

내용 분석을 통한 한국의 학술적 웹 공간 구조 분석

Ascertaining the Structure and Content of a National Scholarly Web Space Based on Content Analysis

정영미(Young-Mee Chung)*

유소영(So-Young Yu)**

초 록

학술적 웹 공간을 대상으로 하는 연구는 페이지와 링크의 역동성 때문에 정량적인 방법과 함께 내용 분석 등의 정성적인 방법을 사용하는 것이 필요하다. 따라서 이 연구에서는 내용 분석의 한 방법으로 한국 학술적 웹 공간 내에서 외부 링크로 연결된 페이지 및 링크의 유형을 분류한 후 이를 네트워크 구조 분석에 반영하여 한국 학술적 웹 공간의 특성을 자세히 살펴보았다. 분석 결과 데이터의 수집 시점을 나타내는 기본 네트워크와 내용 분석 시점을 나타내는 활성 네트워크 사이에 구조적으로 큰 차이가 없었으나, 기관 유형별로 다른 기관들을 링크하는 목적이 다르게 나타났다. 그리고 한국 학술적 웹 공간은 여러 중앙성 지수들과 결속계수 간의 설명력이 유사하게 나타나는 형태의 네트워크임을 확인하였다.

ABSTRACT

Since the Web is dynamic, it is necessary to analyze scholarly Web space with both quantitative and qualitative methods for better understanding of communication characteristics. In this study, we analyzed contents of pages and links to ascertain the characteristics of Korean scholarly Web space in terms of network structure and communication behavior. The result shows that the structure of the original network with all the external links remained is not much different from that of the network with activated external links only. However, the purposes of linking vary among scholarly institutions. The centrality measures correlate more strongly with the clustering coefficient than with the constraint index implying the similar explanatory power of the two types of structural indices.

키워드: 웹계량학, 중앙성, 사회연결망분석, 학술적 웹 공간, 내용 분석

webometrics, centrality, social network analysis, scholarly Web, content analysis

* 연세대학교 문헌정보학과 교수(ymchung@yonsei.ac.kr) (제1저자)

** 연세대학교 대학원 문헌정보학과 박사과정(sweet798@yonsei.ac.kr) (공동저자)

■ 논문접수일자: 2009년 6월 12일 ■ 최초심사일자: 2009년 6월 15일 ■ 게재확정일자: 2009년 6월 25일

■ 정보관리학회지, 26(3): 7-24, 2009. [DOI:10.3743/KOSIM.2009.26.3.007]

1. 서론

웹계량학(Webometrics)에서 학술적 웹 공간의 링크 분석은 주로 대학 사이트들을 중심으로 활발하게 이루어졌다(Li et al. 2005; Li 2004; Thelwall and Tang 2003; Thelwall 2004). 이 연구들은 대학의 웹 사이트들이 받는 링크와 연구성과의 상관관계를 분석하거나, 주고받는 링크 구조를 시각화하여 학술적 웹 사이트들의 연관성을 확인하였다.

그러나 웹 사이트들은 페이지 및 링크의 생성과 소멸이 활발하기 때문에 웹 상의 링크도 유동적으로 변하게 된다. 따라서 웹 링크 분석에서 링크의 수로 표현되는 계량적 측정과 함께 내용 분석 등을 통해 링크와 페이지의 특성을 살펴보는 것이 필요하다. 즉, 소스 페이지와 목표 페이지 및 이들을 연결하는 링크의 유형을 분류함으로써 웹 공간의 구조적 특성을 심층적으로 살펴볼 수 있을 것이다.

페이지 유형을 분류함으로써 웹 공간을 정성적으로 분석해 보고자 하는 시도들은 웹계량학 분야에서 지속적으로 이루어졌다. 초기에는 웹 공간의 페이지를 무작위로 수집하여 페이지 유형 분류체계를 작성하고 이를 토대로 페이지들을 분류하였으며, 최근에는 자동적으로 페이지 유형을 분류하고자 하는 연구들까지 나타나고 있다. 페이지 및 링크 분류와 관련된 선행연구들은 주로 대학이 중심이 되는 학술적 웹 공간에서 소규모 데이터를 이용하여 분류체계를 작성하고 실제 분류를 수행하여 페이지 유형을 확인하는 것이다. 그러나 기존 연구들의 페이지 유형 분류 기준에는 웹 페이지의 내용과 형식 관련 특성이 혼합되어 나타나

고 있음을 볼 수 있다(Bar-Ilan 2004; Bar-Ilan 2005; Haas and Grams 2000; Harris et al. 2004; Wilkinson et al. 2003).

이 연구에서는 내용 분석을 통해 한국의 학술적 웹 공간을 보다 심층적으로 분석하고 그 특성을 살펴보고자 하였다. 이를 위해 보다 확장된 학술적 웹 공간에 적용할 수 있도록 기존 연구들과 달리 내용과 형식을 구분한 체계적인 웹 페이지 분류체계를 만들고 이를 기반으로 링크 유형 분류체계를 작성하였다. 그리고 분류된 웹 페이지 및 링크의 유형을 반영하여 한국의 학술적 웹 공간을 시각화하고, 링크 유형에 따른 네트워크의 구조적 특성의 차이를 살펴보았다.

이 연구를 통해 살펴보고자 한 구체적인 연구 문제는 다음과 같다.

- 한국 학술적 웹 공간에서 소스 페이지 및 목표 페이지로 출현하는 페이지 유형에 차이가 있는가?
- 한국 학술적 웹 공간에서 기관 유형에 따라 주고받는 링크 유형에 차이가 있는가?
- 내용 분석을 실시하지 않은 네트워크와 내용 분석을 통해 확인된 링크를 반영한 학술 네트워크는 구조적으로 차이가 있는가?
- 내용 분석을 통해 확인된 학술적 링크에 의한 네트워크와 비학술적 링크에 의한 네트워크는 구조적으로 차이가 있는가?
- 한국 학술적 웹 공간에서 네트워크의 구조적 지수들이 기관 유형별로 차이가 있는가?
- 구조적 지수들은 한국 학술적 웹 공간의 링크 유형에 따라 설명력에 차이가 발생하는가?

다음 문제에 대한 해답을 얻기 위하여 한국의 주요 대학, 전문연구정보센터(이하 NRIC), 정부출연연구원(이하 정출연)으로 구성된 학술기관 사이트들을 한국 학술적 웹 공간으로 설정하였다(Chung et al. 2009).

2. 데이터 수집 및 분석 방법

2.1 데이터 수집

학술적 웹 공간 분석을 위해 크롤러를 이용하여 학술적 웹 공간의 웹 페이지 및 링크를 수집하고 정련하여 각 사이트들 간의 외부링크를 확인하였다. 웹 링크 데이터 및 페이지 데이터는 전문 웹 크롤러인 SocSciBot3와 자체 제작한 크롤러인 RESPAC(Research Space Crawler)을 이용하여 수집하였다. 그리고 크롤링이 불가능했던 도메인들은 수작업으로 데이터를 수집하였다.

데이터 수집 대상으로는 2006년도 중앙일보 대학평가 결과 상위 10위권 내 11개 대학, 과학기술분야의 18개 정부출연연구원, 17개 과학기술분야 전문연구정보센터(NRIC)를 선정하였다(부록 1 참조). 크롤러에 의한 데이터 수집과 데이터 정련 작업은 2007년 3월 13일부터 5월 8일 사이에 진행되었으며, 수작업에 의한 데이터 추가 수집이 5월 20일까지 진행되었다. 데이터 수집은 5명의 연세대학교 문헌정보학과 대학원생이 여러 대의 PC 상에서 동시에 수행하였다.

