

동시링크분석을 이용한 웹정보원의 지적구조 변화에 관한 연구

A Study on Changes of the Intellectual Structure in Web Information Using the Co-links Analysis

이 성 속(Sung-Sook Lee)*

초 록

이 연구는 웹정보원의 지적 구조를 동시링크분석을 이용하여 시기별 변화와 검색엔진별 차이를 중심으로 분석하였다. 지적 구조의 시기별 변화를 분석한 결과, 이차원 지도상에 나타난 웹정보원의 군집과 위치는 6년간의 시간흐름에 따른 지적 구조의 변화를 나타냈다.

AltaVista와 MSN Search 검색엔진을 대상으로 지적구조의 차이를 분석한 결과, 웹정보원 지도에 나타난 전체적인 지적 구조는 유사한 양상을 나타냈지만, 몇몇 웹정보원의 경우 소속 군집이 달라지는 경우가 발생했다. 인쇄 정보원에 적용되던 동시인용기법이 웹정보원에 대한 통시적 분석에도 응용될 수 있음을 확인하였다.

ABSTRACT

This research analyzed changes of the intellectual structure of web information by examining time changes and search engines using the co-links analysis. According to the results, the co-links web information clusters on the two maps appeared to contain changes in the intellectual structure over the two time periods.

The intellectual structure that appeared in the information map for AltaVista and MSN Search engines was relatively similar. However, there were also cases where the clusters of some web information was different. The results of the research revealed that the cocitation analysis could be applied simultaneously to diachronous analysis in the web information.

키워드: 웹정보원, 동시링크분석, 지적구조

www, webometrics, co-links analysis, altavista, msn search

* 충남대학교 사회과학대학 문헌정보학과 강사(inflee@hanmail.net)

■ 논문접수일자 : 2005년 5월 20일

■ 게재확정일자 : 2005년 6월 17일

1. 서론

1.1 연구목적

특정 학문분야에서 생산되는 학술문헌에는 해당학문 자체의 논리에 의해 체계적으로 구축된 독특한 연구방법, 학파, 이론 주제영역 등의 지적구조가 내재하게 된다. 이러한 지적 구조를 파악하는 것은 한 학문분야의 연구활동과 학문적 특성을 이해하는데 도움이 되며 그 분야의 정보이용 패턴을 파악하는 지름길이 되고, 원활한 정보유통체계를 구성하는 기반이 된다(조명희 1993).

계량정보학의 한 연구분야인 동시인용분석은 학문의 지적 구조를 규명하기 위한 유용한 도구로서 1973년 Small과 Marshakova에 의해 각각 제안되었다. 그동안 문헌, 저자, 용어를 대상으로 한 동시인용분석 연구는 특정 학문영역의 구조와 변화, 더 나아가서 학제적 연구영역을 밝히는 성과를 이루었다.

정보통신기술의 발전에 따른 정보환경의 변화는 사회전반에 커다란 영향을 주고 있다. WWW(이하 웹)은 그 이용과 기술면에서 계속적으로 변하고 있으며, 이에 따라 웹정보의 성장은 놀랄만한 속도로 진행되고 있다. 전 세계의 개인과 기관들은 웹을 통해 온라인으로 다양한 정보를 생산하고 있으며, 그 결과 웹상에는 단순한 정보에서부터 개인 혹은 기관에 관한 정보, 논문, 색인, 통계, 진행 중인 연구 등에 이르기까지 다양한 정보들이 존재한다. 오늘날 웹은 새로운 학술커뮤니케이션의 공간으로 인식되고 있다. 연구자들은 그들의 연구 성과를 웹을 통해 발표하며, 새로운 연구의 아이디어를 얻기

위한 의견을 웹공간에서 나누고, 연구동향을 웹 정보를 통해 파악한다.

이러한 학술정보유통의 변화 속에서 1990년대 중반 계량정보학적 연구기법을 웹에 적용하려는 webometrics가 새롭게 등장하였다. webometrics의 연구 활성화를 의미하듯, webometrics와 유사한 cybermetrics, netometrics, webometry, internetrometrics, web bibliometry 등 다양한 용어들이 존재한다. 학문의 지적구조를 파악하는 동시인용분석의 관점에서 보면, webometrics는 현재의 웹정보원을 대상으로 하기 때문에 일정 시간이 지난 과거의 지적 구조를 나타낸다는 동시인용분석의 한계를 극복하고 현재의 지적 구조를 파악할 수 있다는 장점이 있다.

이 연구는 웹정보원의 특성을 고려한 동시링크분석에 관한 연구이며 인쇄 정보원에 적용되던 동시인용기법이 웹정보원에 대한 통시적 분석에도 응용될 수 있는지의 여부를 검토하기 위한 것이다. 이를 위해 이 연구에서는 동시링크 분석을 이용하여 6년간의 시간흐름에 따른 웹정보원의 지적 구조의 변화를 비교하고, 지적 구조의 검색엔진별 차이를 분석함으로써 웹정보원의 특성과 연구방법상의 차이를 고찰하였다. 이 연구의 결과는 webometrics의 통시적 연구에 관한 기초 자료로 활용될 수 있을 것이다.

1.2 연구범위와 방법

이 연구에서는 우선 webometrics에 대한 선행연구를 검토한 후, 동시링크분석의 개념과 특성을 알아보았다. 실험에서는 특정 학문 분야의 지적 구조의 변화를 파악하기 위해서 1999년 정동열과 최윤미의 연구에서 나타난 커뮤니케

이전 분야의 지적 구조와 6년이 지난 2005년 현재 커뮤니케이션 분야의 지적 구조를 비교하였다. 이 연구의 범위와 방법은 두 시기의 지적 구조를 비교하기 위하여 1999년의 틀을 유지하되 웹정보원의 특성을 반영하도록 설계하였다.

실험 대상 선정 방법은 1999년 연구에서 사용된 웹정보원을 사용하되, 각 웹정보원의 링크수를 조사하고, 폐쇄여부, 동시링크빈도를 고려하여 웹정보원을 최종적으로 선정하였다. 즉 2005년 현재 폐쇄되었거나 동시링크빈도가 현저히 감소한 11개의 웹정보원을 제외하고, 15개의 웹정보원을 추가하여 36개의 웹정보원을 분석 대상으로 선정하였다.

이 연구에서는 커뮤니케이션 분야의 지적 구조의 변화를 파악함과 동시에 검색엔진의 차이를 분석하기 위하여, 선정된 웹정보원의 동시링크빈도를 AltaVista와 MSN Search 검색엔진에서 각각 조사하여 2개의 36차 동시링크행렬을 작성하였다. 동시링크행렬은 통계 프로그램(SPSSWIN 12.0)을 사용하여 상관계수행렬로 변환하였고, 상관관계를 분석하기 위하여 군집분석과 다차원축적기법을 사용하였다.

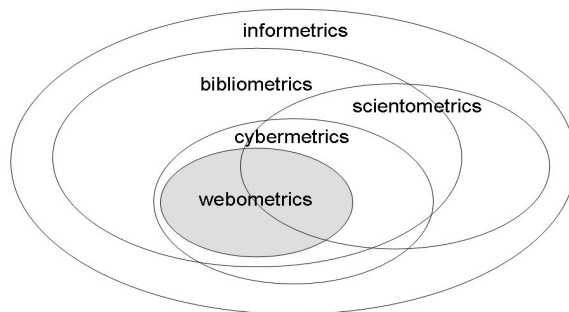
2. 이론적 배경

2.1 Webometrics

webometrics는 계량정보학의 접근 방법을 기반으로 정보자원의 구조와 이용, 웹의 구조와 기법을 수량적으로 연구하는 학문이라고 정의할 수 있다. webometrics는 계량정보학에 그 기원을 두고 있으며, 둘의 관계는 <그림 1>에서와 같이 webometrics가 계량정보학에 포함된다(Thelwall, Vaughan, and Bjorneborn 2003).

webometrics라는 용어는 Almind와 Ingwersen(1997)에 의해 제안되었으며, 저널명이기도 한 cybermetrics라는 용어와 함께 가장 일반적으로 사용되고 있다. webometrics의 동의어로는 netometrics(Bossy 1995), webometry(Abramham 1996), internetrometrics(Almind and Ingwersen 1996), web bibliometry(Chakrabarti et al. 2002) 등이 있다.

현재 진행되고 있는 webometrics의 연구 주요 테마는 ① 웹페이지의 내용 분석, ② 웹 링크 구조 분석, ③ 이용자의 이용형태를 분석하



<그림 1> webometrics와 계량정보학의 관계
(출처: Thelwall, Vaughan, and Bjorneborn 2003)

기 위한 로그파일 분석과 같은 웹 사용법 분석, ④ 검색엔진의 성능을 포함한 웹 기술 분석 등이다. 이런 연구 테마를 바탕으로 webometrics 연구는 웹상에서의 학술커뮤니케이션에 관한 연구(Smith 1999; Harter and Ford 2000; Vaughan and Hysen 2002; Ingwersen 1998; Thelwall 2003; 2004; Thomas and Willett 2000), 일반적·상업적인 웹사이트 사용에 관한 분석(Silverstein et al. 1999; Thelwall 2001; Vaughan and Wu 2003), 웹데이터 마이닝과 웹 전체 위상에 관한 연구(Ball 2000; Adamic and Huberman 2001)가 주로 이루어지고 있다.

