

tnet과 WNET의 가중 네트워크 중심성 지수 비교 연구*

A Comparison Study on the Weighted Network Centrality Measures of tnet and WNET

이재윤 (Jae Yun Lee)**

초 록

이 연구에서는 공개된 가중 네트워크 분석용 소프트웨어인 Opsahl의 tnet과 이재윤의 WNET에서 지원하는 가중 네트워크 중심성 지수를 비교 분석해보았다. tnet은 가중 연결정도중심성, 가중 근접중심성, 가중 매개중심성을 지원하고, WNET은 최근접이웃중심성, 평균연관성, 평균프로파일연관성, 삼각매개중심성을 지원한다. 가상 데이터를 대상으로 한 분석에서 tnet의 중심성 지수는 링크 가중치의 선형 변화에 민감한 반면 WNET의 중심성 지수는 선형 변화에 영향을 받지 않았다. 실제 네트워크 6종을 대상으로 가중 네트워크 중심성을 측정하고 결과를 비교하여 두 소프트웨어의 가중 네트워크 중심성 지수들의 특징을 파악하고 중심성 지수 간 관계를 살펴보았다.

ABSTRACT

This study compared and analyzed weighted network centrality measures supported by Opsahl's tnet and Lee's WNET, which are free softwares for weighted network analysis. Three node centrality measures including weighted degree, weighted closeness, and weighted betweenness are supported by tnet, and four node centrality measures including nearest neighbor centrality, mean association, mean profile association, triangle betweenness centrality are supported by WNET. An experimental analysis carried out on artificial network data showed tnet's high sensitiveness on linear transformations of link weights, however, WNET's centrality measures were insensitive to linear transformations. Seven centrality measures from both tools, tnet and WNET, were calculated on six real network datasets. The results showed the characteristics of weighted network centrality measures of tnet and WNET, and the relationships between them were also discussed.

키워드: 네트워크 분석, 중심성 지수, 가중 네트워크, tnet, WNET
network analysis, centrality measures, weighted networks, tnet, WNET

* 이 논문은 2013년도 명지대학교 교내연구비 지원사업에 의하여 연구되었음.

** 명지대학교 문헌정보학과 부교수(memexlee@mju.ac.kr)

■ 논문접수일자 : 2013년 11월 30일 ■ 최초심사일자 : 2013년 12월 10일 ■ 게재확정일자 : 2013년 12월 15일
■ 정보관리학회지, 30(4), 241-264, 2013. [http://dx.doi.org/10.3743/KOSIM.2013.30.4.241]

1. 서론

네트워크 분석에 대한 관심이 증가하면서 여러 학문 분야에서 네트워크 분석을 적용한 연구가 활발히 발표되고 있다. 최근에는 논문 발표를 넘어서 고고학(Knappett, 2013), 정책학(Vedres & Scotti, 2012), 건축학(Pryke, 2012), 경제학(Knoke, 2012) 등의 분야에서 방법론과 적용 사례를 정리한 단행본이 발간되는 상황에 이르렀다.

네트워크 분석 기법의 적용이 확산되면서 대상 데이터도 사회학에서 다루어졌던 이진 네트워크 데이터를 벗어나서 관계의 강도가 구분되는 가중 네트워크 데이터로 넓어지고 있다. 인터넷 시대에 들어서 인터넷 서점이나 포털 사이트, 디지털도서관과 같은 환경에서 네트워크 데이터의 수집이 자동화되면서 동시출현에 근거한 관계의 강도가 손쉽게 측정됨에 따라 가중 네트워크를 분석해야 하는 경우는 더 증가하고 있다. 전통적으로 동시인용을 비롯한 가중 관계를 다루어왔던 계량서지적 네트워크에 대한 관심도 여러 분야로 확산되고 있다.

전통적인 이진 네트워크에서 각 노드의 입지를 분석하는 지수를 이진 네트워크 중심성 지수라고 한다면, 가중 네트워크를 대상으로 각 노드의 입지를 분석하는 지수는 가중 네트워크 중심성 지수라고 할 수 있다. 이진 네트워크 중심성 지수는 다양한 지수가 제안되어 있고 기존의 여러 네트워크 분석 도구에서 손쉽게 측정할 수 있는 반면, 가중 네트워크 중심성 지수는 아직까지 제안된 지수의 종류도 적고 손쉽게

측정할 수 있는 도구도 드물다. 현재까지 공개된 가중 네트워크 중심성 지수 측정 도구는 tnet과 WNET이 있다.

tnet¹⁾은 T. Opsahl이 통계처리용 공개 프로그래밍 언어인 R의 패키지로 공개한 가중 네트워크 분석용 함수 모음이다. tnet에서 분석 가능한 가중 네트워크 중심성 지수는 가중 연결정도 중심성, 가중 근접중심성, 가중 매개중심성이 있다(Opsahl, Agneessens, & Skvoretz, 2010). tnet의 가중 네트워크 중심성 지수는 기존의 이진 네트워크 중심성 지수를 일반화하여 가중 네트워크에서 사용할 수 있도록 변형하는 방식으로 개발되었다. tnet은 고고학 분야(Peebles & Roberts Jr., 2013), 경제학 분야(Oatley, Winecoff, Pennock, & Danzman, 2013), 의학 분야(Gómez, Nunn, & Verdú, 2013) 등의 여러 영역에서 활용되고 있다. 국내에서는 축산학 분야에서 이형진, 서교, 정남수, 이인복, 서일환, 문운경, 이정재(2012)가 tnet을 이용해서 전북 김제 지역의 39개 가금 농가 사이의 질병 확산 네트워크를 분석하였다.

WNET²⁾은 이재윤이 가중 네트워크 분석 도구 모음으로 개발한 WNETs Tools에 포함된 프로그램이다. 정방대칭행렬 형태로 표현된 네트워크 데이터를 WNET의 데이터로 입력하면 복잡한 가중 네트워크를 축약하여 시각화하기 위한 패스파인더 네트워크(Schvaneveldt, 1990; 이재윤, 2006a) 링크 목록 파일(출력 파일명이 "PFNET-"으로 시작), 병렬 최근접 이웃 연결 클러스터링(이재윤, 2006c) 결과 파일(출력 파일명이 "PNNC-"로 시작), 가중 네트워크 중심

1) <http://toreopsahl.com/tnet/software/>에서 획득 가능.

2) memexlee@mju.ac.kr로 연락하여 획득 가능.

성 측정 결과 파일(출력 파일명이 "WCENT-"로 시작)의 세 가지 파일을 생성한다. WNET에서 산출하는 가중 네트워크 중심성은 이재운(2006b)이 소개한 최근접이웃중심성, 평균연관성, 평균프로파일연관성, 삼각매개중심성의 네 가지 지수이다. WNET은 독서(김관준, 2011), 전기공학(변지혜, 정은경, 2011), 계량정보학(Lee & Choi, 2011), 폭소노미(이재운, 2012b), 의료정보학(허고은, 송민, 2013) 등 여러 분야의 지적 구조 분석에서 활용되었고 대학 학과(최상희, 2013), 개인 연구자(이재운, 2012a; 정선영, 2013) 등의 연구 활동을 분석하는 데에도 적용되었다. 삼각매개중심성과 최근접이웃중심성은 공저 네트워크(김원진, 정영미, 2010), 유사성 네트워크(이재운, 2008; 이기황, 이재운, 백두현, 2010) 등을 분석하는데도 활용되었다.

tnet의 가중 네트워크 중심성 지수는 기존의 이진 네트워크 중심성 지수를 가중 네트워크에 적용할 수 있도록 일반화하여 확장하는 방식을 취한 반면, WNET의 가중 네트워크 중심성 지수는 기존 중심성 지수와 무관하게 개발되었다. 본질적으로 tnet의 가중 네트워크 중심성 지수는 네트워크의 밀도가 낮은 경우를 고려하고 있는 반면, WNET의 가중 네트워크 중심성 지수는 네트워크의 밀도가 매우 높은 경우를 전제로 개발되었다. 이는 양쪽 지수를 제안한 논문에서 제시된 가상 사례에서도 드러난다. tnet에 사용되는 중심성 지수를 소개한 Opsahl과 Agneessens, Skvoretz(2010)은 6개의 노드와 6개의 링크로 구성되어 밀도가 0.4인 가상 사례를 제시하면서 지수를 설명한 반면에 WNET에서 사용되는 중심성 지수를 제안한 이재운(2006b)은 4개의 노드와 4개의 링크로 구성되어 밀도가 1인 가상

사례를 제시하면서 지수를 설명하였다.

이 논문에서는 이처럼 가중 네트워크의 중심성 분석에 동일하게 적용될 수 있지만 개발 배경과 접근 방식이 다른 tnet과 WNET의 가중 네트워크 중심성 지수를 비교해보고자 한다. 이를 위해서 두 도구에서 산출하는 가중 네트워크 중심성 지수의 개념과 특징을 먼저 살펴보고, 가상 네트워크와 실제 네트워크에 대해 tnet과 WNET를 적용해서 결과를 비교 분석하였다.

2. 가중 네트워크 중심성 지수의 정의와 특징

2.1 tnet 가중 네트워크 중심성 지수

2.1.1 가중 연결정도중심성

이진 네트워크에서 노드 i 의 연결정도중심성(Degree Centrality; 이하 C_D , 또는 C_d 로 약칭)은 다음 공식으로 나타낼 수 있다.

$$k_i = C_D(i) = \sum_{j \neq i}^N x_{ij}$$

이 공식에서 x_{ij} 는 노드 i 와 노드 j 가 연결되어 있으면 1, 아니면 0이다. 가중 네트워크를 다룰 때 연결정도 중심성은 링크의 가중치를 합산하는 것으로 변경하여 적용되는 경우가 흔하다(Newman, 2004; Opsahl, Colizza, Panzarasa, & Ramasco, 2008). 노드 i 의 링크 가중치 합계는 흔히 노드 강도(Node Strength; 이하 NS로 약칭)라고 불리며(Barrat, Barthelemy, Pastor-Satorras, & Vespignani, 2004) 공식으로 나타내면 다음과 같다.

$$S_i = C_D^w(i) = \sum_{j \neq i}^N w_{ij}$$

그러나 가중치를 단순히 합산하는 것은 연결된 링크의 수를 반영하지 못하기 때문에 둔감한 척도가 되며(Opsahl, Agneessens, & Skvoretz, 2010), 사실상 지역 중심성이 아닌 전역 중심성을 측정하는 결과가 되어 지수의 성격이 바뀌게 된다(이재윤, 2006b)는 비판이 있다. Opsahl, Agneessens, Skvoretz(2010)는 연결정도 중심성과 노드 강도를 결합하는 가중 연결정도중심성(Weighted Degree Centrality; 이하 C_{WD} , 또는 C_{wd} 로 약칭)을 다음과 같이 제안하면서 두 값의 상대적인 중요도를 파라미터 α 로 조절하도록 하였다.

$$C_{D\alpha}^w(i) = k_i \times \left(\frac{s_i}{k_i}\right)^\alpha = k_i^{(1-\alpha)} \times s_i^\alpha$$

가중 연결정도중심성 공식에서 α 가 0이면 전통적인 연결정도 중심성과 같아지고, 1이면 노드 강도와 같은 공식이 된다. 만약 α 가 1보다 큰 값이 되면 연결정도가 낮을수록 오히려 더 높은 중심성을 가지게 된다(Opsahl, Agneessens, & Skvoretz, 2010). 이는 연결정도가 높을수록 중심성이 높다는 직관에 반하는 성향이므로 가중 연결정도중심성을 측정할 때에는 α 를 1보다 낮은 값으로 설정하는 것이 바람직하다. 일반적으로는 연결정도와 노드 강도를 균등하게 반영하도록 α 를 0.5로 설정하는 것이 무난하다.

