상의 인용을 받고 있는 연구논문의 수가 h개이고, (N_p-h) 개의 논문은 최대한 h 회의 인용을 받고 있는 것이다”라고 해석할 수 있다. 예를 들면, 어떤 연구자의 h-index가 5라면 그 연구자가 다른 연구자들로부터 5회 이상의 인용을 받는 연구논문수가 5개인 것을 의미한다.

SJR과 h-index도 JCR의 지수들과 마찬가지로 2006년과 2007년의 지수값을 평균하여 사용하였다.

4.4 상관관계 분석

Y대학교 생명공학과와 학술지 인용정보로부터 도출된 동시인용 네트워크 지수들과, JCR, EigenFactor, Journal-Rankin, SCImagoJR 등의 여러 서비스에서 제공하는 지수들간의 상관관계를 SPSS를 이용하여 분석하였다.

5. 분석

5.1 자관 상호인용 네트워크의 구조적 속성

<표 4>는 상호인용 네트워크 전체의 구조를 요약한 것이다. 2006-2007년의 Y대학교 생명공학과와 상호인용 네트워크는 총 924개 학술

지를 포함하였으며 방향성이 있는 네트워크로 상호인용에 의한 연결수는 총 45,404개이다. 이 중 인용빈도가 1인 연결이 38,200개이며, 7,204개가 빈도 2이상인 인용이었다. 그리고 자기인용은 924개 모든 학술지에서 이루어졌다.

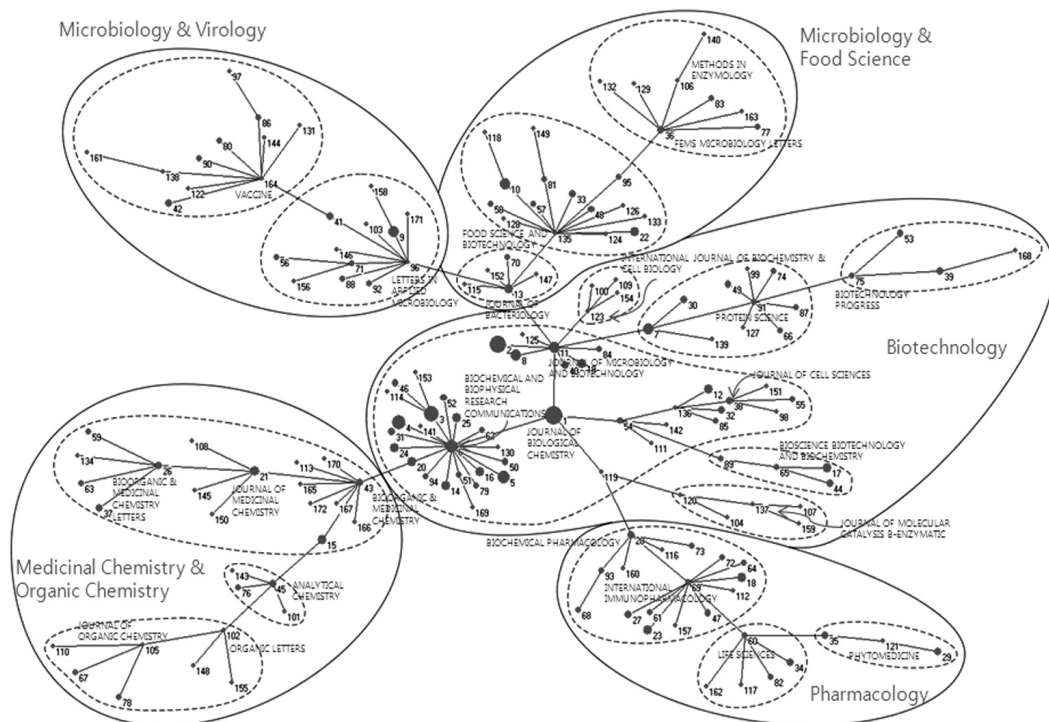
이 상호인용 네트워크의 밀도는 약 0.053으로 낮으며, 인용 및 피인용을 모두 포함하여 계

산한 중심화 지수는 약 0.60이다. 중심화 지수는 한 네트워크가 전체적으로 얼마나 중앙집중적인 구조를 가지고 있는지를 알려주는 지수이다(김용학 2007). 따라서 이 상호인용 네트워크는 연결정도를 통해 볼 때 59.9% 정도 중앙집중적으로 결속되어 있다고 볼 수 있다.

〈그림 2〉는 Y대학교 생명공학과와 2006-2007

〈표 4〉 자관 상호인용 네트워크의 구조적 속성

총 학술지 수	924
연결 수	45,404
자기인용(회귀) 수	924
밀도(회귀 포함)	0.0531
연결정도 중심화 지수(인용/피인용 모두 포함)	0.5992



〈그림 2〉 CBnet-wave로 표현한 Y대학교 생명공학과와 2006-2007

년 데이터를 이용한 학술지간의 상호인용 네트워크를 시각화한 것이다. 924개 학술지 중 피인용빈도가 4이상인 172개 학술지만을 이용하여 맵을 작성하였다. 네트워크 생성방법으로는 CBnet-wave(김관준, 이재윤 2007; 이재윤 2007a) 방식을 이용하였고, 학술지 군집은 CBnet-wave 방식과 부합되는 집단내평균연결 기법으로 구분하였다.

〈그림 2〉에서 각 노드는 학술지를 나타내며 노드의 크기는 자기인용을 포함한 피인용빈도를 반영하였다. 따라서 노드의 크기가 클수록 그 노드에 해당하는 학술지의 피인용빈도가 높음을 나타낸다. 학술지명이 표기된 노드는 '허브(주요학술지)'를 나타내는데, 이 네트워크에서의 허브는 학내 연구자들이 학술논문을 많이 발표하였거나 많이 인용한 학술지이다. 즉, 노드의 크기가 큰 허브 노드는 피인용빈도가 높은 학술지이고 크기가 작은 허브 노드는 발표논문이 많은 학술지를 나타낸다.

자관의 상호인용 네트워크를 살펴보면 크게 5개의 연구 영역이 비교적 잘 구분되어 드러남을 알 수 있다. 5개 연구 영역의 주제를 살펴보면 중앙에 위치한 'Biotechnology' 영역을 중심으로 'Microbiology & Virology', 'Microbiology & Food Science', 'Medicinal Chemistry & Organic Chemistry', 'Pharmacology' 분야로 구분된다.