Bossy(1995)와 Crinin and McKim(1996)은 학문에서 웹이 중요한 커뮤니케이션 수단이라고 보고, 계량정보학적 방법의 웹 적용가능성을 제시함으로써 계량정보학의 연구영역을 웹정보원으로 확장하였다. Rousseau(1997)는 인터넷 사이트간의 링크수를 측정하기 위하여 웹페이지를 대상으로 도메인의 수와 도메인별 인용의 수를 조사한 결과 웹문서의 분포가 로트카의 법칙과 일치함을 밝혀냈다. Vaughan과 Shaw(2003)는 서지 인용과 웹 인용의 유사성을 연구하고, 학문적 영향력 평가에서 웹 인용과 서지 인용이 함께 조사되어야 한다고 주장하였다. 그러나 웹 인용은 서지 인용이 갖는 시간의 문제는 해결해 주지만, 비영속적인 단점이 있다고 지적하였다.

Ingwersen(1998)은 저널의 영향력을 평가하는 영향지표(Impact Factor)를 응용한 Web Impact Factor(이하 WIF)를 제안하였고, 그 적용 가능성을 연구하였다. Smith(1999)는 Ingwersen이 제안한 WIF의 방법론을 분석한 후, 웹공간에서 대학과 전자저널 웹정보원의 영향력을 조사하였다. Zhang(2001)은 연구자들의

정보 검색과 커뮤니케이션 과정에서 웹이 어떠한 영향을 미치는지 분석하였다. 특히 전자정보원의 이용 비율이 크게 증가하였고, 도서관이 이를 위한 적절한 대비를 해야 한다고 하였다. Park과 Barnett, Nam(2002)은 사이버 공간에서 웹정보원이 다른 웹정보원을 링크하는 요인을 분석하고, 중심적인 위치를 차지하는 웹정보원을 밝혀냈다.

Larson(1996)은 지구과학분야의 중요 웹페이지를 선정한 후, 그 웹페이지에 링크된 웹페이지에서 동시링크빈도를 도출하고, 이를 기초로 지구과학분야의 지도를 작성하여 웹을 대상으로 지구과학분야의 지적구조를 확인하였다. Faba-Perez와 Guerrero-Bote, Moya-Anegon(2003)은 웹정보원의 링크 빈도와 학문의 지적구조를 분석하고, 비슷한 패턴을 공유하는 요인으로 주제외에도 지역적인 요인, 관련 프로젝트 등이 영향을 미친다고 하였다.

국내의 경우 정동열과 최윤미(1999)는 커뮤니케이션 분야를 대상으로 웹정보원에 동시인용분석 기법을 적용하여 커뮤니케이션 분야의 지적구조를 파악하였다.

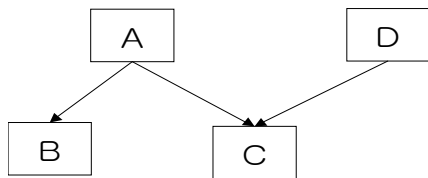
2.2 동시링크분석

webometrics에서 동시링크분석은 계량정보학의 동시인용분석 기법을 응용한 것이다. 동시인용분석(cocitation analysis)은 제3의 논문이나 저자가 두 편의 논문이나 두 명의 저자를 동시 인용하였을 때, 그 두 편의 논문이나 두 명의 저자는 주제적으로(학문적으로) 서로 관련이 있다는 가설을 전제로 한다(Small 1973; White and Griffith 1981). 동시인용기법의 분석 대상

은 주로 문헌, 용어, 저자 등이다.

웹정보원에 계량정보학적 연구방법이 적용될 수 있는 것은 웹정보원의 '하이퍼링크'라는 개념이 인쇄문헌의 인용과 유사한 개념으로 이해될 수 있기 때문이다. 제3의 웹문서가 두 개의 웹문서를 동시 인용하였을 때, 그 두 개의 웹문서는 역시 주제적으로 서로 관련이 있다는 것이다 (Arasu et al. 2001; Larson 1996; Fabra-Perez, Guerrero-Bote and Moya-Anegon 2003).

문서 또는 웹페이지간의 링크는 인용에서의 'citation'과 'reference'와 같이 웹문서가 다른 문서를 링크하고 있을 때의 링크를 outlink라고 하며, outlink는 인용문헌(reference)과 유사한 개념이라고 할 수 있다. 또한 다른 웹문서로부터 링크를 받고 있을 때의 링크를 inlink로 구분하며 피인용문헌(citation)과 유사한 개념이다. <그림 2>에서 웹문서 A의 B와 C는 동시링크(co-links)되었고, 이는 하나의 문헌으로부터 2개의 문헌이 동시인용된 상태와 유사하다. A와 D는 동시에 C를 링크하고 있으므로 동시인용분석에서 2개의 문헌이 서지 결합된 상태와 유사하다. 동시링크는 위의 두 개념을 포괄하는 개념이다 (Thelwall, Vaughan, and Bjerneborn 2003).



<그림 2> 웹의 동시링크 관계

동시인용빈도는 두 문헌이 함께 인용되는 회

수를 말하며, 동시인용빈도는 문헌 a를 인용한 문헌집합 A와 문헌 b를 인용한 문헌집합 B의 교집합의 수($n(A \cap B)$)가 되는 것과 같이, 동시링크빈도는 두 웹문서가 함께 링크되는 회수를 말하며, 웹문서 c를 링크한 웹문서 집합 C와 웹문서 d를 링크한 웹문서집합 D의 교집합의 수($n(C \cap D)$)가 된다.

동시인용분석을 위한 데이터는 ISI의 인용색인 DB를 이용해서 실제적인 동시인용빈도가 얻어지지만, 동시링크분석에서 웹 데이터 수집 방법은 상업적인 검색엔진을 통한 데이터 수집 방법, 독자적인 웹크롤러를 개발해 진행하는 방법, 인터넷의 로그파일 분석을 통한 데이터 수집 방법 등으로 나누어 볼 수 있다. webometrics 연구에 적합한 검색엔진의 특성으로는 대규모 DB 이어야 하고, 최신자료를 유지해야 하고, 특정 웹 페이지를 링크한 모든 웹 페이지를 검색할 수 있어야 하고, 불리언 연산이 사용 가능해야 하고, 검색 결과수의 일관성이 유지되어야 한다 (Smith 1999, 579).

webometrics 연구를 위해 가장 많이 사용되는 검색엔진으로는 AltaVista가 있다 (Larson 1996; Ingwersen 1998; Smith 1999; 정동열, 최윤미 1999; Thelwall 2003). AltaVista는 비록 검색결과 수의 비일관성 문제를 가지고 있지만, 위의 조건을 비교적 충족시킨다. 그동안 AltaVista에서 제공되는 'link', 'host'와 불리언 연산자를 이용한 다양한 연구가 진행되었다. 또한 최근 사회과학 분야를 대상으로 링크정보를 수집하는 웹크롤러인 SocSciBot3 (<http://socscibot.wlv.ac.uk>)가 웹에 공개되었다.

3. 동시링크분석을 위한 실험설계

3. 1 분석대상의 선정

이 연구의 대상 선정 과정은 다음과 같다. 1999년 정동열과 최윤미의 연구에서 사용된 32개 웹정보원 중 2005년 조사 시점에 폐쇄된 4개의 웹정보원을 제외한 28개의 웹정보원을 1차 선정후보로 선정하였다. 1차 선정후보에 새로운 주요 웹정보원의 추가 선정 방법은 1999년 연구의 주요 웹정보원을 동시링크하고 있는 대표적인 웹정보원을 선정하고(〈표 1〉 참고), 이들이 링크하고 있는 웹정보원들 중 커뮤니케이션 분야에 대한 60개의 웹정보원을 2차 선정후보로 선정하였다. 선정된 2차 후보를 대상으로 AltaVista에서 각각의 링크 수를 조사하고, 커뮤니케이션의 하위주제를 고려하여 상위 순위를 차지하는 15개의 웹정보원을 추가하여, 43개의 웹정보원을 2차 분석대상으로 선정하였다.

선정된 2차 분석대상 웹정보원을 AltaVista와 MSN Search 검색엔진에서 각각 동시링크빈도를 조사하여 동시링크빈도행렬을 작성하고, 1차 실험을 실시하였다 1차 실험 결과 의미있

는 분석을 위하여 1999년 분석 대상 웹정보원 중 폐쇄되지는 않았지만 링크빈도가 현저히 감소하였거나, 다른 웹정보원과의 동시링크빈도가 10회이하인 웹정보원을 추가로 제외하여 최종적으로 〈표 2〉와 같이 36개의 웹정보원을 분석대상으로 선정하였다.