2.1.2 가중 근접중심성과 가중 매개중심성 가중 근접중심성(Weighted Closeness Cen-

trality; 이하 C_{WC} , 또는 C_{wc} 로 약칭)과 가중 매개중심성(Weighted Betweenness Centrality; 이하 C_{WB} , 또는 C_{wb} 로 약칭)은 가중 네트워크에서 최단 경로거리를 산출하는 방법을 기반으로 정의되었다. 이진 네트워크 중심성인 근접중심성과 매개중심성이 노드 간 최단 경로거리에 기반하여 산출하는 지표이기 때문이다. 이진 네트워크에서 노드 간 경로거리는 연결 단계의 수로 측정되지만 tnet에서는 해당 경로를 구성하는 각 링크 가중치의 역수를 합산한 값을 가중 네트워크에서의 경로거리로 간주하였다. 노드 i 와 j 사이의 최단 경로거리를 공식으로 나타내면 다음과 같다(Opsahl, Agneessens, & Skvoretz, 2010).

$$d^{w\alpha}(i,j) = \min\left(\frac{1}{(w_{ih})^\alpha} + \dots + \frac{1}{(w_{hj})^\alpha}\right)$$

여기서 α 는 가중치를 반영하는 정도를 조절하는 파라미터이다. α 가 0이면 경로의 단계 수가 그대로 노드 사이의 경로 거리가 되는 이진 네트워크 중심성이 되고, α 가 1보다 작으면 약한 가중치를 지닌 짧은 경로가 최단거리로 선호되며, α 가 1보다 큰 값이 되면 강한 가중치를 지닌 긴 경로가 최단거리로 선호되는 성향을 가지게 된다. 이는 마치 도로주행 네비게이션에서 주행 경로를 선택할 때 거리(연결 단계)를 우선할 것인지, 시간(링크 가중치)을 우선할 것인지 고려하는 상황과 유사하다. 가중 네트워크에서의 경로거리를 기반으로 가중 근접중심성과 가중 매개중심성을 정의하면 각각 아래의 공식과 같다(Opsahl, Agneessens, & Skvoretz, 2010).

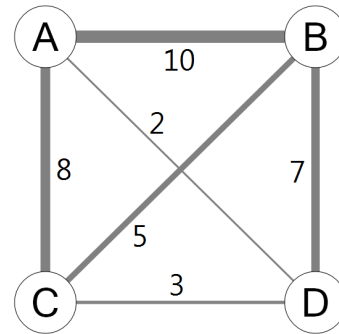
$$C_C^{wa}(i) = \left[\sum_j^N d^{wa}(i,j) \right]^{-1}$$

$$C_B^{wa}(i) = \frac{g_{jk}^{wa}(i)}{g_{jk}^{wa}}$$

가중 매개중심성 공식에서 g_{jk}^{wa} 는 i 를 제외한 다른 노드들 사이의 최단 경로의 수이고 $g_{jk}^{wa}(i)$ 는 그중에서 노드 i 를 통과하는 경로의 수이다. N 개의 노드로 구성된 네트워크에서 특정한 두 노드 사이의 최단 경로가 둘 이상인 경우가 없다면 가중 매개중심성 공식의 분모는 노드 i 자신을 제외한 나머지 노드들의 결합쌍의 수인 $(N-1) \times (N-2) / 2$ 이 되고 분자는 노드 i 를 거치는 최단 경로의 수가 된다. 공식에 나타나 있듯이 가중 근접중심성은 다른 노드들과의 가중 경로거리 합계의 역수이고 가중 매개중심성은 다른 노드들 사이의 최단 가중 경로거리 사이에 해당 노드가 위치하는 비율이다. 가중 매개중심성 공식의 분모는 최단 경로의 총 숫자에 해당하므로 네트워크의 규모를 고려하여 정규화하는 효과가 있는데, 실제 tnet 라이브러리에서 산출되는 가중 매개중심성 지수값은 공식의 분모항을 사용하지 않고 분자항인 해당 노드를 거치는 최단 경로의 수에 해당하는 값이다.

예를 들어 <그림 1>과 같이 노드 4개로 구성된 가상의 비방향성 가중 네트워크에서 tnet의 가중 네트워크 중심성을 측정하기 위해서는 <표 1>과 같이 양방향 링크를 모두 표현한 링크 목록 형태로 네트워크를 표현하여 tnet에 입력하면 된다. tnet에서 중심성 분석을 원활히 수행하기 위해서는 노드명을 기호가 아닌 일련번호와 같은 숫자로 표기하는 것이 좋다. 비방향성 네트워크여서 한쪽 방향 링크만 표현한 경우에

는 링크 목록을 R의 데이터 테이블로 읽어들이고 다음 `symmetrise_w()` 함수를 이용해서 양방향 링크 목록을 변환한 후 분석하면 된다.



<그림 1> 노드 4개로 구성된 가상 네트워크(숫자는 링크 가중치)

<표 1> 노드 4개로 구성된 가상 네트워크의 링크 리스트 표현

A	B	10
A	C	8
A	D	2
B	C	5
B	D	7
C	D	3
B	A	10
C	A	8
D	A	2
C	B	5
D	B	7
D	C	3

<그림 1>의 가상 네트워크에 대한 tnet의 이진 네트워크 중심성과 가중 네트워크 중심성 계산 결과는 <표 2>와 같다. 이진 네트워크로 간주하면 이 네트워크는 모두가 서로 연결된 클릭이므로 이진 네트워크 중심성은 네 노드가 동일하다. 링크 가중치의 합계인 노드 강도는 노드

B가 22로 가장 높고 노드 A가 20으로 두 번째이다. 가중 연결정도중심성($\alpha=0.5$)과 가중 근접중심성($\alpha=0.5, \alpha=1.0$)은 노드 강도의 순서와 동일한 결과를 보여준다(〈표 2〉 참조).

흥미로운 것은 가중 매개중심성 산출 결과이다. $\alpha=0.5$ 인 경우와 1.0인 경우 모두 노드 B만 가중 매개중심성이 1로 측정되고 나머지 노드는 모두 0이 된다. 노드 B의 경우 다른 세 노드들 사이의 최단경로 세 건(A와 C 사이의 최단 경로, A와 D 사이의 최단 경로, C와 D 사이의 최단 경로) 중에서 A와 D 사이의 최단 경로 중간에 위치하게 되므로 가중 매개중심성이 1이 된다. 〈그림 1〉의 가상 네트워크에 속한 네 노드 사이에 가능한 조합 6쌍 사이를 연결하는 경로들의 거리를 모두 계산해보면 〈표 3〉과 같다. 이 표에서 보듯이 A와 D를 연결하는 경우에만 직접 연결된 경로가 최단 경로가 아니고 나머지 5쌍 사이를 연결하는 경우는 모두 직접 연결된 경로가 최단 경로에 해당한다. 두 노드 사이에 직접 연결된 경로는 매개하는 중간 노드가 없기 때문에 직접 경로가 최단 경로일 경우에는 다른 노드가 매개중심성을 높일 수가 없다. 이처럼 가중 네트워크에서 모든 노드가 서로 연결된 클릭의 형태이면서 링크 가중치의 차이가 크지 않은 경우에는 직접 경로가 최단 경로가 되어 각 노드의 매개중심성이 제대로 산출되지 못할

가능성이 있다.

2.2 WNET 가중 네트워크 중심성 지수

2.2.1 최근접이웃중심성

최근접이웃중심성은 다른 노드에 의해 최근접이웃으로 꼽히는 정도를 말한다. 처음 제안되었을 때 (이재윤, 2006b)에는 최근접중심성(Nearest Centrality; 이하 C_N , 또는 C_n 으로 약칭)으로 표기하였으나 WNET 프로그램에서는 의미를 분명하게 하기 위해서 최근접이웃중심성(Nearest Neighbor Centrality)로 표기하고 있다. WNET에서는 “NNC”라는 항목으로 출력된다. 노드 i 의 최근접이웃중심성을 구하는 공식은 다음과 같다.

$$C_{TB}(i) = \sum_j f(r_{ji}),$$

$$f(r_{ji}) = \begin{cases} 1 & \text{if } w_{ji} = \text{MAX}(w_{j1}, \dots, w_{jn}) \\ 0 & \text{else} \end{cases}$$

이 공식에서 r_{ji} 는 노드 j 와 연결된 이웃 노드 중에서 노드 i 가 몇 번째로 가까운 노드인가를 측정한 순위이다. $f(r_{ji})$ 는 r_{ji} 가 1위, 즉 노드 i 가 노드 j 의 최근접이웃일 경우에만 1이고 나머지는 0이므로 이 공식은 결국 노드 i 를 1위로 꼽은 다른 노드의 숫자를 세는 것이다. 최근접중심성

〈표 2〉 노드 4개로 구성된 가상 네트워크의 각 지수별 수치

node	NS	C_D	C_C	C_B	Cwd		Cwc		Cwb	
					$\alpha=0.5$	$\alpha=1$	$\alpha=0.5$	$\alpha=1$	$\alpha=0.5$	$\alpha=1$
A	20	3	0.111	0.00	7.75	0.105	0.122	0.00	0.00	
B	22	3	0.111	0.00	8.12	0.125	0.129	1.00	1.00	
C	16	3	0.111	0.00	6.93	0.104	0.087	0.00	0.00	
D	12	3	0.111	0.00	6.00	0.086	0.079	0.00	0.00	

