

# 국내 인용 데이터베이스에서 저널 페이지랭크 측정 방안\*

## Journal PageRank Calculation in the Korean Science Citation Database

이 재 윤(Jae Yun Lee)\*\*

### 초 록

이 연구는 국내 인용 데이터베이스에서 저널 페이지랭크를 올바르게 측정할 수 있는 방안을 모색하고자 하였다. 국내 인용 데이터베이스는 해외 인용 데이터베이스에 비해서 인용 링크가 외부로 향하는 비율과 자기 인용 비율이 높다는 특성이 있다. 페이지랭크 공식은 반복 계산을 통해서 수렴하는 알고리즘이기 때문에 이런 특성을 감안하지 않으면 신뢰할만한 저널 페이지랭크 측정이 이루어질 수 없다. 따라서 국내 인용 데이터베이스에 적합한 저널 페이지랭크 측정 방안을 살펴보고 KSCD를 대상으로 측정된 결과를 분석하였다. 산출된 페이지랭크 지수에 대해서 상관분석과 회귀분석으로 검증해본 결과 SCImago 저널 랭크 공식을 적용하고 cr2 방식으로 저널 자기 인용을 조정하는 것이 국내 인용 색인 데이터베이스에서 저널 영향력을 평가하기에 적합한 방안인 것으로 확인되었다.

### ABSTRACT

This paper aims to propose the most appropriate method for calculating the journal PageRank in a domestic citation database. Korean journals show relatively high journal self-citation ratios and have many outgoing citations to external journals which are not included in the domestic citation database. Because the PageRank algorithm requires recursive calculation to converge, those two characteristics of domestic citation databases must be accounted for in order to measure the citation impact of Korean journals. Therefore, two PageRank calculation methods and four formulas for self-citation adjustment have been examined and tested for KSCD journals. The results of the correlation analysis and regression analysis show that the SCImago Journal Rank formula with the cr2 type self-citation adjustment method seems to be a more appropriate way to measure the relative impact of domestic journals in the Korean Science Citation Database.

키워드: 저널 평가, 인용 영향력, 저널 인용 네트워크, 페이지랭크 알고리즘, 저널 페이지랭크, 저널 자기인용  
Journal Evaluation, Citation Impact, Journal Citation Network, PageRank Algorithm,  
Journal PageRank, Journal Self-Citations

\* 본 연구는 행정안전부가 지원하는 한국정보화진흥원의 국가DB사업의 일환으로 한국과학기술정보연구원이 구축하여 제공하는 과학기술 참고문헌 인용색인 DB정보를 이용하였음.

\*\* 경기대학교 문헌정보학과 부교수(memexlee@kgu.ac.kr)

논문접수일자 : 2011년 12월 2일 논문심사일자 : 2011년 12월 3일 게재확정일자 : 2011년 12월 11일

## 1. 서론

### 1.1 연구의 배경

학술연구 활동에 대한 객관적 평가의 중요성이 점차 강조되면서 계량정보학과 계량정보분석 기법에 대한 관심이 지속적으로 증대되고 있다. 2000년대의 계량정보학 분야 연구 동향을 분석한 최근 연구(Lee & Choi 2011)에서는 연구 활동이 상승 추세인 키워드로 1위가 'h-index', 2위가 'Google Scholar', 3위가 'Scopus', 4위가 'Output'으로 나타나서 계량적인 실적 평가에 쏠린 학계의 관심이 드러난 바 있다.

특히 인용분석을 위한 데이터베이스 측면에서는 독점적인 지위를 누리던 Web of Science에 대해서 Scopus와 Google Scholar가 경쟁 상대로 부상하기에 이르렀다. 평가를 위한 인용 지표 측면에서도 평균 인용빈도를 뜻하는 저널 영향력 지수(Garfield 1965) 이외에 저널에 대한 h-지수(Hirsch 2005; 김관준, 이재운 2010; 조은성, 송재도 2011)나 네트워크 분석 방식의 다양한 기법이 검토되고 있다.

네트워크 분석<sup>1)</sup>은 사회 구조에 대한 분석 수단으로 문화인류학을 비롯한 사회학 영역에서 발전되어왔으며 2000년대에 들어선 이후 사회과학과 자연과학의 다양한 분야에서 폭넓게 연구되거나 활용되고 있다. 국내에는 김용학(1987a; 1987b)에 의해 사회학 분야에서 먼저 연구가 시작되었으며 문헌정보학계에서도 비교적 이른 시

점인 1990년대 초반 유사라(1991)가 처음으로 소개하였다. 문헌정보학 분야에서는 이후 관련 연구가 한동안 다루어지지 않았으나 2000년대 중반부터 계량서지학 분야에서 관련 연구(박한우 2003; 이재운 2006a, 2006b)가 다시 발표되기 시작하였다. 계량서지학 분야에서 네트워크 분석 기법을 다룬 국내 최근 연구는 지적 구조를 분석하는 경우(김희전, 조현양 2010; 김희정 2011; 박옥남 2011; 이재운 2008; 정영미, 유소영 2009)와 공동연구 네트워크를 분석하는 경우(김원진, 정영미 2010; 박옥남 2011; 이수상 2011; Chung 2011; Chung & Han 2009)가 대부분이다. 김성희와 장로사(2011)도 지적했듯이 비록 네트워크 분석과 관련된 연구가 계량서지학 분야에서 활발하지만 최근에는 이용자 연구(이수상, 장임숙 2008; 이용재, 배회숙 2008; Lee 2010), 이용 기록 분석(이수상, 위성광 2009; 최일영, 이용성, 김재경 2010), 교육과정 분석(송기호, 김태수 2008), 정보검색(김은희, 정영미 2010) 등의 다양한 주제로 네트워크 분석 기법의 적용이 확산되고 있다.

네트워크 분석 기법 중에서 가장 널리 활용되고 있는 것은 중심성 지수이다. 여러 가지 중심성 지수가 제안되어 있으나 영향력 분석을 위한 용도로 가장 대표적인 것이 페이지랭크(PageRank) 알고리즘(Brin & Page 1998)이다. 이는 각 웹페이지를 인용하는 웹페이지의 중요도를 고려하여 중요도를 산출하는 방식으로 구글의 검색결과 순위 기법으로 사용되어

1) 국내 상당수 연구에서 '사회 연결망 분석', 또는 '사회 네트워크 분석'이라는 표현을 사용하고 있으나, 관련 기법이 반드시 '사회적'인 네트워크에 대해서만 적용되는 것은 아니다. 최근에는 오히려 자연계의 '비사회적'인 대상에 대해 적용되는 경우가 늘어나면서 '사회'라는 단어가 생략되는 경우가 흔하다. 따라서 이 논문에서는 폭넓은 의미에서 '네트워크 분석'이라는 표현을 사용하였다.

널리 알려졌다.

원래의 페이지랭크 알고리즘은 한 사이트에서 다른 사이트로 향한 링크가 하나씩 있는 경우를 모델링한 것이었다. 이와 같은 인용 네트워크 모델은 문헌 간의 인용 관계와 유사하므로 페이지랭크 공식은 파라미터의 수정과 같은 사소한 보완을 거쳐서 문헌 인용 네트워크에 적용될 수 있으며, 이를 이용한 연구가 다수 발표되었다(Chen et al. 2007; Herskovic & Bernstam 2005; 이재윤 2011).

페이지랭크 값이 높은 사이트나 문헌이 모든 탐색자의 모든 검색 질의에 다 적합한 것은 아니므로, 검색을 위해서 페이지랭크 알고리즘을 적용할 때에는 개인화를 고려하거나 탐색 주제를 고려한 변형이 필요하다(AI-Saffar & Heileman 2007). 즉 페이지랭크 순위는 데이터베이스 전체를 고려한 전역 순위(global ranking)를 산출하는 것이지만 검색 상황에서는 탐색자나 탐색질문을 고려한 지역 순위(local ranking)를 산출해야 한다.