동시링크 수를 조사하기 위해서는 링크에 관한 정보를 검색할 수 있는 검색엔진을 사용해야 한다. 본 연구에서는 검색엔진 별 웹정보원의 지적구조를 비교하기 위하여, AltaVista와 MSN Search를 비교 대상 검색엔진으로 선정하였다 현재 웹정보원들의 동시링크를 분석할 수 있는 검색엔진은 AltaVista와 Gigablast, MSN Search가 있다(Search engine showdown 2005). 이 중에서 검색 결과 수가 높고 그 결과가 비교적 안정적인 AltaVista와 MSN Search를 선정하였다.

선정된 웹정보원의 주제와 웹정보원의 성격은 1999년의 적용 기준을 그대로 따랐다. 선정된 웹정보원은 American Communication Association과 ALA의 Directory of Electronic Journals Newsletters and Academic List Serve에서 제시한 주제명을 참조하여, 하위주제명을

〈표 1〉 추가 웹정보원 선정을 위한 검색도구로 사용된 정보원

정보원	URL
American Communication Association	http://www.americancomm.org/
Advertizing World	http://advertising.utexas.edu/world/
Yahoo's list: Communication	http://dir.yahoo.com/social__science/Communications/
Communication Resources(Kennesaw State University)	http://www.kennesaw.edu/communication/resources.shtml
Communication Resources(Alma College)	http://newmedia.alma.edu/communication/resources.html
American Rhetoric: Communication, Mass Media, and Rhetoric	http://www.americanrhetoric.com/communicationassociations.htm

General Communication[GC], Computer-mediated Communication[CMC], Speech Communication[SC], Tele-communication [TEC], Mass Communication[MC], Cultural Study[CS], Media and Culture[M&C], Me-

dia and Technology[M&T], Journalism[JR], Advertising[AD], Public Relations[PR] 등으로 구분하였다. 또한 웹정보원에 수록된 웹사이트의 기능과 성격은 단체 혹은 개인 홈페이지, 색인 정보원, 전자잡지로 구분하였다.

〈표 2〉 커뮤니케이션 분야 분석대상 웹정보원 (2005년)

일련 번호	웹정보원 고유번호	정보원	1999년 링크수	2005년 링크수	주제 분야	성격	비고
1	S01[G]	American Communication Association http://www.americancomm.org/ http://cavern.uark.edu/comminfo/www/ACA.html	551	3574	GC	단체	
2	S02[G]	Association for Educational communication and Technology http://www.aect.org/	1034	27350	GC	단체	
3	S03[G]	Communication Institute for Online Scholarship http://www.cios.org/	724	19300	GC	단체	
4	S04[G]	ICA International Communication Association http://www.icahdq.org/ 또는 http://www.io.com/~icahdq/ica/	316	7830	GC	단체	
5	S05[P]	The Public Relations Society of America http://www.prsa.org/	1014	90100	PR	단체	
6	S06[G]	Southern States Communication Association http://ssca.net/	163	430	GC	단체	
7	S07[T]	Analysis Telecoms Virtual Library http://www.analysys.com/vlib/	475	1790	TEC	색인	
8	S08[C]	Computer Mediated Communication Magazines http://www.december.com/cmc/mag	2183	5490	CMC	전자	
9	S09[C]	Journal of Computer-Mediated Communication http://jcmc.huji.ac.il/ 또는 http://www.ascusc.org/jcmc http://shum.huji.ac.il/jcmc/jcmc.html	997	35257	CMC	전자	
10	S10[G]	Advertisizing World http://advertising.utexas.edu/world/	294	1531	GC	색인	
11	S11[M]	Vanderbilt Television News Archive http://tvnews.vanderbilt.edu/	1682	29900	MC	색인	
12	S16[J]	CampuServe's Journalism Forum http://www.jforum.org/	199	198	JR	단체	
13	S17[J]	FACSNET http://www.facsnet.org/	815	3610	JR	단체	
14	S18[M]	Media and Communication Studies Site http://www.aber.ac.uk/media/	794	66	MC	색인	
15	S20[J]	Megasources http://www.ryerson.ca/~dtudor/megasources.htm	340	462	JR	색인	
16	S21[J]	Reproter.org http://www.reporter.org/	942	5690	JR	단체	
17	S22[J]	WWW Virtual Library: Journalism http://vlib.org/Communication	757	612	JR	색인	
18	S26[S]	European Speech Communication Association http://www.wesca.ikp.uni-bonn.de/home.html	529	322	SC	단체	
19	S29[J]	Electronic Journ@list.The http://www.spj.org/	468	38000	JR	단체	
20	S30[J]	Online Journalism Review http://www.ojr.org/	163	117000	JR	단체	
21	S32[J]	JournalismNet http://www.rjca/	159	22900	JR	단체	
22	S33[A]	American Advertising Federation http://www.aaf.org/	-	13276	AD	단체	추가

23	S34[A]	ADCouncil http://www.adcouncil.org/	-	8547	AD	단체	추가
24	S35[C]	JOURNAL OF ONLINE BEHAVIOR http://www.behavior.net/JOB/	-	810	CMC	저널	추가
25	S36[G]	World Catholic Association for Communication http://www.signis.net/	-	12835	GC	단체	추가
26	S37[G]	World Association for Christian Communication http://www.wacc.org.uk/	-	8204	GC	단체	추가
27	S38[J]	The Poynter Institute http://www.poynter.org/	-	32992	JR	단체	추가
28	S39[J]	Reporters Without Borders http://www.rsf.org/	-	22974	JR	단체	추가
29	S40[Mc]	Media Research Center http://www.mediaresearch.org/	-	12299	M&C	단체	추가
30	S41[Mc]	University of Minnesota: Media History Project http://www.mediahistory.umn.edu/	-	8469	M&C	단체	추가
31	S42[M]	National Association of Broadcasters http://www.nab.org/	-	30217	MC	단체	추가
32	S43[M]	Radio-Television News Directors Association http://www.rtnda.org/	-	3496	MC	단체	추가
33	S44[S]	American Speech-Language-Hearing Association http://www.asha.org/	-	6073	SC	단체	추가
34	S45[S]	International Speech Communication Association http://www.isca-speech.org/	-	2725	SC	단체	추가
35	S46[T]	International Telecommunication Union http://www.itu.int/	-	13949	TEC	단체	추가
36	S47[T]	National Telecommunications and Information Administration http://www.ntia.doc.gov/	-	23331	TEC	단체	추가

3. 2 분석대상 웹정보원의 동시링크행렬 생성

최종적으로 선정된 36개의 웹정보원을 가지고 본 실험이 진행되었다. 우선 선정된 분석대상에 관한 AltaVista와 MSN Search 검색엔진의 36차 동시링크행렬을 각각 작성하였다. 동시링크빈도는 AltaVista와 MSN Search 검색엔진에서 'link:URL link:URL'이라는 검색식을 이용하여 구하였다. 동시링크빈도는 검색 시간과 검색식의 사소한 차이에 의해서도 검색 결과가 달라진다. 1차 예비 실험 결과 가장 안정적인 검색결과를 보여준 오후 시간대에 동시링크빈도를 조사하였고, <표 2> 웹정보원의 URL을 그대로 입력하여 검색식을 만들었다.

위의 방법으로 작성된 동시링크행렬을 통계 프로그램(SPSSWIN 12.0)의 기본 데이터로

사용하였다. 통계 프로그램의 분석 방법은 동시링크빈도 행렬의 대각선 값을 상위 수치 3개를 합하여 2로 나눈 값으로 하향조절한 후 상관관계수행렬로 변환하고, 웹정보원들간의 상관관계와 유사성을 구체적으로 분석하기 위하여 다차원축적과 군집분석을 사용하였다. 다차원축적은 변수들간의 정보를 단순화시켜 숨어있는 구조를 밝히는데 사용되는 기법으로 둘씩 짝지은 개체간의 유사성의 정도를 측정하여 m차원의 공간상에 개체들을 상대적으로 좌표화하여 형상화하는 것이다. 군집분석은 개체들 사이의 유사성 또는 비유사성의 정도를 측정하여 개체들을 가까운 순서대로 연결해 나가는 분석방법이다.

4. 동시링크분석 결과 비교 및 분석

4. 1 검색엔진간 커뮤니케이션분야의 지적구조 비교

4. 1. 1 웹정보원의 동시링크빈도

36개의 분석 대상 웹정보원으로부터 630개의 웹정보원 쌍이 발생되어 AltaVista는 〈부록 1〉, MSN Search는 〈부록 2〉와 같이 동시링크빈도 행렬이 작성되었다.

〈부록 1〉과 〈부록 2〉에서 가장 높은 동시링크빈도 쌍을 살펴보면 AltaVista의 경우 각 정보원들의 동시링크빈도의 범위는 0에서 27,900회까지이며, 가장 높은 동시링크빈도를 나타낸 쌍은 S30번(Online Journalism Review)과 S38번(The Poynter Institute)이다. MSN Search의 경우 동시링크빈도의 범위는 0에서 7,508회까지이며, 가장 높은 동시링크빈도를 나타낸 쌍은 S46번(International Telecommunication Union)과 S47번(National Telecommunications and Information Administration)이었다.