〈표 3〉 노드 4개로 구성된 가상 네트워크에서 각 노드 사이의 가능한 경로와 경로 거리

출발 노드~도착 노드	경로 번호	경로 구성 링크	경로 거리		비고
			계산식	값	
A~B	경로 1	(A,B)	1/10	0.100	직접 경로 겸 최단 경로
	경로 2	(A,C), (C,B)	1/8 + 1/5	0.325	
	경로 3	(A,D), (D,B)	1/2 + 1/7	0.643	
	경로 4	(A,C), (C,D), (D,B)	1/8 + 1/3 + 1/7	0.601	
	경로 5	(A,D), (D,C), (C,B)	1/2 + 1/3 + 1/5	1.033	
A~C	경로 6	(A,C)	1/8	0.125	직접 경로 겸 최단 경로
	경로 7	(A,B), (B,C)	1/10 + 1/5	0.300	
	경로 8	(A,D), (D,C)	1/2 + 1/3	0.833	
	경로 9	(A,B), (B,D), (D,C)	1/10 + 1/7 + 1/3	0.576	
	경로 10	(A,D), (D,B), (B,C)	1/2 + 1/3 + 1/5	1.033	
A~D	경로 11	(A,D)	1/2	0.500	직접 경로
	경로 12	(A,B), (B,D)	1/10 + 1/7	0.243	최단 경로
	경로 13	(A,C), (C,D)	1/8 + 1/3	0.458	
	경로 14	(A,B), (B,C), (C,D)	1/10 + 1/5 + 1/3	0.633	
	경로 15	(A,C), (C,B), (B,D)	1/8 + 1/2 + 1/7	0.768	
B~C	경로 16	(B,C)	1/5	0.200	직접 경로 겸 최단 경로
	경로 17	(B,A), (A,C)	1/10 + 1/8	0.225	
	경로 18	(B,D), (D,C)	1/7 + 1/3	0.476	
	경로 19	(B,A), (A,D), (D,C)	1/10 + 1/2 + 1/3	0.933	
	경로 20	(B,D), (D,A), (A,C)	1/7 + 1/2 + 1/8	0.768	
B~D	경로 21	(B,D)	1/7	0.143	직접 경로 겸 최단 경로
	경로 22	(B,A), (A,D)	1/10 + 1/2	0.600	
	경로 23	(B,C), (C,D)	1/5 + 1/3	0.533	
	경로 24	(B,A), (A,C), (C,D)	1/10 + 1/8 + 1/3	0.558	
	경로 25	(B,C), (C,A), (A,D)	1/5 + 1/8 + 1/2	0.825	
C~D	경로 26	(C,D)	1/3	0.333	직접 경로 겸 최단 경로
	경로 27	(C,A), (A,D)	1/8 + 1/2	0.625	
	경로 28	(C,B), (B,D)	1/5 + 1/7	0.343	
	경로 29	(C,A), (A,B), (B,D)	1/8 + 1/10 + 1/7	0.368	
	경로 30	(C,B), (B,A), (A,D)	1/2 + 1/10 + 1/2	1.100	

은 각 이웃 노드마다 상대적으로 가장 강한 링크만을 남겼을 때 해당 노드가 가지게 되는 링크의 수를 세는 것이므로 가중 네트워크에서 지역 중심성을 측정하는 지표로 사용된다. 김관준(2011)은 저자 프로파일링을 이용해서 독서분야를 연구하는 저자들의 패스파인더 네트워크를 제시하면서 각 저자의 크기를 지역 중심성에

비례해서 표현하기도 하였다.

최근접이웃중심성은 전체 노드 수가 N개일 때 자신을 제외한 다른 모든 노드로부터 최근접 이웃으로 꼽힐 경우에 N-1로 최댓값이 된다. 따라서 0에서 1 사이의 값으로 정규화할 때에는 최근접이웃중심성 값을 이론적으로 가능한 최댓값인 N-1로 나누어서 상대적 최근접이웃중

심성으로 사용한다. 상대적 최근접이웃중심성은 WNET에서 “rNNC(0~1)”라는 항목으로 출력된다.

2.2.2 평균연관성

평균연관성(Mean Associations: 이하 C_M , 또는 C_m 으로 약칭)은 다른 노드들과 가진 링크 가중치의 평균값을 의미한다(이재윤, 2006b). 연결이 안된 경우는 가중치를 0으로 간주한다. N개의 노드로 구성된 가중 네트워크에서 노드 i의 평균연관성을 산출하는 공식은 다음과 같다.

$$C_M(i) = \frac{\sum_{j \neq i} w_{ji}}{N-1}$$

이 공식에서 w_{ji} 는 노드 j와 노드 i 사이의 링크 가중치이다. 연결된 이웃 노드 중에서 노드 i가 몇 번째로 가까운 노드인가를 측정된 순위이다. 평균연관성은 해당 노드가 다른 노드들과 평균적으로 얼마나 가까운가를 나타내는 척도이다. 네트워크의 밀도가 높은 경우에는 전체 네트워크에서의 중심성을 측정하는 효과가 있으며, 밀도가 낮은 경우에는 주변 노드들과의 관계만 측정하는 결과가 된다. 평균이 아닌 합계를 사용할 경우에는 tnet에서 사용하는 지수인 노드 강도와 같은 값이 된다. 평균연관성은 노드 강도를 노드 수로 정규화한 지수이므로 두 지수는 동일하다고 볼 수 있다. WNET에서 평균연관성은 “AVGSIM”이라는 항목으로 출력된다.

2.2.3 평균프로파일연관성

자신을 제외한 다른 노드들과의 링크 가중치

가 높고 낮은 패턴을 이용해서 가중 네트워크에서 다른 노드와의 상대적인 거리가 중립적인 정도를 측정하는 지수로 평균프로파일연관성(Mean Profile Associations: 이하 C_{MP} , 또는 C_{mp} 로 약칭)을 쓸 수 있다(이재윤, 2006b). 평균프로파일연관성은 가중 네트워크를 표현한 각 노드 사이의 연관성 행렬을 입력 데이터로 하여 산출한 2차 피어슨 상관계수 행렬에서 산출한다. N개의 노드로 구성된 가중 네트워크에서 노드 i의 평균프로파일연관성을 산출하는 공식은 다음과 같다.

$$C_{MP}(i) = \frac{\sum_{j \neq i} cor_{ji}}{N-1}$$

이 공식에서 cor_{ji} 는 노드 j와 노드 i 사이의 피어슨 상관계수로서 다음과 같이 링크 가중치의 평균과 각 링크 가중치와의 차이를 이용하여 산출된다. $C_M(i)$ 는 노드 i의 평균연관성이다.

$$cor_{ji} = \frac{\sum_{k=1}^N (w_{jk} - C_M(j))(w_{ik} - C_M(i))}{\sqrt{\sum_{k=1}^N (w_{jk} - C_M(j))^2 \sum_{k=1}^N (w_{ik} - C_M(i))^2}}$$

평균프로파일연관성은 피어슨 상관계수의 평균이므로 -1에서 1 사이의 값을 가진다. 평균프로파일연관성은 노드 자체의 특성에 기인하여 높게 산출될 수도 있지만 노드가 속한 집단의 입지가 중심적인 경우에도 높은 값이 산출된다. 예를 들어 저자동시이용 네트워크의 경우에 전체 저자 집단에서 상대적으로 중립적인 분야를 연구하는 저자는 다른 중심성이 낮더라도 평균프로파일연관성이 높게 측정되는 무임승차 효

과가 나타날 수도 있다. 이는 이진 네트워크에서 근접중심성이 가지는 특성과 비슷하다. 평균프로파일연관성은 네트워크를 구성하는 노드가 수십 개 이상이며 밀도도 높은 네트워크에 적용했을 때에만 의미있는 결과가 산출된다. WNET에서 평균프로파일연관성은 “Cmp(-1~1)”이라는 항목으로 출력된다.

2.2.4 삼각매개중심성

가중 네트워크에서 한 노드가 다른 노드들 사이를 매개시켜주는 능력을 측정하는 지수인 삼각매개중심성(Triangle Betweenness Centrality: 이하 C_{TB} , 또는 C_{tb} 로 약칭)은 가중 네트워크에서 노드의 상대적인 전역 중심성을 측정하는 지수이다(이재운, 2006b). WNET에서는 “TBC”라는 항목으로 출력된다. 노드 간 링크 가중치로 구성된 연관성 행렬에서 노드 i 의 삼각매개중심성은 다음 공식과 같이 다른 두 노드 j 와 k 사이의 링크 가중치가 i 와 j , i 와 k 사이의 링크 가중치보다 작은 경우의 수로 측정한다.

$$C_{TB}(i) = \sum_j \sum_k f(b_{jk}),$$

$$f(b_{jk}) = \begin{cases} 1 & \text{if } w_{jk} < \text{MIN}(w_{ij}, w_{ik}) \\ 0 & \text{else} \end{cases}$$

〈그림 1〉의 가상 네트워크에서 노드 B의 삼각매개중심성을 산출해보면 다음과 같다. B를 제외한 나머지 (A,C), (A,D), (C,D)의 세 쌍 중에서 링크 (A,C)의 가중치 $w(A,C)$ 는 8이어서 $\text{MIN}(w(B,A), w(B,C)) = 5$ 보다 크지만, $w(A,D)$ 는 2로 $\text{MIN}(w(B,A), w(B,D)) = 7$ 보다 작고 $w(C,D)$ 는 3으로 $\text{MIN}(w(B,C), w(B,D)) = 5$ 보다 작다. 따라서 노드 B는 노드 A와 D 사이, 그

리고 노드 C와 D 사이를 매개해줄 수 있어서 삼각매개중심성이 2가 된다.

삼각매개중심성은 전체 노드 수가 N 개인 네트워크에서 자신을 제외한 다른 모든 노드들 사이의 관계를 모두 매개해줄 수 있을 때 최댓값 $(N-1) \times (N-2) / 2$ 이 된다. 따라서 0에서 1 사이의 값으로 정규화할 때에는 삼각매개중심성 값을 이론적으로 가능한 최댓값으로 나누어서 상대적 삼각매개중심성으로 사용한다. 상대적 삼각매개중심성은 WNET에서 “rTBC(0~1)”라는 항목으로 출력된다.