## 1.2 연구 문제 및 목적

문헌이나 저널 인용 네트워크에 대해서 페이지랭크 공식을 적용하여 인용 영향력을 측정할 때에는 별도의 탐색질의가 존재하는 상황이 아니므로 데이터베이스 전체를 고려한 전역 순위를 산출하는 것으로 충분하다. 다만 오랫동안 인용빈도에 의한 영향력 측정이 이루어져온 저널의 경우에는 페이지랭크 공식을 그대로 적용할 수 없다. 저널 인용 네트워크는 문헌 사이의 경우와는 달리, 저널과 저널 사이에 다수의 인용이 존재하는 가중 네트워크이기 때문이다. 이

에 따라 Bollen 등(2006)은 링크가 여러 개인 경우, 즉 가중 네트워크를 대상으로 하는 지수로 개발된 가중 페이지랭크 알고리즘(weighted PageRank; WPR)을 저널 인용 네트워크에 적용하여 저널의 중요성을 측정하는 시도를 하였다. 그 후 인용 데이터베이스의 양대 산맥인 WoS와 Scopus에 페이지랭크 방식의 인용 지수인 논문 영향력 지수(Article Influence Score: AIS)와 SJR 지수(SCImago Journal Ranking Indicator)가 각각 새로 도입되었다. 원래 AIS 지수와 SJR 지수는 Thomson ISI나 Elsevier와는 별개의 연구 기관에서 개발된 지표이지만, 그 유용성이 인정되면서 해당 데이터베이스에 공식 지수로 도입되기에 이르렀다(이재윤, 최선희 2010). 페이지랭크 방식의 인용 지수는 인용하는 저널의 중요도를 반영한다는 점, 인용하는 저널의 참고문헌 수를 반영한다는 점, 그리고 인용 네트워크 전체를 활용한다는 점(Franceschet 2010a)에서 장점을 가지고 있다.

현재 서비스되고 있는 국내 인용색인 데이터베이스에서는 아직까지 논문 당 평균 인용빈도 방식의 인용지수 정보만 제공하고 있을 뿐이다. 자체적으로 국내 컴퓨터공학 분야 논문들의 인용 데이터베이스를 구축한 최근 연구(양현, 이승도, 권영빈 2011)와 국내 인용 지수의 문제점을 지적한 연구(조은성, 송재도 2011)에서도 페이지랭크 방식의 지수는 향후 연구과제로 남기고 있다.

국내 저널 인용 네트워크에 대해서 페이지랭크 방식을 적용하여 저널의 영향력을 측정하기 위해서는 국제적인 인용 데이터베이스와 달리 국내 인용 데이터베이스가 가지고 있는 다음과 같은 두 가지 특징을 고려해야 한다.

첫째는 국내 저널의 높은 자기 인용(self-citing) 비율이다. Biglu(2007)에서는 2005년 WoS 데이터를 대상으로 분석한 결과 저널의 자기 인용 비율 평균이 7.81%로 나타났다. 이에 비해서 KSCD 데이터를 분석한 최근 국내 연구(Choi et al. 2011)에서는 자기 인용 비율 평균이 50%를 넘는 것으로 보고하고 있다. 국제 인용 데이터베이스에 비해서 국내 인용 데이터베이스는 수록 저널 종수가 적고 동일 주제를 연구하는 저널도 많지 않으므로 저널 단위의 자기 인용 비율이 높을 수밖에 없다. 페이지랭크 공식은 자신에 대한 링크, 즉 자기인용이 높을 경우에 반복 계산을 통해서 영향력이 과도하게 부풀려지는 문제점을 가지고 있다. 이에 따라 WoS 데이터를 대상으로 하는 논문 영향력 지수에서는 자기 인용 빈도를 아예 배제하고 있으며 Scopus 데이터를 대상으로 하는 SJR 지수에서는 자기 인용빈도를 전체 인용빈도의 1/3로 제한하고 있다. 그러나 자기 인용 비율이 높을 수밖에 없는 국내 학술지에 이와 같은 방식을 그대로 적용하게 되면 발행되는 저널이 적은 주제분야에 불리할 가능성이 높다.

둘째는 데이터베이스 외부 저널, 특히 해외 저널에 대한 인용 비중이 높은 문제이다. KSCD의 경우에 해외 저널을 인용하는 비율이 전체 인용 중에서 75.35%에 달하는 것으로 보고된 바 있다(Kim et al. 2011). 페이지랭크 알고리즘은 인용 네트워크에 포함된 저널들 사이에 인용을 통해서 영향력이 반복적으로 전달되면서 특정한 값으로 수렴되는 구조이다. 그런데 데이터베이스 외부 저널에 대한 인용은 다시 데이터베이스 내부 저널로 되돌아오는 경로가 파악되어 있지 않아서 전체 영향력이 일부 손

실되는 결과를 가져오므로 반복을 통한 영향력의 수렴을 곤란하게 만든다. 이재윤과 최선희(2010)의 연구에서는 이 문제를 피하기 위해서 데이터베이스 외부 저널에 대한 인용을 모두 제외하고 분석하였으므로 본격적인 해결책을 제시하지 못하였다.

이 연구에서는 이와 같은 두 가지 문제점을 고려하여 국내 인용색인 DB인 KSCD를 대상으로 페이지랭크 방식의 인용 지수를 적절히 측정할 수 있는 방안을 모색해보고자 한다. 이를 위해서 가상의 인용 네트워크를 대상으로 저널에 대한 여러 페이지랭크 공식을 적용해보고 인용 데이터베이스 외부 저널에 대한 자기 인용 비율이 높은 국내 상황에 적합한 공식을 파악하였다. 또한 이재윤과 최선희(2010)가 제안한 저널의 자기 인용 비율 조정 방식을 검토하였다. 검토 결과 국내 인용색인 DB에 적합하다고 판단된 방식을 실제 KSCD 데이터에 적용하여 특성을 파악하고 전통적인 평균 인용빈도 방식의 지표와 비교해보았다.

## 2. 국내 인용DB에 적합한 페이지랭크 측정 방식 검토

### 2.1 저널 인용 네트워크에 적합한 페이지랭크 공식

웹이 보편화되기 시작한 20세기 말, 키워드 출현빈도에 근거한 검색엔진은 웹 공간의 기하급수적인 성장과 날로 교묘해지는 스팸 사이트 때문에 한계에 부딪힌 상황이었다(Langville & Meyer 2006). 1998년, 인기도(popularity)를 고려

한 두 가지 알고리즘, 페이지랭크(Brin & Page 1998; Brin et al. 1999)와 HITS(Kleinberg 1998; 1999)가 발표되어 인터넷 검색의 새로운 장을 열었다.

흔히 인기가 높은 웹 페이지라고 하면 많은 링크를 받는 웹 페이지를 연상하지만, 페이지랭크 알고리즘은 받는 링크의 수를 단순히 세지 않고, 링크를 내보내는 웹 페이지의 인기도에 비례하여 차별화하여 합산하는 방식을 사용하였다. Kleinberg의 HITS 알고리즘은 많은 링크를 받는 웹페이지를 authority, 많은 링크를 내보내는 웹페이지를 hub라고 정의하고 좋은 authority를 링크하는 페이지에게는 높은 hub 점수를 부여하고, 좋은 hub 페이지로부터 링크를 받는 페이지에게 높은 authority 점수를 부여하는 방식을 사용하였다. 페이지랭크 알고리즘은 검색엔진 Google에 곧바로 적용되었고, HITS 알고리즘은 뒤에 검색엔진 Teoma에서 채택하였다. 페이지랭크 알고리즘은 저널 인용 네트워크에도 활발히 적용되고 있으나 HITS 알고리즘을 적용한 사례는 매우 드물다. HITS 알고리즘을 저널 인용 네트워크에 적용한 사례 연구에서는 인용을 두드러지게 많이 받는 저널 한 종이 있으면 그 저널과 인용을 주고 받는 같은 주제 분야의 다른 저널들이 인용빈도가 높지 않더라도 덩달아 상위권을 차지해버리는 문제가 발생하는 것으로 나타났다(Cheng et al. 2009). 이와 같이 HITS 알고리즘은 여러 분야를 포괄하는 인용 네트워크에 적용하기에는 적절하지 않으므로 이 연구에서는 페이지랭크 알고리즘만 다루었다.