다른 웹정보원과의 동시링크빈도 총수가 가장 높은 웹정보원은 AltaVista의 경우 S38번(The Poynter Institute)이 50,965회로 가장 높았고, 동시링크빈도의 총수가 가장 낮은 웹정보원은 S36번(World Catholic Association for Communication)이 72회로 나타났다. MSN Search의 경우에는 S38번(The Poynter Institute)이 36,075회로 가장 높고, S18번(Media and Communication Studies Site)이 281회로 가장 낮은 것으로 나타났다.

하나의 웹정보원과 다른 35개 웹정보원의 동시링크빈도의 평균은 AltaVista 경우 130.80 회

이고, 가장 높은 평균 동시링크빈도의 웹정보원인 S38번(The Poynter Institute)의 평균동시링크빈도는 1,456회였고, 가장 낮은 평균 동시링크빈도의 웹정보원은 S36번(World Catholic Association for Communication)이 2.06회로 나타났다. MSN Search의 경우에 동시링크된 빈도의 평균은 235.43회이고, S38번(The Poynter Institute)의 평균동시링크빈도는 1,031회로 가장 높았고, 가장 낮은 평균 동시링크빈도의 웹정보원은 S18번(Media and Communication Studies Site)이 8.03회로 나타났다.

다른 웹정보원과 폭넓게 동시링크된 웹정보원은 AltaVista와 MSN Search 두 경우 모두 웹정보원 S01번(American Communication Association), S03번(Communication Institute for Online Scholarship), S05번(The Public Relations Society of America), S42번(National Association of Broadcasters)이 35개의 다른 모든 웹정보원과 동시링크되었고, S09번(Journal of Computer-Mediated Communication)은 AltaVista의 경우에만, S06번(Southern States Communication Association)은 MSN Search 경우에만 다른 웹정보원과 모두 동시링크되었다. 다른 정보원과 동시링크되는 수가 적은 정보원은 AltaVista와 MSN Search 경우 모두 S36번(World Catholic Association for Communication)이 16회로 나타났다.

AltaVista와 MSN Search에서 630쌍의 동시링크의 총 건수를 보면, MSN Search가 296,639회, AltaVista가 164,820회로 나타났다. MSN Search가 AltaVista보다 평균 2배에 가까운 검색 동시링크 수를 찾아 주었다. 그러나 AltaVista의 경우 S38번과 S30번과 같이 월등하게

많은 동시링크빈도를 산출하는 경우도 있었다.

이와 같이 총 검색 건수가 2배 가까운 차이를 보이므로, MSN Search와 AltaVista의 개별 쌍의 검색수의 차이를 살펴보기 위하여 MSN Search와 AltaVista의 630 쌍의 동시링크행렬을 비교하였다. 두 행렬의 표준편차와 평균을 비교해보면(〈표 4〉참고), 평균은 MSN이 AltaVista보다 2배 가깝게 높았고, 표준편차는 MSN Search보다 AltaVista가 크게 나타났다. 두 검색엔진의 검색 건수 중 검색 결과수가 높은 경우를 보면 AltaVista는 224건, MSN Search는 237건, 동일한 검색건수의 경우가 169건으로 나타나서, 양쪽이 비슷한 것으로 판단된다.

4. 1. 2 웹정보원의 다변량분석

선정된 36개 웹정보원간의 상관관계를 보다 자세히 분석하기 위하여 다차원축적 프로그램인 ALSCAL을 이용하여 웹정보원들의 위치를 2차원 공간상에 점으로 표시한 정보원지도를 산출하였다. 다차원 분석 결과, AltaVista 경우 스트레스 값은 0.146, RSQ값은 0.901로, MSN Search의 경우 스트레스 값은 0.137, RSQ값은 0.907로 나타났다.

AltaVista의 경우 웹정보원들을 동질성이 강한 몇 개의 소군집으로 나누는 〈그림 3〉의 텐드로그램에서 7개의 군집으로 나누고, 이미 부여된 웹정보원의 주제분야를 근거로 형성된 군집에 따라 정보원지도에 군집간의 경계선을 그

리면 〈그림 5〉와 같다. MSN Search의 경우 〈그림 4〉에서 7개의 군집으로 나누고, 주제분야를 근거로 형성된 군집에 따른 군집간의 경계선을 그려보면 〈그림 6〉과 같다.

AltaVista 경우 제 1군집은 지도의 오른쪽 아래에 위치하고 있으며 커뮤니케이션 일반에 관한 정보원과 수사학에 관한 정보원을 포함하고 있으므로 '커뮤니케이션 일반 & 수사학'이라는 군집명을 부여하였다.

제 2군집은 지도의 왼쪽에 위치하며 주로 저널리즘과 메스커뮤니케이션에 대한 정보원들로 구성되어 있어 '저널리즘 & 메스커뮤니케이션'이라는 군집명을 부여하였다. 저널리즘에 관한 웹정보원으로 S16번과 S17번, S20번, S21번, S22번, S29번, S30번, S32번, S38번, S39번이 포함되어 있고, 메스커뮤니케이션을 다루고 있는 S11번과 S42번, S43번이 포함되어 있다.

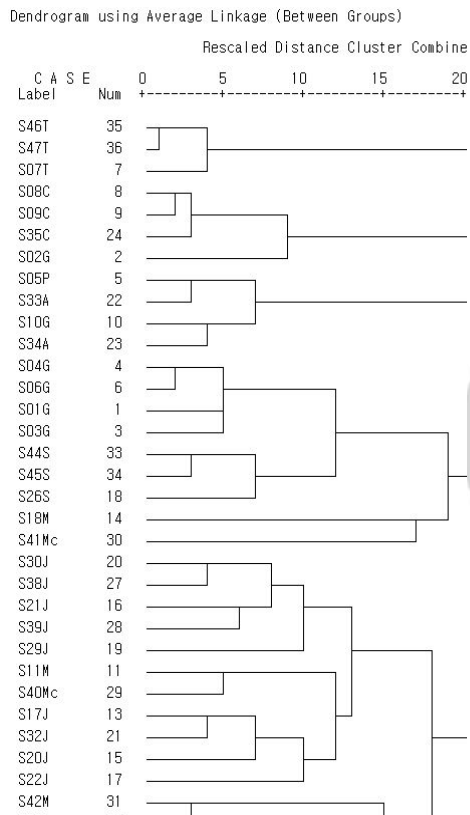
제 3군집은 'AD & PR'이라는 군집명을 가지며, S05번과 S10번, S33번, S34번이 포함되어 있고, 일반 커뮤니케이션을 다루고 있지만 'AD & PR'의 성격이 강한 S10번도 포함되어 있다.

제 4군집은 지도의 오른쪽 위에 위치하며 컴퓨터-매개 커뮤니케이션에 관한 정보원을 포함하고 있으므로 'CMC'라는 군집명을 부여하였다.

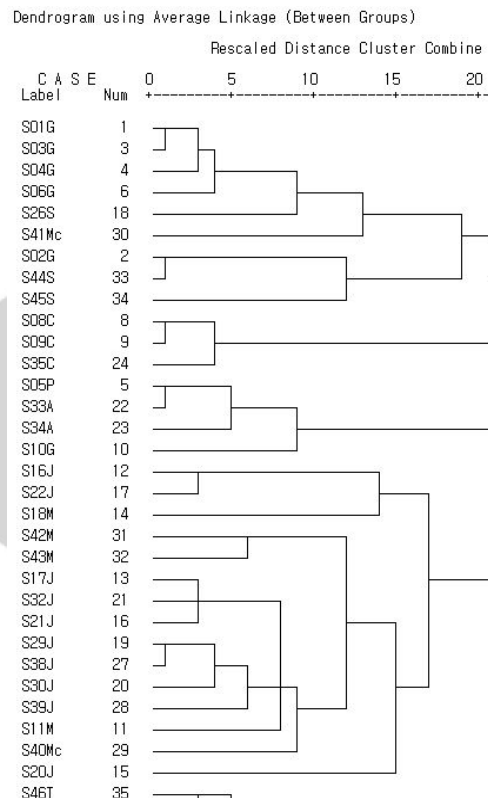
제 5군집은 'M & C'라는 군집명을 가지며, S18번과 S41번이 포함되어 있다.