WNET으로 가중 네트워크 중심성을 분석하기 위해서는 〈그림 1〉의 가상 네트워크를 〈표 4〉와 같이 정방대칭행렬로 표현하여 입력해야 한다.

〈표 4〉 노드 4개로 구성된 가상 네트워크의 행렬 표현

	A	B	C	D
A	0	10	8	2
B	10	0	5	7
C	8	5	0	3
D	2	7	3	0

가상 네트워크에 대한 WNET의 가중 네트워크 중심성 계산 결과는 〈표 5〉와 같이 출력된다. 네트워크를 구성하는 네 노드 중에서 삼각매개중심성과 평균연관성으로는 노드 B가 가장 높고 최근접이웃중심성은 노드 A와 노드 B가 공동 1위임을 알 수 있다. 노드 4개짜리 작은 네트워크이기 때문에 두드러지게 중립적인 노드를 찾을 수가 없어서 평균프로파일연관성은 0에 가깝거나 0보다 낮은 값을 보였다. 〈표 5〉의 WNET 출력 결과에서 마지막 항목인 “NNs”

〈표 5〉 노드 4개로 구성된 가상 네트워크에 대한 WNET의 가중 네트워크 중심성 산출 결과

SN	NODE	TBC	rTBC(0~1)	AVGSIM	Cmp(-1~1)	NNC	rNNC(0~1)	NNs
1	A	1	0.33333	6.66667	-0.20778	2	0.66667	B
2	B	2	0.66667	7.33333	-0.48722	2	0.66667	A
3	C	1	0.33333	5.33333	-0.03066	0	0	A
4	D	0	0	4	0.02480	0	0	B

〈표 6〉 924종 학술지 동시인용 네트워크에서 측정된 WNET 중심성 지수와 이진 네트워크 중심성 지수 사이의 순위 상관관계

	Ctb	Cm	Cmp	Cn
Cd	0.996	0.948	0.728	-0.144
Cc	0.996	0.947	0.727	-0.144
Cb	0.956	0.847	0.529	-0.197

(출처: 유소영, 이재윤, 2008)

는 각 노드의 최근접이웃(Nearest Neighbor) 노드를 알려주는 항목으로서 노드 A와 D는 노드 B가 최근접이웃이며 노드 B와 C는 노드 A가 최근접이웃임을 알 수 있다(〈표 5〉 참조).

삼각매개중심성을 비롯한 WNET의 가중 네트워크 중심성 지수를 일반적인 이진 네트워크 중심성 지수와 비교한 선행연구로는 유소영과 이재윤(2008)이 있다. 이들은 924종의 학술지와 동시인용 링크 45,404건으로 구성된 동시인용 네트워크를 구축하고 이 네트워크에서 각 학술지의 중심성을 측정하였다. 구축된 학술지 동시인용 네트워크의 밀도는 0.053으로 매우 낮은 수준이었다. 측정된 924종 학술지의 중심성 지수 사이의 순위상관계수는 〈표 6〉과 같이 보고되었다. 이 결과에서 삼각매개 중심성이 연결정도 중심성 및 근접 중심성과 매우 상관성이 높은 것으로 나타났는데, 이는 대상 네트워크의 밀도가 매우 낮아서 링크, 즉 동시인용이 1회라도 있었는지 여부가 중심성 산출에 큰 영향을

끼쳤기 때문이라고 판단된다.

3. 가상 네트워크에 대한 가중 네트워크 중심성 측정

tnet과 WNET의 가중 네트워크 중심성 지수를 비교하기 위해서 실험적으로 〈그림 2〉와 같이 노드 9개와 링크 13개로 구성된 가상 네트워크 1을 준비하였다. 이 네트워크는 공저 네트워크나 동시인용 네트워크와 같은 비방향성 가중 네트워크이다. 또한 네트워크의 변화에 따른 가중 네트워크 중심성 지수의 변화도 살펴보기 위해서 가상 네트워크 1에서 링크와 가중치가 추가된 가상 네트워크 2, 3, 4를 준비하였다. 가상 네트워크 2는 〈표 8〉의 가상 네트워크 1에 존재하는 13개 링크의 가중치가 모두 2배로 증가한 경우이다. 이때는 기존에 존재하는 링크의 가중치만 2배로 되고 관계가 없던 노드 사이는

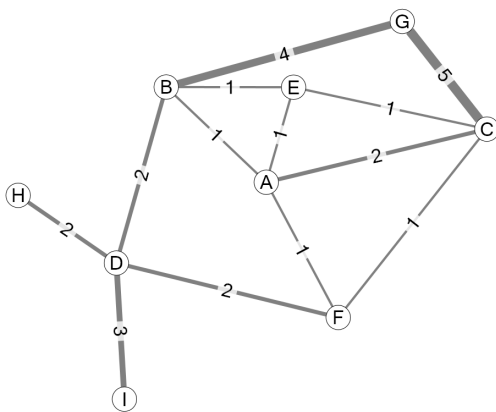
그대로 두었으므로 네트워크의 모양은 <그림 2>와 동일하고 링크 가중치 숫자만 <표 9>와 같이 두 배가 된다. 가상 네트워크 3은 <그림 3>과 같이 가상 네트워크 1에 있는 9개 노드들 사이의 관계가 모두 동일하게 1씩 증가한 경우이다. 행렬로 표현한 <표 10>을 보면 링크가 없던 노드 사이에는 가중치가 1인 링크가 추가되고, 기존에 있던 링크에는 가중치가 1씩 더해졌다. 가상 네트워크 4는 가상 네트워크 1에 있는 9개 노드들 사이의 관계가 모두 동일하게 10씩 증가한 경우이다. 행렬로 표현한 <표 11>을 보면 <표 8>

에서 링크가 없던(가중치가 0인) 노드 사이에는 가중치가 10인 링크가 생성되고, 기존에 있던 링크에는 가중치가 10씩 더해졌으므로 네트워크의 형태는 <그림 3>의 가상 네트워크 3과 동일하고 <표 10>에 비해서 링크 가중치 숫자만 9씩 증가하게 된다. 준비된 4개의 가상 네트워크의 특징을 정리하면 <표 7>과 같다.

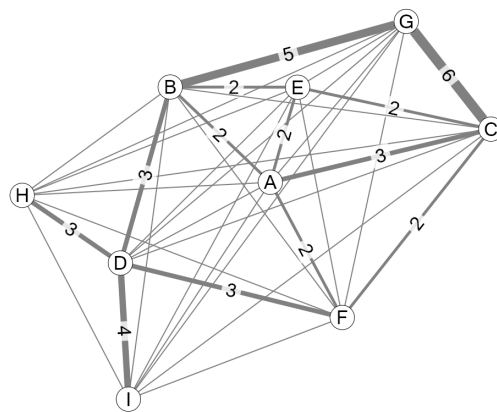
네 개의 가상 네트워크 자료를 대상으로 tnet과 WNET를 실행하여 각 노드의 노드 강도(NS), 이진 네트워크 중심성(Cd, Cc, Cb), tnet의 가상 네트워크 중심성(Cwd, Cwc, Cwb), WNET

<표 7> 가상 네트워크 4개의 특징

	가상 네트워크 1	가상 네트워크 2	가상 네트워크 3	가상 네트워크 4
노드 수	9	9	9	9
링크 수	13	13	36	36
밀도	0.361	0.361	1.000	1.000
최고 가중치와 최저 가중치 차이	5 (5-0)	9 (10-0)	5 (6-1)	5 (15-10)
비고		링크 가중치가 모두 두 배로 증가 (×2)	링크 가중치가 모두 조금씩 증가 (+1)	링크 가중치가 대폭 증가 (+10)



<그림 2> 가상 네트워크 1의 구성
(선 위의 숫자는 가중치)



<그림 3> 가상 네트워크 3의 구성
(선 위의 숫자는 2 이상인 가중치)

〈표 8〉 가상 네트워크 1의 노드 간 근접성 행렬

	A	B	C	D	E	F	G	H	I
A		1	2	0	1	1	0	0	0
B	1		0	2	1	0	4	0	0
C	2	0		0	1	1	5	0	0
D	0	2	0		0	2	0	2	3
E	1	1	1	0		0	0	0	0
F	1	0	1	2	0		0	0	0
G	0	4	5	0	0	0		0	0
H	0	0	0	2	0	0	0		0
I	0	0	0	3	0	0	0	0	

〈표 9〉 가상 네트워크 2의 노드 간 근접성 행렬

	A	B	C	D	E	F	G	H	I
A		2	4	0	2	2	0	0	0
B	2		0	4	2	0	8	0	0
C	4	0		0	2	2	10	0	0
D	0	4	0		0	4	0	4	6
E	2	2	2	0		0	0	0	0
F	2	0	2	4	0		0	0	0
G	0	8	10	0	0	0		0	0
H	0	0	0	4	0	0	0		0
I	0	0	0	6	0	0	0	0	

〈표 10〉 가상 네트워크 3의 노드 간 근접성 행렬

	A	B	C	D	E	F	G	H	I
A		2	3	1	2	2	1	1	1
B	2		1	3	2	1	5	1	1
C	3	1		1	2	2	6	1	1
D	1	3	1		1	3	1	3	4
E	2	2	2	1		1	1	1	1
F	2	1	2	3	1		1	1	1
G	1	5	6	1	1	1		1	1
H	1	1	1	3	1	1	1		1
I	1	1	1	4	1	1	1	1	

〈표 11〉 가상 네트워크 4의 노드 간 근접성 행렬

	A	B	C	D	E	F	G	H	I
A		11	12	10	11	11	10	10	10
B	11		10	12	11	10	14	10	10
C	12	10		10	11	11	15	10	10
D	10	12	10		10	12	10	12	13
E	11	11	11	10		10	10	10	10
F	11	10	11	12	10		10	10	10
G	10	14	15	10	10	10		10	10
H	10	10	10	12	10	10	10		10
I	10	10	10	13	10	10	10	10	

의 가상 네트워크 중심성(Ctb, Cn)을 산출한 결과는 〈표 12〉~〈표 15〉와 같다. tnet의 가상 네트워크 중심성 중에서 가상 연결정도 중심성은 파라미터 α 를 0.5로 설정하여 적용하였다. 가상 연결정도 중심성에서 α 가 0인 경우는 이진 네트워크 중심성의 연결정도 중심성(Cd)과 같고, α 가 1인 경우는 노드 강도(NS)와 같으므로 생략하였다. 가상 근접중심성(Cwc)과 가상 매개중심성(Cwb)은 파라미터 α 를 0.5와 1.0인 경우를 적용해보았다. WNET의 가상 네트워크 중심성 지수 중에서는 전역 중심성인 삼각매개

중심성(Ctb)와 지역 중심성인 최근접이웃중심성(Cn)을 적용해보았다. 평균연관성은 노드 강도를 전체 노드 수 $N - 1$ 로 나눈 것과 같으므로 제외하였고 평균프로파일연관성은 앞의 정의 부분에서 본 바와 같이 너무 작은 네트워크에서는 의미가 없으므로 역시 제외하였다.