웹페이지에 대해서 적용하기 위해서 제안된 원래의 페이지랭크 공식은 다음과 같다.

$$PR_i = \frac{(1-d)}{N} + d \cdot \sum_{j=1}^N \frac{PR_j}{O_j}$$

이 공식에서 각 항의 의미는 다음과 같다.

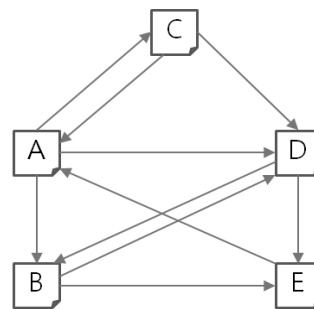
$PR_i$  - 웹 페이지 i의 페이지랭크

$d$  - 상수, 보통은 0.85

$O_j$  - 웹 페이지 j의 나가는 링크 수

$N$  - 전체 웹 페이지 수

<그림 1>의 가상의 5개 웹 페이지 간 링크를 통해 전파된 영향력을 페이지랭크 알고리즘으로 측정하면 웹 페이지 A는 0.257, B는 0.205, C는 0.096, D는 0.232, E는 0.211이 되어서 A가 가장 높고 C가 가장 낮게 측정된다.



<그림 1> 웹 페이지 간 링크 사례

인용빈도가 링크에 부여되는 가중 네트워크에 적용할 수 있도록 페이지랭크 공식을 수정한 가중 페이지랭크(WPR) 공식은 Xing과 Ghorbani (2004)가 제안하고 Bollen 등(2006)이 저널 인용 네트워크에 최초로 적용하였다. 가중 페이지랭크 공식은 다음과 같다.

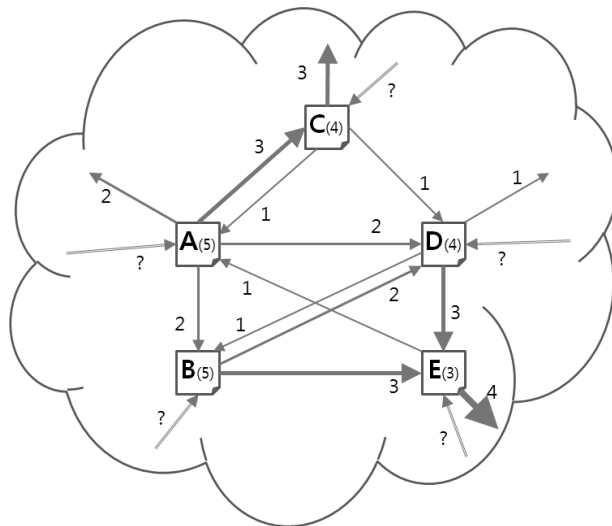
$$JPR_i = \frac{(1-d)}{N} + d \cdot \sum_{j=1}^N \frac{JPR_j \cdot C_{ji}}{C_j}$$

이 공식에서 각 항의 의미는 다음과 같다.

- $JPR_i$  - 저널 i의 페이지랭크
- $C_{ji}$  - 저널 j로부터 저널 i에 대한 인용 빈도
- $C_j$  - 저널 j의 참고문헌 수(나가는 인용 빈도)
- $d$  - 상수, 보통은 0.85
- $N$  - 저널 종수

가중 페이지랭크 공식은 원래의 페이지랭크 공식을 변형해서 저널j에서 저널i로 전달되는

중요도를 저널j의 모든 나가는 인용 중에서 저널i로 향하는 인용의 비율만큼 전달하도록 고안한 것이다. 가상으로 5개 저널 간 인용 관계가 <그림 2>와 같이 설정된 경우에 가중 페이지랭크를 측정된 결과가 <표 1>에 제시되어 있다. <표 1>의 5개 저널별 WPR 값을 인용빈도와 함께 비교해보면 인용빈도가 2회로 가장 적은 저널 A가 영향력이 가장 높게 측정된다. 그 이유는 인용빈도가 6회로 가장 많은 저널 E가 DB내로 내보내는 인용이 단 1회인데 그 1회의



<그림 2> 가상 DB의 5개 저널 사이의 인용 관계도  
(화살표 옆 숫자는 인용 빈도, 괄호 안 숫자는 논문 수)

<표 1> 가상의 5개 저널 간 인용 데이터에 대한 인용 영향력 측정 결과

저널	논문 수 (P)	나가는 인용 건수		저널 단위 영향력			평균 논문 단위 영향력		
		DB내	DB외	인용빈도(TC)	WPR	SJR	TC/P	WPR/P	SJR/P
A	5	7	2	2	0.285	0.181	0.40	0.0570	0.0363
B	5	5	0	3	0.143	0.154	0.60	0.0286	0.0309
C	4	2	3	3	0.134	0.113	0.75	0.0335	0.0283
D	4	4	1	5	0.205	0.221	1.25	0.0512	0.0553
E	3	1	4	6	0.233	0.390	2.00	0.0778	0.1299

인용이 저널 A를 향하기 때문이다. 즉, 많은 인용을 받는 저널 E의 높은 영향력이 고스란히 저널 A로 전달된 것이다. 그런데 저널 E는 사실 DB 내/외를 통틀어서 5건의 인용을 내보내고 있으므로 저널 A로는 저널 E가 가진 영향력의 1/5만 전달되어야 하지만 DB 내부 저널에 대한 인용만 고려하였으므로 저널 A로 전달되는 영향력이 과대 평가되는 문제가 발생하였다. 이런 경우는 해외 저널에 대한 인용이 많은 분야의 국내 저널에서 흔히 발생하게 된다.

인용 DB 외부 저널에 대한 인용까지 고려한 SCImago 저널 랭크 공식은 다음과 같다(SCImago Research Group 2007).

$$SJR_i = \underbrace{\frac{(1-d-e)}{N}}_{\text{①}} + e \cdot \underbrace{\frac{Art_i}{\sum_{j=1}^N Art_j}}_{\text{②}} + d \cdot \underbrace{\sum_{j=1}^N \frac{C_{ji} \cdot SJR_j}{C_j}}_{\text{③-1}} \cdot \underbrace{\frac{1 - \left( \frac{\sum_{k \in \{Dangling-nodes\}} SJR_k}{\sum_{k=1}^N \frac{C_{kh} \cdot SJR_k}{C_k}} \right)}{\sum_{k=1}^N \frac{C_{kh} \cdot SJR_k}{C_k}}}_{\text{③-2}} + d \cdot \underbrace{\left[ \frac{\sum_{k \in \{Dangling-nodes\}} SJR_k}{\sum_{j=1}^N Art_j} \right]}_{\text{④}} \cdot \frac{Art_i}{\sum_{j=1}^N Art_j}$$

이 공식에서 각 항의 의미는 다음과 같다.

- $SJR_i$  - 저널 i의 SCImago Journal Rank
- $C_{ji}$  - 저널 j로부터 저널 i에 대한 인용 빈도

- $C_j$  - 저널 j의 참고문헌 수
- $d$  - 상수, 보통은 0.85
- $e$  - 상수, 보통은 0.10
- $N$  - 저널 종수
- $Art_j$  - 저널 j의 논문 수

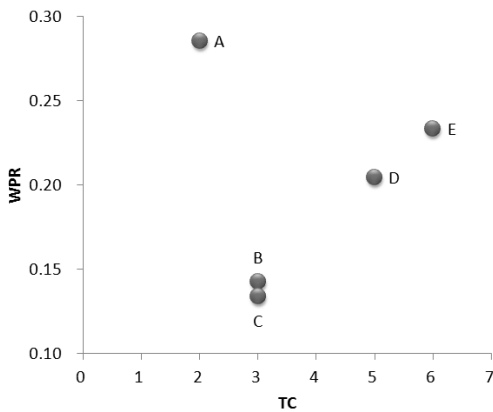
SCImago 저널 랭크 공식은 네 부분으로 구성되는데 이중에서 세 번째 부분은 다시 둘로 나눌 수 있다. 각 부분이 가지는 의미를 풀이하면 다음과 같다. 첫째 부분은 각 저널이 가지는 최소 중요도를 의미하며 전체 저널 종수에 따라 달라진다. 둘째 부분은 각 저널에 게재된 논문수에 비례하여 가지는 중요도이다. 셋째 부분은 앞의 직접 받는 인용을 통해 전달되는 중요도에 해당하며, 뒤는 DB 외부 저널로 향하는 인용을 통해 빠져나가는 중요도를 DB 내부 저널 각각에 균등하게 분배하는 역할을 한다. 넷째 부분은 인용을 내보내지 않는 저널(dangling node)의 중요도를 각 저널에 균등하게 분배하는 역할을 한다.