그림의 중앙에 위치한 가장 큰 노드인 *Journal of Biological Chemistry*를 중심으로 하는 'Biotechnology' 영역은 Y대학교가 세포학을 포함하는 생화학분야의 연구가 활발함을 보여준다. 이 영역에 속한 *Journal of Biological Chemistry*는 피인용빈도가 69회로 자관 인용 데이터 중

에서 가장 높은 피인용빈도를 보였다. 이 연구 영역에서는 허브 노드는 아니지만 피인용빈도가 높은 학술지들이 눈에 띈다. *Nature*(2번 노드, 피인용빈도 50), *Proceedings of the National Academy of Sciences of the USA*(3번 노드, 피인용빈도 48), *Science*(4번 노드, 피인용빈도 35), *Cancer Research*(5번 노드, 피인용빈도 30), *Biochemical and Biophysical Research Communications*(6번 노드, 피인용빈도 29)가 그것이다. 즉, 이 학술지들은 자관 내 피인용빈도는 높으나 발표 논문수는 많지 않은 노드들이다. 이 학술지들은 대부분 과학일반 내용을 다루는 권위가 높은 학술지이거나 *Cancer Research*와 *Biochemical and Biophysical Research Communications*처럼 생명공학의 연구분야 전반에서 관련있는 연구 분야의 학술지들이다.

왼쪽 위에 위치한 'Microbiology & Virology' 분야의 허브로는 *Vaccine*과 *Letters in Applied Microbiology*를 들 수 있다. 세부적으로 바이러스학에서는 *Vaccine*(피인용빈도 4)이, 미생물학에서는 *Letters in Applied Microbiology*(피인용빈도 5)가 허브임을 알 수 있다.

오른쪽 위에 위치한 'Microbiology & Food Science' 분야의 허브로는 *Journal of Bacteriology*(피인용빈도 18), *Food Science And Biotechnology*(피인용빈도 5), *Fems Microbiology Letters*(피인용빈도 11), *Methods in Enzymology*(피인용빈도 5)가 확인되었다.

왼쪽 아래에 위치한 'Medicinal Chemistry & Organic Chemistry'분야에서는 의화학분야의 *Bioorganic & Medicinal Chemistry*(피인용빈도 10), *Journal of Medicinal Chemistry*(피인용빈도 16), *Bioorganic & Medicinal*

Chemistry Letters(피인용빈도 13)와 화학분야의 *Analytical Chemistry*(피인용빈도 9), 조직화학분야의 *Organic Letters*(피인용빈도 5), *Journal of Organic Chemistry*(피인용빈도 5)가 허브임을 확인하였다.

오른쪽 아래에 위치한 'Pharmacology' 분야에서는 *Biochemical Pharmacology*(피인용빈도 13), *International Immunopharmacology*(피인용빈도 7), *Life Sciences*(피인용빈도 8), *Phytomedicine*(피인용빈도 5)가 허브 저널로 나타났다.

이처럼 학과 소속 연구자들의 실적에서 도출한 상호인용 네트워크의 분석을 통해서 학내 주요 연구 영역과 영역별 주요 저널을 파악할 수 있었다.

5.2 동시인용 네트워크 지수들 간의 상관관계

5.2.1 자관 동시인용 네트워크 지수들간의 상관관계

자관 동시인용 네트워크 지수들은 크게 가중치지수들과 비가중치지수로 나눌 수 있다. 가중치지수는 동시인용빈도를 지수계산에 반영하는 경우로, 최근접중심성(Cn), 평균연관성(Cm), 평균프로파일연관성(Cmp), 삼각매개중심성(Ctb), 가중치페이지랭크(xPR(s), xPR(x)), 에이겐벡터(Eigen), 위세중심성(Bona), 제약성(Constr) 등이 여기에 해당된다. 비가중치지수는 피인용빈도가 아닌 동시인용 여부만이 지수 계산에 반영되는 경우로, 일반적인 네트워크 지수들인 연결정도중심성(Degree), 근접중심성(Close), 매개중심성(Between), 결속계수(CC1, CC2)

등이 여기에 해당된다.

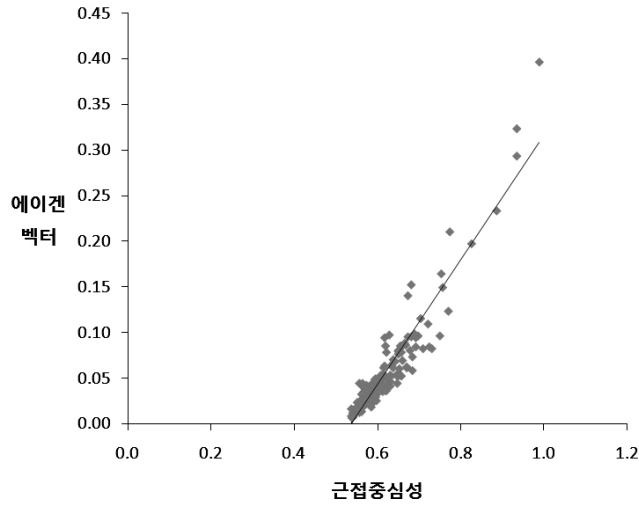
〈표 5〉를 통해 살펴보면 자관 동시인용 네트워크 지수들 간에는 다음과 같은 특성을 갖는다.

첫째, 피인용빈도가 대부분의 네트워크 지수들과 0.6 이상의 높은 상관관계를 보였다. 이는 단순피인용빈도가 동시인용 네트워크의 특성과 유사한 설명력을 가지고 있는 것으로 해석해 볼 수 있다.