〈표 12〉~〈표 15〉에 제시된 지수 산출 결과를 비교해보면, 노드 강도는 네트워크 1이 가장 작고 네트워크 2, 3, 4의 순서대로 커진다. 그러나 각 노드별 상대적인 순서는 그대로 지켜져서 노드 C, D, G가 네 네트워크 모두에서 공동 1위

로 나타난다(〈표 16〉 참조). 이진 네트워크 중심성 세 가지는 네트워크의 형태가 같은 네트워크 1과 2, 3과 4가 각각 동일하게 나타난다.

tnet의 가중 네트워크 중심성 지수 중에서 가중 연결정도중심성($\alpha=0.5$)은 노드 강도가 커짐에 따라 노드 간 순서를 유지하면서 커진다. 가중 근접중심성은 링크 수의 변화 없이 링크 강도만 두 배가 된 네트워크 2에서 네트워크 1과 동일하게 측정된다. 그러나 모든 노드 간 링크 가중치에 1씩 더한 네트워크 3과, 모든 노드 간 링크 가중치에 10씩 더한 네트워크 4에서는 각 노드의 순위가 달라진다(〈그림 4〉, 〈그림 5〉 참조). 네트워크 1과 2에서는 가중 근접중심성 1위인 노드가 B였지만 네트워크 3에서는 $\alpha=0.5$ 일 때 1위인 노드가 C와 D로 바뀌며, 네트워크 4에서는 $\alpha=0.5$ 일 때 1위인 노드가 C와 D로 바뀌고 $\alpha=1.0$ 일 때 1위인 노드가 D로 바뀐다(〈표 16〉 참조). 이로서 가중 근접중심성은 각 노드 간 관계가 일정한 크기로 동일하게 증가할 때 노드 간 지수 순위가 달라지며 α 가 0.5보다 1.0으로 클 때 지수 순위의 변화도

더 심해지는 것임을 알 수 있다.

tnet의 가중 네트워크 중심성 지수 중에서 가중 매개중심성은 링크 가중치의 증가에 따른 변화가 더 심하게 나타난다(〈그림 6〉, 〈그림 7〉 참조). 링크 수의 변화 없이 링크 강도만 두 배가 된 네트워크 2에서도 네트워크 1과 일부 순위가 달라지는 현상을 발견할 수 있다. 네트워크 2에서 $\alpha=1.0$ 으로 설정하였을 때에는 변화가 없으나 $\alpha=0.5$ 로 설정한 가중 매개중심성 측정 결과에서 노드 C의 경우는 가중 매개중심성이 커져서 노드 G와 같아지고, 노드 A의 경우는 가중 매개중심성이 작아져서 0이 된다. 모든 노드 간 링크 가중치에 1씩 더한 네트워크 3에서는 $\alpha=0.5$ 일 때 노드 C와 G만이 가중 매개중심성을 1씩 가지고 나머지 노드는 모두 0이 되는 것으로 나타났다. $\alpha=1.0$ 일 때에는 노드 G와 C의 순위가 역전되며 모든 노드의 가중 매개중심성이 크게 감소하는 것으로 나타났다. 모든 노드 간 링크 가중치에 10씩 더한 네트워크 4에서는 가중 매개중심성이 모두 0이 되어서 노드 간 중심성 차이를 전혀 측정하지 못하는 것으로 나

〈표 12〉 가상 네트워크 1의 각 지수별 수치

node	NS	Cd	Cc	Cb	Cwd		Cwc		Cwb		Ctb	Cn
					$\alpha=0.5$	$\alpha=1$	$\alpha=0.5$	$\alpha=1$	$\alpha=0.5$	$\alpha=1$		
A	5	4	0.071	1.333	4.472	0.065	0.054	0.500	0.333	3	1	
B	8	4	0.083	8.500	5.657	0.089	0.084	11.250	12.333	5	1	
C	9	4	0.071	2.500	6.000	0.071	0.073	3.500	7.333	4	3	
D	9	4	0.083	13.500	6.000	0.082	0.077	14.000	14.333	6	3	
E	3	3	0.067	0.330	3.000	0.057	0.043	0.000	0.000	1	0	
F	4	3	0.077	4.500	3.464	0.072	0.059	0.750	0.000	2	0	
G	9	2	0.063	0.333	4.243	0.074	0.075	4.000	8.000	1	2	
H	2	1	0.053	0.000	1.414	0.053	0.050	0.000	0.000	0	0	
I	3	1	0.053	0.000	1.732	0.057	0.057	0.000	0.000	0	1	

〈표 13〉 가상 네트워크 2의 각 지수별 수치

node	NS	Cd	Cc	Cb	Cwd	Cwc		Cwb		Ctb	Cn
					α=0.5	α=0.5	α=1	α=0.5	α=1		
A	10	4	0.071	1.333	6.325	0.065	0.054	0.000	0.333	3	1
B	16	4	0.083	8.500	8.000	0.089	0.084	11.250	12.333	5	1
C	18	4	0.071	2.500	8.485	0.071	0.073	4.000	7.333	4	3
D	18	4	0.083	13.500	8.485	0.082	0.077	14.000	14.333	6	3
E	6	3	0.067	0.333	4.243	0.057	0.043	0.000	0.000	1	0
F	8	3	0.077	4.500	4.899	0.072	0.059	0.750	0.000	2	0
G	18	2	0.063	0.333	6.000	0.074	0.075	4.000	8.000	1	2
H	4	1	0.053	0.000	2.000	0.053	0.050	0.000	0.000	0	0
I	6	1	0.053	0.000	2.449	0.057	0.057	0.000	0.000	0	1

〈표 14〉 가상 네트워크 3의 각 지수별 수치

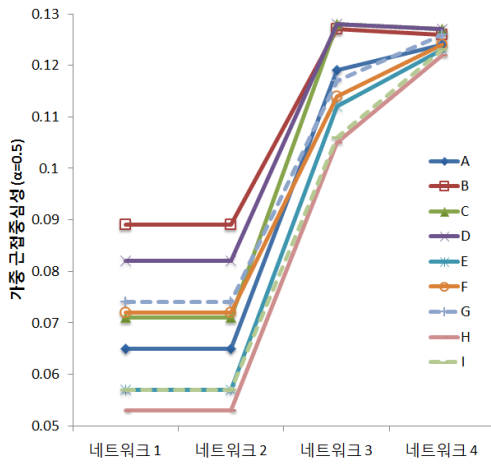
node	NS	Cd	Cc	Cb	Cwd	Cwc		Cwb		Ctb	Cn
					α=0.5	α=0.5	α=1	α=0.5	α=1		
A	13	8	0.125	0.000	10.198	0.119	0.112	0.000	0.333	3	1
B	16	8	0.125	0.000	11.314	0.127	0.152	0.000	7.000	5	1
C	17	8	0.125	0.000	11.662	0.128	0.129	1.000	3.333	4	3
D	17	8	0.125	0.000	11.662	0.128	0.140	0.000	9.000	6	3
E	11	8	0.125	0.000	9.381	0.112	0.097	0.000	0.000	1	0
F	12	8	0.125	0.000	9.798	0.114	0.118	0.000	0.000	2	0
G	17	8	0.125	0.000	11.662	0.117	0.132	1.000	3.000	1	2
H	10	8	0.125	0.000	8.944	0.105	0.095	0.000	0.000	0	0
I	11	8	0.125	0.000	9.381	0.106	0.101	0.000	0.000	0	1

〈표 15〉 가상 네트워크 4의 각 지수별 수치

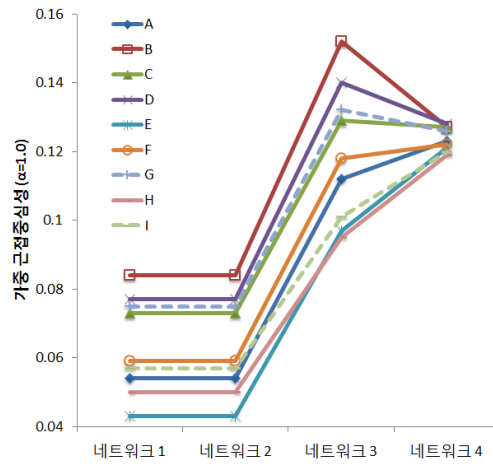
node	NS	Cd	Cc	Cb	Cwd	Cwc		Cwb		Ctb	Cn
					α=0.5	α=0.5	α=1	α=0.5	α=1		
A	85	8	0.125	0.000	26.077	0.124	0.123	0.000	0.000	3	1
B	88	8	0.125	0.000	26.533	0.126	0.127	0.000	0.000	5	1
C	89	8	0.125	0.000	26.683	0.127	0.127	0.000	0.000	4	3
D	89	8	0.125	0.000	26.683	0.127	0.128	0.000	0.000	6	3
E	83	8	0.125	0.000	25.768	0.123	0.121	0.000	0.000	1	0
F	84	8	0.125	0.000	25.923	0.124	0.122	0.000	0.000	2	0
G	89	8	0.125	0.000	26.683	0.126	0.126	0.000	0.000	1	2
H	82	8	0.125	0.000	25.612	0.122	0.119	0.000	0.000	0	0
I	83	8	0.125	0.000	25.768	0.123	0.120	0.000	0.000	0	1