데이터 수집 결과 이 연구에서 정의한 학술적 웹 공간은 총 427,445개의 페이지로 구성되

었으며, 페이지 수준의 링크 수는 총 147,378개로 나타났다.

2.2 웹 페이지 및 링크 유형 분류

웹 페이지 및 링크 데이터를 수집한 후, 46개 기관 간에 외부링크로 연결된 웹 페이지 및 링크에 대해 유형 분류를 수행하였다. 한국 웹 페이지의 특성상 페이지 내의 그림이나 아이콘 등이 많아서 크롤링 시 저장된 페이지로는 분류를 위한 내용을 확인하기 어려운 경우 수집된 외부링크 URL로 직접 접근하여 페이지 유형을 분류하였다.

웹 페이지 유형 분류의 첫째 기준은 페이지의 내용적 측면이며, 이를 통한 분류가 애매한 경우에는 해당 페이지가 위치하는 메뉴명을 참고하였다. 이 연구에서는 패킷 기반 웹 페이지 유형 분류체계를 기존 연구에서 사용한 분류표와 예비실험에서 실제 분류한 결과를 바탕으로 작성하였다.

각 페이지는 자원 형식, 엔티티, 내용 특성, 언어의 4개 패킷의 조합으로 표현되며, 각 패킷 아래 실제 페이지에 부여되는 하위 패킷들이 오도록 하였다. 그리고 모든 패킷에는 '접근불능'의 하위 패킷을 두어 접근이 불가능한 페이지에는 4개 패킷 모두에서 이 하위 패킷을 부여하도록 하였다.

자원 형식 패킷은 페이지의 표현 형식을 나타내는 패킷이며 텍스트, 비텍스트, 혼합, 접근불능의 4개 하위 패킷을 갖는다. 그리고 엔티티 패킷은 웹 페이지의 내용 또는 웹 페이지의 내용의 주체가 되는 객체를 표현하는 패킷이다. 엔티티 패킷에는 대학, 연구원(research institute),

연구실·연구단(laboratory), 학과, 연구센터, 행정부서, 교수, 학생, 연구자, 행정직원, 일반 이용자, 디지털도서관, 커뮤니티, 부설기관, 기타, 접근불능의 총 16개 하위 패킷이 있다. 엔티티 하위 패킷들은 크게 특정 부서 및 기관인 경우, 사람인 경우, 단체 또는 서비스인 경우로 구분된다.

엔티티 하위 패킷 중 일부는 기관의 유형에 따라 자동적으로 부여된다. 예를 들면, 대학의 경우 대학, 학과, 행정부서, 연구센터 등의 하위 패킷을 부여받을 수 있다. NRIC에는 행정부서나 연구센터 하위 패킷만을 부여할 수 있고, 정출연에는 연구원 패킷을 각 사이트의 홈 페이지, 연구실·연구단 패킷을 각 연구분과 및 연구부서에 부여할 수 있다.

내용 특성 패킷은 웹 페이지의 성격, 의도를 나타내는 패킷으로 '일반정보제공', '행정적', '일반학술적', '교육적', '연구중심적', '사회적', '개인적', '향해적', '접근제한적', 기타, 접근불능의 11개의 하위 패킷이 있다. 그리고 언어 패킷은 페이지의 작성 언어로 한글, 영문, 기타 언어, 접근불능의 하위 패킷이 있다.

이 페이지 분류체계를 이용하여 46개 기관 사이트 간에 외부링크로 연결된 51,477개의 소스 페이지와 147,378개의 목표 페이지의 유형을 분류하였다. 분류 작업은 연세대학교 문헌정보학과 대학원생 4명에 의해 2007년 4월부터 2007년 7월에 걸쳐서 이루어졌다.

링크 유형은 소스 페이지와 목표 페이지의 내용 특성에 따라서 분류하는 방식을 사용하였다. 링크 유형은 '관계적', '학술적', '향해적', '인증요청', '비학술적'의 총 5가지이다. 각 링크 유형에 해당되는 소스 페이지와 목표 페이지의

유형은 <표 1>과 같다.

2.3 내용 분석에 따른 네트워크 구조 분석

대부분의 선행연구에서는 수집된 링크 데이터를 그대로 사용하여 웹 공간을 네트워크 형태로 시각화하고 있으나 이 연구에서는 수집된 페이지와 링크에 대한 내용 분석결과를 반영하여 모두 네 가지 유형의 네트워크를 생성한 다음 이들을 비교 분석하였다. 먼저 기본 네트워크는 기관 간의 주고받는 모든 외부 링크로 이루어지며 내용 분석을 실시하지 않아 링크 유형을 알 수 없는 네트워크이다. 반면 활성 네트워크는 링크 유형 분류 결과 소스 페이지나 목표 페이지가 접근불능 페이지인 경우를 제외한 나머지 링크에 의해 연결된 네트워크이다. 또한 기본 네트워크를 <표 1>에서 정한 링크 유형에 따라 학술적 링크 네트워크와 비학술적 링크 네트워크로 분리되어 시각화하였다.

<표 1> 링크 유형 분류체계

링크 유형	소스페이지 → 목표페이지
관계적 (Conventional)	일반정보제공 → 일반정보제공
	일반정보제공 → 연구중심적
학술적 (Scientific)	일반학술적 → 일반정보제공
	일반학술적 → 행정적
	일반학술적 → 일반학술적
	일반학술적 → 교육적
	일반학술적 → 연구중심적
	일반학술적 → 개인적
	일반학술적 → 향해적
	교육적 → 일반정보제공
	교육적 → 연구중심적
	연구중심적 → 일반정보제공

학술적 (Scientific)	연구중심적 → 행정적
	연구중심적 → 일반학술적
	연구중심적 → 연구중심적
	연구중심적 → 개인적
	연구중심적 → 항해적
	개인적 → 일반학술적
	개인적 → 연구중심적
항해적 (Navigational)	일반정보제공 → 항해적
	행정적 → 항해적
	개인적 → 항해적
	항해적 → 일반정보제공
	항해적 → 행정적
	항해적 → 일반학술적
	항해적 → 연구중심적
	항해적 → 개인적
항해적 → 항해적	
인증요청 (Authorized)	일반정보제공 → 접근제한적
	행정적 → 접근제한적
	일반학술적 → 접근제한적
	교육적 → 접근제한적
	연구중심적 → 접근제한적
	개인적 → 접근제한적
	항해적 → 접근제한적
	접근제한적 → 일반정보제공
	접근제한적 → 행정적
	접근제한적 → 연구중심적
접근제한적 → 접근제한적	
비학술적 (Non-Scientific)	일반정보제공 → 사회적
	행정적 → 일반정보제공
	행정적 → 행정적
	행정적 → 일반학술적
	행정적 → 연구중심적
	행정적 → 개인적
	사회적 → 일반정보제공
	개인적 → 일반정보제공
개인적 → 개인적	

링크 유형에 따른 네트워크의 시각화에는 각 노드마다 일정한 수의 최근접 노드를 연결하는 방법으로 그래프를 생성하는 최근접 이웃 그래프 방식을 이용하였다(이재윤 2006, Eppstein

et al. 1997). 이 연구에서는 최근접 이웃 노드의 수를 최대 링크 가중치를 가지는 노드의 수로 하고 최대값을 가지는 최근접 이웃 노드가 2개 이상일 때 이를 모두 연결하였다.

