

저자프로파일링분석과 저자동시인용분석의 유용성 비교 검증*

A Comparison Test on the Potential Utility between Author Profiling Analysis(APA) and Author Co-Citation Analysis(ACA)

유종덕(Jong-duk Ryoo)**

최은주(Eun-Ju Choi)***

초 록

본 연구는 학문의 지적 구조를 분석하는 새로운 분석기법인 저자프로파일링분석과 전통적인 분석기법인 저자동시인용분석을 비교하여 분석함으로써 국내 연구환경에 맞는 지적 구조 분석 방법을 제안하는데 목적을 두고 있다. 이를 위하여 본 연구에서는 인용색인을 이용하지 않고 학문의 지적 구조를 분석할 수 있는 텍스트마이닝을 이용한 저자프로파일링분석을 통하여 새로운 지적 구조 방법의 유용성을 확인하고자 하였다. 분석대상 학술지는 『대한건축학회 논문집 - 계획계』를 대상으로 하였다.

ABSTRACT

This study compared Author Profiling Analysis(APA) to Author Co-Citation Analysis(ACA). The former is a new analytic technique on the intellectual structure of a science whereas the latter is a traditional analytic technique. The purpose of this study was to propose appropriate methods to analyze intellectual structure of a science in the Korean research environment. In order to achieve the goal, this study adopted APA using Text Mining for analysis on the intellectual structure of a science rather than relying on citation index in order to determine a potential utility of the new analytic technique that can identify the intellectual structure.

키워드: 저자프로파일링분석, 저자동시인용분석 지적 구조 분석, 인용색인, 유용성
author profiling analysis(APA), author co-citation analysis(ACA), intellectual structure, citation index, potential utility

* 이 논문은 경기대학교 박사학위 논문의 일부를 요약한 것임.

** 경기대학교 문헌정보학과 강사(jdryoo@kyonggi.ac.kr) (제1저자)

*** 경기대학교 문헌정보학과 교수(ejchoi@kyonggi.ac.kr) (공동저자)

■ 논문접수일자: 2011년 2월 18일 ■ 최초심사일자: 2011년 2월 23일 ■ 게재확정일자: 2011년 3월 10일

■ 정보관리학회지, 28(1): 123-144, 2011. [DOI:10.3743/KOSIM.2011.28.1.123]

1. 서론

1.1 연구의 배경 및 목적

전통적으로 사용하던 인용분석 방법인 동시 인용, 서지결합분석 등은 인용문헌의 데이터가 필요하고 이러한 데이터를 만들기 위해서 많은 노력이 필요하다. 해외의 경우 ISI(Institute for Scientific Information)의 Eugene Garfield가 시작했던 SCI(Science Citation Index) 등의 인용 데이터베이스를 이용하여 비교적 쉽게 동시인용빈도를 구하여 학문의 지적 구조를 확인하는 데 이용하고 있다. 그러나 국내의 경우 아직 포괄적이고 체계적인 인용 데이터베이스가 구축되어 있지 않아 학문의 지적 구조를 확인하기 위해서는 먼저 인용 데이터베이스를 연구자가 직접 구축하여 사용하는 방법이 활용되었다. 2009년부터 KISTI(한국과학기술정보연구원)에서 과학기술관련 학술논문을 대상으로 인용 데이터베이스를 구축하고 있으나 학술지 전체를 포괄하고 있지 못하고 데이터 구축 범위가 전 시기를 아우르지는 못하고 있다. 이러한 국내의 현실을 감안하여 인용색인이 필요하지 않은 지적 구조 분석 방법론의 유용성을 검증하는 것이 본 연구의 목적이다.

학문 분야의 지적 구조와 연구 동향을 분석하기 위하여 국내의 연구자들은 분석 도구로써 동시인용분석이나 저자동시인용분석을 많이 사용하고 있다. 프로파일링 분석 기법의 경우 비교적 최근에 개발된 분석 기법으로 아직도 연구의 성과가 많지 않으며, 김관준과 이재운이 2007년에 제안한 프로파일링 분석 기법에 대한 연구는 국내학회지에 소개한 몇 편의 논문과 해외 학술

지에 게재된 몇 편의 논문을 제외하고는 거의 없다.

저자동시인용분석은 현재의 시점에서 과거의 지적 구조를 관찰하는 것으로 최신의 연구 경향과 신진 연구자 파악이 어렵다는 제한점(서은경 1992)을 가지고 있으며, 인용색인이 반드시 필요하다라는 한계를 가지고 있다. 본 연구에서는 이를 보완하기 위하여 저자를 분석 단위로 하면서 직접 동시인용이나 동시출현을 파악하지 않고, 분석단위와 함께 사용된 단어로 분석단위를 표현하는 방법인 저자프로파일링 기법을 사용하여 지적 구조를 분석하였다.

본 연구는 건축학의 계획계 분야 논문을 대상으로 저자프로파일링분석을 이용한 학문의 지적 구조 분석 기법의 유용성을 밝혀보고자 하였다. 이를 위하여 전통적 분석 기법인 저자동시인용분석을 함께 이용하여 비교 분석함으로써 저자프로파일링분석 기법의 유용성 검증 을 시도하였다.

1.2 연구의 방법

본 연구는 저자프로파일링분석과 저자동시인용분석의 개념과 선행연구에 대하여 이론적으로 고찰한 다음, 두 기법을 적용하여 2003년부터 2008년까지 국내 건축학 분야 계획계의 지적 구조를 파악하고 기법 간의 결과를 비교·분석하였다.

분석 대상은 한국연구재단에 등재된 건축학 분야의 학술지 중 건축학 분야의 주제 영역의 계획 영역을 다루고 있는 학술지인 『대한건축학회 논문집-계획계』 분야이다. 수집 대상 데이터는 2003년부터 2008년까지 『대한건축학회논문집-

계획계』에 게재된 논문과 참고문헌의 저자사항을 포함한 서지사항이며 데이터는 KISTI로부터 제공받아 누락된 데이터에 대한 보정작업을 실시한 후 분석을 실시하였다.

저자프로파일링분석을 위하여 6년간 게재된 논문 2,104편의 전체 저자 3,943명 가운데 투고 논문의 투고 가중치 점수를 기준으로 14점 이상인 상위 저자 총 46명을 선정하였다. 선정된 46명의 저자별로 1차 프로파일 행렬을 구성하고, 코사인 유사도행렬과 피어슨 상관계수행렬을 작성하였다. 저자동시인용분석을 위하여 『대한건축학회논문집-계획계』 2003년~2008년도에 수록된 2,104편 논문의 피인용 저자 3,819명 가운데 피인용 빈도를 기준으로 9회 이상 인용된 저자 총 46명을 선정하고, 저자동시인용빈도행렬과 피어슨 상관계수행렬을 작성하였다.

건축학 분야의 저자 지도는 저자프로파일링 분석과 저자동시인용분석 모두 피어슨 상관계수행렬 기반의 다변량분석 기법중 군집분석과 다차원척도법을 이용하여 작성하였다. 저자군집을 위하여 피어슨 상관계수행렬을 입력물로 작성하였고, 군집화를 위하여 SPSS Statistics ver. 17.0의 군집분석 프로그램을 이용하였다. 저자군집에 사용된 기법은 계층적 응집방법으로 저자프로파일링분석의 경우 WAVE(집단내 평균연결기법), 저자동시인용분석은 Ward 기법을 적용하였다.

건축학 분야의 저자 지도는 계층적 클러스터링 알고리즘을 이용한 CNet(Clustering-Based Network)과 네트워크 시각화 및 분석 프로그램인 Pajek(Wouther, Andrej, Vladimir 2005)을 이용하여 네트워크 방식으로 작성하였다.

두 분석 기법에 기반한 저자 지도와 저자 군

집의 분석 결과를 바탕으로, 한국십진분류법(이하 KDC) 5판과 정만영이 제안한 한국연구재단 <연구분야분류표>(정만영 2008), 한국연구재단이 현재 사용하고 있는 <연구분야분류표>, 건축학 분야의 학술논문과 단행본, 건축도시연구정보센터의 홈페이지(건축도시연구정보센터 2010)를 참조하여 건축학 분야의 하위 주제영역을 명명하였고, 이를 통해 건축학 분야의 지적 구조를 파악하였다.

2. 이론적 배경

2.1 정의 및 특징

2.1.1 저자프로파일링분석의 정의 및 특징

텍스트마이닝분석 기법의 일종인 프로파일링분석 기법은 저자나 저널 또는 디스크립터를 분석 단위로 하면서 직접 동시인용이나 동시출현 단어를 파악하지 않고, 분석단위와 함께 사용된 언어로 분석 단위를 표현하는 방법이다. 분석 단위에 따라서 문헌클러스터링분석 기법과 디스크립터프로파일링분석 기법, 저자프로파일링분석 기법 등으로 나눌 수 있다(김판준, 이재운 2007; Kim and Lee 2008).

지적 구조 분석에 적용되는 최신 텍스트마이닝분석 기법인 문헌클러스터링분석이나 프로파일링분석은 모두 문헌이나 저자 등의 분석단위를 단어 벡터로 표현한 후 벡터 유사도를 산출하여 클러스터나 네트워크를 형성하는 방식을 취하고 있다(이재운, 김희전, 유종덕 2010).