〈표 16〉 개별 가상 네트워크의 각 지수별 1위

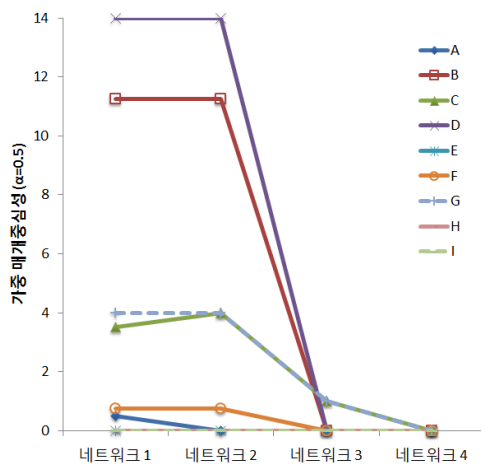
네트워크	지수	NS	Cd	Cc	Cb	Cwd	Cwc		Cwb		Ctb	Cn
						$\alpha=0.5$	$\alpha=0.5$	$\alpha=1$	$\alpha=0.5$	$\alpha=1$		
네트워크 1		C,D,G	C,D,B,A	D,B	D	C,D	B	B	D	D	D	C,D
네트워크 2		C,D,G	C,D,B,A	D,B	D	C,D	B	B	D	D	D	C,D
네트워크 3		C,D,G	·	·	·	C,D,G	C,D	B	C,G	D	D	C,D
네트워크 4		C,D,G	·	·	·	C,D,G	C,D	D	·	·	D	C,D



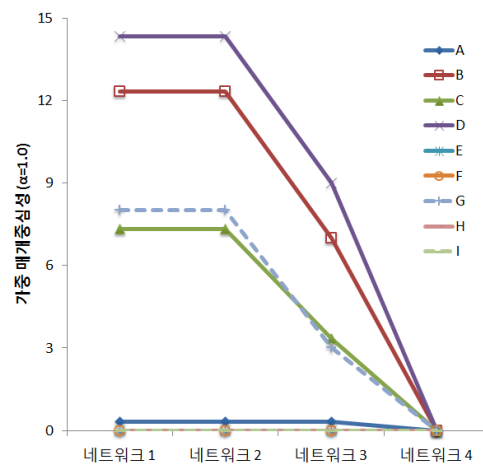
〈그림 4〉 가중 근접중심성($\alpha=0.5$)의 변화



〈그림 5〉 가중 근접중심성($\alpha=1.0$)의 변화



〈그림 6〉 가중 매개중심성($\alpha=0.5$)의 변화



〈그림 7〉 가중 매개중심성($\alpha=1.0$)의 변화

타났다. 결국 가중 매개중심성은 네트워크의 밀도가 높고 링크 가중치 차이가 크지 않을 경우에 제대로 측정되지 못함을 알 수 있다. 그 이유는 앞서서도 살펴보았듯이 클릭에 가깝게 밀도가 높은 네트워크에서는 노드 간 링크 가중치의 차이가 크지 않다면 직접 연결된 경로가 가장 가까운 경로가 되어 매개중심성을 획득하는 노드가 드물어지기 때문이다.

WNET의 가중 네트워크 중심성인 삼각매개 중심성과 최근접이웃중심성은 tnet의 가중 네트워크 중심성과 달리 모든 링크의 가중치가 일정하게 곱해지거나 더해지는 변화에 영향받지 않고 네 가지 네트워크에서 모두 동일한 중심성이 측정되었다(〈표 12〉~〈표 15〉 참조). 이는 WNET의 가중 네트워크 중심성 지수가 링크 가중치의 절대적인 차이를 고려하기 보다는 상대적인 차이, 즉 순위에 좌우되는 지수이기 때문이다.

4. 실제 네트워크에 대한 가중 네트워크 중심성 측정

tnet과 WNET의 가중 네트워크 중심성을 실제 네트워크를 대상으로 측정하여 결과를 비교해보기 위해서 〈표 17〉과 같은 6가지 네트워크 데이터를 준비하였다. 이 중에서 5개 네트워크는 기존 연구에서 사용된 것이고 col-182는 국내 문헌정보학 분야 학술지 5종(정보관리연구, 정보관리학회지, 한국도서관정보학회지, 한국문헌정보학회지, 한국비블리아학회지)에 2002년부터 2011년까지 10년 동안 5편 이상의 논문을 발표한 저자 182명의 공저자 네트워크이다. 6종

네트워크에 대해서 tnet과 WNET를 실행하여 각 노드의 노드 강도(NS), 이진 네트워크 중심성(Cd, Cc, Cb), tnet의 가중 네트워크 중심성(Cwd, Cwc, Cwb), WNET의 가중 네트워크 중심성(Ctb, Cn)을 산출한 후, 각 네트워크마다 각 지수 사이의 피어슨 상관계수를 측정해 보았다. 최종적으로 지수 간 관계를 비교하기 위해서 여러 네트워크의 지수 간 상관계수의 평균을 산출하였다.

이진 네트워크 중심성 지수와 가중 네트워크 중심성 지수를 비교하기 위해서 ps-46을 제외한 5종 네트워크에서 측정된 지수 간 상관계수의 평균은 〈표 18〉과 같다. ps-46은 저자 프로파일링을 통한 저자 유사도 네트워크로서 밀도가 1이어서 이진 네트워크 중심성 측정이 의미 없기 때문에 이진 네트워크 중심성과 가중 네트워크 중심성을 비교할 때에는 제외하였다.

〈표 18〉에서 굵은 실선으로 표시된 이진 네트워크 중심성 지수와 가중 네트워크 중심성 지수 간 상관계수를 살펴보면, 연결정도중심성과 가장 상관성이 높은 가중 네트워크 중심성 지수는 $\alpha=0.5$ 인 가중 연결정도중심성이고 삼각매개중심성이 그 다음이었다. 근접중심성과 가장 상관성이 높은 가중 네트워크 중심성은 $\alpha=0.5$ 인 가중 근접중심성이고 $\alpha=0.5$ 인 가중 연결정도중심성, 삼각매개중심성이 그 다음 순이었다. 매개중심성과 가장 상관성이 높은 가중 네트워크 중심성은 삼각매개중심성이었고 $\alpha=0.5$ 인 가중 매개중심성이 그 다음이었다.

tnet과 WNET의 가중 네트워크 중심성 지수를 비교하기 위해서 6종 네트워크에서 측정된 지수 간 상관계수의 평균을 산출하면 〈표 19〉와 같다. 〈표 19〉에서 굵은 실선으로 표시된 tnet과

〈표 17〉 중심성 측정 대상 가중 네트워크

약칭	col-182	cw-85	cw-98	dc-70	ac-50	ps-46
출처	직접 구축	이재윤, 정주희 (2006)	Lee & Choi (2011)	이재윤 (2012b)	이은숙, 정영미 (2002)	유종덕, 최은주 (2011)
관계유형	공저자	단어 동시출현	단어 동시출현	문헌 동시인용	저자동시인용	저자 프로파일링
주제 분야	문헌정보학	인지과학	계량정보학	복소노미	정보학	건축학
노드 수	182	85	98	70	50	46
링크 수	375	679	2035	1003	948	1035
최대 경로거리	12	3	2	3	2	1
평균 경로거리	4.639	1.879	1.556	1.610	1.202	0.978
밀도	0.023	0.190	0.428	0.415	0.774	1.000
평균 군집화계수	0.334	0.431	0.630	0.759	0.846	1.000
가중평균 군집화계수	0.330	0.392	0.596	0.725	0.843	1.000
연결정도 중심화	0.078	0.269	0.542	0.527	0.214	0.000

〈표 18〉 ps46을 제외한 5개 가중 네트워크에서 측정된 가중 네트워크 중심성 지수 간 상관계수의 평균

	NS	Cd	Cc	Cb	Cwd, a=0.5	Cwc, a=0.5	Cwc, a=1	Cwb, a=0.5	Cwb, a=1	Ctb	Cm	Cmp	Cn
NS	1.000	0.810	0.741	0.701	0.963	0.823	0.776	0.732	0.735	0.829	1.000	0.835	0.487
Cd	0.810	1.000	0.924	0.830	0.934	0.815	0.688	0.654	0.581	0.901	0.810	0.759	0.303
Cc	0.741	0.924	1.000	0.791	0.858	0.892	0.750	0.626	0.553	0.842	0.741	0.704	0.237
Cb	0.701	0.830	0.791	1.000	0.792	0.694	0.567	0.821	0.712	0.906	0.701	0.519	0.361
Cwd, a=0.5	0.963	0.934	0.858	0.792	1.000	0.857	0.768	0.734	0.697	0.905	0.963	0.836	0.426
Cwc, a=0.5	0.823	0.815	0.892	0.694	0.857	1.000	0.938	0.671	0.658	0.797	0.823	0.792	0.321
Cwc, a=1	0.776	0.688	0.750	0.567	0.768	0.938	1.000	0.597	0.647	0.680	0.776	0.772	0.361
Cwb, a=0.5	0.732	0.654	0.626	0.821	0.734	0.671	0.597	1.000	0.919	0.797	0.732	0.486	0.502
Cwb, a=1	0.735	0.581	0.553	0.712	0.697	0.658	0.647	0.919	1.000	0.720	0.735	0.495	0.613
Ctb	0.829	0.901	0.842	0.906	0.905	0.797	0.680	0.797	0.720	1.000	0.829	0.681	0.393
Cm	1.000	0.810	0.741	0.701	0.963	0.823	0.776	0.732	0.735	0.829	1.000	0.835	0.487
Cmp	0.835	0.759	0.704	0.519	0.836	0.792	0.772	0.486	0.495	0.681	0.835	1.000	0.269
Cn	0.487	0.303	0.237	0.361	0.426	0.321	0.361	0.502	0.613	0.393	0.487	0.269	1.000