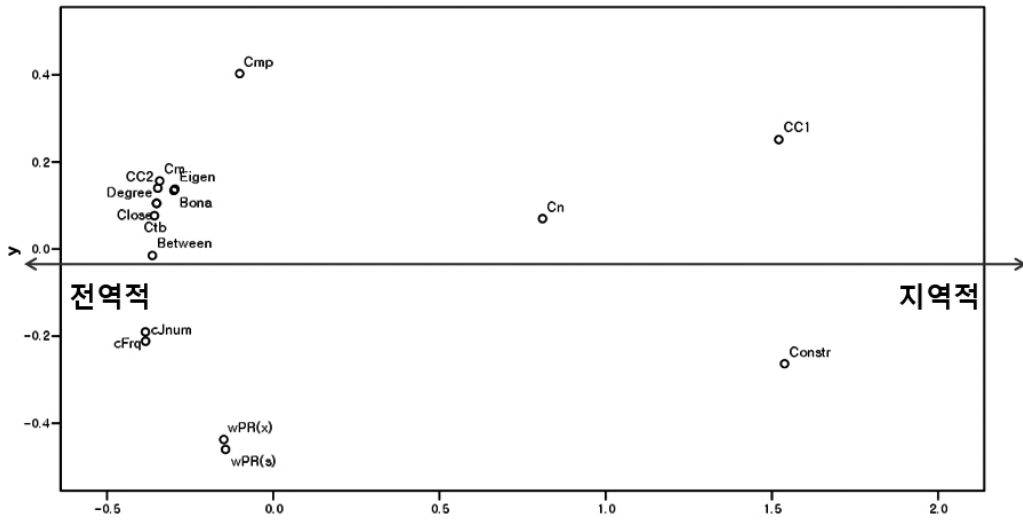
둘째, 피인용빈도가 가중치지수들과 비가중치 지수들간에 모두 높은 상관관계를 보임을 알 수 있다. 이는 피인용빈도에 의한 가중치를 고려하지 않은 구조적 속성 지수만으로도 관계를 파악하는 것도 동시인용네트워크를 설명하는데 의미가 있는 것으로 생각해 볼 수 있다.

예를 들어 에이겐벡터중심성과 근접중심성 값 간의 산포도를 그려보면 〈그림 3〉과 같이 피인용빈도에 의한 가중치를 반영한 에이겐벡터중심성지수와 인용 여부만을 반영하는 근접중심성 지수 사이의 밀접한 관계가 드러난다.

셋째, 자관의 동시인용 네트워크 지수들을 크게 전역적 지수들과 지역적 지수들로 나누어 볼 수 있다. 〈그림 4〉는 동시인용 네트워크 지수들을 이용해 PROXSCAL(이재윤 2007b) 방식으로 MDS 맵을 작성한 것이다. 이를 살펴 보면 '전역적-지역적'의 축을 확인할 수 있다. 삼각매개중심성(Ctb), 가중치페이지랭크(xPR(s)) 등의 맵 상의 좌측에 위치하는 지수들은 네트워크의 전역적인 특성을 반영하는 지수들이다. 그리고 결속계수1(CC1), 제약성(Constraint)과 같은 맵상의 우측에 위치하는 지수들은 네트워크의 지역적인 특성을 반영하는 지수들이다. 따라서 동시인용 네트워크의 전체적 특성을 파악하는데는 전역적인 설명력을 가진 지수들을 이용하는



〈그림 3〉 근접중심성과 에이겐벡터의 관계



〈그림 4〉 자관 동시인용 네트워크 지수들의 MDS 맵(PROXSCAL)

것이 바람직할 것으로 생각된다.

5.2.2 일반적 인용 지수들간의 상관관계

일반적 인용지수들간의 관계에서는 다음과 같은 특징을 보였다.

첫째, 일반적 인용 지수들간에는 대체적으로 모두 유의미한 상관관계를 보였다. 논문수(articles)는 다른 인용 지수들과 유의미한 상관관계가 적게 나타났지만, 그 외의 지수들간에는 모두 유의미한 상관관계를 보였다.

〈표 5〉 자관 인용 네트워크 지수들간의 상관관계(Spearman, 2-tailed)

	cFrq	cJnum	wPR(s)	wPR(x)	Ctb	Cm	Cn	Cmp	Constr	Eigen	Bona	Degree	Close	Between	CC1	CC2	
cFrq	1	0.935**	0.737**	0.732**	0.827**	0.772**	-0.127	0.470**	-0.652**	0.817**	0.825**	0.839**	0.838**	0.855**	-0.801**	0.808**	
cJnum		1	0.803**	0.808**	0.831**	0.783**	-0.142	0.490**	-0.674**	0.809**	0.817**	0.843**	0.842**	0.840**	-0.766**	0.817**	
wPR(s)			1	0.980**	0.555**	0.528**	-0.107	0.242**	-0.375**	0.571**	0.578**	0.556**	0.553**	0.603**	-0.623**	0.522**	
wPR(x)				1	0.571**	0.542**	-0.128	0.270**	-0.396**	0.588**	0.595**	0.573**	0.570**	0.613**	-0.615**	0.541**	
Ctb					1	0.943**	-0.146	0.708**	-0.875**	0.909**	0.915**	0.996**	0.996**	0.956**	-0.795**	0.984**	
Cm						1	0.016	0.786**	-0.742**	0.937**	0.940**	0.948**	0.947**	0.847**	-0.633**	0.964**	
Cn							1	-0.037	0.220**	-0.082	-0.084	-0.144	-0.144	-0.197**	0.230**	-0.103	
Cmp								1	-0.470**	0.847**	0.841**	0.728**	0.727**	0.529**	-0.254**	0.783**	
Constr									1	-0.649**	-0.659**	-0.868**	-0.867**	-0.884**	0.784**	-0.834**	
Eigen										1	1.000**	0.923**	0.923**	0.806**	-0.593**	0.940**	
Bona											1	0.929**	0.928**	0.814**	-0.604**	0.944**	
Degree												1	1.000**	0.946**	-0.774**	0.992**	
Close													1	0.945**	-0.771**	0.992**	
Between														1	-0.921**	0.903**	
CC1															1	-0.699**	
CC2																1	-0.699**

** p < .01

〈표 6〉 일반적 인용 지수들간의 상관관계(Spearman, 2-tailed)