마지막으로 각 저널에 수록된 논문 당 평균 영향력은 SJR 값을 해당 저널의 논문 수로 나누어서 산출한다.

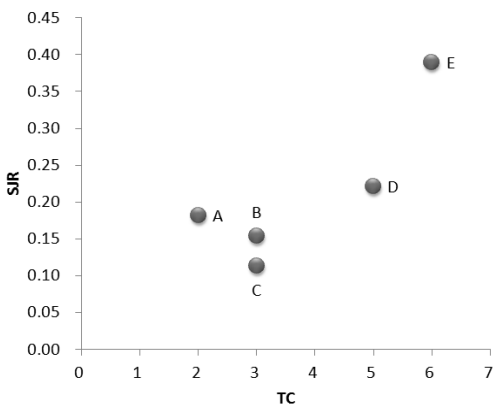
$$SJRQ_i = \frac{SJR_i}{ART_i} = \frac{\text{저널 } i \text{의 } SJR}{\text{저널 } i \text{의 논문 수}}$$

가상의 5개 저널로 구성된 인용 데이터베이스에서 SJR 공식을 적용한 결과가 앞의 WPR 공식 적용 결과와 함께 <표 1>에 포함되어 있다. WPR 공식과 SCImago 저널 랭크 공식을 적용하여 산출된 페이지랭크 값을 인용빈도와 비교해보면 각각 <그림 3>과 <그림 4>로 나타

낼 수 있다. 두 페이지랭크 공식에서 크게 다른 결과를 보이는 저널은 A이다. 저널 E가 저널 A에게 전달하는 영향력이 WPR 기준으로는 과대평가되었으나, 인용 DB 외부 저널에 대한 인용까지 고려하는 SJR 공식에서는 저널 E가 가진 영향력의 1/5만 저널 A로 전달되기 때문이다. 저널 E의 영향력 중에서 나머지 4/5는 전체 저널에게 고르게 분산된다. 따라서 저널 A의 영향력은 대폭 감소하여 3위에 머무르게 된다.



<그림 3> 인용빈도와 가중 페이지랭크 비교



<그림 4> 인용빈도와 SJR 지수 비교

이와 같이 DB 외부 저널에 대한 인용을 고려하는지 여부가 국내 저널의 인용 영향력 계산에 미치는 영향력은 작지 않다. 따라서 국내 인용 데이터베이스에서 저널의 영향력을 페이지랭크 알고리즘으로 측정하려면 가중 페이지랭크를 사용하지 말고 외부 저널에 대한 인용을 고려하는 SCImago 저널 랭크 공식을 적용하는 것이 바람직하다.

## 2.2 자기 인용 비율 제한 방법

저널의 자기 인용이 페이지랭크 방식 인용 지수 산출에 큰 영향을 미치므로 해외 인용 데이터베이스에 대한 인용지수에서도 자기 인용을 배제하거나 제한하는 방법을 쓰고 있다. WoS 데이터를 대상으로 하는 Eigenfactor(Bergstrom 2007)와 AIS 계산에서는 저널의 자기 인용을 모두 배제하고 있으며, Scopus 데이터를 대상으로 하는 SCImago SJR 지수는 저널의 자기 인용 빈도를 총 참고문헌 수의 1/3로 제한하는 방식을 채택하고 있다.

국내 인용 데이터베이스는 해외의 경우에 비해서 동일한 분야의 저널이 많지 않으므로 저널의 자기 인용을 무시하기가 어렵다. 특정한 논문을 인용할 가능성이 있는 동일 주제의 논문은 동일 저널에 게재될 가능성이 높은 것이다. 이에 따라 이재윤과 최선희(2010)는 저널의 자기 인용을 타 저널 인용 비율을 고려하여 상대적으로 삭감하는 방안을 제안한 바 있다. 선행 연구에서 소개된 자기 인용 빈도의 제한 방법을 간단히 정리하면 <표 2>와 같다.

이재윤과 최선희(2010)의 연구에서 cr1 방식은 cr2 방식과 유사한 결과를 보였다. 그러나



〈표 2〉 저널 자기 인용 빈도의 제한 방식

명칭	공식	설명
sjr	$\min\left(RS_j, \frac{RO_j}{2}\right)$	타 저널 인용 빈도의 1/3이 넘는 자기 인용 빈도를 버림. 자기 인용 비율이 최대 33.3% 이내로 제한됨.
half	$\frac{RS_j}{2}$	자기 인용 빈도를 무조건 절반으로 줄임.
cr1	$RS_j \times \frac{RO_j}{RS_j + RO_j}$	자기 인용 빈도를 타 저널을 인용하는 비율에 비례(자기 인용 비율에 반비례)하도록 반영함. 자기 인용 비율이 최대 50% 이내로 제한됨.
cr2	$RS_j \times \left(1 - \left(\frac{RS_j}{RS_j + RO_j}\right)^2\right)$	자기 인용 빈도를 자기 인용 비율의 제곱에 반비례하도록 반영함. 자기 인용 비율이 최대 66.7%로 제한됨.

\* 각 공식에서  $RS_j$ 는 저널 j의 실제 자기인용 빈도,  $RO_j$ 는 저널 j가 다른 저널을 인용한 빈도임.

국내 과학기술분야 저널의 자기 인용 비율이 평균적으로 50%를 넘으므로(Choi et al. 2011) 조정된 자기 인용 비율의 최댓값이 50%에 머무는 cr1 방식을 적용하면 자기 인용 비율 조절 결과가 현실 데이터와 지나치게 다른 결과를 얻게 되는 문제가 발생한다. 따라서 국내 인용 데이터베이스에서는 조정된 자기 인용 비율의 최댓값이 66.7%까지 허용되는 cr2 방식을 사용하는 것이 더 적절하다고 판단된다. cr2 방식을 이용하면 자기 인용 비율이 낮은 저널의 자기 인용 빈도를 cr1 방식을 적용한 경우에 비해서 덜 줄이는 효과도 얻는다. 따라서 이 연구에서는 cr2 방식만을 이후 분석에서 검토하기로 한다.

가상의 5개 저널에 대해서 자기 인용 빈도를 절반으로 줄이는 half 방식, 전체의 1/3까지만 남기는 sjr 방식, 그리고 cr2 방식을 적용해본 결과가 〈표 3〉에 제시되어 있다. 이 표를 보면 half 방식은 원 데이터의 자기 인용 비율이 높을수록 자기인용 비율을 낮추는 효과가 적게 나타났다. sjr 방식은 원 데이터의 자기 인용 비율이 80% 이상으로 매우 높은 수준에서는 지나치게 많은 인용을 삭제하는 결과를 낳는다. 특히 33.3% 이상의 자기 인용은 아무리 많아도 무시하므로 각 저널의 특성을 반영하지 못하게 된다. cr2 방식은 자기 인용 비율이 높은 저널에 대해서 자기 인용을 삭감하는 정도가 sjr 방식보다 덜하며 자기 인용 비율을 일률적으로

〈표 3〉 가상 저널에 대한 자기 인용 빈도의 조절 사례

저널	원 데이터			half 방식		sjr 방식		cr2 방식	
	참고문헌 수	자기인용 빈도	자기인용 비율	자기인용 빈도	자기인용 비율	자기인용 빈도	자기인용 비율	자기인용 빈도	자기인용 비율
F	100	5	5%	2.50	3%	5.00	5%	4.99	5%
G	100	40	40%	20.00	25%	30.00	33%	33.60	36%
H	100	60	60%	30.00	43%	20.00	33%	38.40	49%
I	100	80	80%	40.00	67%	10.00	33%	28.80	59%
J	100	95	95%	47.50	90%	2.50	33%	9.26	65%

유지하지도 않는다.

