

공공도서관의 주제별 자료 이용 현황 분석: 서울특별시 A구 산하 공공도서관을 중심으로*

A Study on Collection Use of an Public Libraries Focused of the Clustering Analysis of Circulation Statistics of the Seoul Borough A Library Users

김완중 (Wan-Jong Kim)**

초 록

본 연구는 서울특별시 A구 산하 9개 도서관 이용자들의 대출 기록을 모두 조사하여 주제별 장서이용 행태를 분석하고 주제 분야 간 어떠한 연관성을 지니고 있는가를 알아보는데 그 목적이 있다. 이를 위하여 2006년 6월부터 2014년 6월까지 발생한 2,723,115건의 대출 기록을 수집하여 분석하였다. 한국십진분류(KDC)에 따라 총류(000)부터 역사(900)까지 10개 대분류와 총류(000)부터 전기(990)까지 100개 중분류로 나누어 주제별 대출빈도를 분석하였으며, 도서관 이용자들의 대출기록 기록을 중분류에 따른 상관계수에 근거한 군집분석을 통하여 주제별 연관성을 분석하였다.

ABSTRACT

The goal of this study is to analyze use patterns of library collections using circulation statistics of 9 public libraries user's of the Seoul borough "A". For this study, the 2,723,115 circulation-related data of 9 public libraries located in borough "A" which were occurred between June 2006 and June 2014 were collected and used. According to the Korea Decimal Classification (KDC), All circulation records is divided into 10 categories from general (000) to history (900) and 100 divisions from general (000) to biography (990), is analyzed the frequency by category and is analyzed by cluster analysis based on thematic relevance.

키워드: 공공도서관, 대출기록, 계량정보분석, 군집분석
public library, circulation records, bibliometric analytics, clustering analysis

* 이 논문은 2014년도 문화체육관광부 도서관정책기획단의 재원을 지원받아 수행된 연구임.

** 한국과학기술정보연구원 NDSL서비스실 선임연구원(wjkim@kisti.re.kr)

■ 논문접수일자: 2014년 8월 26일 ■ 최초심사일자: 2014년 8월 31일 ■ 게재확정일자: 2014년 9월 17일
■ 정보관리학회지, 31(3), 353-369, 2014. [http://dx.doi.org/10.3743/KOSIM.2014.31.3.353]

1. 서론

우리나라의 거의 모든 공공도서관은 지역 주민이 지식을 습득하고, 개인 혹은 단체가 평생 학습을 할 수 있는 장소이자 문화생활을 누리 기 위한 역할 등을 수행하고 있다. 이에 각 지역의 공공도서관은 지역주민들이 공공도서관에 바라는 다양한 요구를 수용하기 위하여 장서 열람 및 대출기록 등의 분석을 통한 수서정책 과 주민들의 평균 학력, 평균 소득, 가구 국성 등을 비롯한 여러 가지 사회 환경적 특징에 부 합한 도서관의 신규 서비스 개발에 많은 노력 과 시간을 투자해야 한다. 하지만 대부분의 공 공도서관은 「도서관 법 시행령」(시행 2010.3.19) 에서 정한 바대로 사서직원을 배치한 경우가 많 지 않고, 법에서 정한 수만큼의 인력이 배치되어 있다 할지라도 데이터 분석 능력을 겸비한 사서 직원이 많지 않아 이러한 업무를 수행하기가 사 실상 어려운 실정이다. 따라서 이러한 종류의 데 이터를 분석하여 도서관 정책에 활용한 사례가 많지 않다.

대부분의 도서관 사서들은 경험적으로 자관 에서 소장 중인 장서가 대출되는 것에 대한 몇 몇 특징적인 사항들을 알고 있다. 예를 들면, 자 관에서 가장 많이 대출되고 있는 주제가 무엇인 지, 연령대나 성별 등의 여러 가지 특성에 따른 주제별 대출 비율에 어떠한 차이가 있는지 등에 대한 대략적인 정보를 갖게 된다. 하지만 공공 도서관에서 장서개발, 기존 서비스 변경, 신규 서비스 개발 등 도서관 경영에 있어 매우 중요 한 의사를 결정하고 시행착오를 최소화하기 위 해서는 관련 데이터에 기반한 정확한 분석이 필 요하다.