articles	cites	JIF	IM	JII	PII	B2	B4	B6	EF	AI	SJR	hindex
articles	1	0.712**	0.073	0.338**	-0.157*	-0.115	-0.13	-0.122	0.415**	-0.197*	-0.039	0.275**
cites	0.712**	1	0.507**	0.493**	0.420**	0.337**	0.329**	0.367**	0.795**	0.366**	0.422**	0.694**
JIF	-0.002	0.507**	1	0.857**	0.682**	0.694**	0.687**	0.690**	0.549**	0.842**	0.805**	0.703**
IM	0.073	0.493**	0.857**	1	0.584**	0.698**	0.687**	0.695**	0.661**	0.849**	0.827**	0.765**
JII	0.338**	0.420**	0.477**	0.584**	1	0.810**	0.708**	0.732**	0.878**	0.628**	0.585**	0.790**
PII	-0.157*	0.420**	0.682**	0.726**	0.810**	1	0.910**	0.934**	0.656**	0.851**	0.779**	0.741**
B2	-0.115	0.337**	0.694**	0.910**	0.910**	0.910**	1	0.994**	0.596**	0.821**	0.730**	0.672**
B4	-0.130	0.329**	0.687**	0.708**	0.994**	0.994**	0.994**	1	0.583**	0.821**	0.726**	0.672**
B6	-0.122	0.367**	0.695**	0.934**	0.934**	0.934**	0.934**	0.994**	1	0.829**	0.730**	0.696**
EF	0.415**	0.795**	0.661**	0.656**	0.656**	0.596**	0.583**	0.607**	1	0.700**	0.650**	0.867**
AI	-0.197*	0.366**	0.849**	0.851**	0.851**	0.821**	0.821**	0.829**	0.700**	1	0.895**	0.789**
SJR	-0.039	0.422**	0.805**	0.585**	0.779**	0.730**	0.726**	0.730**	0.650**	0.895**	1	0.872**
hindex	0.275**	0.694**	0.703**	0.765**	0.741**	0.672**	0.672**	0.696**	0.867**	0.789**	0.872**	1

* p<.05, ** p<.01

또 페이지랭크 방식을 사용하는 SJR와 Eigen Factor간의 상관계수는 SJR과 JIF간의 상관계수보다 낮았는데 이는 SJR과 EigenFactor가 damping factor(다른 노드들과 주고받는 링크가 전혀 없는 노드)를 처리하는 방식이 다르기 때문인 것으로 생각된다. SJR에서는 이를 그대로 페이지랭크 계산에 포함시키지만, EigenFactor에서는 이를 정규화하는 방식을 사용하고 있다.

둘째, 지수가 평균지향지수인지 합산지향지수인지에 따라서 유사한 유형의 지수들끼리 높은 상관관계를 보이는 것을 알 수 있었다.

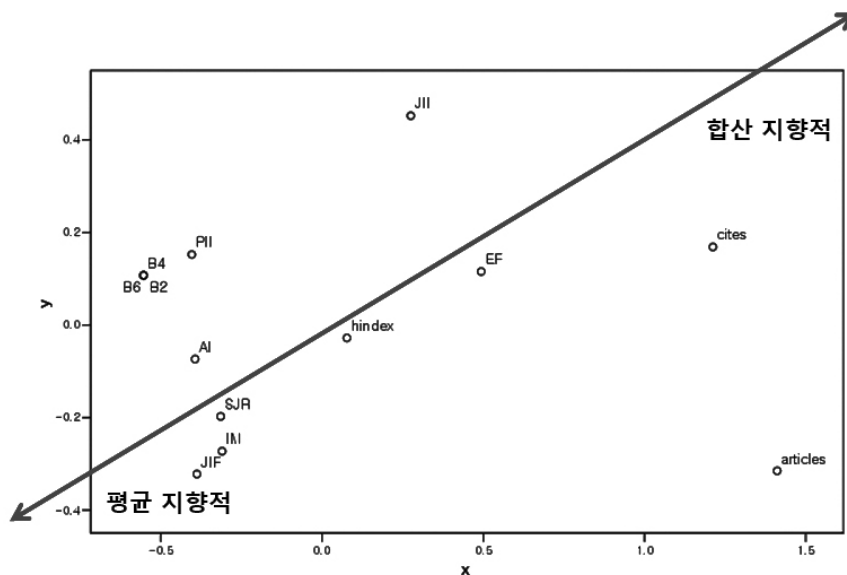
평균지향지수는 피인용빈도와 함께 학술지에서 발표된 학술논문의 수를 공식의 분모에 반영하는 지수들로 JIF, 즉시성 지수(IM), JII, EigenFactor(EF), AI, SJR 등이 여기에 속한다. 반면에 합산지향지수는 피인용빈도 또는 학술논문수만을 반영하는 지수들로 학술논문수(articles), 총인용빈도(cites) 등이 여기에 속

한다.

h-index는 평균지향지수와 합산지향지수의 중간적인 성격을 가졌으며 실제 상관관계분석에서도 이 두 유형의 지수들과 모두 고르게 높은 상관관계를 보였다. JIF는 학술논문수나 총인용빈도와는 낮은 상관관계를 보였으나, 즉시성 지수와는 0.8 이상의 높은 상관관계를 보였다. 또한 자신과 유사한 평균지향지수들과 대체적으로 높은 상관관계를 보이고 있다.

〈그림 5〉는 일반적 인용 지수들을 이용해 PROXSCAL(이재운 2007b) 방식으로 MDS 맵을 작성한 것이다. 이를 살펴보면 ‘평균지향-합산지향’의 축을 확인할 수 있다.

축의 좌측에 가까울수록 평균지향적 지수들로 JIF, SJR 등이 여기에 위치하며, 축의 우측에 가까울수록 합산지향적 지수들로 총인용빈도와 학술논문수 등이 위치하고 있는 것을 알 수 있다. 그리고 h-index는 좌표상 중앙에 위



〈그림 5〉 일반적 인용 지수들의 MDS 맵(PROXSCAL)

치하는 것을 알 수 있다.

또한 SJR 공식을 살펴보면 분모에 학술지의 규모(학술논문 수)를 반영한 JIF와 유사한 평균지향 지수임을 알 수 있다. 따라서 이들 간에도 높은 상관관계를 보이고 MDS 맵 상에서도 가까이 위치하는 것으로 생각된다. 이는 SJR이 합산지향지수인 총인용수와는 0.4 정도의 상관관계를 보인 것에서 다시 한번 확인 할 수 있다.

셋째, 인용 지수들간의 상관관계는 지수 계산에 사용한 인용 색인 데이터베이스보다는 지수의 계산 방식이 더 큰 영향을 미치는 것으로 파악되었다.

상관관계 분석결과 SJR과 JIF는 각각 SCOPUS와 Thomson의 각기 다른 인용 색인 데이터를 사용하여 계산된 지수이지만, 이들 간에는 0.8 이상의 높은 상관관계를 보였다. 따라서 인용색인 데이터베이스에 따라서 분석 결과에 어느 정도 차이가 있을 수 있으나 각 데이터베이스에서 상위 랭크된 학술지에 대한 평가 결과는 크게 달라지지 않는다는 점을 확인할 수 있었다.

5.2.3 자관 네트워크지수들과 일반적 인용 지수들간의 상관관계

자관의 동시인용 네트워크를 통해 파악된 네트워크지수들과 일반적 인용 지수들간에는 유의미한 상관관계를 보이는 것으로 나타났다. 그러나 그 상관의 정도는 대부분 0.5 이하로 낮게 나타났다(표 7 참조).