또한 자기 인용 빈도를 절반으로 줄이는 half 방식은 선행 연구에서 자기 인용을 그대로 포함하는 것과 유사한 결과를 보였으므로 이 방식도 향후 분석에서는 제외하고 sjr 방식과 cr2 방식만을 저널의 자기 인용 빈도 제한 방식으로 검토하기로 한다.

### 3. KSCD에서 페이지랭크 방식 인용 지수 측정

#### 3.1 저널 페이지랭크 측정 방법

KSCD 데이터 중에서 인용 네트워크 분석에 적합한 형태로 2008년 인용 데이터를 입수하였다. SCImago의 경우와 마찬가지로 2008년 이전 2005년부터 2007년까지 3년 동안 발간된 논문이 2008년에 인용된 빈도를 산출하였으며, 해당 기간 동안의 논문 수와 인용 데이터가 모두 파악된 저널은 총 449종이었다.

저널 페이지랭크 계산은 DB 외부 저널에 대한 인용이 많은 국내 저널의 특성을 감안하여 SCImago 저널 랭크 공식을 사용하였으며 각 저널의 논문 수에 따른 차이를 조정하기 위해서 페이지랭크 값을 해당 저널의 3년간 발행 논

문 수로 나눈 논문 당 페이지랭크(PageRank per Paper; 이하 PR/P로 표기)를 산출하고, 3년 간 논문 당 평균 인용빈도(이하 TC/P로 표기)를 함께 산출하여 비교하였다. SCImago 저널 랭크 공식에서 댐핑 팩터(damping factor)인 d는 0.85로, 저널의 논문 수를 고려하는 팩터인 e는 0.10으로 설정하였다.

자기 인용 빈도 조정 방식을 비교하기 위해서 저널의 자기 인용을 그대로 사용한 경우(이하 'CS포함'으로 표기)와 모두 제거한 경우(이하 'CS배제'로 표기), 그리고 sjr 방식과 cr2 방식으로 조정된 경우의 4가지 방식을 적용해보았다.

#### 3.2 저널 페이지랭크 측정 결과 분석

##### 3.2.1 전체 저널에 대한 측정 결과 분석

전체 449종 저널에 대해서 측정된 인용 지수 사이의 Spearman 순위 상관계수는 <표 4>와 같다. 논문 당 평균 인용빈도에는 자기 인용 빈도가 포함되어 있음에도 불구하고 저널 자기 인용을 포함한 페이지랭크보다 자기 인용을 배제하거나 조정된 경우가 오히려 논문 당 평균 인용빈도와 유사한 결과가 나타났다. 이처럼 페이지랭크 계산에는 지나친 자기 인용이 저널의 영향력 계산을 크게 왜곡한다는 점이 다시 확인되었다.

<표 4> 지수 간 Spearman 순위 상관(전체 저널 대상)

	TC/P	PR/P(CS포함)	PR/P(CS배제)	PR/P(sjr)
PR/P(CS포함)	.082			
PR/P(CS배제)	.517**	.447**		
PR/P(sjr)	.495**	.502**	.992**	
PR/P(cr2)	.547**	.504**	.981**	.991**

\*\* 상관 유의 수준이 0.01

해외 인용 데이터베이스에서 논문 당 평균 인용빈도인 Impact Factor와 논문 수로 나눈 페이지랭크 방식 지수를 비교한 선행 연구에서는 순위 상관계수가 WoS 데이터에 대해서 0.90(Franceschet 2010b), Scopus 데이터에 대해서 0.93(González-Pereira, Guerrero-Bote, & Moya-Anegón 2010)으로 매우 높게 나타났다. 반면에 KSCD에서는 sjr 방식으로 자기 인용을 보정한 경우에 0.495, cr2 방식으로 자기 인용을 보정한 경우에 0.547의 순위 상관계수를 보였으므로 상대적으로 매우 낮은 상관성을 보였다. 이는 KSCD 등재 저널의 중수가 해외 인용 데이터베이스의 1/10에 못미칠 정도로 훨씬 적어서이기도 하지만, 국내 학술지는 상대적으로 자기 인용 비율이 높기 때문에 자기 인용에 많이 좌우되는 페이지랭크 산출 결과가

평균 인용빈도와 더 큰 차이를 보이는 것으로 짐작된다.

각 인용 지수가 저널의 통계적 특성과 어떤 상관관계가 있고 어느 정도 영향받는지 살펴 보기 위해서 순위 상관분석과 단순 회귀분석을 실시하였다. <표 5>에서는 인용 지수별 저널 순위와 각 저널의 특성 변수 사이의 Spearman 순위 상관관계를 제시하였고, <표 6>에서는 각 저널의 인용 지수를 종속 변수로 하고 특성 변수를 독립변수로 설정한 단순 회귀분석 결과의 베타 계수를 제시하였다.

논문 수와의 상관관계가 강하면 저널의 규모에 따라 지수별 순위가 달라지는 경향이 강함을 의미한다. 자기 인용을 포함한 페이지랭크 산출 결과는 논문 수와 -0.891이라는 매우 강한 부정적 상관관계를 보이므로 규모가 작은 저널

<표 5> 인용 지수와 변수 간 순위 상관관계(전체 저널 대상)

	논문 수	총 인용빈도 (TC)	자기인용 비율	DB외부저널 인용비율	논문 당 자기인용 빈도	논문 당 참고문헌 수
TC/P	.237**	.789**	.344**	-.605**	.866**	.378**
PR/P(CS포함)	-.891**	-.456**	-.116*	-.196**	.063	.014
PR/P(CS배제)	-.247**	.205**	-.183**	-.365**	.280**	.254**
PR/P(sjr)	-.299**	.160**	-.207**	-.400**	.272**	.214**
PR/P(cr2)	-.268**	.213**	-.158**	-.473**	.336**	.179**

\* 상관 유의 수준이 0.05

\*\* 상관 유의 수준이 0.01

<표 6> 인용 지수에 대한 변수 영향력의 베타 계수(괄호 안은 유의 확률)

	논문 수	총 인용빈도 (TC)	자기인용 비율	DB외부저널 인용비율	논문 당 자기인용 빈도	논문 당 참고문헌 수
TC/P	-.201(.000)	1.693(.000)	-.076(.000)	-.115(.000)	.774(.000)	.350(.000)
PR/P(CS포함)	-.177(.035)	-.037(.876)	.168(.007)	-.482(.000)	-.298(.001)	-.062(.191)
PR/P(CS배제)	-.383(.000)	2.296(.000)	-.037(.381)	-.082(.089)	-.020(.729)	.131(.005)
PR/P(sjr)	-.362(.000)	2.144(.000)	-.008(.839)	-.048(.314)	-.026(.650)	.119(.011)
PR/P(cr2)	-.343(.000)	2.031(.000)	-.021(.613)	-.052(.287)	.033(.565)	.121(.010)

일수록 상위권으로 판단되는 것으로 나타났다. 자기 인용을 조정한 지수 중에서는 cr2 방식이 상관관계가 다소 낮게 나타났다. 회귀분석 결과를 보면 모든 지수가 저널의 규모에 따른 부정적인 영향을 받는 것으로 나타났으며, 평균 인용 빈도보다 자기 인용을 조정한 페이지랭크 방식 지수가 저널의 규모에 따른 영향을 다소 더 받는 것으로 나타났다. 다만 논문 수로 대표되는 저널의 규모는 국내에서는 학문 분야별로 편차가 심하므로 이런 결과가 저널의 특성이 아닌 학문 분야 단위의 특성에 기인하였을 가능성도 있다.

자기 인용 비율과 평균 인용 빈도 순위는 긍정적인 상관관계를, 페이지랭크 방식 지수 순위는 부정적인 상관관계를 보였다. 그러나 회귀분석 결과에서는 자기 인용 비율을 조정한 페이지랭크 방식 지수가 자기 인용 비율에 따른 영향을 받지 않는 것으로 나타났다. 논문 당 자기인용 빈도를 독립변수로 하여 분석한 결과에서도 자기 인용 비율을 조정한 페이지랭크 방식 지수는 영향을 받지 않는 것으로 나타났다.