프로파일링분석 기법은 디스크립터프로파일링분석 기법과 저자프로파일링분석기법으로 나

눌 수 있으며, 디스크립터프로파일링분석 기법은 기본적으로 문헌에 부여된 디스크립터로 주제를 표현하는 한편, 이러한 디스크립터프로파일(디스크립터 벡터)을 생성하는 과정에서 프로파일의 구성요소로 해당 디스크립터가 부여된 문헌의 표제와 초록에 출현한 단어들을 사용하는 것이다. 저자프로파일링분석 기법은 디스크립터로 주제를 표현하는 것이 아니라 저자로 주제를 표현하는 방법을 제외하고는 디스크립터프로파일링분석 기법과 동일한 분석절차를 사용하고 있다. 분석을 위하여 해당 문헌의 저자들을 선정하고 선정된 저자들과 문헌에서 추출한 텍스트 용어들과의 저자-색인어 프로파일을 통하여 군집을 나누고, 연구 영역의 주제를 식별하여 표현하는 방법으로 지적 구조를 분석하는 방법이다.

2.1.2 저자동시인용분석의 정의 및 특징

1981년 White와 Griffith가 저자동시인용분석에 대하여 제안하였는데 이는 Small(1973)과 Marshakova(1973)가 제안한 동시인용분석에 기반하고 있다.

1980년대에 이르러 동시인용분석은 분석 단위가 문헌에서 피인용 저자로 분석 되는 저자동시인용분석으로 발전하게 되었다. 저자동시인용분석은 White와 Griffith(1981)에 의해 소개된 후 국외(Culnan 1990; White, McCain 1998)에서 뿐만이 아니라 국내(김영진 1986; 김도미 1992; 이명옥 1996; 김희전 2010)에서도 특정 학문 분야의 지적 구조를 분석하기 위한 좋은 분석 기법으로 활용되어 다양한 학문 영역에서 계속되어 사용되고 있다. 저자동시인용분석은 분석 단위가 문헌이 아닌 연구자 중심의 지적

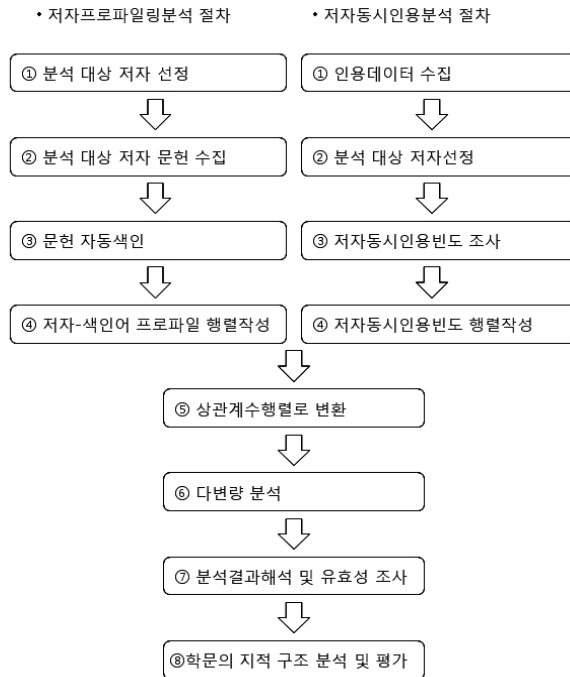
구조 분석 기법으로, 두 저자가 또 다른 문헌의 저자에게 자주동시인용이 될수록 인용된 두 저자는 더욱 밀접한 관계를 갖는다고 가정하고, 각 저자간의 거리를 나타내는 변수로 저자의 동시인용빈도수를 사용하여, 저자동시인용기법에 의해 분류된 저자군집도 문헌동시인용 군집처럼 특정 주제분야의 하위주제영역을 나타냄을 증명하였다.

2.2 지적 구조 분석의 절차

학문 분야의 지적 구조를 규명하기 위해 본 연구에서 사용하는 저자프로파일링분석과 저자동시인용분석의 절차를 단계별로 정리하면 <그림 1>과 같다(McCain 1990; 김관준, 이재운 2007).

저자프로파일링분석과 저자동시인용분석의 분석절차는 1단계에서 4단계까지만 다를뿐 5단계부터는 같은 분석의 절차를 따르게 된다. 저자프로파일링분석의 경우 서지 데이터의 투고 저자를 대상으로 저자를 선정하게 된다. 그러나 저자동시인용분석의 경우 인용데이터를 대상으로 피 인용빈도를 조사해야만 저자가 선정된다. 두 분석 기법은 서지 데이터와 인용 데이터가 각 분석의 원 자료가 된다.

저자프로파일링분석의 3단계와 저자동시인용분석의 3단계는 문헌 자동색인과 저자동시인용빈도 조사로 이 단계를 1차 프로파일 행렬작성 단계(White and Griffith 1981)이고, 4단계는 각 분석 기법별로 저자-색인어 프로파일 행렬작성, 저자동시인용빈도 행렬작성 단계로 2차 연관성 행렬작성 단계이다.



〈그림 1〉 저자프로파일링분석과 저자동시인용분석의 절차

3. 분석 기법별 지적 구조 분석

3.1 지적 구조 규명을 위한 분석 과정

3.1.1 자료 수집

1) 저자프로파일링분석

저자 선정은 저자동시인용분석과는 다르게 복수저자를 고려하여 투고 가중치 점수를 산정하였다. 저자프로파일링분석의 선정 저자로 『대한건축학회논문집 - 계획계』에 2003년부터 2008년까지 게재된 총 분석 대상인 2,104건의 논문의 저자 3,943명 중 투고 가중치 점수 14점 이상 논문을 게재한 저자 46명을 선정하여 분석하였다. 46명의 저자를 선정한 이유는 저자동시인용분석과 비교가 용이한 적절한 숫자의 분석대상자

를 선정하였고, 연구자의 판단에 의하여 적절한 기준에 의해 저자를 선정하여야 한다고 주장한 White와 McCain(1989)의 연구를 근거로 했기 때문이다. 적절한 기준을 마련하기 위하여 기 수행된 연구들의 분석대상자 선정 저자 평균을 확인한 결과 50.54명으로 확인되어 15점 이상인 경우 41명, 13점 이상인 54명으로 저자를 선정하였으나 저자동시인용분석과의 비교를 위하여 46명의 저자를 선정하였다.

2) 저자동시인용분석

저자동시인용분석의 데이터 수집은 저자프로파일링분석과 동일하며, 인용 데이터 분석 시 참고문헌에 게재된 데이터 중 학술지에 게재된 연구논문으로 데이터를 제한하였다. 참고문헌의 데

이터 중, 단행본이나, 석·박사 학위논문, 학술대회 발표 논문 등은 수집 대상에서 제외하였다. 데이터 구축 대상인 『대한건축학회 논문집-계획계』 분야의 경우 2003년부터 2008년까지 논문집에 게재된 논문은 2,104편이며, 참고문헌의 논문 중 학술지 논문의 피인용 저자는 3,819명이다.

3.1.2 저자 선정

1) 저자프로파일링분석

저자프로파일링분석을 위해 선정된 46명의 저자-색인어 프로파일 행렬을 구성하고, 코사인유사도행렬과 피어슨 상관계수행렬을 작성하였다. 저자프로파일링분석에서 가중치를 주

어 두고 가중치 점수를 산정한 이유는 제1저자만을 고려한 동시인용분석에서는 공동저작의 관계가 반영되지 않음으로써 지적 구조에서 일부 하위 주제분야의 군집형성이 서로 다른 결과를 나타내었다. 복수저자를 고려하여 분석하는 방법을 사용할 때 지적 구조가 그 분야의 주제적인 관계를 이해하는 데 보다 적합한 것으로 나타났다(이은숙 2003). 두고 가중치 점수가 14점 이상을 충족하는 저자 중 가장 많은 점수를 획득한 저자는 두고 가중치 점수 63점인 홍원화이며, 이어서 신중진(35점), 최원기(32점), 최재필(30점)로 1위와 2위와의 격차가 큰 것으로 나타났다(표 1 참조).

〈표 1〉 저자프로파일링분석 46명 선정저자 및 두고 가중치 점수

저자 순위	저자	저자 빈도	제1저자 빈도	공저자 빈도	두고 점수	저자 순위	저자	저자 빈도	제1저자 빈도	공저자 빈도	두고 점수
A01	홍원화	48	15	33	63	A24	손장열	19	0	19	19
A02	신중진	19	16	3	35	A25	이동연	19	0	19	19
A03	최원기	17	15	2	32	A26	임채진	16	3	13	19
A04	최재필	16	14	2	30	A27	강병근	14	4	10	18
A05	송승영	20	8	12	28	A28	김광호	11	7	4	18
A06	최무혁	27	1	26	28	A29	김원필	16	2	14	18
A07	김병선	26	1	25	27	A30	이강희	11	6	5	17
A08	김영하	20	7	13	27	A31	이정호	16	1	15	17
A09	이진숙	15	12	3	27	A32	최찬환	14	3	11	17
A10	임석호	15	11	4	26	A33	곽노열	8	8	0	16
A11	황은경	15	11	4	26	A34	김영욱	15	1	14	16
A12	김성아	14	11	3	25	A35	이상현	10	6	4	16
A13	김태영	24	1	23	25	A36	전봉희	11	5	6	16
A14	이경훈	23	2	21	25	A37	조영행	8	8	0	16
A15	서승직	22	2	20	24	A38	최 열	8	8	0	16
A16	석호태	14	10	4	24	A39	김도년	9	6	3	15
A17	박영기	21	2	19	23	A40	이정수	8	7	1	15
A18	최안섭	15	7	8	22	A41	이정재	12	3	9	15
A19	서지은	15	6	9	21	A42	김찬주	8	6	2	14
A20	김병수	17	3	14	20	A43	윤중호	8	6	2	14
A21	하재명	19	1	18	20	A44	이강국	8	6	2	14
A22	김성우	15	4	11	19	A45	조동우	12	2	10	14
A23	배웅규	13	6	7	19	A46	조진균	7	7	0	14

2) 저자동시인용분석

저자 선정은 피인용 빈도 9회 이상을 기준으로 하였으며, 분석 대상 논문 2,104편 논문에서 인용된 6,927문헌의 피인용 저자 3,819명을 대상으로 최종적으로 46명의 저자를 선정하였다. 9회 이상 인용된 저자를 선정하였을 때 49명의 저자가 선정되었으나, 문화재청, 서울특별시, 건설교통부 등의 단체명이 포함되어 있어 이를 제외하였다. 그 이유는 이들 단체명의 경우 주제를 정확히 파악하기에 모호한, 광범위한 주제를 포함하고 있어 본 연구에서 주제 영역을 표현하기 부적합하기 때문이다.