자관 네트워크 지수들과 일반적 인용 지수들간의 상관관계의 특징은 다음과 같다.

첫째, 자관 네트워크 지수들과 일반적 인용 지수들간의 상관관계는 상당수가 통계적으로 유의한 것으로 나타났지만 상관도는 대부분 0.5

이하에 머물러서 그다지 높지 않았다.

둘째, 피인용빈도와 유의미한 상관관계를 가지는 일반적 인용 지수들이 많지 않았다.

셋째, 자관 인용 네트워크 지수 중에서 평균 프로파일연관성이 대체적으로 모든 일반 인용 지수들과 비교적 높은 상관관계를 보이고 있다.

세 가지 주요 특징에 대해서 상세히 설명하면 다음과 같다.

(1) 자관 네트워크 지수와 일반적 인용 지수 사이의 상관관계

자관 네트워크 지수들과 일반적 인용 지수들간의 스피어만 상관분석을 통해 가장 높은 상관관계를 보인 것은 삼각매개중심성과 SJR로 0.507의 상관관계를 보였다($p < .01$). 일반적 인용 지수를 중심으로 살펴보면, JIF가 삼각매개중심성과 가장 높은 상관관계를 보였다(0.381, $p < .01$), JII는 자기인용미포함 가중치페이지랭크(wPR(x))와 가장 높은 상관관계를 보였다(0.446, $p < .01$). 위세지향지수인 EigenFactor는 보나시치 지수와(0.485, $p < .01$), AI는 에이젠벡터(0.304, $p < .01$)와 가장 높은 상관관계를 보였으며, h-index는 에이젠벡터 및 보나시치 지수와 가장 높은 상관관계를 보였다(0.478, $p < .01$). 그 외의 지수들 간에도 대체적으로 0.1에서 0.4 사이의 약한 상관관계를 보였다.

이는 Y대학교의 인용 행태와 일반적 학술지 인용 행태가 서로 관계가 있지만 그 관계는 강하지 않음을 나타낸다. 즉, 일반적 학술지 이용을 나타내는 학술지 영향력 지수들과 함께 자관의 학술지 이용을 나타내는 인용 네트워크의 네트워크 지수들을 동시에 고려하는 것이 필요함을 알 수 있다.

〈표 7〉 일반적 인용 지수들과 자관 인용 네트워크 지수들 간 관계(Spearman, 2-tailed)

	cFrq	cJnum	wPR(s)	wPR(x)	Ctb	Cm	Cn	Cmp	Constr	Eigen	Bona	Degree	Close	Between	CC1	CC2
articles	0.327**	0.337**	0.263**	0.231**	0.295**	0.253**	0.000	0.049	-0.337**	0.216**	0.219**	0.298**	0.296**	0.335**	-0.382**	0.275**
cites	0.336**	0.397**	0.348**	0.340**	0.389**	0.419**	-0.010	0.324**	-0.316**	0.403**	0.402**	0.399**	0.396**	0.360**	-0.328**	0.397**
JIF	0.043	0.131	0.159*	0.182*	0.174*	0.263**	0.016	0.381**	-0.056	0.274**	0.269**	0.186*	0.185*	0.105	0.010	0.214**
IM	0.055	0.141	0.198**	0.212**	0.167*	0.260**	0.030	0.366**	-0.031	0.290**	0.284**	0.180*	0.180*	0.104	0.010	0.209**
JII	0.342**	0.412**	0.430**	0.446**	0.371**	0.410**	-0.043	0.368**	-0.212**	0.443**	0.443**	0.380**	0.380**	0.331**	-0.276**	0.383**
PII	0.138	0.227**	0.283**	0.323**	0.205**	0.269**	-0.044	0.349**	-0.020	0.324**	0.321**	0.211**	0.211**	0.154	-0.066	0.226**
B2	0.071	0.169*	0.259**	0.287**	0.130	0.191*	-0.024	0.244**	0.002	0.213**	0.210**	0.136	0.136	0.102	-0.034	0.146
B4	0.060	0.153	0.241**	0.267**	0.113	0.175*	-0.004	0.229**	0.018	0.197*	0.194*	0.118	0.118	0.087	-0.024	0.128
B6	0.077	0.171*	0.245**	0.275**	0.126	0.193*	0.006	0.241**	0.014	0.214**	0.211**	0.133	0.133	0.100	-0.033	0.143
EF	0.349**	0.429**	0.437**	0.447**	0.395**	0.460**	0.005	0.451**	-0.236**	0.484**	0.485**	0.406**	0.406**	0.334**	-0.257**	0.418**
AI	0.069	0.175*	0.244**	0.280**	0.159*	0.257**	0.011	0.406**	0.012	0.304**	0.301**	0.169*	0.169*	0.088	0.022	0.197*
SJR	0.132	0.228**	0.238**	0.265**	0.257**	0.358**	0.035	0.507**	-0.070	0.405**	0.399**	0.268**	0.268**	0.171*	-0.030	0.305**
hindex	0.324**	0.399**	0.390**	0.401**	0.387**	0.451**	0.023	0.468**	-0.236**	0.478**	0.478**	0.400**	0.399**	0.330**	-0.232**	0.418**