그리고 링크 유형에 따른 네트워크 각각에서 구조적 지수를 산출하여 학술적 웹 공간의 특성을 보다 자세히 살펴보고 지수들 간의 관계를 분석하였다. 구조적 지수로는 중앙성 지수와 구조적 공백을 알 수 있는 제약성 지수, 결속 계수를 이용하였다. 이 연구의 학술적 웹 공간은 방향성 네트워크이기 때문에 중앙성 지수는 인링크, 아웃링크, 전체를 각각 고려하여 산출하였다. 연구에 사용된 지수는 <표 2>와 같다(유소영, 이재윤 2008).

네트워크의 연결 구조에서 발생하는 위세성이나 중앙성과 같은 성질을 측정하기 위한 대표적인 개념으로 중앙성(centrality)이 있다. 중앙성 지수는 어떤 노드가 그 네트워크 내에서 중심에 위치하는 정도를 반영하는 것으로서 다양한 공식들이 제안되어 왔는데 대표적으로 연결정도중앙성(degree centrality), 인접중앙성(closeness centrality), 사이중앙성(betweenness centrality)이 있다(Freeman 1979; Scott 2000; 김용학 2007).

연결정도중앙성은 한 노드가 주변의 인접 노드들과 연결이 많을수록 중심에 위치한다고 보는 지역중앙성(local centrality) 지수이다. 그리고 인접중앙성은 한 노드가 네트워크 내 다른 노드들과 갖는 거리의 가까운 정도를 나타내고, 사이중앙성은 한 노드가 어떤 노드와 다른 노드들 사이에서 중개자 역할을 하는 정도를 나타낸다. 인접중앙성과 사이중앙성은 네트워크 전체 내의 노드 간 연결을 고려하는 전역

〈표 2〉 네트워크의 구조적 지수

유형	지수	기호	공식
중앙성 지수	연결정도중앙성	degree degree_in degree_out	$C_i = \frac{\sum_j Z_{ij} + Z_j}{\sum_i \sum_j g_{ij}}$
	인접중앙성 (Sabidussi 1966)	close close_in close_out	$[\sum_{j=1}^g d(n_i, n_j)]^{-1}$ $d(n_i, n_j)$: 노드 i, j 간의 최단 경로 거리
	사이중앙성 (Freeman 1979)	betw	$\frac{\sum_i \sum_j g_{ij}(m)}{(N^2 - 3N + 2)/2} \cdot i \neq j, i < j$ $g_{ij}(m)$: 노드 i, j 간의 최단 경로 내 노드 m이 위치하는 경우의 수
구조적 공백 지수	제약성	constr	$C_{ij} = (p_{ij} + \sum_{q=1}^n p_{iq}p_{qj})^2, q \neq i, j$ $p_{i,j}$: i가 j에게 투자하는 자본의 양
결속 계수	결속계수 1	CC1	$\frac{2 \times E(G1(i)) }{k(i) \times (k(i) - 1)}$ $k(i)$: i의 이웃 노드 수 $ E(G1(i)) $: 노드 i의 1단계 이웃 노드들 간의 연결선 수
	결속계수 2	CC2	$\frac{ E(G1(i)) }{ E(G2(i)) }$ $ E(G1(i)) $: 노드 i의 1단계 이웃 노드들 간의 연결선 수 $ E(G2(i)) $: 노드 i의 1, 2단계 이웃 노드들 간의 연결선 수

중앙성(global centrality) 지수이다.

연결정도중앙성과 인접중앙성은 방향성 네트워크에도 적용가능하다. 하지만 위의 세 가지 중앙성 지수는 모두 링크가중치를 반영해 주지는 못하기 때문에 이를 보완하기 위해 위세중앙성을 사용할 수 있다(Bonacich 1987).

제약성은 사회적 자본의 분배를 설명하기 위해 도입한 '구조적 공백(structural hole)'을 나타내 주는 지수로 한 노드와 다른 노드 그리고 이들과 제 3자간의 관계를 고려한다. 사회학에서는 사회 조직 네트워크에서 구조적 공백에 위치할수록 일반적으로 좋은 성과를 내고 양질의 정보를 획득할 수 있음을 연구를 통해 확인하였다(Burt 1982; 김용학 2007). 그러나 구조적 공백이라는 지수는 존재하지 않으며 비중복성과 낮은 제약성 값으로 네트워크 내에서 구

조적 공백을 찾아낸다(Burt 1992).

끝으로 결속계수(clustering coefficient)는 한 노드의 지역적 연결성을 살펴보는 지수 중 하나로 제약성과 마찬가지로 삼각관계를 이용해 계산한다. 즉 어떤 노드와 연결된 두 노드가 있을 때 이들 간에도 연결이 존재하는지를 측정한다. Watts와 Strogatz(1998)는 “좁은 세상” 네트워크에서는 결속계수 값이 커진다는 것을 확인하였다(Watts and Strogatz 1998).

3. 학술적 웹 공간 분석 결과

3.1 페이지 유형 분류 결과

페이지 유형 분류 결과 소스 페이지 중 실제

로 접근이 가능한 페이지들은 45% 정도로서 목표 페이지의 80% 수준과 비교할 때 매우 낮은 편이었다. 이는 소스 페이지와 목표 페이지의 유형 차이에서 오는 결과로 생각된다. 실제로 소스 페이지들의 내용 특성 패킷은 비교적 다양한 하위 패킷을 부여받은 반면, 목표 페이지의 내용 특성 패킷은 주로 '일반정보제공'인 경우가 많았다. 즉 소스 페이지와 목표 페이지의 유형이 다르게 나타났는데, 이는 소스 페이지는 학과 내 관련 링크 모음이나 연구 내용 페이지와 같은 페이지가 많은 반면에 목표 페이지는 각 기관 사이트의 홈페이지 또는 학과나 연구소, 연구단위의 홈페이지와 같이 소스 페이지보다는 웹 사이트상에서 깊은 수준에 위치하는 페이지가 많기 때문인 것으로 보인다. 실제로 하위 디렉토리에 위치하는 페이지들은 페이지의 관리 및 갱신이 원활하지 않을 뿐만 아니라 접근성이 메인 페이지에 비해 떨어지는 경향이 있다고 볼 수 있다. 따라서 소스 페이지 중 50% 미만의 페이지만을 실제적으로 분류할 수 있었던 것으로 보인다.

소스 페이지는 자원 형식 패킷이 대부분 텍스트였으며 엔티티 패킷은 연구센터(28.48%), 디지털도서관(15.15%)이 주로 나타났다. 이는 소스 페이지 중 NRIC가 많은 비율을 차지하기 때문으로 생각된다. 그리고 내용 특성 패킷은 '연구중심적'인 경우(32.92%)가 많았는데 이 또한 NRIC가 많은 소스 페이지를 가지고 있기 때문으로 분석된다. 그리고 언어 패킷 분석 결과, 소스 페이지 중 37.01%가 한글로 작성되었음을 알 수 있었다.

