학술지 인용 영향력 측정에서 즉시 인용 포함 방안

Considering immediate citations in journal impact measurements

이재윤, 명지대학교 문헌정보학과, memexlee@mju.ac.kr

Jae Yun Lee, Dept. of Library and Information Science, Myongji University

학술지 인용 영향력 지수 JIF에서는 학술논문이 발행된 당해년도에 즉시 인용된 횟수가 포함되지 않는다는 한계가 있다. 특정 연도에 발표된 이후의 모든 인용을 포함하는 IMP 지수는 논문발표 당해년도의 인용을 포함하지만, 몇 년 이전의 논문에 대한 평가가 된다는 한계가 있다. 이를 극복하기 위한 새로운 방안을 제안하였다.

1. 서 론

JCR, Scopus, KCI 등의 인용 데이터베이스에서 제공하는 학술지 인용 지수 중에서 가장 대표적인 것은 논문 당 평균 인용빈도에 해당하는 지수이다. 이를 JCR에서는 학술지 영향력 지수(Journal Impact Factor), KCI에서는 영향력 지수, Scopus에서는 CiteScore라고 각각 부른다. 이 지수들은 모두 지난 몇 해 동안(보통은 2~3년) 발표된 논문이 특정 연도 한 해 동안 인용된 횟수로 산출된다. 그런데 논문이 발표된 당해년도에 인용된 횟수는 평균 산출에 반영되지 않으며, 별도로 즉시성 지수(Immediacy Index)라는 지수로 보고되지만 거의 활용되지 않고 있다. 당해년도의 즉시 인용이 영향력 지수에 반영되지 않는 이런 한계 때문에 대안으로 제안된 지수가 통시적 영향력 지수 (Diachronous Impact Factor) IMP이다(Ingwersen, Larsen, Rousseau, & Russell, 2001). 최근 한국연구재단에서는 영향력 지수와 함께 통시적 영향력 지수 IMP도 산출하여 공개하고 있으며 KCI 등재지 평가에서도 참고하도록 권고하고 있다.

영향력 지수는 여러 해에 걸쳐 발표된 논문이 특정 연도에 인용된 횟수를 기준으로 한다는 점에서 공시적 영향력 지수라고도 부른다. 이와 달리 통시적 영향력 지수는 특정 연도에 발표된 논문이 이후 여러 해에 걸쳐 인용된 횟수를 측정한다.

IMP는 논문이 발표된 당해년도의 인용을 자연스럽게 포함한다는 장점이 있으나(이재윤, 2014), 직전 연도가 아닌 몇 해 전에 발표된 논문에 대한 평가가 된다는 단점이 있다. 실제로 KCI에서 제공하는 인용지수는 가장 최근의 영향력 지수가 2016년 기준이지만 통시적 영향력 지수 IMP는 2014년 기준으로 산출된 것이 가장 최근 지수이다.

이 연구에서는 영향력 지수처럼 직전 연도까지 발표된 논문까지 포함하되, IMP처럼 논문이 발표된 당해년도의 즉시인용도 포함하여 학술지 영향력을 측정하는 새로운 방안을 제안하고 KCI 데이터를 대상으로 적용해보았다.

2. KCI의 영향력 지수와 즉시성 지수

KCI에서 인용지수가 공개된 학술지 중에서 ISSN을 활용해서 2013년부터 2016년까지 인용지수가 지속적으로 확보된 학술지는 1976종이다. 이들의 2년 단위 영향력지수(IF2Y)와 즉시성 지수(II), 그리고 즉시성 지수가 영향력 지수에 비해 어느 정도 비율인지를 산출해보면 다음 표와 같다.

<표 1> KCI 1976종 학술지의 인용지수 평균과 중앙값

	IF2Y	II	IF2Y 대비 II의 비율
2016 중앙값	0.562	0.143	0.254
2016 평균	0.676	0.182	0.270
2015 중앙값	0.570	0.136	0.239
2015 평균	0.690	0.171	0.248

이 표에서 보듯이 즉시성 지수 평균은 2년 단위 영향력 지수 평균의 1/4 정도에 달한다. 그런데 이 비율은 전체 학술지 평균을 비교한 것이고, 각 학술지 별로 즉시성 지수와 영향력 지수를 비교해보면 학술지나 분야별로 상당한 편차가 있음을 알 수 있다. 2016년 지수로 발표된 KCI 데이터에 대해서 각 학술지별로 영향력 지수 대비 즉시성 지수의 비율을 산출한 결과를 KCI의 8개 대분류별로 분포를 살펴보면 <표 2>와 같다.

이 표에서 보듯이 개별 학술지마다 산출해 보면 공학 분야나 사회과학 분야 최댓값처럼 즉시성 지수가 영향력 지수의 5~6배 이상 되는 경우도 있다. 흔히 최근 논문이 상대적으로 덜 인용되는 것으로 알려져 있는 인문학 분야에서도 즉시성 지수가 영향력 지수의 2.25배나 되는 학술지가 있고, 평균이 0.319이다. 이로서 논문이 발표된 당해년도의 즉시 인용을 학술지 영향력 평가에 반영하지 않는 것은 공정하지 않은 평가라고 할 수 있다.