DB 외부 저널 인용 비율과는 모든 지수의 순위가 부정적인 상관관계를 보였다. 그러나 회귀분석 결과를 보면 자기 인용 비율을 조정한 페이지랭크 방식 지수는 DB 외부 저널 인용 비율로부터 받는 영향이 통계적으로 유의하지 않았다. 국내 인용 DB에서는 수록 범위가 제한되어 있으므로 자기 인용 비율을 조정한 페이지랭크 방식 지수가 DB 외부 저널 인용 비율에 영향을 받지 않는다는 점은 중요한 특성이다.

논문 당 참고문헌 수와의 상관관계는 높을수록 참고문헌이 많은 저널에 유리한 지수라고

해석할 수 있다. 각 학문분야마다 인용을 많이 하고 적게 하는 관행이 상이하므로 이 상관관계가 강한 지수를 사용하면 일반적으로 인용이 활발한 학문분야에 편향된 판단을 내리게 된다. 이 상관관계는 평균 인용 빈도인 TC/P가 가장 높으며 자기 인용을 포함한 페이지랭크가 가장 0에 가깝게 나타났다. 회귀분석 결과에서도 평균 인용 빈도보다 페이지랭크 방식 지수가 약 1/3 정도로 각 학문분야의 인용 관행에 따른 영향을 덜 받는 것으로 나타났다.

### 3.2.2 최상위 저널에 대한 분석

지수별 상위 저널을 살펴보기 위해서 평균 인용빈도 기준 상위 10종과 cr2 공식으로 자기 인용 빈도를 조정한 페이지랭크 기준 상위 10종을 각각 제시하면 <표 7>, <표 8>과 같다. 평균 인용빈도 기준 10위 이내에는 비교적 여러 분야의 저널이 포함되어 있는 반면에 페이지랭크 기준으로는 상위 10종 중에서 8종이 식품영양 분야 저널이었다. 이는 식품영양 분야가 상대적으로 저널 간 상호 인용이 활발하기 때문인 것으로 짐작된다. 페이지랭크 기준 상위 저널 10종은 평균 인용빈도 기준으로도 모두 21위 이내에 포함되는 상위 저널이었다.

10위 이내에 포함된 상위 저널을 대상으로 자기 인용 비율의 평균을 산출해보면 평균 인용 빈도 상위 10종의 자기 인용 비율 평균은 72%인데 반해서 cr2 적용 페이지랭크 기준 상위 10종의 자기 인용 비율 평균은 33%에 불과하였다. 평균 인용 빈도 9위인 설비공학논문집과 10위인 한국양봉학회지는 각각 자기 인용 비율이 97%와 98%에 달할 정도로 높았다. 이로써 페이지랭크 공식을 적용하면 특정 저널이

〈표 7〉 평균 인용빈도(TC/P) 기준 상위 저널

저널명	지수별 순위				
	TC/P	PR/P(CS포함)	PR/P(CS배제)	PR/P(sjr)	PR/P(cr2)
한국식품조리과학회지	1	154	3	1	1
Food Science and Biotechnology	2	337	15	15	15
한국식품과학회지	3	176	1	2	2
Biotechnology and Bioprocess Engineering: BBE	4	282	70	82	71
대한지역사회영양학회지	5	181	6	7	7
Journal of Microbiology and Biotechnology	6	387	38	45	43
한국약용작물학회지	7	88	12	11	6
Macromolecular Research	8	339	368	393	388
설비공학논문집	9	16	180	215	169
한국양봉학회지	10	12	81	95	91

〈표 8〉 cr2 적용 페이지랭크(PR/P(cr2)) 기준 상위 저널

저널명	지수별 순위				
	TC/P	PR/P(CS포함)	PR/P(CS배제)	PR/P(sjr)	PR/P(cr2)
한국식품조리과학회지	1	154	3	1	1
한국식품과학회지	3	176	1	2	2
한국식품저장유통학회지	11	224	4	3	3
한국식품영양학회지	18	184	2	4	4
한국식생활문화학회지	21	262	5	5	5
한국약용작물학회지	7	88	12	11	6
대한지역사회영양학회지	5	181	6	7	7
한국식품영양과학회지	15	311	9	6	8
한국영양학회지	14	242	8	8	9
한국응용생명화학학회지	16	186	7	9	10

자기 인용에만 크게 의존해서 최상위 저널로 판정되는 것을 피할 수 있음이 확인되었다.

평균 인용빈도와 cr2 적용 페이지랭크를 기준으로 전체 측정 대상 저널 중에서 상위 5%에 해당하는 최상위권 저널의 지수를 별도로 비교해보았다. 앞에서 평균 인용빈도와 cr2 적용 페이지랭크의 순위 상관관계가 0.547이어서 통계적으로 유의한 관계로 나타났지만 〈표 7〉과 〈표 8〉에 나타난 최상위권 저널이 상당히 다르게 나타났다기 때문이다.

평균 인용빈도나 cr2 적용 페이지랭크 기준으로 어느 한 쪽에서라도 상위 5%(22.4위) 이내에 포함되는 저널은 총 32종이었다. 최상위권 32종 저널에 대해서 지수 간 순위 상관관계를 구해보면 〈표 9〉와 같다. 이 표에서 보듯이 평균 인용빈도(TC/P)와 cr2 적용 페이지랭크 기준으로 어느 한 지수에서라도 5% 이내에 드는 최상위 저널만 살펴보면, 지수 간 상관관계가 95% 유의수준에서 통계적으로 같지 않을 정도로 크게 달라진다. 이로서 최상위권 저널에 대해서는

평균 인용 빈도 순위와 페이지랭크 순위가 통계적으로 다르다는 것이 확인되었다.

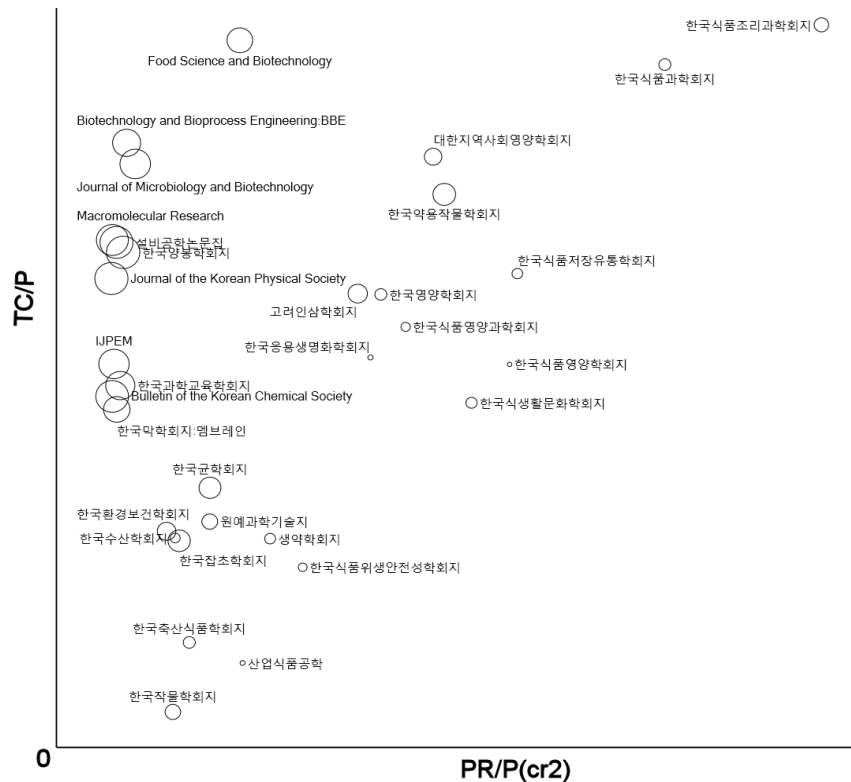
상위 저널 32종의 평균 인용빈도(TC/P)와 cr2 적용 페이지랭크를 산포도로 나타내면 <그림 5>와 같다. 이 그림을 살펴보면 대각선 우측 아래쪽이 비어 있다. 페이지랭크가 높은 저널