목표 페이지는 자원 형식 패킷이 소스 페이지와 마찬가지로 대부분이 텍스트 형태로 되

어 있었다. 그리고 엔티티 패킷은 디지털도서관(32.49%), 연구센터(25.92%), 연구원(research institute, 18.75%) 순으로 많이 나타났다. 목표 페이지의 내용 특성 패킷은 대부분 '일반정보제공'(67.69%) 유형으로 나타나서 각 연구센터 및 연구원, 디지털도서관 등의 홈페이지로의 링크가 많은 것을 알 수 있었다. 마지막으로 언어 패킷은 대부분 한글(77.86%)로 나타나 한글로 된 페이지가 다수임을 알 수 있었다.

3.2 링크 유형 분류 결과

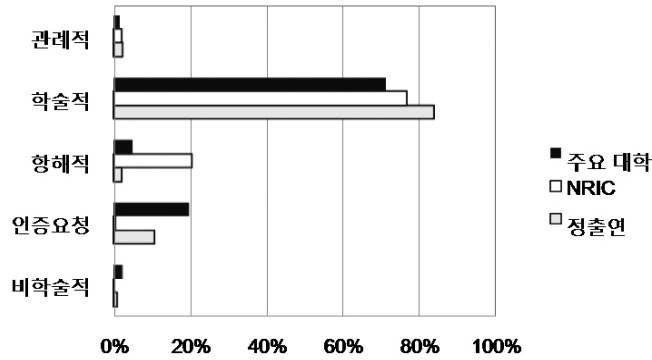
링크를 분류한 결과 총 53,392개의 링크가 링크 유형 분류체계에 따라 분류되었다. 소스 페이지와 목표 페이지 중 어느 한 쪽이 접근 불가능한 경우이거나, 두 페이지 모두 접근 불가능하여 분류되지 못한 경우에는 링크 유형을 분류하지 못했다.

링크 유형 중 가장 많이 나타난 것은 '학술적 링크'로서 총 41,797개(78.28%)가 출현하였고, 그 다음으로는 '인증요청' 링크가 7,476개(14.00%), '항해적' 링크가 2,197개(4.11%), '관계적' 링크가 1,094개(2.05%), '비학술적' 링크가 828개(1.55%)로 분류되었다.

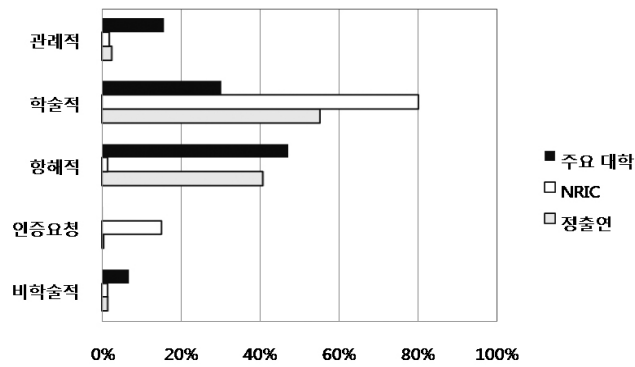
인링크와 아웃링크의 기관 유형별 출현 비율은 <그림 1>, <그림 2>와 같다.

인링크를 볼 때 세 가지 유형의 기관 모두 '학술적'인 링크를 가장 많이 받고 있는 것으로 나타났다. 또 NRIC는 '항해적' 링크를, 정출연과 주요 대학은 '인증요청' 링크를 그 다음으로 많이 받고 있었다.

아웃링크에서는 기관 유형별로 아웃링크 유형의 차이가 있었다. NRIC는 '학술적' 링크를



〈그림 1〉 인링크 유형 비율



〈그림 2〉 아웃링크 유형 비율

가장 많이 보내는 반면, 정출연은 ‘학술적’ 링크와 ‘항해적’ 링크를 많이 보내고 있었으며, 주요 대학은 ‘항해적’ 링크, ‘학술적’ 링크, ‘관례적’ 링크 등을 많이 보내고 있는 것으로 나타났다.

3.3 학술 네트워크 시각화

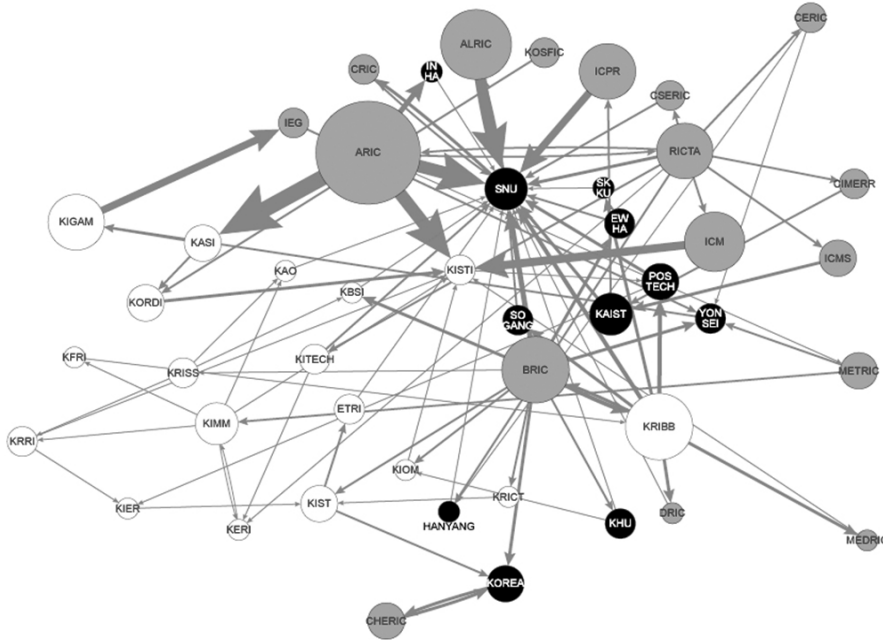
3.3.1 내용 분석에 따른 네트워크 비교

소스 페이지와 목표 페이지의 내용에 기초하여 분류한 링크 유형에 따라 학술 네트워크들을 생성하고 이를 기본 네트워크와 비교하였다. 네트워크 생성에는 최근접 이웃 그래프 방식을

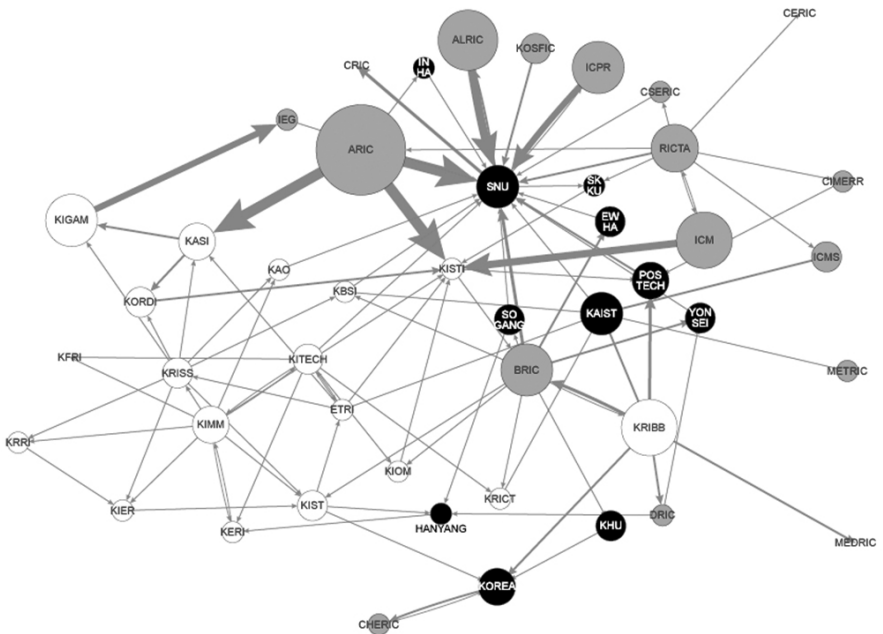
이용하였으며, 각 노드의 크기는 한 기관에서 나가는 아웃링크의 총수를 반영하였다. 또한 기관 유형을 노드의 색상을 달리하여 표시하였는데 검정색은 주요 대학, 회색은 NRIC, 흰색은 정출연을 나타낸다.