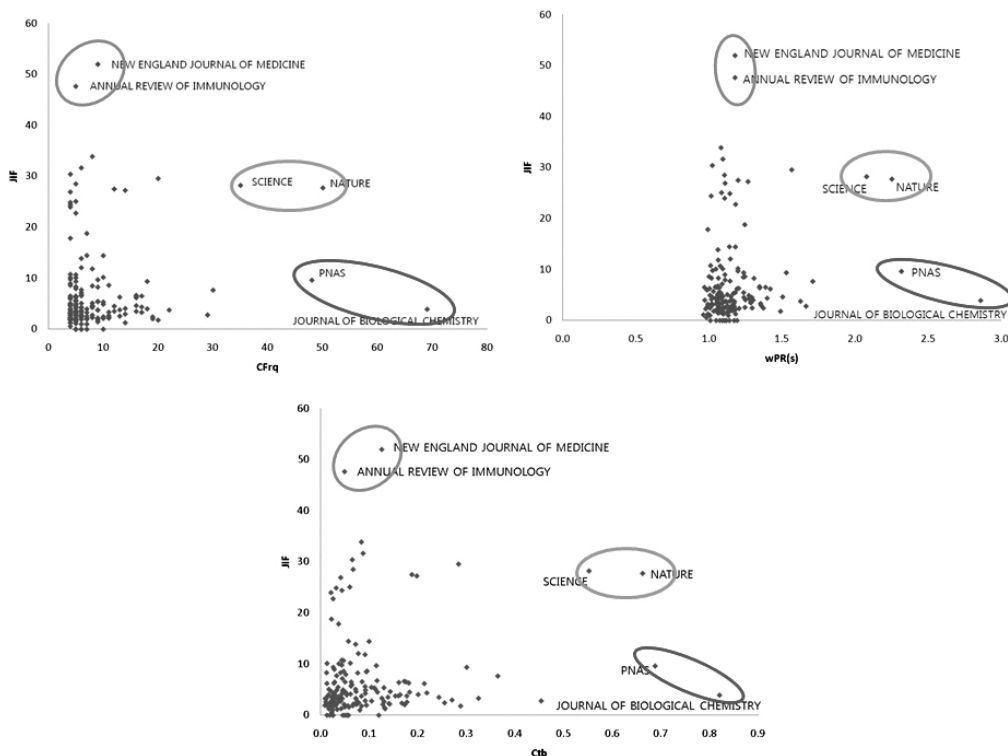
* p < .05, ** p < .01

이는 자관 네트워크지수들과 일반적 인용 지수들간의 산포도를 살펴보면 분명하게 드러난다. <그림 6>과 <그림 7>은 일반적 인용 지수인 JIF, h-index와 자관 인용 네트워크에서 계산된 피인용빈도, 가중치페이지랭크, 삼각매개중심성간의 산포도를 나타낸 것이다.

<그림 6>을 살펴보면 대체적으로 낮은 JIF 값을 갖는 학술지들이 낮은 자관 인용 네트워크 지수값을 갖는 것으로 나타난다. 그러나 세계의 자관 네트워크 인용 지수들과 JIF와의 관계에서 특이한 학술지 인용의 경우가 발생함을 알 수 있다. 각 산포도에서 학술지명이 표기된 *New England Journal of Medicine*, *Annual Review of Immunology*, *Science*, *Nature*, *PNAS*

(*Proceedings of the National Academy of Sciences of the USA*), *Journal of Biological Chemistry*의 6개 학술지의 경우이다.

각 산포도의 좌측 상단에 위치하는 *New England Journal of Medicine*, *Annual Review of Immunology*, *Science*의 경우는 자관 인용 네트워크 지수값은 낮지만 일반 인용 지수값은 높은 경우이다. 반면에 각 산포도의 우측 하단에 위치한 *PNAS*, *Journal of Biological Chemistry*는 자관 인용 네트워크 지수값은 높지만 일반 인용 지수값은 낮은 경우이다. 또 *Science*와 *Nature*도 일반적 인용 지수값에 비해 자관 인용 네트워크의 지수값이 비교적 높은 학술지에 속한다.



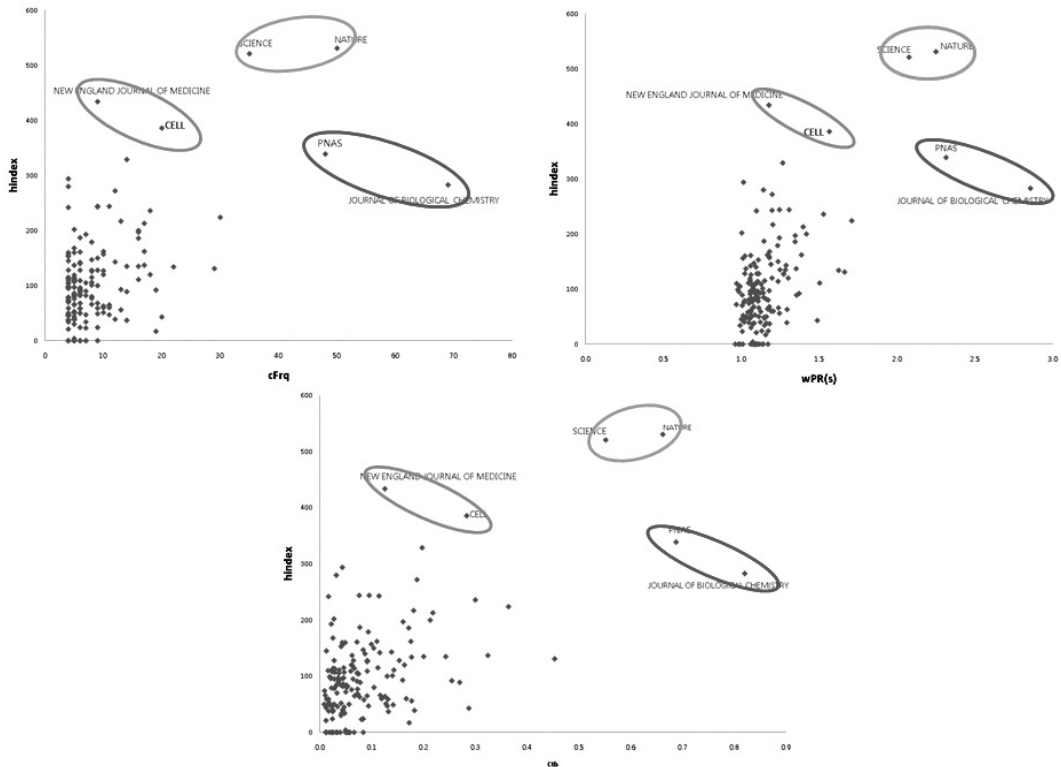
<그림 6> JIF와 자관 동시인용 네트워크 지수들의 관계

즉 Y대학교 생명공학과의 이 학술지들 간의 이용 패턴이 일반적인 학술지 인용 정도와 다르다는 것을 알 수 있다. 구체적으로 Y대학교 생명공학과의 2006년과 2007년에 발표된 학술논문에서는 의학 분야 특화 학술지인 *New England Journal of Medicine*과 리뷰 학술지 *Annual Review of Immunology*의 인용이 일반적 인용지수에 비해 상대적으로 적었다. 반면에 과학 일반분야의 학술지인 *PNAS*와 생화학 분야의 학술지인 *Journal of Biological Chemistry*는 자관의 인용이 일반적 인용 지수에 비해 상대적으로 많았다. 그리고 과학 일반분야의 학술지인 *Science*와 *Nature*는 자관과 일반 인용 지수 모

두에서 중간 정도의 비슷한 인용을 보였다.