저자프로파일링분석과의 용이한 비교를 위하여 이의 기준에 적합한 인용빈도 9이상인 49명에서 단체저자명을 제외하고 46명을 선정하였다. 최종 선정된 저자 46명중 피인용 빈도가 가장 높은 상위 5명의 저자는 피인용 빈도 24회인 신중진, 손장열(23회), 최윤경(21회), 박영기(19회), 이상호, 홍원화(18회) 순으로 나타났다(표 2 참조).

3.1.3 1차 프로파일 행렬

1) 저자프로파일링분석 문헌 자동색인

46명의 저자가 선정된 후, 분석 대상으로 선정된 저자들의 단위를 단어 벡터로 표현하여

<표 2> 저자동시인용분석 46명의 선정 저자 순위 및 피인용 빈도

저자번호	피인용 저자명	피인용 빈도	저자번호	피인용 저자명	피인용 빈도
A01	신중진	24	A24	이규인	12
A02	손장열	23	A25	이한석	12
A03	최윤경	21	A26	이 훈	12
A04	박영기	19	A27	임창복	12
A05	이상호	18	A28	최찬환	12
A06	홍원화	18	A29	김도년	11
A07	최재필	17	A30	김동욱	11
A08	심우갑	16	A31	조동우	11
A09	이경희	16	A32	강부성	10
A10	이진숙	16	A33	송승영	10
A11	김영욱	15	A34	오덕성	10
A12	손세관	15	A35	이강희	10
A13	주남철	15	A36	이정재	10
A14	강인호	14	A37	김수암	9
A15	정무웅	14	A38	김영하	9
A16	조성희	14	A39	민병호	9
A17	김성아	13	A40	박철수	9
A18	김정태	13	A41	여명석	9
A19	김진균	13	A42	윤정숙	9
A20	이경훈	13	A43	이연구	9
A21	임채진	13	A44	이재훈	9
A22	김광우	12	A45	제해성	9
A23	서기영	12	A46	조성기	9

문헌 자동색인을 실시하였다. 이를 토대로 저자-색인어 프로파일 벡터는 먼저 수집한 선정 저자의 문헌을 대표하는 제목이나 초록 등의 텍스트에 대해서 자동색인을 수행하여 문헌-색인어 행렬을 도출한다. 본 연구에서 사용된 대상은 문헌의 한글·영문 제목, 한글·영문 키워드, 한글·영문 초록이 사용되었다. 본 연구에서는 선정 저자 46명의 618편의 대표 논문이 위와 같은 방법으로 추출되어 저자-문헌 프로파일 행렬을 구성하였다.

2) 저자동시인용분석의 저자동시인용빈도 조사

저자동시인용분석을 위하여 선정된 46명을 대상으로 하여 인용한 논문과 피인용 저자 사이의 행렬을 작성하였다. Excel 2007 프로그램의 행에 선정 저자 46명의 이름을 위치하게 하고, 선정 저자들을 인용한 문헌들의 저자-문헌 행렬을 작성한다. 동일한 방법으로 문헌-저자 행렬을 작성하고 이를 토대로 2차 상관계수 행렬인 저자동시인용빈도 행렬을 작성한다.

3.1.4 2차 연관성 행렬

1) 저자프로파일링분석 저자-색인어 프로파일 행렬작성

저자-문헌 프로파일 벡터를 작성한 후 이를 이용하여 분석 대상인 46명의 저자 간 문헌유사성 행렬을 산출하였다. 이는 저자들의 분석 대상이 되는 해당저자의 학술지 한글·영문 기사 제목, 한글·영문 키워드, 한글·영문 초록에서 추출한 용어의 유사도가 저자에 포함된 벡터를 대상으로 행렬을 산출한 것이다.

2) 저자동시인용분석의 저자동시인용빈도 행렬 작성

분석 대상 저자가 선정되면 선정된 저자와 다른 한 명의 저자가 나머지 저자들과 갖는 동시인용빈도를 계산할 수 있도록 저자동시인용빈도행렬을 작성한다.

동시인용빈도의 대각선 값을 처리하는 방식은 행렬값의 가공 방법에 따라 1차 프로파일 행렬과 2차 연관성 행렬로 나누어 볼 수 있다. 1차 프로파일 행렬값은 빈도값을 그대로 이용하거나, 저자의 출현빈도(보정한 대각선값)를 감안하여 코사인계수 등으로 정규화한 값으로 구성된 행렬이다(이재윤 2006).

2차 연관성 행렬은 White와 Griffith(1981)가 제안한 바와 같이 1차 프로파일 행렬을 다시 가공하여 생성되는 연관성 행렬을 말한다. 본 연구에서는 대각선값의 보정과 결측치 처리 없이 자기인용을 포함한 모든 값을 사용하여 대각선 행렬값을 구하여 사용하였다.

3.1.5 피어슨 상관계수행렬

1) 저자프로파일링분석

3.1.4에서 2차 상관계수 행렬을 바탕으로 코사인 유사도행렬을 이용하여 저자 간 문헌유사성 행렬을 산출하였다. 본 연구에서는 코사인 유사도행렬 외에 벡터 유사도 공식인 피어슨 상관계수 공식을 적용한 피어슨 상관계수행렬을 작성하였다. 분석 결과, 상관관계가 가장 높은 것으로 나타난 저자쌍은 최원기-서승직(0.993)이며, 유사도 값이 0.8 이상인 저자쌍은 서지은-이정호(0.991), 송승영-석호태(0.879), 이진숙-김병호(0.878), 배웅규-김도년(0.868), 김병선-김광호(0.868) 순으로 나타났다. 상관관계가 가