〈그림 3〉과 〈그림 4〉를 살펴보면 기본 네트워크와 활성 네트워크에서 링크를 많이 주고받는 주요 기관 노드들은 유사한 것으로 나타나 있다.

두 네트워크 모두에서 ARIC은 SNU, KASI, KISTI로 아웃링크를 많이 보내고 있는 최근접 이웃이며, SNU는 ARIC, ALRIC, ICPR로부



<그림 3> 기본 네트워크의 구조



<그림 4> 내용 분석을 반영한 활성화 네트워크의 구조

터 아웃링크를 많이 받고 있는 최근접 이웃이다. SNU는 두 네트워크 모두에서 다른 기관 노드들에게 아웃링크를 많이 받는 최근접 이웃인 경우가 가장 빈번하고 BRIC과 RICTA는 다른 기관 노드들에게 인링크를 많이 보내는 최근접 이웃인 경우가 빈번한 노드이다.

기관 유형별로 링크의 특징을 살펴보면 정출연은 두 네트워크 모두에서 아웃링크를 많이 보내는 최근접 이웃 노드의 기관 유형이 대부분 정출연이다. 예외적인 경우는 KIGAM과 KRIBB인데 KRIBB는 주요 대학인 POSTECH, NRIC인 BRIC, DRIC 등에 아웃링크를 주로 보내고 있고, KIGAM도 NRIC인 IEG로 많은 아웃링크를 보내고 있다. 활성 네트워크에서는 KRIBB가 SNU보다 KOREA의 최근접 이웃이 된다는 차이가 있기는 하지만 대체적으로 주요 노드들은 동일하였다. 그리고 KASI와 KISTI는 NRIC인 ARIC으로부터 많은 아웃링크를 받고 있다.

반면에 NRIC는 두 네트워크 모두에서 아웃링크를 많이 보내는 최근접 이웃이 주요 대학이거나 정출연이다. NRIC 간에 최근접 이웃인 경우는 두 네트워크에서 RICTA를 중심으로 출연하고 있는데 기본 네트워크에서는 RICTA와 ICM, RICTA와 CIMERR, RICTA와 ICMS가 있고 활성 네트워크에서는 RICTA와 ICM, RICTA와 CIMERR, RICTA와 CSERIC, RICTA와 CERIC이 있다. 그리고 ARIC은 원 네트워크에서는 METRIC과, 활성 네트워크에서는 IEG와 최근접 이웃 관계를 보인다.

그리고 주요 대학은 SNU가 ARIC, ALRIC, ICPR, BRIC과 같은 대부분의 NRIC와 정출연인 KRIBB과 최근접 이웃으로 연결되어 있다.

주요 대학 중에는 KAIST, POSTECH 등이 SNU와 최근접 이웃 노드로 연결되어 있다.

따라서 수집 시점의 기본 네트워크와 내용 분석을 통해 활성 링크를 확인한 시점의 네트워크에서 주요 노드들은 크게 차이가 없는 것으로 보인다. 하지만 주고받는 링크의 수에서는 차이를 보였다. 즉 내용 분석을 통해 활성 링크를 확인함으로써 주고받는 링크의 규모를 보다 심층적으로 파악할 수 있었다.

학술적 웹 공간에서 링크 유형에 따른 네트워크 구조에 차이가 있는지 살펴보기 위하여 활성 링크 중 학술적 웹 공간에서 의미가 있다고 할 수 있는 학술적 링크와 비학술적 링크만을 구분하여 네트워크를 비교해 보았다. 학술적 링크와 비학술적 링크는 최근접 이웃 그래프 방식을 사용하지 않고 모든 링크를 전부 표현하였다. 여기에서 링크 가중치는 최대값이 11,362이고 최소값이 1로 그 차이가 크므로 시각화를 용이하게 하기 위하여 모든 링크가중치에 제곱근을 취하여 링크의 굵기를 표현하였다. 그리고 노드의 색상은 기관 유형을 나타내는데 검정색은 주요 대학, 회색은 NRIC, 흰색은 정출연이다.

<그림 5>와 <그림 6>을 비교하여 보면 학술적 링크에 의한 네트워크와 비학술적 링크에 의한 네트워크 간에 차이가 있음을 알 수 있다.

먼저 학술적 링크 네트워크의 경우, 기본 네트워크에서와 같이 SNU가 ARIC과 ALRIC, ICPR로부터 많은 수의 인링크를 받고 있으며 KISTI는 ARIC과 ICM으로부터 많은 수의 인링크를 받고 있다. 그리고 ARIC이 SNU, KISTI, KASI로 많은 수의 아웃링크를 집중적으로 내 보내고 있고, ICM, ICPR, ALRIC도 이와 유사

하게 KISTI 또는 SNU를 대상으로 많은 수의 아웃링크를 보내고 있다. BRIC은 다양한 기관에 아웃링크를 고르게 내보내고 있다.

그러나 비학술적 링크 네트워크의 경우 원 네트워크에서 보이던 경향이 다소 약화되었다. 이는 비학술적 링크에 의한 네트워크가 828개의 적은 수의 링크로 이루어졌기 때문인 것으로 보인다. 그러나 학술적 링크 네트워크에서와 마찬가지로 BRIC은 다양한 기관에 아웃링크를 고르게 내보내고 있는 것으로 나타났다. 반면에 ALRIC, ARIC, ICM의 아웃링크 수는 현저히 감소하여서 네트워크 내에서 강한 링크를 보이지 않고 있다. 이를 통해 볼 때 ALRIC, ARIC, ICM은 학술적인 아웃링크를 많이 내보내고 있는 것으로 파악되었다.

학술적 링크 네트워크와 비학술적 링크 네트워크의 차이는 각 네트워크에서 인링크 및 아웃링크가 전혀 없는 고립노드를 살펴보면 더욱 확실해진다.

학술적 링크 네트워크에서 고립노드는 총 8개로 이 중 KIER, KRIS, KRRI, KIOM, KAO, KIMM, KITECH는 정출연이고, CERIC은 NRIC이다. 반면에 비학술적 링크 네트워크의 고립노드는 총 6개로 이들은 모두 NRIC이다. 또 비학술적 링크 네트워크에서 정출연은 KRIBB을 통하여 모두 연결되어 있는 반면에 학술적 링크 네트워크에서는 이러한 연결이 보이지 않는다.