〈표 4〉 AltaVista와 MSN Search의 표준편차와 평균

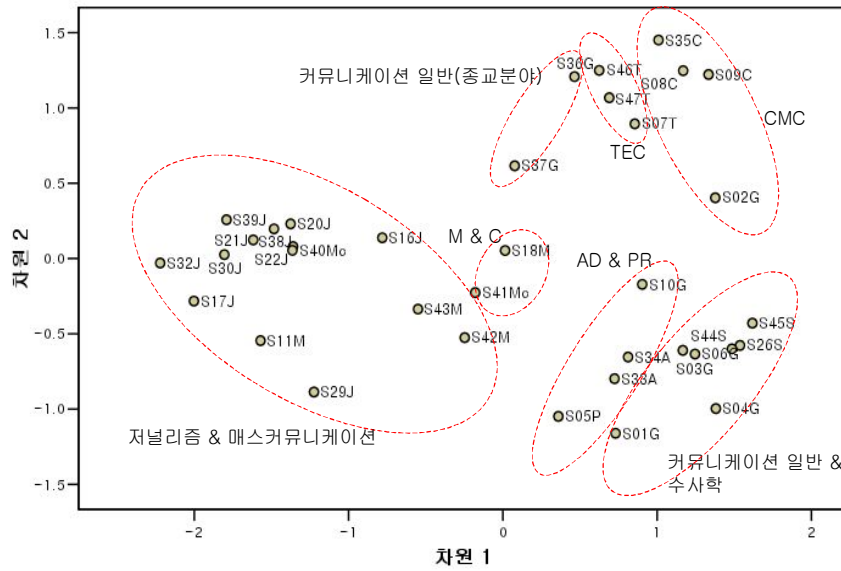
	MSN Search	AltaVista
표준편차	654	1124
평균	180	94.1



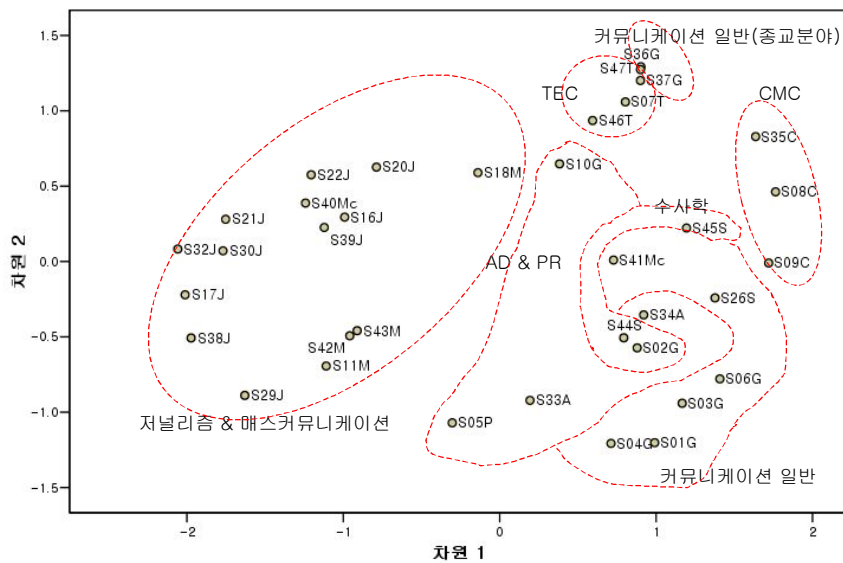
〈그림 3〉 2005년 커뮤니케이션 분야의 웹정보원의 군집형성 과정(AltaVista)



〈그림 4〉 2005년 커뮤니케이션 분야의 웹정보원의 군집형성 과정(MSN Search)



〈그림 5〉 군집분석에 의한 커뮤니케이션 분야의 웹정보원지도(AltaVista)-2005년



〈그림 6〉 군집분석에 의한 커뮤니케이션 분야의 웹정보원지도(MSN Search)-2005년

제 6군집은 텔레커뮤니케이션에 관한 정보원이 S07번과 S46번, S47번이 포함되어 있으므로 'TEC'라는 군집명을 부여하였다.

제 7군집은 S36번과 S37번이 포함되어있는데, 두 웹정보원은 커뮤니케이션 일반에 해당하지만 종교라는 특수한 공통점을 가지고 있는 정

보원이므로, '커뮤니케이션 일반(종교분야)'라는 군집명을 부여하였다.

MSN Search 경우 제 1군집은 지도의 오른쪽에 위치하고 있으며 커뮤니케이션 일반에 관한 정보원을 포함하고 있으므로 '커뮤니케이션 일반'이라는 군집명을 부여하였다. 여기에는 커뮤니케이션 일반에 관한 정보원인 S01번, S03번, S04번, S06번과 수사학에 관한 S26번, 미디어와 문화에 대한 S41번도 포함된다.

제 2군집은 지도의 왼쪽에 위치하며 주로 저널리즘과 매스커뮤니케이션에 대한 정보원들로 구성되어 있어 '저널리즘 & 매스커뮤니케이션'이라는 군집명을 부여하였다. 저널리즘에 관한 웹정보원으로 S16번과 S17번, S20번, S21번, S22번, S29번, S30번, S32번, S38번, S39번이 포함되어 있고, 매스커뮤니케이션을 다루고 있는 S11번과 S18번, S42번, S43번이 포함되어 있다.

제 3군집은 'AD & PR'이라는 군집명을 가지며, 포함되는 개별 정보원은 AltaVista의 경우와 동일하다.

제 4군집은 AltaVista 경우와 같이 지도의 오른쪽 위에 위치하며 컴퓨터-매개 커뮤니케이션에 관한 정보원을 포함하고 있으므로 'CMC'라는 군집명을 부여하였다. 포함되는 정보원으로는 CMC에 관한 S08번, S09번, S35번이다.

제 5군집은 '수사학'이라는 군집명을 가지며, S41번과 S45번이 포함되고, 커뮤니케이션 일반을 다루고 있는 S28번도 포함된다.

제 6, 7군집은 AltaVista 경우와 같이 각각 'TEC'와 '커뮤니케이션 일반(종교분야)'라는 군집명을 부여하였다.

AltaVista와 MSN Search 지도를 비교해

보면, 커뮤니케이션의 큰 주제분야는 전체적으로 비슷한 양상을 보인다. 지도의 오른쪽에 커뮤니케이션 일반과 왼쪽에 저널리즘, 매스커뮤니케이션이 있고, 위쪽에 CMC, TEC, 커뮤니케이션 일반(종교분야)이 포함되어 있다. 저널리즘 & 매스커뮤니케이션과 커뮤니케이션 일반 사이에는 AD & PR이 위치한다. 다만 세부 주제에 포함되는 개별 웹정보원이 대부분 일치하지만 몇몇 웹정보원은 소속 군집이 달라지는 경우도 발생했다. S02번의 경우 AltaVista에서는 CMC 군집에 MSN Search에서는 수사학 군집에 포함된다.

또한 S41번과 S18번이 AltaVista에서는 'M & C'에 포함되어 있으나 MSN Search에서는 S18번은 저널리즘 & 매스커뮤니케이션 군집에 포함되어 있고, S41번은 커뮤니케이션 일반 군집에 포함되어 있다. S44번과 S45번은 AltaVista에서는 수사학에, MSN Search에서는 커뮤니케이션 일반 군집에 포함되어 있다.

그러나 MSN Search 지도에서 '커뮤니케이션 일반'과 '수사학'이 분리되어 있지만, 제 6군집을 기준으로 보면 하나의 군집으로 묶어지며, AltaVista 지도에서 '커뮤니케이션 일반'과 'M & C'가 분리되어 있지만, 제 6군집을 기준으로 보면 하나로 합쳐진다. 이와 같은 점을 종합하면, AltaVista와 MSN Search의 웹정보원 지도에서 커뮤니케이션의 지적 구조는 대략적으로 비슷한 양상을 보인다고 볼 수 있다.

webometrics의 동시링크분석에서 AltaVista와 MSN Search 중 어떤 검색엔진을 쓰더라도 특정 주제의 지적구조는 비슷하다고 볼 수 있다. 다만 통계 분석면에서 자료가 많은 경우 분석이 잘된다고 본다면 상대적으로 MSN Search의

값이 일정하게 높기 때문에 더 유용하다고 할 수 있다.

두 행렬 중 검색수가 한쪽이라도 0인 경우의 총수가 109건으로, AltaVista와 MSN Search 모두 20건의 차이가 발생하였다. 즉 MSN Search가 찾지 못했는데 AltaVista가 찾은 경우는 20건이 되고, 반대로 AltaVista가 찾지 못했는데 MSN Search가 찾은 경우도 20건이었다. 그러므로 동시링크분석에서 AltaVista와 MSN Search 검색엔진을 모두 활용하는 방안을 모색하는 것이 바람직한 것으로 여겨진다.

4. 2 시기에 따른 지적 구조의 변화 비교

4. 2. 1 웹정보원의 동시링크빈도

1999년과 2005년의 커뮤니케이션 분야의 지적 구조와 웹정보원의 특징을 비교하기 위하여, 1999년의 연구가 AltaVista를 이용한 결과이기 때문에 앞 절의 2005년 AltaVista의 결과와 비교하였다. 1999년에는 32개의 웹정보원을 대상으로 하였으나, 2005년에는 1999년의 32개 웹정보원 중 폐쇄되었거나 동시링크빈도가 현저히 감소한 11개의 웹정보원을 제외하였고, 15개의 웹정보원을 추가하여 36개의 웹정보원을 분석대상으로 하였다.