<표 3> 저자프로파일링분석 46명의 피어슨 상관계수 행렬일부

Author	최종환	신중진	최민기	최재필	추송영	최문영	김병진	권영하	이진욱	임희복	홍은경	권성안	권미영	이경표	서종국	최보희	박성기	최인철	서기근	김병수	황재명	김성우	백종규
최종환	1.000	-0.030	0.253	-0.047	0.229	0.111	0.357	0.091	-0.054	-0.034	-0.093	-0.129	-0.093	-0.034	0.200	0.182	0.008	0.050	0.004	0.051	0.050	0.042	0.011
신중진	-0.030	1.000	-0.230	0.065	-0.274	0.186	-0.187	0.554	-0.202	0.088	0.175	-0.045	0.235	0.189	-0.250	-0.255	0.232	-0.145	0.131	-0.282	0.465	0.082	0.732
최민기	0.253	-0.230	1.000	-0.148	0.391	-0.117	0.396	-0.175	0.126	-0.086	-0.211	-0.154	-0.194	0.093	0.499	-0.209	0.213	0.000	-0.205	0.355	-0.257	0.012	-0.206
최재필	-0.047	0.065	-0.148	1.000	-0.224	0.171	-0.082	-0.148	-0.129	0.040	0.050	0.127	0.101	0.212	-0.161	-0.209	0.213	0.000	-0.167	-0.205	0.163	0.060	0.066
추송영	0.229	-0.274	0.391	-0.224	1.000	-0.159	0.440	-0.273	0.090	0.010	-0.162	-0.233	-0.188	-0.276	0.438	0.079	-0.169	-0.016	-0.273	0.299	-0.327	0.020	-0.255
최문영	0.111	0.186	-0.117	0.396	-0.092	1.000	0.139	0.000	-0.108	0.074	0.111	-0.004	0.143	0.389	-0.107	0.183	0.425	-0.054	0.348	-0.161	0.417	0.156	0.157
김병진	0.357	-0.197	0.396	-0.092	0.440	0.139	1.000	-0.115	0.059	-0.039	-0.117	-0.137	-0.113	-0.105	0.455	0.377	0.013	0.074	-0.177	0.214	-0.153	0.144	-0.152
권영하	0.091	0.554	-0.175	0.156	-0.273	0.290	-0.115	1.000	-0.215	0.068	0.211	-0.030	0.268	0.212	-0.196	-0.252	0.255	0.104	0.146	-0.269	0.655	0.086	0.672
이진욱	-0.054	-0.202	0.125	-0.289	0.090	-0.108	0.059	-0.215	1.000	-0.124	0.196	-0.030	0.268	0.212	-0.196	-0.252	0.255	0.104	0.146	-0.269	0.655	0.086	0.672
이경표	-0.034	0.068	-0.086	0.040	0.100	0.074	-0.089	0.068	-0.124	1.000	0.637	0.655	0.025	-0.005	-0.079	0.048	0.011	-0.012	-0.131	0.018	-0.191	-0.211	-0.218
박성기	-0.003	0.175	-0.211	0.060	-0.162	0.111	-0.117	0.201	-0.198	0.637	1.000	0.015	0.042	0.145	-0.214	-0.162	0.139	0.165	0.036	-0.245	0.165	0.117	0.141
김병수	-0.139	-0.045	-0.154	0.127	-0.223	0.004	-0.137	-0.030	-0.125	-0.056	-0.015	1.000	-0.016	0.056	-0.175	-0.216	0.082	-0.063	0.082	-0.122	-0.021	-0.005	-0.050
황재명	-0.053	0.236	-0.134	0.101	-0.188	0.143	-0.113	0.288	-0.161	0.026	0.042	-0.018	1.000	0.111	-0.154	-0.190	0.088	0.153	0.168	-0.229	0.270	0.136	0.212
김성우	0.200	-0.250	0.093	-0.161	0.438	-0.107	0.455	-0.196	0.143	-0.079	0.145	0.026	1.000	0.111	-0.208	-0.274	0.406	-0.063	0.357	-0.209	0.394	0.028	0.165
백종규	0.082	-0.265	0.332	-0.209	0.079	-0.189	0.377	-0.252	0.086	-0.048	-0.182	-0.216	-0.180	0.406	0.371	1.000	-0.139	-0.031	-0.264	0.251	-0.323	0.018	-0.250
최인철	0.008	0.232	-0.172	0.213	-0.169	0.426	0.113	0.086	-0.162	0.011	0.139	0.062	0.098	0.406	-0.177	0.371	1.000	0.252	-0.229	0.340	0.100	0.203	0.203
서기근	0.050	-0.145	0.124	0.000	-0.015	-0.054	0.074	-0.104	0.292	-0.117	-0.165	-0.063	-0.153	-0.063	0.134	-0.031	1.000	0.102	0.228	-0.068	-0.065	-0.130	0.130
김병수	0.004	0.131	-0.205	0.157	-0.273	0.348	-0.177	0.146	-0.017	-0.012	0.036	0.082	0.168	0.357	-0.219	-0.254	0.252	1.000	-0.136	0.342	0.034	0.063	0.063
황재명	0.051	-0.282	0.355	-0.205	0.249	0.151	0.214	-0.209	0.078	-0.131	-0.245	-0.182	0.229	0.206	0.384	0.251	0.229	0.228	-0.136	1.000	-0.305	-0.105	-0.286
김성우	0.050	0.485	-0.257	0.163	-0.307	0.417	-0.153	0.555	-0.191	0.118	0.166	-0.021	0.270	0.434	-0.276	-0.323	0.340	-0.083	0.342	-0.305	1.000	0.102	0.511
김재필	0.042	0.082	0.012	0.060	0.200	0.155	0.144	0.086	-0.121	0.240	0.117	-0.003	0.138	0.133	0.023	-0.018	0.100	-0.065	0.034	-0.105	0.102	1.000	0.013
백종규	0.011	0.732	-0.205	0.065	-0.265	0.157	-0.152	0.672	-0.218	0.012	0.141	-0.050	0.212	0.165	-0.225	-0.250	0.285	-0.130	0.053	-0.286	0.511	1.000	0.013
최인철	0.044	-0.255	0.346	-0.183	0.420	-0.064	0.525	-0.225	0.116	-0.001	-0.111	-0.173	-0.189	-0.163	0.396	0.362	-0.107	0.040	-0.250	0.259	-0.287	0.036	-0.256
김병수	-0.197	-0.053	-0.148	0.022	-0.185	-0.056	-0.166	-0.025	-0.095	-0.097	-0.085	0.150	0.117	-0.001	-0.168	0.362	-0.107	-0.099	0.056	-0.157	-0.006	0.066	-0.160
황재명	-0.125	0.078	-0.247	0.270	-0.280	0.216	-0.216	0.095	-0.106	-0.083	-0.021	0.051	0.170	0.434	-0.283	-0.259	0.163	-0.066	0.416	-0.215	0.217	-0.028	0.043
김성우	-0.010	0.139	-0.196	0.065	-0.236	0.402	-0.109	0.200	-0.160	0.009	0.171	-0.019	0.035	0.319	-0.203	-0.214	0.278	-0.133	0.185	-0.204	0.354	0.013	0.129
김재필	0.222	-0.160	0.319	-0.060	0.367	0.075	0.858	-0.073	0.098	-0.037	-0.093	-0.123	-0.082	-0.070	0.372	0.321	0.004	0.052	-0.106	0.211	-0.094	0.153	-0.134
백종규	0.054	0.455	-0.154	0.159	-0.209	0.349	-0.019	0.598	-0.157	0.075	0.162	-0.003	0.247	0.286	-0.175	-0.193	0.322	-0.084	0.234	-0.236	0.525	0.136	0.592
최인철	0.107	-0.007	0.003	-0.022	0.111	0.045	0.052	-0.045	-0.042	0.649	0.438	-0.084	-0.049	0.010	0.022	0.049	0.071	-0.064	-0.005	-0.031	-0.028	0.185	-0.078
김성우	0.026	0.145	-0.205	0.145	-0.261	0.345	-0.177	0.168	-0.029	-0.033	0.022	0.085	0.190	0.365	-0.219	-0.255	0.245	0.094	0.991	-0.147	0.357	0.030	0.069
김재필	-0.021	0.270	-0.115	0.062	-0.215	0.089	-0.137	0.519	-0.159	0.018	0.132	-0.075	0.197	0.067	-0.134	-0.168	0.076	-0.111	0.010	-0.226	0.344	0.025	0.389
백종규	0.035	-0.109	0.003	-0.130	-0.023	0.029	0.008	-0.140	-0.060	0.060	0.040	-0.064	-0.169	-0.098	0.013	-0.007	-0.037	-0.059	-0.036	-0.054	-0.141	-0.080	-0.134
김병수	-0.107	0.101	-0.244	0.081	-0.318	0.204	-0.223	0.166	-0.127	-0.144	-0.094	0.101	0.143	0.288	-0.266	-0.268	0.323	0.086	0.309	-0.234	0.231	0.016	0.110
황재명	-0.139	-0.010	-0.159	0.265	-0.228	0.094	-0.153	0.032	-0.147	-0.015	0.005	0.480	0.019	0.041	-0.179	-0.221	0.099	-0.035	0.066	-0.209	0.031	-0.001	-0.013
김성우	-0.137	0.082	-0.143	0.021	-0.162	-0.024	-0.149	0.060	-0.096	-0.064	-0.061	-0.001	0.378	-0.043	-0.165	-0.158	-0.062	-0.144	-0.006	-0.173	0.048	0.100	0.021
최인철	-0.106	-0.003	-0.222	-0.051	-0.226	0.120	-0.183	0.006	-0.124	-0.090	0.014	-0.066	0.001	0.253	-0.233	-0.215	0.109	-0.166	0.051	-0.166	0.130	-0.096	-0.011
백종규	-0.079	0.157	-0.135	0.092	-0.168	0.167	-0.049	0.307	-0.124	0.020	0.025	-0.064	0.096	0.288	-0.141	-0.151	0.255	-0.101	0.104	-0.177	0.260	0.104	0.205
최인철	0.011	0.857	-0.216	0.060	-0.275	0.138	-0.171	0.689	-0.199	0.001	0.148	-0.047	0.192	0.128	-0.237	-0.259	0.229	-0.091	0.059	-0.274	0.478	-0.030	0.868
김성우	0.040	0.059	-0.127	0.120	-0.150	0.137	-0.087	0.071	-0.126	-0.031	0.051	0.007	-0.007	0.133	-0.134	-0.152	0.123	0.113	-0.166	0.228	0.002	-0.001	0.000
김재필	0.307	-0.250	0.361	-0.208	0.428	-0.118	0.525	-0.238	0.071	-0.068	-0.178	-0.164	-0.169	-0.189	0.218	0.415	0.397	0.194	0.068	-0.221	0.229	0.025	-0.218
백종규	-0.037	0.095	-0.188	0.453	-0.212	0.236	-0.156	0.088	-0.107	-0.075	-0.045	-0.003	0.049	0.190	-0.199	-0.187	0.365	-0.033	0.248	-0.167	0.165	-0.003	0.036
김성우	0.153	-0.284	0.517	-0.239	0.440	-0.179	0.255	-0.283	0.432	-0.066	-0.222	-0.197	-0.210	-0.254	0.545	0.348	0.348	0.032	-0.252	0.706	-0.341	-0.049	0.261
이경표	0.781	0.014	0.229	-0.110	0.170	0.015	0.232	0.108	-0.065	-0.123	-0.119	-0.140	-0.069	-0.129	0.252	0.120	-0.109	0.036	-0.114	0.039	0.048	-0.035	0.106
최종환	0.336	-0.181	0.458	-0.143	0.429	0.102	0.545	-0.107	0.134	0.023	-0.087	-0.202	-0.166	-0.117	0.486	0.351	0.010	0.246	-0.179	0.282	-0.177	0.079	-0.122
최진규	0.345	-0.172	0.323	-0.086	0.264	-0.006	0.700	-0.129	-0.049	-0.039	-0.112	-0.071	-0.163										

장 높은 것으로 나타난 저자쌍인 최원기-서승직은 유사도값이 0.993으로, 최원기와 서승직은 동일한 주제분야를 연구할 가능성이 높음을 알 수 있다. 다시 말하면 상관관계가 높은 쌍은 동일하거나 유사한 주제 분야를 연구하고 있다고 해석할 수 있다.

상관계수행렬에서 음의 상관관계가 큰 저자쌍은 하재명-윤종호(-0.341), 송승영-하재명(-0.327)이며, 그 밖에 -0.300 이하의 유사도값을 갖는 저자쌍은 석호태-하재명(-0.323), 송승영-김영옥(-0.318), 김병수-하재명(-0.305) 순으로 나타났다. 원태연, 정성원(1998)의 상관관계 해석 기준에 따르면 뚜렷한 음적 상관관계를 나타내는 것으로 보이며, 이들은 뚜렷한 음적 상관관계를 가지고 있으므로 각 저자쌍의 저자들은 연구 주제가 뚜렷하게 차이가 있음을 알 수 있다.