따라서 내용 분석 결과 한국 학술적 웹 공간에서는 기관 유형에 따라 주고받는 구체적인 링크 유형에 차이가 있음을 알 수 있었다. 즉 46개 기관 내에서 정출연은 상대적으로 학술적 링크를 주고받는 기관 수가 적고, 반대로 NRIC은 비학술적 링크를 주고받는 기관 수가 적은

것으로 나타났다.

3.4 학술 네트워크의 구조적 지수 분석

3.4.1 구조적 지수 간 상관관계 분석

한국 학술적 웹 공간의 링크 유형에 따른 구조적 지수 간의 설명력 차이를 보기 위하여 Spearman 검정을 통해 지수 간 상관관계를 분석하였다(구조적 지수 기호는 <표 2> 참조). 분석 결과 기본 네트워크(표 3 참조), 활성 네트워크(표 4 참조), 학술적 링크 및 비학술적 링크 네트워크 모두에서 인접중앙성과 연결정도중앙성 간에 0.8 이상의 높은 유의미한 상관관계를 보였다.

일반적으로 인접중앙성은 네트워크의 전역적 중앙성을 나타내고, 연결정도중앙성은 지역적 중앙성을 나타내는 것으로 본다. 그러나 기본 네트워크 및 내용 분석에 의한 세 개의 네트워크에서 모두 인접중앙성과 연결정도중앙성이 높은 상관관계를 보였다. 이처럼 전역적 중앙성 지수와 지역적 중앙성 지수의 설명력이 매우 유사하게 나타난 것은 한국 학술적 웹 공간이 지역적 중앙성이 높은 노드가 전역적 중앙성도 높은 중심가(downtown) 형태이기 때문인 것으로 분석된다. 또한 기본 네트워크와 활성 네트워크에서 사이중앙성도 연결정도중앙성 및 인접중앙성과 0.9 이상의 높은 상관관계를 보임으로써 이러한 특성을 재확인 할 수 있었다.

그리고 결속계수는 기본 네트워크와 내용 분석에 의한 세 개의 네트워크에서 제약성 지수를 제외한 지수들과 대체적으로 유의미한 상관관계를 보여 다른 중앙성 지수들과 설명력이 유사

〈표 3〉 기본 네트워크의 지수 간 상관관계(Spearman, 양측)

	bet	close_in	close_out	close	degree_in	degree_out	degree	cc1	cc2	constr
bet	1	0.690**	0.703**	0.862*	0.706**	0.703**	0.915*	-0.883*	0.862*	0.020
close_in	0.690*	1	0.198	0.665*	0.993**	0.203	0.728*	-0.590*	0.675*	-0.094
close_out	0.703*	0.198	1	0.781*	0.216	0.999**	0.779*	-0.726*	0.778*	-0.183
close	0.862*	0.665**	0.781**	1	0.688**	0.782**	0.966*	-0.926*	0.997*	-0.162
degree_in	0.706*	0.993**	0.216	0.688*	1	0.220	0.739*	-0.624*	0.697*	-0.085
degree_out	0.703*	0.203	0.999**	0.782*	0.220	1	0.782*	-0.727*	0.780*	-0.188
degree	0.915*	0.728**	0.779**	0.966*	0.739**	0.782**	1	-0.893*	0.969*	-0.184
cc1	-0.883*	-0.590**	-0.726**	-0.926*	-0.624**	-0.727**	-0.893*	1	-0.910*	0.078
cc2	0.862*	0.675**	0.778**	0.997*	0.697**	0.780**	0.969*	-0.910*	1	-0.152
constr	0.020	-0.094	-0.183	-0.162	-0.085	-0.188	-0.184	0.078	-0.152	1

* p < .05 ; ** p < .01

〈표 4〉 활성 네트워크의 지수 간 상관관계(Spearman, 양측)

	bet	close_in	close_out	close	degree_in	degree_out	degree	cc1	cc2	constr
bet	1	0.665*	0.745**	0.931**	0.642**	0.762**	0.949*	-0.895*	0.903*	-0.143
close_in	0.665**	1	0.178	0.665**	0.960**	0.204	0.676*	-0.536*	0.691*	-0.255
close_out	0.745**	0.178	1	0.770**	0.124	0.992**	0.802*	-0.757*	0.736*	-0.250
close	0.931**	0.665*	0.770**	1	0.647**	0.783**	0.980*	-0.919*	0.989*	-0.300*
degree_in	0.642**	0.96*	0.124	0.647**	1	0.150	0.644*	-0.502*	0.671*	-0.196
degree_out	0.762**	0.204	0.992**	0.783**	0.15	1	0.822*	-0.774*	0.746*	-0.312*
degree	0.949**	0.676*	0.802**	0.980**	0.644**	0.822**	1	-0.917*	0.961*	-0.341*
cc1	-0.895**	-0.536*	-0.757**	-0.919**	-0.502**	-0.774**	-0.917*	1	-0.866*	0.224
cc2	0.903**	0.691*	0.736**	0.989**	0.671**	0.746**	0.961*	-0.866*	1	-0.311*
constr	-0.143	-0.255	-0.250	-0.300*	-0.196	-0.312*	-0.341*	0.224	-0.311*	1

* p < .05 ; ** p < .01

한 것으로 나타났다. 그러나 비학술적 링크 네트워크에서 CC1은 일부 중앙성 지수(사이중앙성, 외향 연결정도중앙성, 외향 인접중앙성)와는 유의미한 상관관계를 보이지 않았는데 이는 네트워크의 특성에 기인한 것으로 생각된다.

반면에 제약성 지수는 기본 네트워크와 학술적 링크 네트워크, 비학술적 링크 네트워크에서 다른 구조적 지수들과 유의미한 상관관계를 보이지 않았고, 활성 네트워크에서만 인접중앙성, 외향 연결정도중앙성, 연결정도중앙성, 결속계수와 0.35 이하의 약한 상관관계를 보였다.

즉, 제약성 지수는 대체적으로 다른 구조적 지수들과 유사하지 않으며 설명하는 부분에 차이가 있는 것으로 해석된다.

한국 학술적 웹 공간은 링크 유형에 관계없이 중앙성 지수들과 결속계수들이 유사한 구조적 설명력을 가지는 반면 제약성 지수는 이들과는 설명하는 부분이 다르다는 것을 확인하였다.

3.4.2 기관 유형에 따른 구조적 지수 차이 분석

내용 분석 이전의 기본 네트워크와 내용 분

석 결과를 반영한 네트워크에서 각각 기관 유형별로 구조적 지수 값에 유의미한 차이가 있는지를 Kruskal-Wallis 검정을 통해 살펴보았다(표 5 참조). 분석 결과, 내향 인접중앙성 및 인접중앙성, 내향 연결정도중앙성 및 연결정도중앙성이 한국 학술적 웹 공간에서 링크의 유형에 관계없이 기관 유형별로 유의미한 차이를 보였고 그 외 구조적 지수들은 링크 유형에 따라 다르게 나타났다. 즉, 한국 학술적 웹 공간에서 인링크만을 고려하거나 인링크/아웃링크를 모두 고려할 때 각 기관 사이트들의 가까운 정도나 직접 연결 정도가 기관 유형에 따라 차이를 보였다. 그리고 이는 기본 네트워크와 활성 네트워크, 학술적 링크 네트워크, 비학술적 링크 네트워크 모두에서 동일하였다.