동시링크행렬을 비교하면 다음과 같다. 가장 높은 동시링크빈도의 범위는 1999년에는 0-238회까지였으나, 2005년에는 0-27,900회까지로 급격하게 증가되었다는 것을 확인할 수 있다. 가장 높은 동시링크빈도 쌍은 1999년에는 S08번(Computer Mediated Communication Magazines)과 S09번(Journal of Computer-Mediated Communication)이었으나, 2005년에

는 S30번(Online Journalism Review)과 S38번(The Poynter Institute), 이었다. 가장 많이 동시 인용된 웹정보원은 1999년에는 622회의 S22번(WWW Virtual Library: Journalism)이었으나, 2005년에는 50,965회의 S38번(The Poynter Institute)이었다. 반면에 동시링크된 총 건수가 적은 웹정보원은 1999년도에는 S14번(Center for Media Education)이었으나, 2005년도에는 S36번(World Catholic Association for Communication)이었다. 1999년도에 동시링크 건수가 적은 웹정보원은 2005년도에는 폐쇄되었거나 동시링크수가 10이하인 것으로 조사되어, 분석대상 리스트에서 제외되었다. 또한 평균 동시링크빈도는 1999년에는 4.90회였으나, 2005년도에는 130.80회로 급격히 성장하였고, 가장 높은 평균동시인용빈도도 1999년에는 21.4회에서 2005년에는 1,456회로 크게 증가하였다.

다른 웹정보원과 폭넓게 동시링크된 웹정보원에 대해 비교해 보면, 1999년에는 S01번(American Communication Association)이 30개의 웹정보원과 동시링크된 것으로 나타났고, 2005년도에는 S01번(American Communication Association)과 S03번(Communication Institute for Online Scholarship), S04번(ICA International Communication Association), S05번(The Public Relations Society of America), S09번(Journal of Computer-Mediated Communication), S42번(National Association of Broadcasters)이 분석 대상의 모든 웹정보원과 동시링크 되었다.

1999년과 2005년의 동시링크빈도 자체의 변화를 비교하면, 동시링크빈도의 큰 성장을 확인

할 수 있다. 동시링크빈도의 범위가 0-238회에서 0-27,900회로 증가하였고, 가장 많이 동시 인용된 웹정보원의 동시링크빈도가 622회에서 50,965회로 증가하였고, 평균 동시빈도도 1999년에는 4.90회였으나 2005년에는 130.80회로 증가하였다.

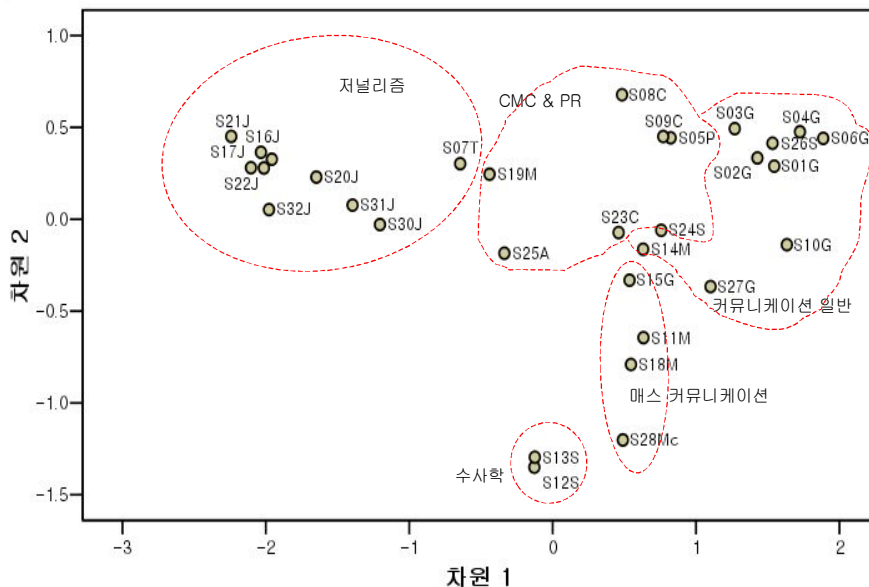
또한 폭넓게 동시링크된 정도를 보면 1999년에는 하나의 웹정보원이 32개의 분석대상 웹정보원 중 31개와 동시링크되었으나, 2005년도에는 6개의 웹정보원이 분석대상 웹정보원 모두와 동시링크되었다. 1999년도에 비해 2005년도의 동시링크되는 범위가 넓어졌다고 볼 수 있다. 이와 같은 통계수치는 예외의 웹정보원도 있지만, 웹정보원들이 1999년보다는 2005년도에 자주 동시링크되었고 서로 밀접하게 연결되었음을 시사한다.

4. 2. 2 웹정보원의 다변량분석

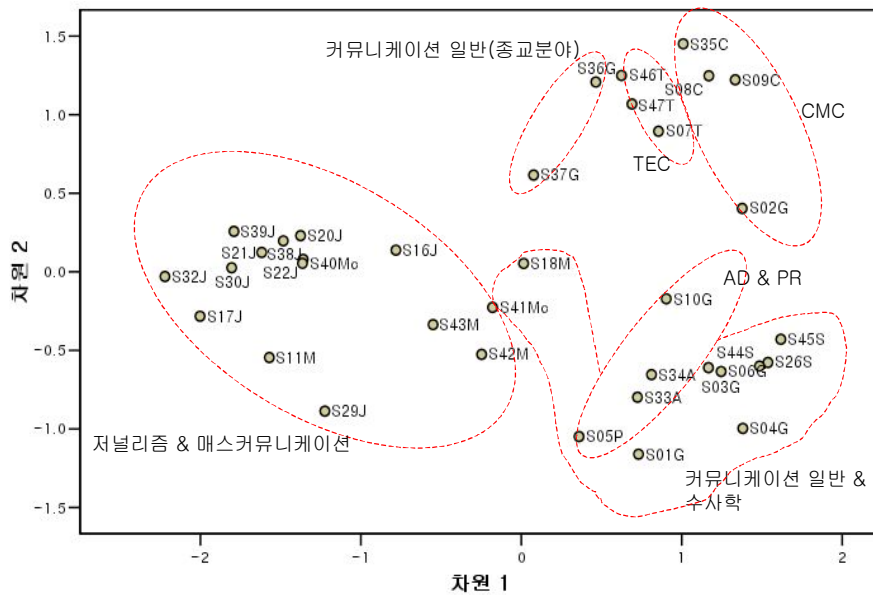
웹정보원을 통한 커뮤니케이션 분야의 지적 구조의 변화를 분석하기 위하여, 1999년의 웹정보원 지도인 <그림 7>(정동열, 최윤미 1999)과 2005년의 웹정보원 지도인 <그림 8>을 비교하였다.

두 시기의 웹정보원 지도를 통하여 가장 안정적이고 확고한 군집형태를 유지하고 있는 영역은 커뮤니케이션 일반과 저널리즘으로, 양시기에 걸쳐 거의 일치성을 보이고 있다. 즉 지도에서 커뮤니케이션 일반과 저널리즘 군집은 각각 오른쪽과 왼쪽에 위치한 큰 축을 유지하면서, 나머지 군집이 변화되었다

1999년에는 수사학에 속하는 웹정보원은 다른 커뮤니케이션 분야의 웹정보원의 동시링크수가 없어서 독립된 군집을 형성하였는데, 2005년도에는 커뮤니케이션 일반의 군집에 포함되



<그림 7> 군집분석에 의한 커뮤니케이션분야의 웹정보지도(AltaVista)-1999년



〈그림 8〉 군집분석에 의한 커뮤니케이션분야의 웹정보지도(AltaVista)-2005년

었다. 이는 수사학이 다른 커뮤니케이션 웹정보원과 동시링크 수가 증가하였다고 볼 수 있고, 수사학이 커뮤니케이션의 주류에 포함되었다고 볼 수 있다.

1999년도에는 저널리즘과 매스커뮤니케이션 부분이 독립된 군집을 형성하였으나, 2005년도에는 저널리즘과 매스커뮤니케이션이 한 군집을 이루고 있다. 2005년도에는 매스커뮤니케이션에 속하는 웹정보원은 저널리즘에 속하는 웹정보원과의 동시링크빈도가 상당히 빈번하여 한 군집을 이루었다.

또한 1999년에는 CMC와 PR이 한 군집을 형성하였으나, 두 주제는 복합적인 연구 경향에서 벗어나 2005년도에는 CMC와 PR이 각각 독자적인 군집을 형성하게 되었다. 이는 커뮤니케이션 분야에도 컴퓨터를 매개로 하는 주제 분야가 활발함을 의미한다. 또한 1999년에는 가

시화되지 않았던 TEC 영역이 2005년도에는 별도의 영역으로 분리되었다.

4. 2. 3 웹정보원의 지적구조 변화의 특징

이 절에서는 앞 절에서 나타난 웹정보원의 지적 구조의 변화가 분석 대상의 차이에 의한 것인지 혹은 대상 웹정보원 자체의 주제 성격 변화에 의한 것인지를 분석하기 위하여, 2005년을 기준으로 폐쇄된 4개의 웹정보원을 제외하고 1999년과 2005년에 공통으로 존재하는 28개의 웹정보원을 중심으로 변화를 살펴보았다.