2) 저자동시인용분석

동시인용빈도행렬이 작성되면 저자쌍의 상대적 유사성 혹은 상이성을 나타내는 지수 상관계수행렬을 작성한다. 본 연구에서도 피어슨 상관계수를 이용하여 상관계수행렬을 작성하였다.

〈표 4〉의 분석 결과, 상관관계가 가장 높은 것으로 나타난 저자쌍은 김광우와 여명석(0.938)이며, 매우 강한 상관관계를 가지는 0.7 이상의 저자쌍은 최윤경-임채진(0.821)로 나타났으며, 뚜렷한 상관관계를 가지는 저자쌍을 나열하면, 김영옥-최윤경(0.662), 신중진-임창복(0.607), 손장열-윤정숙(0.570), 최윤경-최재필(0.535), 이경훈-최윤경(0.533), 서기영-신중진(0.527), 박철수-강부성(0.521), 김영옥

-임채진(0.514), 민병호-최윤경(0.491), 박철수-조성희(0.471), 김영옥-최재필(0.470), 강부성-조성희(0.467), 서기영-임창복(0.464), 강부성-김진균(0.447), 서기영-오덕성(0.447), 조성기-최찬환(0.462), 윤정숙-이규인(0.427), 임창복-정무웅(0.415), 민병호-이경훈(0.407) 등이다.

상관계수행렬에서 약한 음의 상관관계를 갖는 가장 큰 저자쌍은 서기영-최윤경(-0.161), 김광우-김진균(-0.153), 강부성-이정재(-0.153), 강부성-손장열(-0.150) 등이 있다.

3.2 지적 구조 분석 결과

3.2.1 군집분석

군집을 위한 대표적인 응집방식으로는 Ward 기법과 집단내 평균연결방식인 WAVE 기법 등이 있으며, 모두 계층적 응집방식을 이루기 위한 기법 중 하나이다. 저자프로파일링분석의 경우 WAVE 기법을 이용하여 8군집을 나누었고, 저자동시인용분석의 경우 Ward 기법을 이용하여 7군집으로 군집화 하였으나 군집내의 적절한 저자군집을 이루기 위한 방법으로 두 기법간의 결과적 차이는 없다.

3.2.2 다차원척도법에 의한 분석

1) 저자프로파일링분석

본 연구에서는 저자 간의 관계를 2차원 공간상에 나타내기 위하여 피어슨 상관계수행렬을 입력 데이터로 두 저자 사이의 유사성 또는 비유사성에 따라 좌표상에 저자의 위치를 점으로 표시하여 시각화 하는 기법으로 분석한다(표 5 참조).

지도상에 나타나는 점 즉 좌표값은 공간상의

〈표 5〉 저자프로파일링분석에서 각 저자의 좌표값

저자번호	저자	1차원	2차원	저자번호	저자	1차원	2차원
A1	홍원화	0.743	-0.062	A24	손장열	0.906	-0.051
A2	신중진	-0.610	-0.137	A25	이동언	-0.457	0.367
A3	최원기	0.921	0.067	A26	임채진	-0.668	0.104
A4	최재필	-0.629	0.105	A27	강병근	-0.605	-0.056
A5	송승영	0.921	-0.068	A28	김광호	0.842	0.033
A6	최무혁	-0.508	-0.053	A29	김원필	-0.588	-0.113
A7	김병선	0.863	-0.005	A30	이강희	0.462	-0.319
A8	김영하	-0.594	-0.134	A31	이정호	-0.601	0.104
A9	이진숙	0.824	0.263	A32	최찬환	-0.560	-0.199
A10	임석호	-0.066	-0.420	A33	곽노열	0.607	-0.239
A11	황은경	-0.397	-0.255	A34	김영욱	-0.663	0.133
A12	김성아	-0.485	0.347	A35	이상현	-0.546	0.284
A13	김태영	-0.624	-0.017	A36	전봉희	-0.450	0.236
A14	이경훈	-0.638	0.006	A37	조영행	-0.547	0.066
A15	서승직	0.926	0.050	A38	최열	-0.512	-0.158
A16	석호태	0.924	-0.067	A39	김도년	-0.605	-0.157
A17	박영기	-0.576	-0.040	A40	이정수	-0.440	0.036
A18	최안섭	0.681	0.288	A41	이정재	0.917	-0.021
A19	서지은	-0.593	0.102	A42	김찬주	-0.624	0.159
A20	김병수	0.912	0.162	A43	윤종호	0.937	0.121
A21	하재명	-0.635	-0.070	A44	이강국	0.716	-0.032
A22	김성우	-0.041	-0.145	A45	조동우	0.888	-0.018
A23	배용규	-0.589	-0.164	A46	조진균	0.858	-0.030

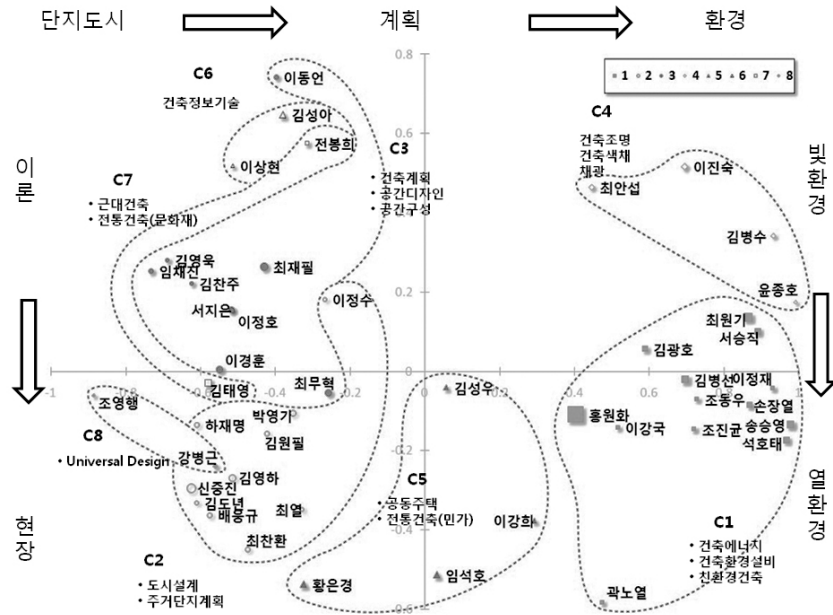
지도에 대상들의 상대적 위치를 나타낸다. 즉, 두 저자 간의 유사도가 높고 비유사도가 낮을수록 두 점은 공간상에서 가까이 배열되고, 반대로 유사도가 낮고 비유사도가 높은 저자의 경우 멀리 떨어져서 나타나며, 다른 저자들과 유사성이 높은 저자들은 지도의 중심에 위치하고, 유사성이 낮은 저자들은 지도의 주변에 위치하게 된다.

1차원 축 값 가운데 최고값을 갖는 저자는 윤종호(0.937)이며, 최저값(-0.668)을 갖는 저자는 임채진으로 나타났다. 2차원 축 값 가운데 최고값(0.367)을 가지는 저자는 이동언이며, 최저

값(-0.420)을 갖는 저자는 임석호이다.

〈표 5〉에 나타난 저자별 좌표값에 따라 상관 계수행렬 데이터를 입력데이터로 하여 SPSS의 PROXCAL 명령을 통하여 저자들의 위치를 2차원 공간상에 지도로 나타내었고 〈그림 2〉의 다차원척도 지도와 같다.

저자프로파일링분석의 경우 1차원 값 중 최고값을 갖는 저자는 윤종호이며 0.937의 좌표값을 나타내었으며, 다차원척도 지도상의 가장 우측에 출현하였고 건축학 분야 가운데 건축에너지, 건축조명, 채광 관련연구를 하고 있는 것으로 밝혀졌다. 최저값을 갖는 저자는 임채진



〈그림 2〉 저자프로파일링분석의 다차원척도 지도

(-0.668)이며, 다차원척도 지도상의 좌측에 출현하였으며 건축학 분야 연구 중에서 공공건물 건축, 공간구성, 공간디자인등을 연구하는 연구자로 확인되었다. 각 저자의 좌표값은 저자 간의 유사도가 낮고 상이도가 높은 저자의 경우 멀리 떨어져서 나타나므로 종축을 기준으로 지도상의 좌측과 우측에 나타나는 윤중호와 임채진은 두 저자가 연구하는 주제의 유사도가 가장 낮은 경우로 해석할 수 있다.

2) 저자동시인용분석

다차원분석을 위하여 피어슨 상관계수행렬을 입력 데이터로 하여 SPSS 프로그램을 실행하였다. 다차원척도분석은 앞에서와 같이 저자프로파일링분석과 동일한 분석을 위하여 실행하였다. 그 결과 각 저자의 유사도에 따라 다음의 <표 6>과 같이 1차원(횡축)과 2차원(종축)

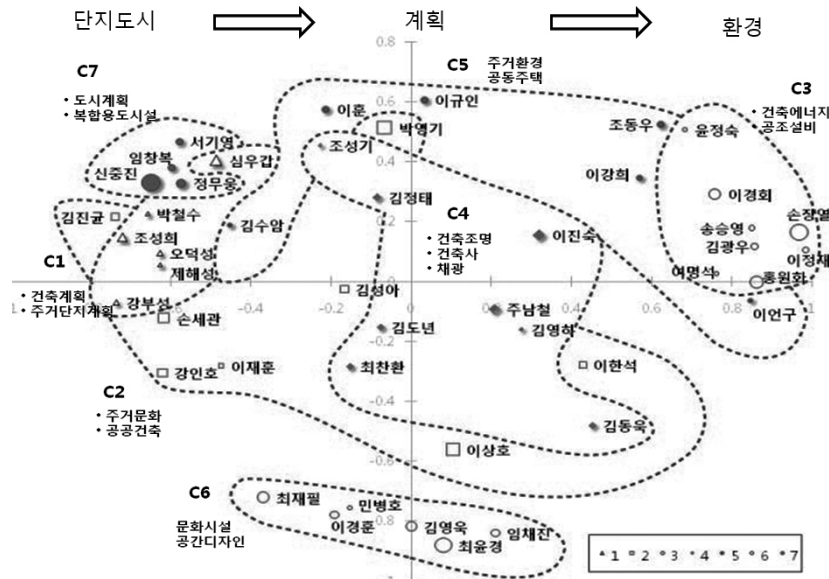
의 좌표값이 나타났다. 1차원 값 가운데 최고값은 0.984로 해당 저자는 이정재이며, 최저값(-0.738)을 갖는 저자는 김진균이다. 2차원 값 가운데 최고값을 가지는 저자는 이규인(0.605)이며, 최윤경은 -0.881로 최저값을 갖게 되었다.