〈표 19〉 6개 가중 네트워크에서 측정된 가중 네트워크 중심성 지수 간 상관계수의 평균

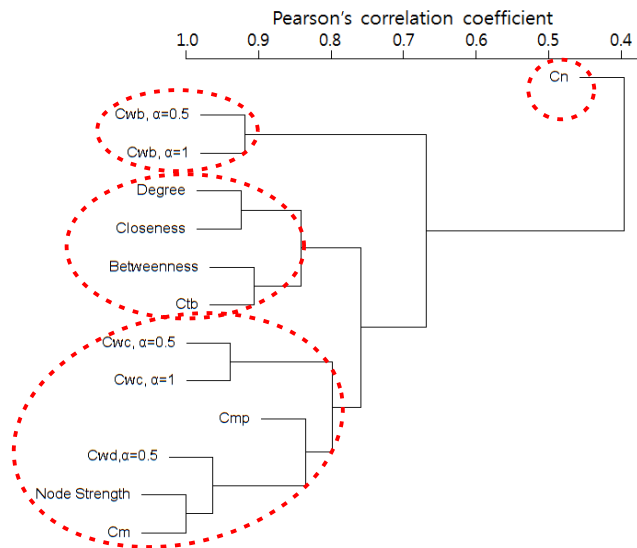
	Cwd, a=0.5	Cwc, a=0.5	Cwc, a=1	Cwb, a=0.5	Cwb, a=1	Ctb	Cm	Cmp	Cn
Cwd, a=0.5	1.000	0.872	0.798	0.664	0.676	0.904	0.969	0.849	0.426
Cwc, a=0.5	0.872	1.000	0.942	0.620	0.652	0.828	0.845	0.815	0.327
Cwc, a=1	0.798	0.942	1.000	0.555	0.639	0.722	0.806	0.801	0.354
Cwb, a=0.5	0.664	0.620	0.555	1.000	0.899	0.719	0.666	0.444	0.453
Cwb, a=1	0.676	0.652	0.639	0.899	1.000	0.701	0.713	0.489	0.580
Ctb	0.904	0.828	0.722	0.719	0.701	1.000	0.843	0.717	0.379
Cm	0.969	0.845	0.806	0.666	0.713	0.843	1.000	0.848	0.477
Cmp	0.849	0.815	0.801	0.444	0.489	0.717	0.848	1.000	0.272
Cn	0.426	0.327	0.354	0.453	0.580	0.379	0.477	0.272	1.000

WNET의 가중 네트워크 중심성 지수 간 상관 관계를 보면 tnet의 가중 연결정도중심성과 가중 근접중심성이 WNET의 평균연관성과 가장 상관관계가 높음을 알 수 있다. 평균연관성이 노드 강도와 같은 의미이기 때문에 노드 강도가 반영되는 tnet의 중심성 지수와 상관관계가 높은 것은 당연한 결과이다. 평균연관성을 제외하면 WNET의 삼각매개중심성과 평균프로파일 연관성이 tnet의 가중 연결정도중심성 및 가중 근접중심성과 상관관계가 높게 나타났다. tnet의 가중 매개중심성은 WNET의 삼각매개중심성과는 상관관계가 높은 편이지만 WNET의 평균연관성은 상대적으로 낮은 상관관계를 보였다. WNET의 최근접이웃중심성은 tnet의 가중 네트워크 중심성 지수 모두와 상대적으로 낮은 상관관계를 보였다.

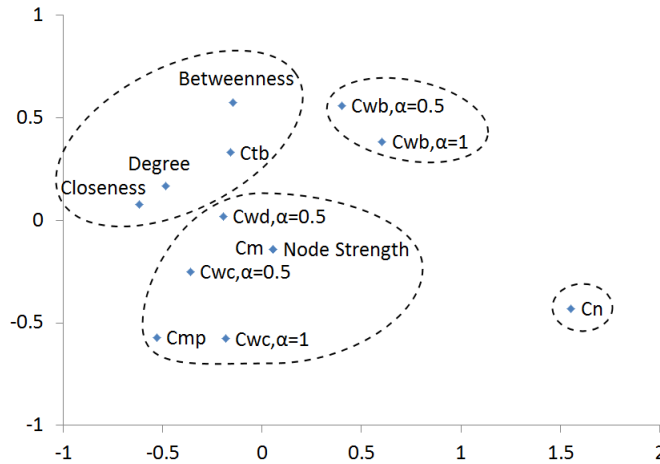
〈표 18〉에 제시된 13종 지수 사이의 평균 상관계수를 데이터로 하여 13종 지수를 평균연결

기법으로 군집화하면 〈그림 8〉과 같은 결과를 얻는다. 동일한 데이터를 SPSS의 다차원척도법 알고리즘 중 시각화 성능이 좋은(이재윤, 2007) PROXSCAL로 배치하고, 평균연결기법에 의한 군집을 표시한 결과는 〈그림 9〉와 같다.

군집화결과와 PROXSCAL 지도를 보면 가중 네트워크 중심성 지수 중에서는 WNET의 삼각매개중심성이 이진 네트워크 중심성 지수, 특히 매개중심성 지수와 가장 가까워서 같은 군집에 소속된 것으로 나타났다. tnet의 가중 연결정도중심성과 가중 근접중심성은 노드 강도, WNET의 평균연관성 및 평균프로파일 연관성과 함께 군집을 이루었다. tnet의 가중 매개중심성은 다른 지수들과의 관계가 상대적으로 낮아서 독립된 군집을 이루었다. WNET의 최근접이웃중심성은 다른 지수들과의 관계가 가장 멀어서 가장 독특한 지수인 것을 알 수 있다. 특히 〈그림 9〉의 PROXSCAL 지도에서 보면 이진



〈그림 8〉 지수 간 상관계수에 따른 평균연결 군집화



〈그림 9〉 지수 간 상관계수로 도출한 다차원척도법(PROXSCAL) 배치도 (평균연결 4군집 표시)

네트워크 중심성 지수와 최근접이웃중심성 지수가 반대편에 자리한 것으로 나타나서 이진 네트워크 중심성 분석 결과와는 가장 다른 결과를 도출하는 지수임을 알 수 있다. 결과적으로 〈그림 9〉의 PROXSCAL 지도에서 왼쪽 상단은 네트워크의 구조를 많이 반영하는 지수가 배치되고 오른쪽 아래는 네트워크의 구조를 적게 반영하는 지수가 배치된 것으로 해석된다. 이는 tnet의 가중 근접중심성 지수와 가중 매개중심성 지수에서 파라미터 α 가 0.5인 경우가 1.0인 경우보다 상대적으로 왼쪽 위에 배치된 것으로도 알 수 있다.

5. 결론

네트워크 데이터 수집과 분석이 여러 분야로 확산되면서 점차 가중 네트워크를 분석하게 될 경우가 늘어나게 되었다. 이에 따라 가중 네트워크 중심성을 측정할 수 있는 공개 소프트웨어

인 tnet과 WNET의 가중 네트워크 중심성 지수를 비교해본 결과는 다음과 같다.

첫째, 가상 데이터를 대상으로 한 분석에서 tnet의 가중 네트워크 중심성 지수는 링크 가중치에 똑같은 값이 곱해지거나 더해지는 선형 변화가 있을 경우에 각 노드의 순위가 다르게 측정된 반면, WNET의 중심성 지수는 전혀 변화가 없었다. tnet의 가중 네트워크 중심성 지수 중에서는 가중 매개중심성이 링크 가중치의 선형 변화에 가장 크게 영향을 받았으며, 파라미터 α 가 클 경우에 더 큰 영향을 받았다. 특히 가중 네트워크의 밀도가 높고 링크 가중치의 차이가 적은 경우에는 tnet의 가중 매개중심성이 제대로 측정되지 못하는 것으로 나타났다. 이는 노드 간 링크 가중치의 차이가 크지 않다면 직접 연결된 경로가 가장 가까운 경로가 되어 매개중심성을 획득하는 노드가 드물어지기 때문이다. 따라서 네트워크의 밀도가 매우 높거나 1로 완전 연결된 클리크인 경우에는 WNET의

중심성 지수를 측정하는 것이 바람직할 것이다.

둘째, 실제 네트워크를 대상으로 중심성 지수를 측정하고 상관관계를 비교해본 결과 이진 네트워크 중심성 지수처럼 네트워크 구조를 가장 많이 반영하는 지수는 WNET의 삼각매개중심성 지수인 것으로 나타났다. 반대로 WNET의 최근접이웃중심성 지수는 네트워크 구조를 가장 적게 반영하는 지수였다. tnet의 가중 연결정도 중심성과 가중 근접중심성은 노드 강도, WNET의 평균연관성, 평균프로파일연관성과 가까운 상관관계를 보였다. 반면에 tnet의 가중 매개중심성은 tnet의 다른 중심성 지수와의 거리가 다소 멀어서 다른 결과를 도출하는 지수임을 알 수

있었다. WNET의 최근접이웃중심성 지수는 다른 지수들과 가장 거리가 멀어서 가장 독특한 지수인 것으로 나타났다.

전체적으로 밀도가 1이거나 1에 가까우면서 링크 가중치의 차이가 크지 않은 네트워크일 경우에는 tnet이 아닌 WNET을 사용하는 것이 의미있는 측정 결과를 얻는 방법일 것으로 판단된다. 밀도가 높지 않은 네트워크에서는 tnet이나 WNET 모두 의미있는 측정 결과를 보여주며, tnet의 가중 매개중심성이나 WNET의 최근접이웃중심성은 여타 지수와 구분되는 측면을 보여주는 지수이므로 상호 보완적으로 분석에 활용할 필요가 있다.

참 고 문 헌

- 김원진, 정영미 (2010). 과학기술분야 국제협력 증진을 위한 아시아 국가 간 공동연구 현황 분석. 정보관리학회지, 27(3), 103-123. <http://dx.doi.org/10.3743/KOSIM.2010.27.3.103>
- 김판준 (2011). 저자 프로파일링 기법을 이용한 국내 독서 연구 영역 분석. 한국비블리아학회지, 22(4), 21-44.
- 변지혜, 정은경 (2011). 저자서지결합분석에 의한 국내 전기공학 분야 지적구조에 관한 연구. 정보관리연구, 42(4), 75-94. <http://dx.doi.org/10.1633/JIM.2011.42.4.075>
- 유소영, 이재윤 (2008). 학제적 분야의 정보서비스를 위한 학술지 인용 분석에 관한 연구: Y대학교 생명공학과를 중심으로. 정보관리학회지, 25(4), 283-308. <http://dx.doi.org/10.3743/KOSIM.2008.25.4.283>
- 유종덕, 최은주 (2011). 저자프로파일링분석과 저자동시인용분석의 유용성 비교 검증. 정보관리학회지, 28(1), 123-144. <http://dx.doi.org/10.3743/KOSIM.2011.28.1.123>
- 이기황, 이재윤, 백두현 (2010). 계량적 접근에 의한 조선시대 필사본 조리서의 유사성 분석. 언어와 정보, 14(2), 131-157.
- 이은숙, 정영미 (2002). 복수저자를 고려한 동시인용분석 연구: 정보학과 컴퓨터과학을 대상으로. 지식처리연구, 3(2), 1-26. Retrieved from <http://jkpm.yonsei.ac.kr/fulltext/v3n2a1.pdf>