3. 통시적 학술지 영향력 지수 IMP

영향력 지수 JIF와 달리 당해연도의 즉시 인용을 반영하는 기존 지수로는 IMP라고 약칭하는 통시적 영향력 지수(Diachronous Impact Factor)가 있다(Ingwersen et al., 2001). KCI에서는 학술지 인용지수를 공표하는 페이지에서 '통시적 저널 영향력 지수'라는 이름으로 IMP 측정 결과를 공개하고 있다. IMP는 특정 연도에 발행된 논문이 이후 여러 해에 걸쳐 인용된 것으로 영향력을 측정하는 방식이다. IMP 산출 공식은 다음과 같다(Ingwersen et al., 2001).

$$IMP_n(Y) = \frac{\sum_{i=k}^{k+n-1} CIT(Y+i, Y)}{PUB(Y)}$$

여기서 k는 0이나 1로 설정할 수 있으며 0이면 발행년을 포함하고 1이면 제외하게 된

<표 2> KCI 학술지 2016년 영향력 지수 대비 즉시성 지수 비율의 분야별 분포

	최솟값	1사분위수	중앙값	평균	3사분위수	최댓값
공학	0.000	0.178	0.288	0.355	0.414	6.471
농수해양학	0.000	0.182	0.293	0.321	0.395	1.426
복합학	0.000	0.158	0.272	0.289	0.390	1.750
사회과학	0.000	0.146	0.252	0.300	0.372	5.140
예술체육학	0.000	0.122	0.236	0.293	0.331	2.189
의약학	0.000	0.097	0.184	0.249	0.324	1.829
인문학	0.000	0.155	0.283	0.319	0.439	2.250
자연과학	0.000	0.143	0.244	0.302	0.386	1.259

다. KCI에서는 발행년도의 즉시 인용을 포함하기 위해서 k를 0으로 설정한 IMP 지수가 발행 후 2년부터 4년까지의 인용에 대해서 산출되어 발표되고 있다.

학술지 영향력 지수 IF와 IMP는 <표 3>과 같은 발행-인용 행렬(Publication-Citation Matrix; Ingwersen et al., 2001)로 산출 과정의 차이를 비교할 수 있다.

<표 3> 2012~2016년 발행된 정보관리학회지 논문의 발행-인용 행렬

	발행년				
	2012	2013	2014	2015	2016
논문수	57	55	50	56	53
2012인용횟수	9				
2013인용횟수	56	22			
2014인용횟수	45	72	19		
2015인용횟수	50	113	62	16	
2016인용횟수	44	110	87	41	15

영향력 지수는 <표 3>의 발행-인용 행렬에서 특정한 인용년도에 해당하는 가로줄의 숫자를 합산하여 평균을 산출한 것이다(이재윤, 2014). 예를 들어 정보관리학회지의 2016년 영향력 지수는 <표 3>의 마지막 행에서 2014년 논문과 2015년 논문이 인용된 횟수를 논문 수로 나눈 결과가 된다.

$$IF_2(2016) = \frac{87+41}{50+56} = 1.208$$

그러나 이와 같은 산출 방식에서는 <표 3>의 발행-인용 행렬에서 대각선에 위치한 당해 연도의 인용횟수(9, 22, 19, 16, 15)는 모두 영향력 산출에서 배제된다.

2014년에 발행된 정보관리학회지 논문이 2016년까지 인용된 평균 빈도인 통시적 저널 영향력 지수 $IMP_3^{(0)}(2014)$는 <표 3>에서 발행년 2014년 열의 세로줄 숫자인 2014년, 2015년, 2016년 인용횟수를 합산하여 논문 수로 나눈 결과가 된다.

$$IMP_3^{(0)}(2014) = \frac{19+62+87}{50} = 3.36$$

이처럼 IMP는 2014년에 발행된 논문이 2014년에 인용된 횟수를 포함한다. 따라서 <표 3>의 대각선에 위치한 당해년도 인용이 모두 지수 산출에 반영된다는 장점이 있다. 반면에 2016년 현재의 시점에서 최근 논문인 2015년 논문이 인용된 횟수는 평가에 반영하지 않는다는 한계가 있다.

4. 당해년도 인용과 최근 논문 인용을 모두 포함하는 지수 산출 방안

이 연구에서는 당해년도 인용과 직전 연도 발행된 논문의 인용까지 포함하는 지수를 새롭게 정의해보기로 한다.