<표 6>에 나타난 저자별 좌표값에 따라 상관계수 행렬 데이터를 입력데이터로 하여 SPSS의 PROXCAL 명령을 통하여 저자들의 위치를 2차원 공간상에 지도로 표현할 수 있고 이는 <그림 3>과 같다.

1차원 값 중 최고값을 갖는 이정재(0.984)는 지도의 가장 우측에 나타났으며 최저값(-0.738)을 갖는 김진균은 지도의 가장 좌측에 위치하고 있다. 2차원 값 중 최고값(0.605)을 가지는 저자 이규인은 지도의 가장 상단에 위치하며, 최저값(-0.881)을 갖는 저자 최윤경은 종축을 기준으로 지도의 가장 하단에 위치하고 있다.

〈표 6〉 저자동시인용분석에서 각 저자의 좌표값

저자번호	피인용저자	1차원	2차원	저자번호	피인용저자	1차원	2차원
A1	강부성	-0.736	-0.070	A24	이강희	0.570	0.348
A2	강인호	-0.621	-0.306	A25	이경희	0.757	0.292
A3	김광우	0.856	0.118	A26	이경훈	-0.191	-0.779
A4	김도년	-0.077	-0.153	A27	이규인	0.031	0.605
A5	김동욱	0.452	-0.481	A28	이상호	0.104	-0.561
A6	김정아	-0.167	-0.027	A29	이연구	0.846	-0.062
A7	김수암	-0.452	0.188	A30	이재훈	-0.474	-0.282
A8	김영옥	0.000	-0.818	A31	이정재	0.984	0.104
A9	김영하	0.275	-0.159	A32	이진숙	0.318	0.156
A10	김정태	-0.085	0.281	A33	이한석	0.427	-0.279
A11	김진균	-0.738	0.215	A34	이훈	-0.214	0.573
A12	민병호	-0.154	-0.756	A35	임창복	-0.599	0.380
A13	박영기	-0.067	0.513	A36	임채진	0.210	-0.840
A14	박철수	-0.658	0.225	A37	정무용	-0.574	0.327
A15	서기영	-0.578	0.467	A38	제해성	-0.626	0.055
A16	손세관	-0.616	-0.122	A39	조동우	0.622	0.526
A17	손장열	0.968	0.163	A40	조성기	-0.229	0.455
A18	송승영	0.850	0.178	A41	조성희	-0.721	0.146
A19	신중진	-0.649	0.332	A42	주남철	0.207	-0.092
A20	심우갑	-0.487	0.403	A43	최윤경	0.080	-0.881
A21	여명석	0.761	0.027	A44	최재필	-0.369	-0.720
A22	오덕성	-0.628	0.095	A45	최찬환	-0.153	-0.284
A23	윤정숙	0.683	0.506	A46	홍원화	0.861	-0.003



〈그림 3〉 저자동시인용분석의 저자 지도

3.2.3 CBNet를 이용한 분석

1) 저자프로파일링분석

저자프로파일링분석의 결과로 나타난 저자 지도를 계층적 클러스터링 알고리즘을 이용한 CBNet-WAVE 기법으로 네트워크로 표현하였다. 네트워크로 저자 지도를 시각화 하기 위하여 공개 소프트웨어인 Pajek을 사용하였다. 네트워크로 표현한 저자 지도를 살펴보면 중심 저자를 확인할 수 있다(그림 4 참조).

이처럼 CBNet 지도는 다차원척도법과는 달리 저자 간의 연관성을 보다 직관적이고 개괄적으로 파악할 수 있으며, 군집 간의 허브역할을 하는 저자와 중심 저자 및 주변 저자의 구분이 직관적으로 가능하다.

2) 저자동시인용분석

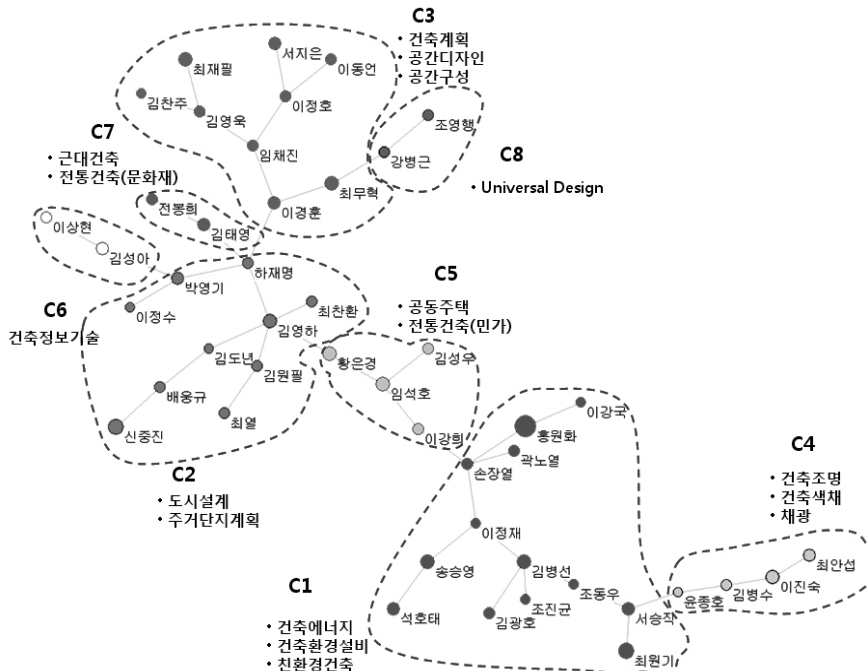
다차원척도가 비교적 복잡한 지도의 양상을 보여 네트워크로 표현한 지도가 훨씬 명확하게 군집을 표현하고 있음을 알 수 있다(그림 5 참조).

4. 구성요소에 의한 지적 구조 비교

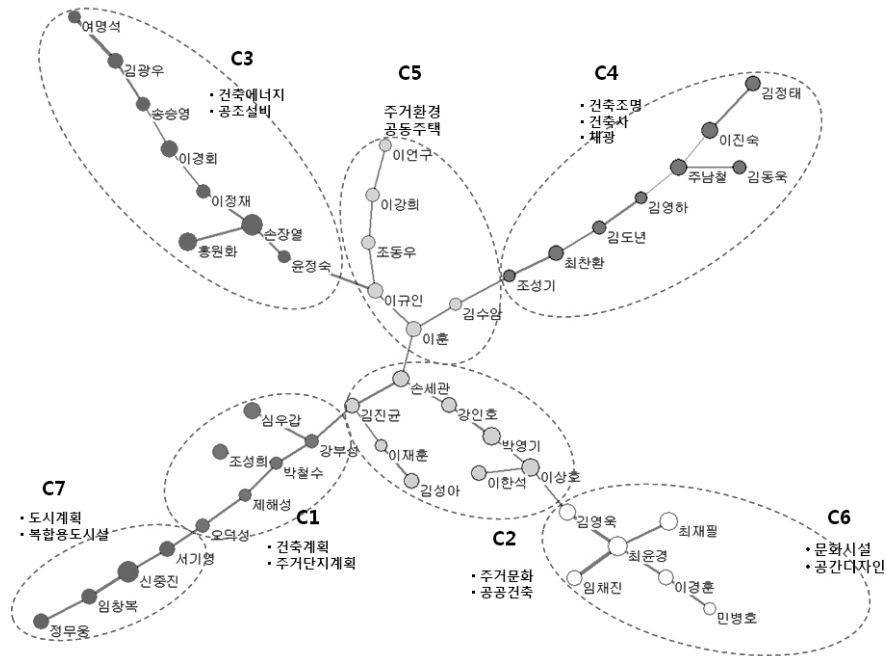
4.1 분석 기법별 구성요소 비교

4.1.1 연령

저자프로파일링분석을 위해 선정된 선정 저자 46명의 평균 연령은 50.5세이고, 저자동시



<그림 4> 저자프로파일링분석에 기반한 CBNet-WAVE



〈그림 5〉 저자동시인용분석에 기반한 CBNet-Ward

인용분석을 위해 선정된 저자 46명의 평균 연령은 55.7세로 두 기법간에는 5.2세 정도의 차이가 있다. 저자프로파일링분석을 이용하여 분

석된 연구의 저자들이 저자동시인용분석에 선정된 저자들보다 상대적으로 젊다는 것을 보여 준다(표 7 참조).