은 평균 인용 빈도도 높은 편이지만 반대로 평균 인용 빈도가 높은 저널 중에는 페이지랭크가 매우 낮은 저널도 상당수 포함되어 있기 때문이다. 원의 크기는 저널의 자기 인용 비율에 비례하도록 표현되어 있는데 페이지랭크가 낮고 평균 인용 빈도가 높은 왼쪽 상단의 저널은

<표 9> 지수 간 Spearman 순위 상관 - 상위 5% 이내 저널

	TC/P	PR/P(CS포함)	PR/P(CS배제)	PR/P(sjr)
PR/P(CS포함)	-.098			
PR/P(CS배제)	.133	.027		
PR/P(sjr)	.145	.017	.994**	
PR/P(cr2)	.163	.064	.986**	.991**

\*\* 상관 유의 수준이 0.01



<그림 5> 최상위권 저널 32종의 TC/P 지수와 PR/P(cr2) 지수 비교 (원의 크기는 자기 인용 비율에 비례함)

〈표 10〉 인용 지표와 변수 간 순위 상관관계(상위 5% 저널)

	논문 수	총 인용빈도 (TC)	자기인용 비율	DB외부저널 인용비율	논문 당 자기인용 빈도	논문 당 참고문헌 수
TC/P	.441*	.703**	.376*	-.222	.635**	.603**
PR/P(CS포함)	-.720**	-.598**	.047	-.475**	-.053	-.446*
PR/P(CS배제)	-.053	.006	-.787**	-.290	-.587**	-.207
PR/P(sjr)	-.030	.022	-.782**	-.304	-.563**	-.181
PR/P(cr2)	-.054	.014	-.746**	-.346	-.528**	-.198

\* 상관 유의 수준이 0.05

\*\* 상관 유의 수준이 0.01

모두 원의 크기가 커서 자기 인용 비율이 높음을 알 수 있다. 이는 평균 인용 빈도만을 기준으로 하여 저널을 평가하면 국내 인용 데이터베이스에서는 자기 인용에 주로 의존하는 저널에 뚜렷하게 유리한 평가가 이루어짐을 뜻한다. 페이지랭크를 기준으로 저널의 영향력을 평가하면 다른 인용 지표와 극단적으로 상반된 결과를 얻게될 위험은 줄어든다.

최상위권 저널 32종에 대해서 인용 지표와 변수 간 순위 상관관계를 측정된 결과인 〈표 10〉을 보면 페이지랭크 방식 지수는 논문 수나 총 인용빈도와 같은 저널의 규모와 통계적으로 상관관계가 없는 것으로 나타났다. 또한 논문 당 참고문헌 수와의 상관관계도 통계적으로 유의하지 않았으므로 분야별로 다를 수 있는 인용 관행에도 거의 영향받지 않는 것으로 나타났다.

#### 4. 결론

국내 인용색인 데이터베이스인 KSCD는 저널 등재 범위가 제한되어 있고 수록 저널의 자기 인용 비율이 해외의 경우보다 매우 높다는

특징이 있다. 이를 고려하여 네트워크 분석 방식의 저널 페이지랭크 지수를 측정하는 적절한 방안을 모색해보았다.

제한된 저널 등재 범위의 문제는 페이지랭크 공식 중에서 SCImago 저널 랭크 공식을 사용하는 것이 가장 적절한 대안인 것으로 판단되었다. 높은 자기 인용 비율의 문제에 대처하기 위해서는 cr2와 같은 자기 인용 조정 공식을 채택하였다. 분석 결과 DB 외부 저널에 대한 인용을 고려하도록 SCImago의 저널 페이지랭크 공식을 적용하고 자기 인용을 cr2 방식으로 조정된 페이지랭크 지수는 다음과 같은 특징이 있는 것으로 나타났다.

첫째, 저널 자기 인용의 영향을 받지 않는다. 저널의 자기 인용 비율 및 논문 당 자기 인용 빈도에 따른 영향이 모두 통계적으로 유의하지 않았다. 따라서 자기 인용이 상대적으로 흔한 국내 저널에 대한 평가에 적합하다.

둘째, DB 외부 저널 인용 비율에 영향 받지 않는다. 해외 저널에 대한 인용 비율이 높으며 분야별 편차도 큰 국내 인용 DB에서는 중요한 특징이다. 특히 KSCD와 같이 국내 저널 중에서도 수록 범위가 제한되어 있는 인용 DB에서는 더 필요한 특징이다.

셋째, 분야별로 다를 수 있는 인용 관행(인용 수준)의 영향을 덜 받는다. 논문 당 참고문헌 수에 따른 영향이 평균 인용 빈도에 비해서 약 1/3 정도로 적은 것으로 나타났다. 이는 상이한 주제분야의 저널을 비교할 때 유리한 특성이다.

넷째, 저널 규모(논문 수)에 따른 영향은 평균 인용 빈도보다 페이지랭크 측정 결과가 다소 더 받는 것으로 나타났다. 규모가 작은 저널일수록 더 높은 페이지랭크값이 산출되었다. 국내에서 물리학이나 화학과 같이 특정 학문 분야 내에 대규모 종합 저널이 존재하고 이 저널을 인용할 다른 저널이 드문 분야에는 불리한 특성이다. 다만 자기 인용을 조정하는 방식 중에서는 미미하나  $cr^2$  공식이 영향을 덜 받는 것으로 나타났다. 그러나 저널의 규모는 학문분야별로 매우 다른 특성을 보이는 변수이므로 개별 학문분야로 비교 범위를 제한하면 다른 결과가 나타날 여지도 있다.

측정 결과에서 상위권 저널 목록을 검토한 결과 제안된 방식을 적용하면 특정 저널이 자

기 인용에만 크게 의존해서 최상위 저널로 선정되는 것을 피할 수 있음이 확인되었다. 다만 특정 분야의 다수 저널이 최상위 10위권을 과점하는 현상이 발견되었는데, 이들 저널은 평균 인용빈도 기준으로도 상위 20위 내외에 해당하였기 때문에 페이지랭크 공식을 적용한 것이 결정적으로 작용한 것은 아니라고 추정된다. 국내에는 한 두 종의 대규모 저널이 주도하는 학문 분야가 있는가 하면 다수의 중소규모 저널이 공존하는 분야도 존재한다. 저널의 규모나 종수 뿐만 아니라 학문 분야별로 인용 관행도 상이하므로, 향후 분야별 특성이 국내 저널 인용 지수에 미치는 영향을 심층적으로 분석해 볼 필요가 있다.

나아가서 국내 종합 인용 DB를 운영하는 기관들이 최근 추진하고 있는 해외 인용 DB와의 연계가 본격화된다면 외부 저널로 향하는 인용뿐만 아니라 외부 저널에서 들어오는 인용을 함께 고려함으로써 국내 저널에 대한 더 완전한 인용 영향력 평가가 가능하도록 추진해야 할 것이다.