이에 본 연구는 서울특별시 A구 산하 2개 공 공도서관과 7개 작은도서관 이용자들의 대출기 록을 모두 조사하여, 주제별 장서이용 행태를 분 석하고자 한다. 이를 위하여 2006년 6월부터 2014 년 6월까지 9개 도서관에서 발생한 2,723,115건 의 대출기록을 수집하여 분석하였다. 대출빈도 분석은 도서관별 대출빈도 분석과 성별 대출빈 도를 분석하였다. 또한 9개 도서관에서 소장 중인 장서의 주제를 한국십진분류(KDC)에 따 라 총류(000)부터 역사(900)까지 10개 대분류 와 총류(000)부터 전기(990)까지 100개 중분 류로 나누어 주제별 대출빈도를 분석하였다. 또한 10개 대분류와 100개 중분류에 따라 주제 별 상관계수에 근거한 군집분석을 통해 각 주 제들이 서로 어떠한 연관성을 지니고 있는지 분석하였다.

본 연구에서 사용한 용어 및 연구의 제한점 은 다음과 같다. 첫째, 본 연구에서는 A구 산하 9개 도서관은 모두 공공도서관이나 도서관의 규모에 따른 개념적 구분을 위하여 일반 공공 도서관은 구립도서관으로 작은도서관은 작은 도서관으로 표현하였다. 둘째, 9개 도서관에서 소장 중인 다양한 형태의 매체에 대한 대출기 록을 분석함으로써 발생할 수 있는 분석 결과 의 편향이나 오류를 최소화 하고 분석의 정확 도를 높이기 위하여, 대출 가능한 인쇄형 단행 본에 대한 대출기록만을 분석하였다. 셋째, 9개 도서관의 통합 회원정보 가운데 출생연도를 기 준으로 한 연령대별 분석과 거주지 우편번호를 통한 지역별 분석도 수행하고자 관련 데이터를 수집하였으나 이 항목들에 대한 회원들의 정보 보유율이 낮아 이 두 가지 분석은 수행하지 않 았다.

2. 이론적 연구

도서관의 대출기록을 이용한 장서이용 및 장서개발과 관련된 선행연구를 살펴보면 다음과 같다. 먼저 Trueswell(1969)은 계량정보분석 방법을 이용하여 전체 원인의 20%가 전체 결과의 80%에 영향을 미친다는 현상, 즉 80:20 법칙이라고 칭하는 파레토 법칙(Paleto principle)이 도서관 대출기록에서도 발견된다는 것을 경험적으로 밝혀냈다. Trueswell은 전체 장서의 20%만으로 전체 대출의 80%가 발생하고 있다는 사실과 전체 장서의 50%만으로 전체 대출의 90%가 발생한다는 사실을 발견했다.

Trueswell(1969)의 연구의 기초가 된 파레토 법칙(Paleto principle)은 파레토 분포(Pareto distribution)라고도 칭하며, 일반적으로 자연과학 분야보다는 사회과학 분야에서 자주 확인되는 확률분포이다. Pareto(1964)는 '이탈리아 인구의 20%가 이탈리아 전체 부의 80%를 가지고 있다고 주장하면서 소수의 원인이 다수의 결과에 영향을 미친다는 사회적 불평등에 대한 논의를 확률적으로 이끌어내었다.

유경종과 박일종(2009)은 도시지역과 군지역에 위치한 공공도서관 2곳씩을 선정하여 자료이용 특성에 관한 비교연구를 수행하였다. 장서유형을 일반도서와 아동도서로 구분하여 대출책수와 대출빈도에 대한 이용지수를 비교한 결과 도시지역 공공도서관의 이용지수가 군지역에 위치한 공공도서관 보다 편차가 큰 것으로 나타나 도시지역에 위치한 공공도서관이 자료구입에 대한 수서정책을 군지역에 위치한 공공도서관 보다 더 적극적으로 재검토 할 필요성을 제기하였다. 또한 지역적 위치에 따른 이

용자들의 주제별 이용지수를 비교한 결과 지리적 위치보다는 도서관 특성에 맞는 주제별 수서정책을 추진할 필요성도 함께 제시하였다.