〈표 7〉 선정 저자의 연령별 구분

구 분	연령구분(세)	인 원(명)	백분율(%)
선정 저자 (저자프로파일링분석) 평균연령 (50.5세)	30 이하	0	0.0
	31~40	5	10.9
	41~50	17	37.0
	51~60	18	39.1
	61~70	6	13.0
	71 이상	0	0.0
	합 계		46
선정 저자 (저자동시인용분석) 평균연령 (55.7세)	30 이하	0	0.0
	31~40	0	0.0
	41~50	13	28.2
	51~60	19	41.3
	61~70	11	23.9
	71 이상	3	6.6
	합 계		46

4.1.2 연구활동기간

저자동시인용분석의 선정 저자들 보다 저자프로파일링분석을 위해 선정된 저자들의 연구활동기간이 짧아 저자프로파일링분석의 선정 저자들이 저자동시인용분석의 선정 저자들에 비하여 비교적 신진 연구자들인 것으로 판단된다(표 8 참조).

4.1.3 소속기관 · 학위 수여교와 주제 군집간 상관성 분석

선정저자들의 소속기관과 학위수여교간의 상관관계를 분석하기 위하여 두 개의 범주형 변수가 서로 상관이 있는지 독립적인지를 판단하는 통계적 검정방식인 카이제곱검정을 실시하려 하였으나 교차표분석 χ^2 는 셀값이 1인칸이 너무 많아서 적용이 불가하였다. 따라서 교차표에서 명목 변수간 일치도를 계산하는 공식으로 상대적인 유사도 비교만을 수행하였다. 클러스

터링 성능 평가를 위한 비편향적 단일 척도로 가중 평균 클러스터 유사도인 WACS(Weighted Average Cluster Similarity) (정영미, 이재운 2001) 척도와 F척도가 대표적이므로 두 가지를 사용하였다. WACS척도를 산출하기 위하여 범주와 클러스터 사이의 유사도를 계산하여 범주와 클러스터의 크기를 고려한 가중 평균으로 전체의 유사도를 계산한 결과 다음과 같은 결과를 얻었다(표 9 참조).

저자프로파일링분석 군집 - 선정저자 학위수여교간의 WACS척도 값은 0.318로 제일 높은 값을 나타내었고, F척도값은 0.412로 나타났으며, 역시 가장 높은 척도값을 나타내었다. 이러한 결과로 볼 때 저자프로파일링분석 군집 - 선정저자 학위수여교 간의 관계는 학문적 배경과 저자간 상관관계가 있다는 검증결과가 입증되었다.

〈표 8〉 선정 저자의 연구 활동 기간

구 분	연구기간(년)	인원(명)	백분율(%)
선정 저자 (저자프로파일링분석) 평균 연구활동기간 (14.9년)	1 ~ 10	12	26.1
	11 ~ 15	12	26.1
	16 ~ 20	11	23.9
	21 ~ 25	9	19.6
	26 ~ 30	2	4.3
	31년 이상	0	0.0
	합 계	46	100.0
선정 저자 (저자동시인용분석) 평균 연구활동기간 (19.5년)	1 ~ 10	0	0.0
	11 ~ 15	8	17.4
	16 ~ 20	19	41.4
	21 ~ 25	14	30.4
	26 ~ 30	3	6.5
	31년 이상	2	4.3
	합 계	46	100.0

〈표 9〉 소속기관과 학위 수여교간의 상관관계분석

비교 대상	WACS	F척도
저자동시인용분석 군집 - 선정저자 소속기관	0.271	0.302
저자동시인용분석 군집 - 선정저자 학위수여교	0.284	0.319
저자프로파일링분석 군집 - 선정저자 소속기관	0.296	0.327
저자프로파일링분석 군집 - 선정저자 학위수여교	0.318	0.412

4.2 분석 기법별 지적 구조 해석과 비교

4.2.1 분석 기법별 지적 구조의 비교 분석

군집별 주제명에 대하여 전문가와의 면담을 통하여 분석 기법별로 포괄성, 적절성, 최신성에 대하여 자문을 구한 결과 저자프로파일링분석에 대하여 2인의 저자 모두 세 개의 항목에 대하여 긍정적 답변을 표시하였다. 특히 주제의 포괄성과 최신성이 저자동시인용분석 보다 매우 뛰어나다는 의견을 피력하였다. 저자동시인용분석의 제4 군집과 제5 군집의 군집 저자들이 군집내 다른 저자들과 군집되기에는 연구분야가 너무 상이하다는 의견을 보였고, 전체적인 군집의 결과가 저자프로파일링보다 매우 부정확하다는 답변을 제시하였다.

두 분석 기법별 지적 구조의 비교 분석을 한 결과를 살펴보면 다음과 같다.

첫째, 주제분야의 세부적 영역 표현은 저자프로파일링분석이 더 명확하게 표현되었다. 저자동시인용분석과의 주제 영역과 비교를 통해 저자프로파일링분석에 새롭게 등장한 학문의 영역인 '건축정보기술', 'Universal Design'이 새로운 군집으로 군집화 되었으며, 저자동시인용분석에서 확인되지 않았던 '친환경 건축', '공간구성'이 등장하였고, '전통건축(민가)', '전통건축(문화재)', '근대건축'이 '건축사'에서 주제

가 더욱 세분화 되어 나타남으로써 연구 동향을 파악할 수 있었다. '건축정보기술', 'Universal Design'은 2000년대 들어서 많이 연구가 진행되고 있는 것으로 보았을 때 피인용을 분석하는 저자동시인용분석에서는 주제로 표현되지 않고, 저자프로파일링분석에서 주제가 표현되는 결과를 보인 것으로 해석된다.

둘째, 분석 기법별로 다차원척도에 나타난 차원의 비교에서 저자프로파일링분석이 저자동시인용분석 보다 명확하게 해석 되었다. 저자동시인용분석에서 횡축 방향의 '단지·도시'에서 건물내부의 세부적인 '환경'의 해석만 할 수 있었던 것에 비하여, 저자프로파일링분석의 다차원척도 지도에서는 횡축방향의 '단지·도시'에서 '환경'이라는 특성을 확인하였고 종축을 기준으로 좌측상에서 하방향으로 '이론·의미'에서 '현장'의 특성을 확인하였고, 우측의 상하에서 '빛환경', '열환경'이라는 해석이 가능한 결과를 보였다.

셋째, 분석 기법별로 CBNet의 저자 지도를 이용하여 저자프로파일링분석과 저자동시인용분석을 통하여 중심저자와 주변저자를 각각 확인할 수 있었고, 분석 기법별 비교를 통하여 저자들의 이동을 파악할 수 있었다. 이러한 결과는 저자프로파일링분석과 저자동시인용분석을 동시에 분석하였을 때 확인 가능한 결과이다.

저자프로파일링분석을 시기별로 나누어서 분석을 시도한다면 비슷한 결과를 얻을 수 있을 것으로 기대된다. 또한 저자프로파일링분석의 네트워크 링크 가지의 수가 19개, 저자동시인용분석 링크 가지의 수가 13개로 저자프로파일링분석이 주제의 구분을 더 세세하게 표현하고 있는 것으로 확인되었다. 이러한 결과는 전문가들의 의견 수렴과정에서 확인이 되었는데 주제를 좀 더 포괄적이면서도 세밀하게 표현한다는 의견을 피력하였다.

넷째, 분석 기법별로 선정 저자의 연령과 연구 활동 기간을 비교하여 보았을 때, 저자프로파일링분석대상 저자들의 평균 연령이 50.5세, 연구 활동 기간이 14.9년, 저자동시인용분석대상의 평균 연령이 55.7세, 연구 활동 기간이 19.5년으로 저자프로파일링분석을 통하여 신진 연구자들의 파악이 비교적 용이하였다. 또한 50세 이하 연구자가 저자프로파일링분석의 경우 47.9%인데 반해 저자동시인용분석의 경우 28.2%밖에 차지하지 않아 신진 연구자들의 포함 범위를 잘 확인해 주었고, 군집별 주제명에 포함된 저자프로파일링분석의 제6 군집 '건축정보기술'의 저자들 평균연령이 46.5세 연구 활동 기간이 12.0년, 제8 군집의 'Universal Design'의 평균 연령이 43.5세, 연구 활동 기간이 5.5년으로 분석 기법별 15개 군집 중 가장 젊은 연구자들이 소속되어 있었고, 전문가 의견수렴 과정에서도 최신 주제에 신진 연구자들이 명확히 포함되어 있음이 확인되었다.

다섯째, 분석 기법별로 교차분석을 실시한 결과 저자프로파일링분석의 학위 수여교와 주제 군집간 교차분석에서 각각 학위 수여교와 주제 군집간 상관성이 확인되었고, 저자동시인

용분석의 소속기관-군집별 주제명의 교차분석에서는 주제 군집과 소속기관간 상관성이 확인되기도 하였다. 이러한 결과는 저자프로파일링분석의 학위 수여교가 학문적 배경으로 작용하여 주제의 특성으로 작용한다는 결과를 보여주는 것으로 판단되고, 소속기관과 주제 군집간에도 일부 상관관계가 존재한다는 것을 보여주고 있다. 또한 분석 기법별로 소속기관-군집별 주제명, 학위 수여교-군집별 주제명의 WACS 척도 분석과 F척도 분석을 통하여 저자프로파일링분석의 학위 수여교-군집별 주제명의 교차분석에서 그 상관성이 확인되기도 하였다.

5. 결론

본 연구는 학문의 지적 구조를 분석하는 새로운 분석기법인 저자프로파일링분석과 전통적인 분석기법인 저자동시인용분석을 비교하여 분석함으로써 국내의 연구환경에 맞는 지적 구조 분석 방법을 제안하는 데 목적을 두고 있다. 이를 통하여 텍스트마이닝을 이용한 저자프로파일링분석을 사용함으로써 인용색인을 이용하지 않고 학문의 지적 구조를 분석할 수 있는 새로운 지적 구조 방법의 유용성을 확인하고자 하였다. 또한 저자동시인용분석과 새로운 분석기법인 저자프로파일링분석을 실시하여 건축학 분야의 전통적인 지적 구조와 최신의 연구동향 및 근래 활발한 연구활동을 하는 핵심연구자를 파악하고자 하는 데 목적을 두고 있다.