- 이재운 (2006a). 지적 구조의 규명을 위한 네트워크 형성 방식에 관한 연구. *한국문헌정보학회지*, 40(2), 333-355. <http://dx.doi.org/10.4275/KSLIS.2006.40.2.333>
- 이재운 (2006b). 계량서지적 네트워크 분석을 위한 중심성 척도에 관한 연구. *한국문헌정보학회지*, 40(3), 191-214. <http://dx.doi.org/10.4275/KSLIS.2006.40.3.191>
- 이재운 (2006c). 지적 구조 분석을 위한 새로운 클러스터링 기법에 관한 연구. *정보관리학회지*, 23(4), 215-231. <http://dx.doi.org/10.3743/KOSIM.2006.23.4.215>
- 이재운 (2007). 지적 구조 분석을 위한 MDS 지도 작성 방식의 비교 분석. *한국문헌정보학회지*, 41(2), 335-357. <http://dx.doi.org/10.4275/KSLIS.2007.41.2.335>
- 이재운 (2008). 연구자의 투고 학술지 현황에 근거한 국내 학문분야 네트워크 분석. *정보관리학회지*, 25(4), 327-345. <http://dx.doi.org/10.3743/KOSIM.2008.25.4.327>
- 이재운 (2012a). 자기 인용 네트워크와 인용 정체성을 이용한 연구자의 연구 이력 분석에 관한 연구. *정보관리학회지*, 29(1), 157-174. <http://dx.doi.org/10.3743/KOSIM.2012.29.1.157>
- 이재운 (2012b). 폭소노미 연구 문헌에 대한 자아 중심 주제 인용 분석. *정보관리학회지*, 29(4), 295-312. <http://dx.doi.org/10.3743/KOSIM.2012.29.4.295>
- 이재운, 정주희 (2006). 연구자 소속과 표제어 분석을 통한 국내 인지과학 분야의 학제적 구조 파악. 제13회 한국정보관리학회 학술대회 논문집, 127-134.
- 이형진, 서교, 정남수, 이인복, 서일환, 문운경, 이정재 (2012). 가중 HPAI 확산 네트워크에서 중심성 분석: 2008년 한국 김제 지역의 HPAI 발병 사례를 중심으로. *농촌계획*, 18(4), 79-89. <http://dx.doi.org/10.7851/ksrp.2012.18.4.079>
- 정선영 (2013). T. D. Wilson의 연구주제 네트워크 분석. 제20회 한국정보관리학회 학술대회 발표논문집, 51-54.
- 최상희 (2013). 대학도서관 서비스를 위한 서지분석기반 학과의 주제적 특성 분석 연구. *정보관리학회지*, 30(1), 237-258. <http://dx.doi.org/10.3743/KOSIM.2013.30.1.237>
- 허고은, 송민 (2013). 저자동시인용 분석과 동시출현단어 분석을 이용한 의료정보학 저널의 지적구조 분석. *정보관리학회지*, 30(2), 207-225. <http://dx.doi.org/10.3743/KOSIM.2013.30.2.207>
- Barrat, A., Barthelemy, M., Pastor-Satorras, R., & Vespignani, A. (2004). The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences*, 101(11), 3747-3752. <http://dx.doi.org/10.1073/pnas.0400087101>
- Gómez, J. M., Nunn, C. L., & Verdú, M. (2013). Centrality in primate-parasite networks reveals the potential for the transmission of emerging infectious diseases to humans. *Proceedings of the National Academy of Sciences of the United States of America*, 110(19), 7738-7741. <http://dx.doi.org/10.1073/pnas.1220716110>
- Knappett, C. (Ed.). (2013). *Network analysis in archaeology: New approaches to regional interaction*.

- Oxford: Oxford University Press.
- Knoke, D. (2012). *Economic networks*. Cambridge, UK: Polity Press.
- Lee, J. Y., & Choi, S. (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.
<http://dx.doi.org/10.3743/KOSIM.2011.28.2.011>
- Newman, M. E. J. (2004). Analysis of weighted networks. *Physical Review E*, 70(5), 056131.
- Oatley, T., Winecoff, W. K., Pennock, A., & Danzman, S. B. (2013). The political economy of global finance: A network model. *Perspectives on Politics*, 11(1), 133-153.
<http://dx.doi.org/10.1017/S1537592712003593>
- Opsahl, T., Agneessens, F., & Skvoretz, J. (2010). Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks*, 32(3), 245-251.
<http://dx.doi.org/10.1016/j.socnet.2010.03.006>
- Opsahl, T., Colizza, V., Panzarasa, P., & Ramasco, J. J. (2008). Prominence and control: The weighted rich-club effect. *Physical Review Letters*, 101(16), 168702.
<http://dx.doi.org/10.1103/PhysRevLett.101.168702>
- Peeples, M. A., & Roberts Jr., J. M. (2013). To binarize or not to binarize: Relational data and the construction of archaeological networks. *Journal of Archaeological Science*, 40(7), 3001-3010.
<http://dx.doi.org/10.1016/j.jas.2013.03.014>
- Pryke, S. (2012). *Social network analysis in construction*. Chichester: Wiley-Blackwell.
- Schvaneveldt, R. W. (Ed.). (1990). *Pathfinder associative networks: Studies in knowledge organizations*. Norwood, N.J.: Ablex Pub. Corp.
- Vedres, B., & Scotti, M. (Eds.). (2012). *Networks in social policy problems*. New York: Cambridge University Press.

• 국문 참고문헌에 대한 영문 표기
(English translation of references written in Korean)

- Byun, Ji-Hye, & Chung, Eun-Kyung (2011). Domain analysis on electronic engineering in Korea by author bibliographic coupling analysis. *Journal of information management*, 42(4), 75-94.
<http://dx.doi.org/10.1633/JIM.2011.42.4.075>
- Choi, Sanghee (2013). Bibliometric analysis to analyze topic areas of faculty for academic library service. *Journal of the Korean Society for Information Management*, 30(1), 237-258.
<http://dx.doi.org/10.3743/KOSIM.2013.30.1.237>

- Heo, Go Eun, & Song, Min (2013). Examining the intellectual structure of a medical informatics journal with author co-citation analysis and co-word analysis. *Journal of the Korean Society for Information Management*, 30(2), 207-225. <http://dx.doi.org/10.3743/KOSIM.2013.30.2.207>
- Jung, SunYoung (2013). Network analysis for research subject of T.D. Wilson. *Proceedings of the 20th Annual Conference of the Korean Society for Information Management*, 51-54.
- Kim, Pan Jun (2011). Domain analysis of reading research in Korea using author profiling. *Journal of the Korean Biblia Society*, 22(4), 21-44.
- Kim, Won-Jin, & Chung, Young-Mee (2010). A study on research collaboration among Asian countries in science and technology. *Journal of the Korean Society for Information Management*, 27(3), 103-123. <http://dx.doi.org/10.3743/KOSIM.2010.27.3.103>
- Lee, Eun Suk, & Chung, Young Mee (2002). A co-citation analysis of multiple authorship in the subject fields of information science and computer science. *Journal of Knowledge Processing and Management*, 3(2), 1-26. Retrieved from <http://jkpm.yonsei.ac.kr/fulltext/v3n2a1.pdf>
- Lee, Hyungjin, Suh, Kyo, Jung, Namsu, Lee, Inbok, Seo, Ilhwan, Moon, Woonkyung, & Lee, Jeong-Jae (2012). Centrality measure in weighted HPAI transmission network: The case of the highly pathogenic H5N1 avian influenza Virus in Gimje, South Korea in 2008. *Journal of Korean Society of Rural Planning*, 18(4), 79-89. <http://dx.doi.org/10.7851/ksrp.2012.18.4.079>
- Lee, Jae Yun (2006a). A study on the network generation methods for examining the intellectual structure of knowledge domains. *Journal of the Korean Society for Library and Information Science*, 40(2), 333-355. <http://dx.doi.org/10.4275/KSLIS.2006.40.2.333>
- Lee, Jae Yun (2006b). Centrality measures for bibliometric network analysis. *Journal of the Korean Society for Library and Information Science*, 40(3), 191-214. <http://dx.doi.org/10.4275/KSLIS.2006.40.3.191>
- Lee, Jae Yun (2006c). A novel clustering method for examining and analyzing the intellectual structure of a scholarly field. *Journal of the Korean Society for Information Management*, 23(4), 215-231. <http://dx.doi.org/10.3743/KOSIM.2006.23.4.215>
- Lee, Jae Yun (2007). A comparison analysis of various approaches to multidimensional scaling in mapping a knowledge domain's intellectual structure. *Journal of the Korean Society for Library and Information Science*, 41(2), 335-357. <http://dx.doi.org/10.4275/KSLIS.2007.41.2.335>
- Lee, Jae Yun (2008). Analyzing the network of academic disciplines with journal contributions of Korean researchers. *Journal of the Korean Society for Information Management*, 25(4), 327-345. <http://dx.doi.org/10.3743/KOSIM.2008.25.4.327>
- Lee, Jae Yun (2012a). Exploring a researcher's personal research history through self-citation

- network and citation identity. *Journal of the Korean Society for Information Management*, 29(1), 157-174. <http://dx.doi.org/10.3743/KOSIM.2012.29.1.157>
- Lee, Jae Yun (2012b). Ego-centered topic citation analysis on folksonomy research documents. *Journal of the Korean Society for Information Management*, 29(4), 295-312. <http://dx.doi.org/10.3743/KOSIM.2012.29.4.295>
- Lee, Jae Yun, & Jung, Ju Hee (2006). Examining the interdisciplinary structure of Korean cognitive science through analyzing author affiliations and title words. *Proceedings of the 13th Annual Conference of the Korea Society for Information Management*, 127-134.
- Lee, Kihwang, Lee, Jae Yun, & Paek, Doohyun (2010). A quantitative approach to a similarity analysis on the culinary manuscripts in the Chosun Periods. *Language and Information*, 14(2), 131-157.
- Ryoo, Jong-duk, & Choi, Eun-Ju (2011). A comparison test on the potential utility between author profiling analysis(APA) and author co-citation analysis(ACA). *Journal of the Korean Society for Information Management*, 28(1), 123-144. <http://dx.doi.org/10.3743/KOSIM.2011.28.1.123>
- Yu, So Young, & Lee, Jae Yun (2008). Journal citation analysis for library services on interdisciplinary domains: A case study of Department of Biotechnology, Y University. *Journal of the Korean Society for Information Management*, 25(4), 283-308. <http://dx.doi.org/10.3743/KOSIM.2008.25.4.283>