그러나 외향 인접중앙성과 외향 연결정도중앙성은 학술적 링크 네트워크에서만 기관유형별로 유의미한 차이를 보여서, 아웃링크에 의한 각 기관 사이트들의 가까운 정도나 직접 연결 정도가 학술적 링크 네트워크에서만 기관 유형별로 차이가 있음을 알 수 있었다.

그리고 사이중앙성은 원 네트워크와 활성 네트워크에서만 기관유형별로 유의미한 차이를 보였다. 즉, 한 기관 사이트가 다른 두 사이트들 사이의 최단 경로에 위치하는 정도는 기본 네트워크와 활성 네트워크에서 기관 유형별로 차이가 있었다. 또 제약성 지수와 결속계수(CC1, CC2)는 학술적 링크 네트워크에서만 기관 유형에 따른 차이가 보이지 않았다.

4. 결론

이 연구에서는 웹 공간의 내용 분석을 통해 페이지 및 링크 유형을 분류한 다음 링크 유형을 반영하여 학술적 웹 공간을 네트워크 형태로 시각화하고, 이들 네트워크의 구조적 특성을 분석하였다. 연구 결과 한국 학술적 웹 공간의 다음과 같은 특성을 갖는 것으로 나타났다.

먼저 외부링크에 의해 연결된 소스 페이지 및 목표 페이지는 모두 한글로 작성된 텍스트 형태가 대부분이었으나 두 유형의 페이지가 담

〈표 5〉 기관 유형에 따른 구조적 지수의 차이 검정(Kruskal-Wallis, 양측)

네트워크 유형	K-W	betw	close_in	close_out	close	degree_in	degree_out	degree	CC1	CC2	constr
원 네트워크	카이제곱	6.941	25.637	4.326	10.79	25.813	4.203	11.553	6.855	11.541	9.113
	자유도	2	2	2	2	2	2	2	2	2	2
	asympt. sig.	0.031*	0.000*	0.115	0.005*	0.000*	0.122	0.003*	0.032*	0.003*	0.010*
활성 네트워크	카이제곱	7.567	20.685	0.314	9.66	18.992	0.547	9.574	4.777	11.092	11.382
	자유도	2	2	2	2	2	2	2	2	2	2
	asympt. sig.	0.023*	0.000*	0.855	0.008*	0.000*	0.761	0.008*	0.092	0.004*	0.003*
학술적 링크 네트워크	카이제곱	0.202	9.062	12.006	7.005	11.742	9.745	5.872	2.164	2.019	2.117
	자유도	2	2	2	2	2	2	2	2	2	2
	asympt. sig.	0.904	0.011*	0.002*	0.030*	0.003*	0.008*	0.053	0.339	0.364	0.347
비학술적 링크 네트워크	카이제곱	4.139	21.747	2.152	23.131	29.803	2.422	7.03	18.228	14.756	7.764
	자유도	2	2	2	2	2	2	2	2	2	2
	asympt. sig.	0.126	0.000*	0.341	0.000*	0.000*	0.298	0.030*	0.000*	0.001*	0.021*

* p < .05

고 있는 내용 특성에 있어서는 차이를 보였다. 소스 페이지는 주로 '연구중심적' 페이지인 반면 목표 페이지는 '일반정보제공적'인 페이지가 많았다. 그리고 소스 페이지와 목표 페이지 모두 디지털도서관과 연구센터 페이지가 많은 것으로 나타났다.

링크 분석 결과 주요 아웃링크 유형은 기관 유형별로 차이를 보였는데 주요 대학은 '항해적' 링크를, NRIC는 '학술적' 링크를, 그리고 정출연은 '학술적' 링크 및 '항해적' 링크를 가장 많이 내보내고 있었다. 반면에 주요 인링크 유형은 모두 '학술적' 링크로 나타나 모든 기관 유형에서 차이가 없었다.

한국 학술적 웹 공간을 기본 네트워크와 활성 네트워크로 나누어 구조화하였을 때 두 네트워크의 주요 노드와 주요 링크 구조는 매우 유사하게 나타났다. 즉 데이터 수집 시점을 나타내는 기본 네트워크와 내용 분석 시점을 나타내는 활성 네트워크에서는 링크를 주고받는 주요 기관 간에 큰 차이가 없었다. 그러나 학술적 링크에 의한 네트워크와 비학술적 링크에 의한 네트워크 구조에서는 고립노드가 되는 기관 유형이 다르게 나타났으며, 이를 통해 기관

유형에 따라 한 기관이 다른 기관들을 링크하는 의도나 목적이 다르다는 것을 알 수 있었다.

그리고 중앙성 지수, 제약성 지수, 결속계수를 통해 학술적 네트워크의 구조적 특성을 분석하고, 구조적 지수 간 설명력의 차이가 있는지 살펴본 결과 한국의 학술적 웹 공간은 인접 중앙성과 연결정도중앙성 간의 상관관계가 높은 중심가(downtown) 형태를 띠고 있음을 알 수 있었다. 또한 기본 네트워크와 내용 분석에 의한 네트워크 모두에서 연결정도중앙성, 인접 중앙성, 사이중앙성, 결속계수 간에는 설명력이 유사하게 나타난 반면, 구조적 공백을 나타내는 제약성 지수는 이 지수들과는 설명력이 다르게 나타났다.

이 연구는 웹 공간 분석에서 페이지 및 링크의 계량적 측정과 함께 내용 분석을 병행함으로써 학술적 웹 공간의 구조적 특성을 심층적으로 살펴보았다는 점에서 의미가 있다. 앞으로 이 연구에서 사용된 네트워크의 구조적 지수들 이외에 링크 가중치를 반영하는 지수들을 이용하여 한국 학술적 웹 공간의 특성을 보다 상세하게 파악할 필요가 있을 것이다.

참 고 문 헌

- 김용학. 2007. 『사회 연결망 분석』. 개정판. 서울: 박영사.
- 유소영, 이재운. 2008. "학제적 분야의 정보서비스를 위한 학술지 인용 분석에 관한 연구: Y대학교 생명공학과를 중심으로." 『정보관리학회지』, 25(4): 283-308.
- Bar-Ilan, J. 2004. "A microscopic link analysis of academic institutions within a country - the case of Israel." *Scientometrics*, 59(3): 391-403.