〈그림 7〉은 h-index 와 자관 인용 네트워크 지수들간의 산포도이다. h-index도 JIF와 마찬가지로 자관 특성적인 학술지 인용의 경우가 발생함을 알 수 있다. 각 산포도에서 학술지명이 표기된 *New England Journal of Medicine*, *Cell*, *Science*, *Nature*, *PNAS*(*Proceedings of the National Academy of Sciences of the USA*), *Journal of Biological Chemistry*의 6개 학술지의 경우이다.

h-index는 JIF와 차이를 보이는데 우선 *Annual Review of Immunology* 대신 *Cell*이 자관 특성을 나타내는 학술지로 출현하고 있다. 그리



〈그림 7〉 h-index와 자관 동시인용 네트워크 지수들의 관계

고 자관특성을 보이는 학술지의 분포가 JIF와 는 다름을 알 수 있다. JIF의 경우에는 JIF 값 은 높지만 자관 인용 네트워크 지수가 낮거나 또는 그 반대인 경우가 비교적 확연히 구분되 었으나, h-index의 경우에는 이들의 위치가 서 로 더 가까워져 있다. 이를 통해 볼 때, 평균지 향 지수인 JIF 값보다는 평균지향과 합산지향 의 중간 성격을 가진 h-index 값이 자관 인용 네트워크 지수와의 차이가 크지 않다고 볼 수 있다.

결론적으로 자관 인용 네트워크는 일반적 인 용 지수와 대체적으로 비슷한 경향을 보이는 동시에 일반적 인용색인 데이터베이스에서는 확인할 수 없는 자관 고유의 인용 특성을 설명 해 주고 있으며, 일반적 인용 지수의 유형에 따 라서 자관 인용 네트워크 지수들과의 설명력이 차이가 있음을 알 수 있다.

(2) 피인용빈도와 일반적 인용 지수의 상관 관계

피인용빈도와 유의미한 상관관계를 가지는 일반적 인용 지수들은 많지 않았다. 특히 피인 용빈도는 평균지향지수인 JIF, SJR 등과 유의 미한 상관관계를 보이지 않았으며 총인용빈도 수와 유의미한 상관관계를 보였다. 이에 비하 여 다른 자관 네트워크지수들은 대부분의 일반 적 인용 지수들과 유의미한 상관관계를 보이고 있다.

따라서 비록 자관의 동시인용 네트워크 지수 들이 피인용빈도와 높은 상관관계를 보였지만 (표 5 참조), 피인용빈도와는 다른 네트워크의 구조적 설명력을 가지고 있으므로 자관의 피인 용빈도와 함께 자관의 동시인용 네트워크 분석

도 유용하다고 할 수 있을 것이다.

(3) 평균프로파일연관성과 일반적 인용 지수의 상관관계

자관 인용 네트워크 지수 중에서 평균프로파 일연관성이 대체적으로 모든 일반 인용 지수들 과 비교적 높은 상관관계를 보였다. JIF를 중심 으로 자관의 동시인용 네트워크 지수와의 관계 를 살펴보면, 평균프로파일연관성과 삼각매개중 심성에서 상관계수값에 차이를 보이는 것을 알 수 있다. 실제로 JIF와 삼각매개중심성간의 스 피어만 상관계수값보다 JIF와 평균프로파일연 관성의 스피어만 상관계수값이 높게 나타났다.

평균프로파일연관성과 삼각매개중심성 모두 전역중심성을 측정하는 지수들이지만 삼각매 개중심성이 인용 범위를 반영하여 전체적으로 고르게 인용되는 학술지일수록 그 값이 크게 비하여, 평균프로파일연관성은 중립적인 영역 에서 자주 인용되는 학술지일수록 그 값이 큰 차이를 가지고 있다. 대체적으로 평균프로파일 연관성 지수가 삼각매개중심성 지수에 비해 다 른 일반적 학술지 영향 지수들과의 상관이 높 게 나타난 것은 이와 같은 지수의 특성 때문인 것으로 생각된다.

6. 결론 및 제언

이 연구에서는 한 기관의 인용 패턴을 네트 워크의 구조적 속성을 나타내는 지수를 통해 분석하여 이를 일반적인 학술지 인용 지수들과 비교함으로써 자관의 인용 패턴 특성을 파악하 고자 하였다. 자관 인용 네트워크 및 일반적 인

용 지수들의 분석 결과는 다음과 같다.

첫째, 자관의 상호인용 네트워크를 통해 주요 연구 분야와 분야 내 주요 학술지를 확인할 수 있었다. Y대학교 생명공학과와 2006-2007 발표 논문을 통해 확인된 연구 분야는 크게 'Biotechnology', 'Microbiology & Virology', 'Microbiology & Food Science', 'Medicinal Chemistry & Organic Chemistry', 'Pharmacology'의 다섯 개 영역이었다.

둘째, 자관의 동시인용 네트워크 내 지수들은 가중치지수들과 비가중치지수들간에 모두 높은 상관관계를 보였다. 즉, 이진 네트워크의 구조적 속성으로도 피인용빈도에 의한 인용 네트워크와 유사한 구조적 속성을 파악할 수 있었다. 다만 이는 분석 자료 수집 기간이 2년으로 짧은 기간의 적은 분량이어서 인용 행렬에서 0인 칸이 많았기 때문이며, 만약 분석 자료를 더 긴 기간에서 확보하여 학술지간 인용이 0인 경우가 적어진다면 가중치지수와 비가중치지수 사이의 차이는 커질 것으로 예상된다.

셋째, 학술지의 일반적인 영향력 지수들간에는 인용 데이터베이스의 수록범위에 차이가 있음에도 불구하고 서로 높은 상관관계를 보였다. 또한 평균지향지수와 합산지향지수별로 유사한 유형의 지수들끼리 높은 상관관계를 보였다. 그리고 빈도를 중시하는 JIF와 위세성을 고려하는 EigenFactor, AI, SJR 등과도 비교적 높은 상관관계를 보였으며, h-index는 대부분의 일반적 영향력 지수들과 높은 상관관계를 나타내었다. 특히 SJR이 JIF와 높은 상관관계를 보여 이 지수가 평균지향 지수임을 확인할 수 있었다. 또 SJR과 JIF의 분석 대상이 되는 학술지에 차이가 있지만 학술지의 상위 순위에는

크게 영향을 미치지 않는 것을 유추해 볼 수 있었다. 이를 통해 볼 때 JIF로 대표되는 평균지향지수뿐 아니라 합산지향지수이면서 위세지향지수를 학술지 평가에 함께 사용하는 것이 필요할 것으로 생각된다. 무엇보다도 인용 색인 데이터베이스의 수록범위보다는 지수의 유형에 따른 설명력 차이가 크게 나타났다는 점을 학술지 평가에서 고려해야 할 것이다.