1999년 28차 동시링크행렬의 다차원축적 결과 스트레스 값은 0.093, RSQ 값은 0.970이었고, 2005년 28차 동시링크행렬의 스트레스 값은 0.097, RSQ 값은 0.963으로 만족할 만한 수준이다.

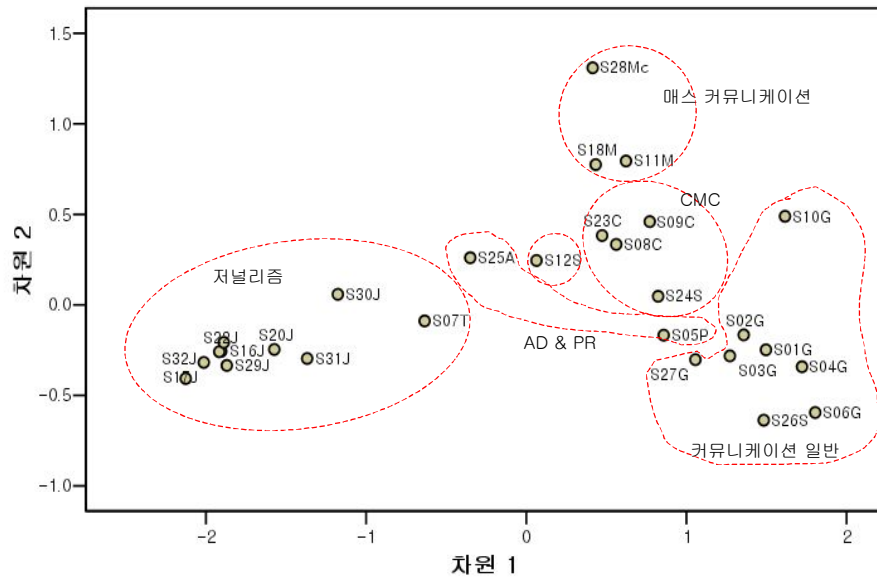
〈그림 9〉와 〈그림 10〉을 비교해보면, 두 시

기에 동일한 웹정보원을 대상으로 도출된 웹정보원의 지도는 상이한 지적구조를 나타내고 있음을 확인할 수 있다. 이는 앞 절에서 살펴본 두 시기의 지적구조 변화는 웹정보원 자체의 변화에 기인하였다고 볼 수 있다. 두 웹정보원 지도에서 변화가 없는 군집은 대략적으로 커뮤니케이션 일반과 저널리즘, CMC 분야였다. 나머지 군집은 차이가 있었는데, 그 차이를 살펴보면 다음과 같다.

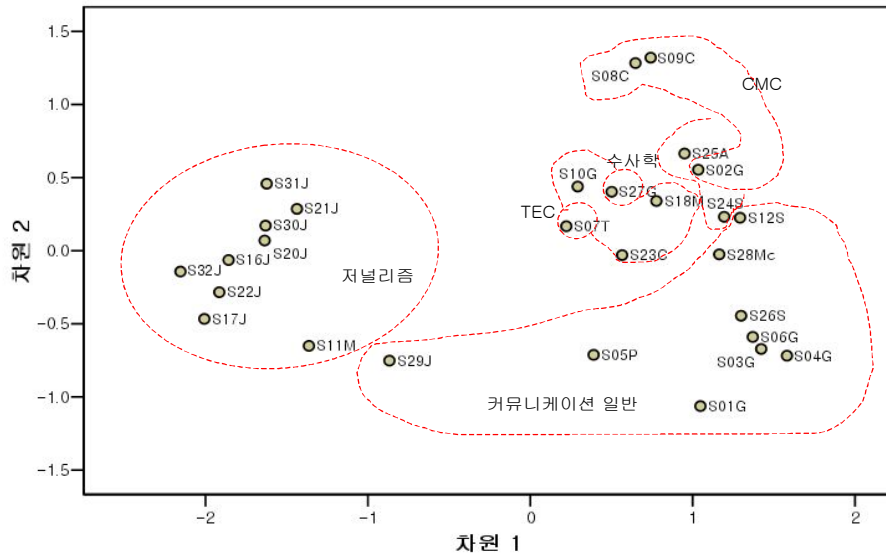
S12번은 1999년에는 독립된 군집을 형성하고 있었으나, 2005년도에는 커뮤니케이션 일반에 속하였다. 앞 절 <그림 8>의 지적구조에서 커뮤니케이션 분야와 수사학 분야가 통합된 것은 새로운 웹정보원의 추가에 기인한 것이 아니라, S12번의 성격이 변하여 커뮤니케이션일반의 웹정보원과 동시링크빈도가 높아졌기 때문

이라고 해석될 수 있다. 이런 변화는 앞절 <그림 8>의 지적구조에서 발견된 수사학 분야가 커뮤니케이션 일반 분야와 통합되는 양상과 유사하다.

1999년에 독립된 군집이었던 매스커뮤니케이션은 2005년도에는 다른 군집으로 흩어지는 현상을 보이고 있다. 이는 매스커뮤니케이션 영역이 주제의 독립 정도가 약해졌다고 볼 수 있다. 앞 절 <그림 8>의 지적 구조에서는 이러한 현상이 매스커뮤니케이션 관련 웹정보원이 저널리즘 군집에 속하는 결과로 나타난다. 또한 S07번의 경우 1999년 웹정보원 지도에서는 저널리즘에 속하였으나, 2005년도에는 TEC라는 독립된 군집을 이루고 있다. 이 경우도 앞 절 <그림 8>에서 TEC가 독립된 군집을 이루는 결과와 같다.



<그림 9> 동일 웹정보원을 대상으로 한 커뮤니케이션분야의 웹정보지도(AltaVista)-1999년



〈그림 10〉 동일 웹정보원을 대상으로 한 커뮤니케이션분야의 웹정보지도(AltaVista)-2005년

CMC는 앞 절 〈그림 7〉의 지도에서는 PR과 묶였으나, 2005년 폐쇄된 4개의 웹정보원을 제외한 〈그림 9〉의 28차 웹정보원 지도에서 분리됨을 알 수 있다. 즉 앞 절 〈그림 7〉에서 CMC가 PR과 묶인 것은 일부사이트가 두 군집을 묶어 주는 역할을 했으나, 그 사이트를 제외함으로써 독립된 영역으로 나타나고 있다. 이런 양상은 2005년의 웹정보원 지도에서 나타났고, 앞 절 〈그림 8〉의 웹정보원 지도에도 그대로 반영되었다.

이와 같은 분석을 토대로 1999년과 2005년의 시기별 웹정보원의 영향력을 보면, 1999년 분석 당시 주제가 명확한 웹정보원이 살아남는 것으로 나타났다. 여기서 주제가 명확하다는 것의 의미는 같은 주제 사이트와 동시링크빈도가 높은 웹정보원을 의미하며 또한 지도에서 자신의 웹정보원 주위에 유사한 주제의 웹정보원이

많이 분포해 있다는 것이다. 예를 들어 저널리즘의 군집에 해당하는 웹정보원 S16번, S17번, S20번, S22번, S29번, S30번, S31번, S32번이 이에 해당된다. 반면에 1999년에 관련 주제와 동시링크수가 적은 웹정보원은 주제가 명확하지 않았던 것을 의미한다. 이런 웹정보원들은 2005년도에 와서는 폐쇄되었거나 동시링크빈도가 현저히 감소하였다. 예를 들어 웹정보원 S14번, S15번, S19번이 해당된다.

이런 점에서 보면 2005년 분석 대상 웹정보원도 동일 주제와 링크가 높은 웹정보원은 향후 그 영향력이 계속될 것이며 계속적으로 살아남을 가능성이 높다고 볼 수 있다. 또한 동일 주제와 연결되어 동시링크빈도가 높은 웹정보원은 다른 주제와 연결된 웹정보원보다 그 웹정보원에 관심있는 이용자가 이용할 가능성이 높아질 것이다. 그러므로 향후 상대적으로 성장 가능성

이 높다고 볼 수 있다. 즉 커뮤니케이션 분야의 지적 구조 분석에서 발견된 이런 웹정보원의 특성은 특정 웹정보원의 성장 가능성을 예측할 때도 도움이 될 것이다.

또한 웹정보원을 대상으로 한 동시링크분석의 대상 후보 선정시 단순히 피링크를 기준으로 숫자가 높은 것을 기준으로 삼거나 다른 웹정보원과의 동시링크빈도가 높은 것을 후보로 선정하는 것에 추가해서 유사 주제와의 동시링크빈도가 높은 웹정보원을 가급적 선정하는 것이 주제를 분석하기에 용이할 것이다.

이상과 같이 1999년과 2005년의 웹정보원 지도의 비교를 통해 커뮤니케이션 분야의 세분화와 통합화에 따른 지적 구조의 변화를 살펴본다. 이 연구에서는 동시링크분석을 이용하여 웹정보원의 지적 구조의 변화를 파악할 수 있었고, 이런 결과는 인쇄정보원을 중심으로 한 동시인용기법이 웹정보원에 대한 통시적 분석에도 적용될 수 있음을 시사하고 있다.