- Bar-Ilan, J. 2005. "What do we know about links and linking? A framework for studying inlinks in academic environments." *Information Processing and Management*, 41: 973-986.
- Bonacich, P. 1987. "Power and centrality: A family of measures." *American Journal of Sociology*, 92: 1170-1182.
- Burt, R. S. 1982. *Toward a Structural Theory of Action: Network Models of Social Structure, Perception and Action*. San Francisco: Academic Press.
- Chung, Y. M., S. Y. Yu, Y. K. Kim, and S. Y. Kim. 2009. "Characteristics and link structure of a national scholarly Web space: The case of South Korea." *Scientometrics*, 80(3). [in press].
- Freeman, L. C. 1979. "Centrality in social networks: Conceptual clarification." *Social Networks*, 1: 215-239.
- Haas, S. W. and E. S. Grams. "Readers, authors, and page structure: A discussion of four questions arising from a content analysis of Web pages." *Journal of the American Society for Information Science* 51(2):181-192.
- Harries, G., D. Wilkinson, L. Price, R. Fairclough, and M. Thelwall. 2004. "Hyperlinks as a data source for science mapping." *Journal of Information Science*, 30(5): 436-447.
- Li, X., M. Thelwall, D. Wilkinson, and P. Musgrove. 2005. "National and international university departmental Web site interlinking Part 1: Validation of departmental link analysis." *Scientometrics*, 64(2): 151-185.
- Thelwall, M. 2004. *Link Analysis: An Information Science Approach*. Amsterdam: Elsevier.
- Thelwall, M. and R. Tang. 2003. "Disciplinary and linguistic considerations for academic Web linking: An exploratory hyperlink mediated study with Mainland China and Taiwan." *Scientometrics*, 58(1): 155-181.
- Sabidussi, G. 1966. "The centrality index of a graph." *Psychometrika*, 31: 581-603.
- Scott, J. 2000. *Social Network Analysis: A Handbook*, 2nd ed. London: Sage Publications.
- SocSciBot3. <<http://socscibot.wlv.ac.uk/>>.
- Watts, D. and S. H. Strogatz. 1998. "Collective dynamics of 'small-world' networks." *Nature*, 393(6684): 440-442.
- Wilkinson, D., G. Harries, M. Thelwall, and E. Price. 2003. "Motivations for academic Web site interlinking: Evidence for the Web as a novel source of information on informal scholarly communication." *Journal of Information Science*, 29(1): 59-66.

<부록 1> 한국 학술적 웹 공간

기관 유형	기관명	기관 약자/기호	URL
주요 대학	KAIST	KAIST	http://www.kaist.ac.kr
	서울대	SNU	http://www.snu.ac.kr
	POSTECH	POSTECH	http://www.postech.ac.kr
	고려대	KOREA	http://www.korea.ac.kr
	연세대	YONSEI	http://www.yonsei.ac.kr
	성균관대	SKKU	http://www.skku.ac.kr
	한양대	HANYANG	http://www.hanyang.ac.kr
	서강대	SOGANG	http://www.sogang.ac.kr
	경희대	KHU	http://www.khu.ac.kr
	이화여대	EWHA	http://www.ewha.ac.kr
인하대	INHA	http://www.inha.ac.kr	
NRIC	건설연구정보센터	CRIC	http://www.cric.or.kr
	기계공학연구정보센터	METRIC	http://www.metric.or.kr
	농생명과학연구정보센터	ALRIC	http://www.alric.org
	물리화학연구정보센터	ICPR	http://icpr.or.kr
	산업공학연구정보센터	CIMERR	http://iems.net
	생물학연구정보센터	BRIC	http://bric.postech.ac.kr
	섬유이류연구정보센터	RICTA	http://www.ricta.or.kr
	수리과학연구정보센터	ICMS	http://www.mathnet.or.kr
	의약품연구정보센터	DRIC	http://dric.sookmyung.ac.kr
	의학연구정보센터	MEDRIC	http://www.medric.or.kr
	재료공학연구정보센터	ICM	http://icm.re.kr
	컴퓨터연구정보센터	CSERIC	http://cseric.cau.ac.kr
	토목연구정보센터	CERIC	http://www.ceric.net
	항공우주연구정보센터	ARIC	http://www.aric.or.kr
	해양수산연구정보센터	KOSFIC	http://kosfic.chonnam.ac.kr
	화학공학연구정보센터	CHERIC	http://www.cheric.org
	환경지질연구정보센터	IEG	http://ieg.or.kr
정출연	한국과학기술연구원	KIST	http://www.kist.re.kr
	한국과학기술정보연구원	KISTI	http://www.kisti.re.kr
	한국기계연구원	KIMM	http://www.kimm.re.kr
	한국기초과학지원연구원	KBSI	http://www.kbsi.re.kr
	한국생명공학연구원	KRIBB	http://www.kribb.re.kr
	한국생산기술연구원	KITECH	http://www.kitech.re.kr
	한국식품개발연구원	KFRI	http://www.kfri.re.kr
	한국에너지기술연구원	KIER	http://www.kier.re.kr
	한국전기연구원	KERI	http://www.keri.re.kr
	한국전자통신연구원	ETRI	http://www.etri.re.kr
	한국지질자원연구원	KIGAM	http://www.kigam.re.kr
	한국천문연구원	KAO	http://www.kao.re.kr
	한국철도기술연구원	KRRI	http://www.krri.re.kr
	한국표준과학연구원	KRISS	http://www.kriss.re.kr
	한국한의학연구원	KIOM	http://www.kiom.re.kr
	한국항공우주연구원	KASI	http://www.kari.re.kr
	한국해양연구원	KORDI	http://www.kordi.re.kr
한국화학연구원	KRICT	http://www.kRICT.re.kr	

<부록 2> 링크 유형별 출현 빈도

링크 유형(비율)	소스페이지 → 목표페이지	빈도	유형내 비율(%)
관례적(2.05%)	일반정보제공 → 일반정보제공	1,094	100.00
	전체(합계)	1,094	100.00
학술적(78.28%)	일반정보제공 → 연구중심적	161	0.39
	일반학술적 → 일반정보제공	4,114	9.84
	일반학술적 → 행정적	6	0.01
	일반학술적 → 일반학술적	13	0.03
	일반학술적 → 교육적	1	0.00
	일반학술적 → 연구중심적	99	0.24
	일반학술적 → 개인적	1	0.00
	일반학술적 → 항해적	2	0.00
	교육적 → 일반정보제공	43	0.10
	교육적 → 연구중심적	16	0.04
	연구중심적 → 일반정보제공	34,445	82.41
	연구중심적 → 행정적	1	0.00
	연구중심적 → 일반학술적	17	0.04
	연구중심적 → 연구중심적	2,683	6.42
	연구중심적 → 개인적	29	0.07
	연구중심적 → 항해적	163	0.39
	개인적 → 일반학술적	1	0.00
	개인적 → 연구중심적	2	0.00
	전체(합계)	41,797	100.00
항해적(4.11%)	일반정보제공 → 항해적	2	0.09
	행정적 → 항해적	2	0.09
	개인적 → 항해적	1	0.05
	항해적 → 일반정보제공	2,144	97.59
	항해적 → 행정적	1	0.05
	항해적 → 일반학술적	5	0.23
	항해적 → 연구중심적	18	0.82
	항해적 → 개인적	11	0.50
	항해적 → 항해적	13	0.59
	전체(합계)	2,197	100.00
인증요청(14.00%)	일반정보제공 → 접근제한적	155	2.07
	행정적 → 접근제한적	275	3.68
	일반학술적 → 접근제한적	144	1.93
	교육적 → 접근제한적	16	0.21
	연구중심적 → 접근제한적	2,553	34.15
	개인적 → 접근제한적	1	0.01
	항해적 → 접근제한적	13	0.17
	접근제한적 → 일반정보제공	4,054	54.23
	접근제한적 → 행정적	3	0.04
	접근제한적 → 연구중심적	168	2.25
	접근제한적 → 접근제한적	94	1.26
	전체(합계)	7,476	100.00
비학술적(1.55%)	일반정보제공 → 사회적	1	0.1
	행정적 → 일반정보제공	457	55.2
	행정적 → 행정적	6	0.7
	행정적 → 일반학술적	1	0.1
	행정적 → 연구중심적	280	33.8
	행정적 → 개인적	9	1.1
	사회적 → 일반정보제공	5	0.6
	개인적 → 일반정보제공	64	7.7
	개인적 → 개인적	5	0.6
	전체(합계)	828	100.00