본 논문의 분석 기법별 유용성 검증을 위한 연구결과를 요약하면 다음과 같다.

첫째, 저자프로파일링분석의 지도 해석을 통하여 새롭게 등장한 학문의 영역을 저자동시인용분석의 지도와 비교하여 파악함으로써 건축 분야의 연구 동향을 확인할 수 있었다. 주제 분야의 세부적 영역 표현은 저자프로파일링분석이 저자동시인용분석 보다 명확하게 표현되었다.

둘째, 분석 기법별로 다차원척도에 나타난 차원의 비교에서 저자프로파일링분석이 저자동시인용분석보다 명확한 해석의 결과를 보였다. 저자동시인용분석에서 횡축 방향의 해석만 할 수 있었던 것에 비하여, 저자프로파일링분석의 다차원척도 지도에서는 횡축, 종축을 기준으로 좌와 우의 상하에 대한 해석이 가능한 결과를 보였다.

셋째, 분석 기법별로 CBNet의 저자 지도를 이용하여 저자프로파일링분석과 저자동시인용분석을 통하여 중심저자와 주변저자를 각각 확인할 수 있었고, 분석 기법별 비교를 통하여 저자들의 이동을 파악할 수 있었다. 이러한 결과는 저자프로파일링분석과 저자동시인용분석을 동시에 분석하였을 때 확인 가능한 결과이다. 저자프로파일링분석을 시기별로 나누어서 분석을 시도한다면 비슷한 결과를 얻을 수 있을 것으로 기대된다.

넷째, 분석 기법별로 선정 저자의 연령과 연구 활동 기간을 비교 하여 보았을 때, 저자프로파일링분석이 저자동시인용분석 보다 신진 연구자 파악이 용이하였으며, 군집별 주제명의 전문가 자문에서도 이러한 결과는 확인되었다.

다섯째, 분석 기법별로 소속기관과 군집별 주

제명, 학위 수여교와 군집별 주제명의 WACS 척도 분석과 F척도 분석을 통하여 저자프로파일링분석의 학위 수여교와 군집별 주제명 간 교차분석을 통한 상관관계를 밝히는 것이 타당하다는 결과를 보였다. 이러한 결과는 저자프로파일링분석의 결과에서 나타나는 저자들의 학위 수여교 분포는 저자들의 학문적 배경이 반영되어 특정 주제를 구성하는 요인으로 작용하는 것으로 판단된다.

본 연구를 통하여 전통적으로 사용하고 있던 분석방법인 저자동시인용분석 기법이나, 서지결합법 등의 지적 구조 분석방법들이 인용색인이라는 데이터를 이용하여 분석을 시도하고 있으나 국내의 경우 인용색인 구축의 어려움, 구축된 인용색인의 불완전성으로 인하여 국내 학문의 지적 구조를 분석하기에는 많은 어려움이 있다. 이러한 어려움을 해결하기 위하여 본 연구에서 제안한 저자프로파일링분석 기법을 이용한 지적 구조 분석의 장점은 다음과 같다.

첫째, 저자동시인용분석 기법에 비하여 최근의 연구 경향을 잘 반영하여 준다. 둘째, 주제 군집에 응집된 저자들과 주제의 관계가 보다 명확하게 표현된다. 셋째, 신진 연구자들의 파악이 용이하다. 넷째, 문헌의 출판과 함께 즉시 분석이 가능하다. 다섯째, 인용색인을 사용하는 분석 방법보다 분석을 위한 데이터 준비가 간단하다.

저자프로파일링분석 기법은 인용색인의 작성성이 잘 이루어지지 않는 국내 문헌들의 환경하에서 더욱 활용 가능성이 높은 유용한 분석 기법임이 확인되었다.

참 고 문 헌

- 김도미. 1992. 『저자동시인용 분석과 인용한 문헌의 색인어 분석에 의한 지적 구조의 규명: 경제학 분야를 대상으로』. 석사학위논문, 연세대학교 대학원.
- 김영준. 2008. 『한국 언론학의 지적 구조 변화에 관한 연구: 저자 동시인용 분석 및 저자간 교류 분석』. 박사학위논문, 중앙대학교 대학원.
- 김영진. 1981. 『논문의 동시인용을 통한 지적 구조의 규명에 관한 연구』. 석사학위논문, 연세대학교 대학원.
- _____. 1986. 논문의 동시인용을 통한 지적 구조의 규명에 관한 연구. 『정보관리학회지』, 3(1): 103-135.
- 김관준, 이재운. 2007. 연구영역분석을 위한 디스크립터 프로파일링에 관한 연구. 『정보관리학회지』, 24(4): 285-303.
- 김희진. 2010. 『저자동시인용분석과 저자서지결합분석에 의한 지적구조분석: 사회복지학 분야를 중심으로』. 박사학위논문, 경기대학교 대학원.
- 노형진. 2010. 『Excel 및 SPSS를 활용한 다변량분석 원리와 실전』. 서울: 한울출판사.
- 민윤경. 1992. 『저자동시인용분석에 의한 경영학 분야의 지적구조 규명에 관한 연구』. 석사학위논문, 이화여자대학교 대학원.
- 서은경. 1992. 정보검색분야의 지적 구조와 변화에 관한 연구: 영어문화권 저자들을 중심으로. 『정보관리학회지』, 9(1): 55-82.
- 원태연, 정성원. 1998. 『한글 SPSS 통계조사분석』. 서울: SPSS아카데미.
- 이명옥. 1996. 저자동시인용분석을 이용한 여성연구의 특성 및 동향. 『한국여성학』, 12(1): 180-203.
- 이은숙. 2003. 『복수저자를 고려한 저자동시인용 분석 연구: 정보학과 컴퓨터과학을 대상으로』. 석사학위논문, 연세대학교 대학원.
- 이재운. 2006. 지적구조의 규명을 위한 네트워크 형성 방식에 관한 연구. 『한국문헌정보학회지』, 40(2): 333-355.
- _____. 2007a. 지적구조 분석을 위한 MDS 지도 작성 방식의 비교 분석. 『한국문헌정보학회지』, 41(2): 335-357.
- _____. 2007b. 국내 광역 과학지도 생성 연구. 『정보관리학회지』, 24(3): 363-383.
- _____. 2008. 서지적 저자결합분석: 연구동향 분석을 위한 새로운 접근. 『정보관리학회지』, 25(1): 173-190.
- 이재운, 김희진, 유종덕. 2010. 저자 프로파일링과 요인분석을 이용한 국내 주거학 분야의 지적구조분석. 『한국문헌정보학회지』, 44(2): 286-308.
- _____. 2009. 텍스트마이닝을 이용한 건축학 분야의 지적 구조 분석. 『경기대학교 대학원논문집』, 39: 1-21.
- 정만영. 2008. 건축학 분류의 헤더토포아. 『건축』, 08(08): 72-73.
- 정영미, 이재운. 2001. 지식 분류의 자동화를 위한 클러스터링 모형 연구. 『정보관리학회지』, 18(2): 203-239.
- Chen, C. 1999. "Visualizing semantic spaces and author co-citation networks in dig-

- ital libraries." *Information Processing and Management*, 35(3): 401-420.
- Culnan, M. J. 1986. "The intellectual development of management information systems, 1972-1982: A co-citation analysis." *Management Science*, 32(2): 156-172.
- Gmür, M. 2003. "Co-citation analysis and the search for invisible colleges: A methodological evaluation." *Scientometrics*, 57(1): 27-57.
- Kim, H. and J. Y. Lee. 2009. "Archiving research trends in LIS domain using profiling analysis." *Scientometrics*, 80(1): 75-90.
- Kim, Heejung and Jae Yun Lee. 2008. "Exploring the emerging intellectual structure of archival studies using text mining: 2001-2004." *Journal of Information Science*, 34(3): 356-369.
- Marshakova, I. V. 1973. "System of document connections based on references." *Nauchno-Tekhnicheskaya Informatsiya Series*, 6: 3-8.
- McCain, K. W. 1983. "The author cocitation structure of macroeconomics." *Scientometrics*, 5(5): 277-289.
- _____. 1989. "Mapping authors in intellectual space: Population genetics in the 1980s." *Communication Research*, 16(5): 667-681.
- Small, H. G. 1973. "Co-citation in the scientific literature: A new measure of the relationship between two documents." *Journal of the American Society for Information Science*, 24(4): 265-269.
- White, H. D. 1983. "A cocitation map of the social indicators movement." *Journal of the American Society for Information Science*, 34(5): 307-312.
- _____. 1990. Author co-citation analysis: overview and defense. In Christine L. B., ed. *Scholarly Communication and Bibliometrics*. Newbury Park, CA: Sage Publication, 84-106.
- White, H. D. and B. C. Griffith. 1981. "Author cocitation: a literature measure of intellectual structure." *Journal of the American Society for Information Science*, 32(3): 163-171.
- White, H. D. and K. W. McCain. 1989. "Bibliometrics." *Annual Review of Information Science and Technology*, 24: 119-186.
- _____. 1998. "Visualizing a discipline: an author co-citation analysis of information science, 1972-1995." *Journal of the American Society for Information Science*, 49(4): 327-355.
- Wouter de N, M. Andrej and B. Vladimir. 2005. *Exploratory Social Network Analysis with Pajek*. Cambridge University Press.
- Zhao, D. and A. Strotmann. 2008. "Comparing all-author and first-author co-citation analysis of information science." *Journal of Informetrics*, 2(3): 229-239.