5. 결론

이 연구에서는 동시링크분석을 이용하여 6년간의 시간흐름에 따른 웹정보원의 지적 구조와 변화를 비교하고, 검색엔진별 지적 구조의 차이를 분석함으로써 웹정보원의 특성과 연구방법상의 차이를 고찰하였다. 이 연구의 결과를 요약하면 다음과 같다.

첫째 웹정보원의 동시링크분석을 통해 살펴본 커뮤니케이션 분야의 지적 구조는 1999년에서 2005년으로 시간이 경과함에 따라 세부 영역의 입지가 변화한 것으로 나타났다. 커뮤니케

이션 일반과 저널리즘의 큰 축은 유지하되, 커뮤니케이션일반에는 수사학이, 저널리즘에는 메스커뮤니케이션이 통합되었다. 또한 CMC와 PR, TEC는 각각 독립된 군집을 형성하였다. 이와 같이 동시링크분석을 통하여 웹정보원의 지적 구조의 변화를 파악할 수 있었다.

둘째 AltaVista와 MSN Search 검색엔진을 대상으로 지적구조의 차이를 분석한 결과, 웹정보원 지도에 나타난 전체적인 지적 구조는 유사한 양상을 나타냈지만, 몇몇 웹정보원의 경우 소속 군집이 다른 경우도 있었다. 동시링크 분석에서 두 검색엔진을 동시에 활용하는 방안을 모색하는 것이 바람직한 것으로 여겨진다.

셋째 동시링크빈도의 큰 성장과 함께 동시링크되는 범위가 넓어졌다. 이와 같은 결과는 웹정보원들이 1999년보다는 2005년도에 자주 동시링크되었고 서로 밀접하게 연결되었음을 시사한다.

넷째 1999년과 2005년 웹정보원의 영향력을 비교해보면, 영향력을 계속 유지한 정보원은 1999년 분석 당시 같은 주제 사이트와 동시링크빈도가 높은 웹정보원 또는 지도상에 관련 주제 웹정보원이 주위에 많이 분포해있었던 웹정보원이었다. 이런 결과는 2005년 분석 대상 웹정보원에도 동일하게 적용되어, 동일 주제와의 링크가 높은 웹정보원은 향후 그 영향력이 계속될 것이라고 예측할 수 있다. 향후 일정 시기 이후에 동일한 분석을 해보는 것도 의미가 있을 것이다.

다섯째 동시링크분석의 대상 후보 선정시 유사 주제와의 동시링크빈도가 높은 웹정보원을 선정하는 것이 주제를 분석하는데 용이할 것이다.

이 연구는 통시적 webometrics 연구의 출발

점으로 의미를 갖는다. webometrics는 시작된 지 얼마 되지 않은 분야이며, 국내에서는 그 연구 성과가 미미하다. 특히 동시링크분석은 계량 정보학적 연구기법을 사용한다는 전제에는 이의가 없지만, 분석 대상의 선정, 데이터의 수집

과 분석 등의 매 단계마다 웹정보원의 특성을 반영한 효율적인 방법들이 지속적으로 논의되어야 한다. 또한 웹 자체가 가지고 있는 가변성과 다양성에 대한 지속적인 연구가 이루어져야 할 것이다.

참 고 문 헌

- 서은경. 1992. 정보검색분야의 지적 구조와 변화에 관한 연구. 『정보관리학회지』 9(1): 55-82.
- 정동열, 최윤미. 1999. 웹정보원의 동시인용분석에 관한 실험적 연구. 『정보관리학회지』, 16(2): 7-25.
- 조명희. 1993. 저자공인용매핑과 학문의 지적구조변화. 『정보관리학회지』, 10(1): 65-96.
- Abraham, R. H. 1996. *Webometry: measuring the complexity of the World Wide Web*. [cited 2005. 2. 22]. <<http://www.ralph-abraham.org/vita/redwood/vienna.html>>.
- Adamic, L. A. and Huberman, B. A. 2001. "The Web's hidden order." *Communications of the ACM*, 44(9): 55-59.
- Almind, T. C. and Ingwersen, P. 1996. *Informetric Analysis on the World Wide Web: A Methodological Approach to 'Internetometrics'*. Centre for Informetric Studies, Royal School of Library and Information Science.
- Almind, T. C. and Ingwersen, P. 1997. "Informetric analysis on the World Wide Web: A methodological approach to 'webometrics'." *Journal of Documentation*, 53(4): 404-426.
- Arasu, A., Cho, J., Garcia-Molina, H., Paeckke, A. and Raghavan, S. 2001. "Searching the Web." *ACM Transactions on Internet Technology*, 1(1): 2-43.
- Ball, P. 2000. "The art of networking." *Nature*. [cited 2005. 4. 9]. <<http://www.nature.com/nsu/001026/001026-7.html>>.
- Bossy, M. J. 1995. *The last of the litter: "Netometrics"*. [cited 2005. 4. 9]. <<http://bibliofr.info.unicaen.fr/bnum/jelec/Solaris/d02/dbossy.html>>.
- Chakrabarti, S., Joshi, M. M., Punera, K. and Pennock, D. M. 2002. *The structure of broad topics on the Web*. [cited 2005. 5. 1]. <<http://www2002.org/CDROM/refereed/338>>.
- Crinin, B. and McKim, G. 1996. "Science and scholarship on the World Wide Web: A North American Perspective." *Jour-*

- nal of Documentation*, 52(2): 163-171.
- Faba-Perez, C., Guerrero-Bote, V. P., and Moya-Anegon, F.D. 2003. "Data mining in a closed Web environment." *Scientometrics*, 58(3): 623-640.
- Harter S. P. and Ford, C. E.. 2000. "Web-based analyses of E-journal impact: Approaches, problems, and issues." *JASIS*, 51(13): 1159-1176.
- Ingwersen, P. 1998. "The calculation of Web Impact Factors." *Journal of Documentation*, 54(2): 236-243.
- Larson, R. 1996. *Bibliometrics of the World Wide Web*. [cited 2003. 12. 3]. <<http://sherlock.berkeley.edu/asis96/asis96.html>>.
- Park, H. W., Barnett, G. A., and Nam, I. 2002. "Hyperlink-affiliation network structure of top Web sites: examining affiliates with hyperlink in Korea." *JASIST*, 53(7): 592-601.
- Rousseau, R. 1997. "Sitations: An exploratory study." *Cybermetrics*, 1(1), [cited 2003. 11. 30]. <<http://www.cindoc.csic.es/cybermetrics/articles/v2i1p2.html>>.
- Search engine showdown. [cited 2005. 3. 7]. <<http://www.searchengineshowdown.com/>>.
- Silverstein, C., Henzinger, M., Marais, H. and Moricz, M. 1999. "Analysis of a very large Web search engine query Log." *SIGIR Forum*, 33(1): 6-12.
- Small, H. G. 1973. "Co-citation in the Scientific Literature: A Measure of the Relationship Between Two Documents." *JASIST*, 24(4): 265-269.
- Smith, A. G. 1999. "A tale of two web space: comparing sites using web impact factors." *Journal of Documentation*, 55(5): 577-592.
- Thelwall, M. 2001. "A Web crawler design for data mining." *Journal of Information Science*, 27(5): 319-325.
- Thelwall, M. 2003. "A layered approach for investigating the topological structure of communities in the Web." *Journal of Documentation*, 59(4): 410-429.
- Thelwall, M. 2004. *Link analysis : an Information Science Approach*. Library and information science. Amsterdam : Boston : Elsevier Academic Press.
- Thelwall, M., Tang, R., and Price, L. 2002. "Linguistic patterns of academic Web use in Western Europe." *Scientometrics*, 56(3): 417-432.
- Thelwall, M., Vaughan, L. and Bjorneborn, L. 2003. "ARIST Webometrics." *Annual review of information science and technology. v. 37*. Blaise Cronin, editor. Medford, N.J. : Information Today.
- Thomas, O. and Willett, P. 2000. "Webometric analysis of departments of librarianship and information science." *Journal of Information Science*, 26(6): 421-428.
- Vaughan, L. and Hysen, K. 2002. "Relation-

- ship between links to journal web sites and Impact Factors.” *Aslib Proceedings*, 54(6): 356-361.
- Vaughan, L. and Shaw, D. 2003. “Bibliographic and web citation: what is the difference?.” *JASIST*, 54(14): 1313-1322.
- Vaughan, L. and Wu, G. 2003. “Link counts to commercial web sites as a source of company information,” *Proceedings the 9th International Conference of Scientometrics and Informetrics*: 321-329. Beijing, China, Aug. 25-29, 2003.
- White, H. D. and Griffith, B. C. 1981. “Author Cocitation: A Literature Measure of Intellectual Structure.” *JASIST*, 32(3): 163-171.
- Zhang, Y. 2001. “Scholarly use of Internet-based electronic resources.” *JASIST*, 52(8): 628-654.

K C I

K C I

K C I