## 참 고 문 헌

김성희, 장로사. 2010. 사회 연결망 분석 연구동향 및 정보학 분야에서의 활용가능성에 관한 연구. 『정보관리학회지』, 27(4): 71-87.  
 김용학. 1987a. 사회연결망 이론의 기초개념. 『인문과학』, 58: 141-163.  
 김용학. 1987b. 사회 연결망 분석의 이론틀: 구

조와 행위 연결을 중심으로. 『한국사회학』, 21: 31-68.  
 김원진, 정영미. 2010. 과학기술분야 국제협력 증진을 위한 아시아 국가 간 공동연구 현황 분석. 『정보관리학회지』, 27(3): 103-123.  
 김은희, 정영미. 2010. 사용자 태그와 중심성 지수를 이용한 블로그 검색 성능 향상에 관한



- 연구. 『정보관리학회지』, 27(1): 61-77.
- 김관준, 이재윤. 2010. 학술지 영향력 측정을 위한 h-지수의 응용에 관한 연구. 『정보관리학회지』, 27(1): 269-287.
- 김희전, 조현양. 2010. 저자동시인용분석과 저자지결합분석에 의한 지적 구조 분석. 『정보관리학회지』, 27(3): 283-306.
- 김희정. 2011. 네트워크 분석을 기반으로 한 웹 아카이빙 주제영역 연구. 『한국비블리아학회지』, 22(2): 235-248.
- 박옥남. 2011. 정보조직 지식구조에 대한 연구. 『한국비블리아학회지』, 22(3): 247-267.
- 박한우. 2003. 월드와이드웹에 나타난 국제 학술 커뮤니케이션 네트워크에 대한 탐사적 연구. 『한국문헌정보학회지』, 37(2): 153-168.
- 송기호, 김태수. 2008. 교과 연계성 강화를 위한 학습주제 중심의 통합 정보활용교육과정에 관한 연구. 『정보관리학회지』, 25(3): 41-64.
- 양현, 이승도, 권영빈. 2011. 과학 문헌 인덱스 처리 시스템 구현: 정보과학회 논문지를 중심으로. 『정보과학회논문지: 컴퓨팅의 실제 및 레터』, 17(5): 289-298.
- 유사라. 1991. 네트워크 연구의 사회적 접근과 분석이론. 『도서관학』, 20: 297-330.
- 이수상, 위성광. 2009. 디지털 도서관 이용자의 검색행태 연구: 검색 로그 데이터의 네트워크 분석을 중심으로. 『한국도서관·정보학회지』, 40(4): 139-158.
- 이수상, 장임숙. 2008. 이주노동자의 사회적 서비스에 대한 접근성과 사회연결망. 『한국문헌정보학회지』, 42(4): 243-268.
- 이수상. 2011. 공저빈도에 따른 공저 네트워크의 속성 연구. 『한국도서관·정보학회지』, 42(2): 105-125.
- 이용재, 배화숙. 2008. 결혼이민자의 사회서비스 및 정보의 접근성에 관한 연구. 『한국도서관·정보학회지』, 39(4): 479-504.
- 이재윤. 2006a. 지적 구조의 규명을 위한 네트워크 형성 방식에 관한 연구. 『한국문헌정보학회지』, 40(2): 333-355.
- 이재윤. 2006b. 계량서지적 네트워크 분석을 위한 중심성 척도에 관한 연구. 『한국문헌정보학회지』, 40(3): 191-214.
- 이재윤. 2008. 연구자의 투고 학술지 현황에 근거한 국내 학문분야 네트워크 분석. 『정보관리학회지』, 25(4): 327-345.
- 이재윤. 2011. 인용 네트워크 분석에 근거한 문헌 인용 지수 연구. 『한국문헌정보학회지』, 45(2): 119-143.
- 이재윤, 최선희. 2010. 국내 학술지의 인용 네트워크 지수 측정. 『제17회 한국정보관리학회 학술대회 발표논문집』, 15-20.
- 정영미, 유소영. 2009. 내용 분석을 통한 한국의 학술적 웹 공간 구조 분석. 『정보관리학회지』, 26(3): 7-24.
- 조은성, 송재도. 2011. 국내외 마케팅 학술지의 영향력: Kor-Factor와 Impact Factor의 문제점을 중심으로. 『마케팅관리연구』, 16(2): 53-82.
- 최일영, 이용성, 김재경. 2010. 사회 네트워크 분석에 기반한 도서관 학술DB 이용 패턴 연구. 『정보관리학회지』, 27(1): 25-40.
- Al-Saffar, S. and G. Heileman. 2007. "Experimental bounds on the usefulness of

- personalized and topic-sensitive Page-Rank." *Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence*, 671-675.
- Bergstrom, C. T. 2007. "Eigenfactor: Measuring the value and prestige of scholarly journals." *C&RL News*, 68(5): 314-316.
- Biglu, M. H. 2008. "The influence of references per paper in the SCI to Impact Factors and the Matthew Effect." *Scientometrics*, 74(3): 453-470.
- Bollen, Johan, Marko A. Rodriguez, and Herbert Van de Sompel. 2006. "Journal status." *Scientometrics*, 69(3): 669-687.
- Brin, S. and L. Page. 1998. "The anatomy of a large-scale hypertextual Web search engine." *Computer Networks and ISDN Systems*, 30(1-7): 101-117.
- Brin, S., L. Page, R. Motwani, and T. Winograd. 1999. *The PageRank citation ranking: Bringing order to the Web*. Technical Report 1999-66, Stanford InfoLab.
- Chen, P., H. Xie, S. Maslov, and S. Redner. 2007. "Finding scientific gems with Google's PageRank algorithm." *Journal of Informetrics*, 1(1): 8-15.
- Choi, Seon-Heui, Byung-Kyu Kim, MuYeong Kang, Beom-Jong You, Jongwook Lee, and Jae-Won Park. 2011. "A study of citing patterns of Korean scientists on Korean journals." *Journal of the Korean Society for Information Management*, 28(2): 97-115.
- Chung, EunKyung. 2011. "Interdisciplinary collaborations in the domain of digital libraries." *Journal of the Korean Society for Information Management*, 28(2): 37-51.
- Chung, Young-Mee and Ji-Yeon Han. 2009. "Mapping knowledge structure of science and technology based on university research domain analysis." *Journal of the Korean Society for Information Management*, 26(2): 195-210.
- Franceschet, M. 2010a. "Journal influence factors." *Journal of Informetrics*, 4(3): 239-248.
- Franceschet, M. 2010b. "Ten good reasons to use the Eigenfactor™ metrics." *Information Processing and Management*, 46(5): 555-558.
- González-Pereira, B., V. P. Guerrero-Bote, and F. Moya-Anegón. 2010. "A new approach to the metric of journals' scientific prestige: The SJR indicator." *Journal of Informetrics*, 4(3): 379-391.
- Herskovic, Jorge R. and Elmer V. Bernstam. 2005. "Using incomplete citation data for MEDLINE results ranking." *Proceedings of the AMIA Annual Symposium*, 316-320.
- Kim, Byung-Kyu, MuYeong Kang, Seon-Heui Choi, Soon-Young Kim, Beom-Jong You, and Jae-Do Shin. 2011. "Citing behavior of Korean scientists on for-

- eign journals in KSCD." *Journal of the Korean Society for Information Management*, 28(2): 117-133.
- Kleinberg, J. 1998. "Authoritative sources in a hyperlinked environment." *Proceedings of the 9th ACM-SIAM Symposium on Discrete Algorithms*, 668-677.
- Kleinberg, J. 1999. "Authoritative sources in a hyperlinked environment." *Journal of the ACM*, 46(5): 604-632.
- Langville, Amy N. and Carl D. Meyer. 2005. "A survey of eigenvector methods for Web information retrieval." *SIAM Review*, 47: 135-161.
- Langville, Amy N. and Carl D. Meyer. 2006. *Google's PageRank and Beyond: The Science of Search Engine Rankings*. Princeton, NJ: Princeton University Press.
- Lee, Jae Yun and Sanghee Choi. 2011. "Intellectual structure and infrastructure of informetrics: Domain analysis from 2001 to 2010." *Journal of the Korean Society for Information Management*, 28(2): 11-36.
- Lee, Jeongsoo. 2010. "Social network approach for sharing knowledge." *Journal of the Korean Society for Information Management*, 27(2): 61-74.
- Scimago Research Group. 2007. *Description of SCImago Journal Rank Indicator*. [cited 2009.5.21].  
<<http://www.scimagojr.com/SCImagoJournalRank.pdf>>.
- Su, Cheng, YunTao Pan, JunPeng Yuan, Hong Guo, ZhengLu Yu, and ZhiYu Hu. 2009. "PageRank, HITS and Impact Factor for journal ranking." *Proceedings of the 2009 World Congress on Computer Science and Information Engineering*, 285-290.
- Teoma Web Search.  
<<http://www.teoma.com/>>.
- Xing, W. and A. Ghorbani. 2004. "Weighted pagerank algorithm." *Proceedings of the Second Annual Conference on Communication Networks and Services Research (CNSR'04)*, 305-314.