2년 단위의 지수 산출로 전제하면, 직전 2년 동안 발행된 논문에 대해서 차년도까지 인용된 횟수, 차차년도 인용된 횟수의 합계를 분자로 하고, 이를 논문 수의 합계로 나눠준다. 이때 2년 전 발행 논문은 차년도까지 인용된 횟수와 차차년도 인용된 횟수가 함께 분자에 포함되므로 분모의 논문 수도 2배로 반영한다. 1년 전 발행 논문은 차년도까지 인용된 횟수만 분자에 포함되므로 분모의 논문 수는 그대로 반영한다.

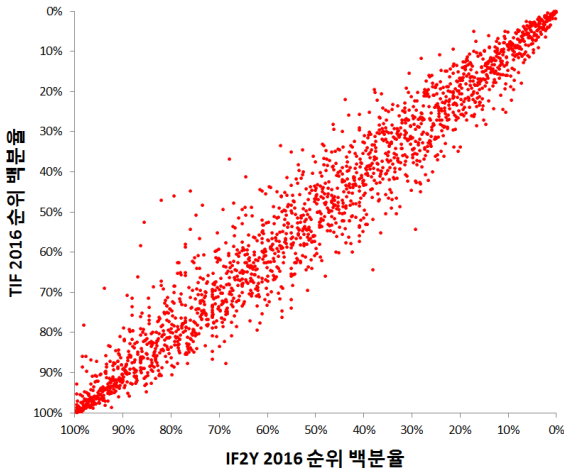
<표 3>의 정보관리학회지를 예로 들면, 2016년을 기준으로 2년 전인 2014년 발행된 50편의 논문이 차년도인 2015년까지 인용된 횟수(19+62)와 차차년도인 2016년에 인용된 횟수(87), 그리고 2015년 발행된 56편의 논문이 차년도인 2016년까지 인용된 횟수(16+41)을 합산하여 분자로 삼고, 2014년 발행된 논문수(50)를 두 번 더하고 2015년 발행된 논문수(56)를 더하여 분모로 삼아서 연평균 인용횟

수를 산출한다. 그 결과 2016년 정보관리학회지의 직전 2년 동안 발행논문의 연평균 인용빈도는 다음과 같이 1.442로 산출된다.

$$\text{연평균 인용빈도} = \frac{19+62+87+16+41}{50+50+56} = \frac{225}{156} = 1.442$$

기존 영향력 지수와 달리 당해년도의 즉시 인용까지 포함하므로 이를 총 영향력 지수(Total Impact Factor)라고 부르기로 한다.

시험적으로 KCI에서 반출이 가능하도록 공개된 인용지수 데이터에서 여러 해에 걸쳐 지속적인 지수 산출이 가능한 1,976종의 학술지를 대상으로 2년 단위 2016년 영향력 지수(IF2Y2016)와 총 영향력 지수(TIF2016)를 각각 구해보았다. 두 지수 간 피어슨 상관계수는 0.975, 스피어맨 순위상관계수는 0.974로서 매우 유사한 순위로 나타났다. 각 학술지가 전체 학술지에서 차지하는 지수별 순위의 백분율 비교 결과를 그림으로 나타내면 다음과 같다.



<그림 1> KCI 학술지의 2016년 영향력 지수 순위와 2016년 총 영향력 지수 순위의 비교

<그림 1>을 보면 영향력 지수와 총 영향력 지수가 뚜렷하게 비례하는 경향을 보이지만, 일부 학술지의 경우는 추세에서 벗어나 크게 순위가 달라지기도 한다는 것을 알 수 있다.

예를 들어 <고문화>는 2016년 영향력 지수 순위 753위(상위 38.1%)에서 총 영향력 지수 순위 1,271(상위 64.3%)로 크게 하락하는 반면에, <비교민주주의연구>는 2016년 영향력 지수 순위 1,622위(상위 82.1%)에서 총 영향력 지수 순위 929(상위 47.0%)로 크게 상승하였다. 이처럼 즉시 인용을 포함할 경우에 개별 학술지의 영향력 순위가 크게 달라지는 경우도 있음이 확인되었다.

5. 결 론

즉시 인용을 포함하는 총 영향력 지수는 차년도까지의 인용 횟수를 반영하는 과정에서 연초에 발행된 논문이 연말에 발행된 논문에 비해서 인용될 기회를 더 많이 가지게 되므로 완벽한 지수라고 할 수는 없다. 그러나 기존의 영향력 지수는 당해년도에 인용된 횟수를 배제함으로써 오히려 연초에 발행된 논문에 불리하다는 한계를 가진다. 따라서 총 영향력 지수와 기존 영향력 지수를 상호 보완적으로 활용하는 것이 바람직하다.

참고문헌

이재윤 (2014). 통시적 저널 영향력 지수에 대한 고찰. 제21회 한국정보관리학회 학술대회 논문집, 3-6.

Ingwersen, P., Larsen, B., Rousseau, R., & Russell, J. (2001). The publication-citation matrix and its derived quantities. Chinese Science Bulletin, 46(6), 524-528.