넷째, 자관의 동시인용에 의한 학술지의 영향력 지수들과 일반적 영향력 지수들간에는 낮지만 통계적으로 유의한 상관관계를 보였다. 이는 자관 이용자의 인용 패턴이 일반적인 학술지 인용 패턴과 유사하지만 부분적으로 다른 특성이 존재함을 뜻한다. 따라서 정보서비스를 위한 학술지의 중요성을 파악하기 위해서는 자관의 인용자료에서 산출한 지수와 일반적 인용 지수를 함께 고려하는 것이 필요하다고 할 수 있을 것이다.

결론적으로 자관 인용 네트워크의 구조적 분석은 관내 연구자의 연구 동향과 학술지 인용에 대해서 단순한 인용 빈도로 알 수 있는 것 이상의 정보를 제공해주는 것으로 나타났다. 또한 Web of Science에서 제공하는 JIF와 다른 인용 데이터베이스에서 제공되는 인용지수들은 각각의 특성과 성향이 다르므로 학술지 평가를 위해서는 여러 일반적 인용 지수들을 함께 고려해야 할 것이다.

이 연구를 통하여 자관의 인용 네트워크 및 일반적 학술 인용 지수로 자관에서 주로 인용되는 상위 학술지를 확인하는 것이 가능함을 확인하였다. 그러나 비교적 이용이 적은 하위 학술지를 선별해 내는 것은 적절하지 않다고 볼 수 있다. 또한 실제로 자관 인용 네트워크

지수들을 서로 비교하거나 이를 일반적 학술지 인용 지수들과 비교하는 것은 가능하지만 이는 비교일 뿐 어떤 지수가 학술지 이용을 가장 잘

나타내는 지수인지를 판단하는 것은 불가능하다. 따라서 각 지수의 신뢰도는 추후 설문조사 등으로 보완하는 작업이 필요할 것으로 보인다.

참 고 문 헌

- 고성순, 최상기. 2005. 인용문헌 분석을 이용한 외국학술잡지 이용행태 분석. 『한국도서관정보학회지』, 36(1): 1-17.
- 김용학. 2007. 사회 연결망 분석(개정판). 서울: 박영사.
- 김관준, 이재운. 2007. 연구 영역 분석을 위한 디스크립터 프로파일링에 관한 연구. 『정보관리학회지』, 24(4): 285-303.
- 김홍렬. 2005. 생명과학 학술지의 인용 분석 연구. 『정보관리학회지』, 22(3): 85-102.
- 노경관, 한상완. 2006. 특허분석을 통한 과학기술자의 과학논문 인용행태에 관한 연구. 『정보관리학회지』, 23(3): 223-239.
- 유재욱. 1993. 도서관학 학술지에 나타난 인용행태 분석. 『정보관리학회지』, 10(2): 145-167.
- 이재운. 2006a. 지적 구조의 규명을 위한 네트워크 형성 방식에 관한 연구. 『한국문헌정보학회지』, 40(2): 333-355.
- 이재운. 2006b. 계량서지적 네트워크 분석을 위한 중심성 척도에 관한 연구. 『한국문헌정보학회지』, 40(3): 191-214.
- 이재운. 2007a. 클러스터링 기반 네트워크 생성 알고리즘. 『제14회 한국정보관리학회 학술대회 논문집』, 147-154.
- 이재운. 2007b. 지적 구조 분석을 위한 MDS 지도 작성 방식의 비교 분석. 『한국문헌정보학회지』, 41(2): 335-357.
- 정영미. 2005. 『정보검색연구』. 서울: 구미무역(주)출판부.
- Ball, P. 2006. "Prestige is factored into journal ratings." *Nature*, 439(16): 770-771.
- Bonacich, P. 1972. "Factoring and weighting approaches to status scores and clique identification." *Journal of Mathematical Sociology*, 2: 113-120.
- Bonacich, P. 1987. "Power and centrality: A family of measures." *American Journal of Sociology*, 92: 1170-1182.
- Bollen, J., M. A. Rodriguez, and H. van de Sompel. 2006. "Journal status." *Scientometrics*, 69(3): 669-687.
- Borgatti, S. P., M. G. Everett, and L. C. Freeman. 2002. *Ucinet for Windows: Software for Social Network Analysis*. Harvard, MA: Analytic Technologies.
- EigenFactor.org.
 <<http://www.eigenfactor.org>>.
 [Retrieved October 23, 2008].
- de Nooy, W., A. Mrvar, and V. Batagelj. 2005. *Exploratory Social Network Analysis*

- with Pajek*. Cambridge: Cambridge University Press.
- Freeman, L. C. 1979. "Centrality in social networks conceptual clarification." *Social Networks*, 1: 215-239.
- Hascall, V. C. and R. W. Hanson. 2008. "JBC on journal ranking." *JBConline*.
<<http://www.jbc.org/misc/journalranking.shtml>>.
- Hirsch, J. E. 2005. "An index to quantify an individual's scientific research output." *Proceedings of the National Academy of Science*, 102(46): 16569-16572.
- Journal-Ranking.com.
<<http://journal-ranking.com>>.
[Retrieved October 23, 2008].
- Leydesdorff, L. 2007. "Betweenness centrality as an indicator of the interdisciplinarity of scientific journals." *Journal of the American Society for Information Science and Technology*, 58(9): 1303-1319.
- Narin, F., M. Carpenter, and N.C. Berlt. 1972. "Interrelationships of scientific journals." *Journal of the American Society for Information Science*, 23: 323-331.
- Sabidussi, G. 1966. "The centrality index of a graph." *Psychometrika*, 31: 581-603.
- SCImago. 2007. *SJR-SCImago Journal & Country Rank*.
<<http://www.scimagojr.com>>.
[Retrieved October 23, 2008].
- Wasserman, S. and K. Faust. 1994. *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press.
- Xing, W. and A. Ghorbani. 2004. "Weighted Pagerank algorithm." In *Proceedings of the Second Annual Conference on Communication Networks and Services Research(CNSR'04)*, pp.305-314.