김선애와 서혜란(2009)은 연구는 공공도서관의 소장자료를 대상으로 공공도서관 장서의 주제별, 연령별 소장현황을 분석하였다. 분석결과, 공공도서관 장서는 주제별 불균형이 심각한 것으로 나타났으며, 특히 문학과 사회과학의 비중이 전체 자료의 54.8%를 차지하는 것으로 나타났다.

Ochola(2009)는 장서개발을 함에 있어 대출기록과 상호대차 기록을 이용하는 것이 이용자의 요구와 관심도를 반영할 수 있는 유용한 방법인 것을 밝혀냈으며, 대출과 상호대차가 많이 발생하지 않는 주제 분야에서는 적극적인 검토가 보다 더 많이 필요하다고 주장하였다.

박영애와 이재윤(2010)은 일반 공공도서관에 비해 장서의 양과 다양성이 부족한 작은도서관을 대상으로 대출과 상호대차 기록을 분석하였다. 그 결과, 문학 및 교양 분야에 대한 이용률보다는 실용, 기능 등 전문도서의 이용률이 높은 것으로 나타났다. 다만, 아직까지 상호대차가 활발하지 않은 상당수의 작은도서관에 대해서는 제안된 장서평가 방법을 그대로 적용하기 어렵다는 제한점도 함께 제시하였다.

이진경과 이지연(2011)은 개인의 배경과 방문 목적이 다양한 공공도서관 이용자들로 인해 모든 이용자의 요구를 충분히 만족시키기 어렵다는 것에 주목하였다. 이진경과 이지연(2011)은 이를 극복하기 위한 노력으로 이용자들의 장서이용행태를 파악하고 이를 바탕으로 서울시구립 공공도서관 자료의 만족 수준을 높일 수 있는 장서개발 방향을 제시하였다.

김선애(2013)는 A대학도서관의 자료의 주제별 그리고 이용자 유형에 따른 자료이용행태 분석을 위해 16개군으로, 이용자 유형은 5개 집단(학부생, 대학원생, 교수, 직원, 기타이용자)으로 구분하였다. 이를 통해 대학도서관이 당면한 매우 시급하고도 중요한 문제인 소장공간 부족 및 이용자의 열람공간 축소 등으로 나타나는 공간과밀화 문제를 해결할 수 있는 방법을 장서 이용평가에 기초하여 수립되어야 한다고 주장하였다.

이처럼 공공도서관에서 소장 중인 장서와 대출기록을 분석한 연구는 많지 않으며, 개관 이후의 공공도서관의 장서와 모든 대출기록을 분석한 사례는 전무한 실정이다. 공공도서관의 대출기록에 대한 분석은 기존의 소장자료에 대한 주제별 장서 평가와 동시에 주제별 유사도를 측정함으로써 공공도서관에서 수행해야 할 다양한 의사결정을 지원하기 위한 근거 데이터로 활용될 수 있을 것이다.

3. 연구방법

본 연구는 2006년 6월부터 2014년 6월까지 서울특별시 A구 산하 9개 도서관에서 발생한 2,723,115건의 대출기록을 수집하여 분석하였다. 대출빈도 분석은 도서관별 대출빈도 분석과 성별 대출빈도를 분석하였다. 또한 9개 도서관에서 소장 중인 장서의 주제를 한국십진분류(KDC)에 따라 총류(000)부터 역사(900)까지 10개 대분류와 총류(000)부터 전기(990)까지 100개 중분류로 나누어 주제별 대출빈도를 분석하였다. 또한 10개 대분류와 100개 중분류에

따라 주제별 상관계수에 근거한 군집분석을 통해 어떤 주제별로 어떠한 연관성을 지니고 있는지 분석하였다. 통계 분석과 네트워크 분석을 위해 PASW 18과 NodeXL을 사용하였다. 단계별 연구방법 및 절차는 <표 1>과 같다.

<표 1> 단계별 연구방법 및 절차

단계	연구 방법 및 절차
1	데이터 수집
2	데이터 전처리
3	주제분야별 동시대출빈도 산출
4	9개 도서관별 대출빈도 분석
5	성별 대출빈도 분석
6	대주제별 분석 - 대주제별 대출빈도 분석 - 대주제별 상관계수 산출(동시대출빈도 이용) - NodeXL을 이용한 군집분석 및 시각화 - 각 군집별 의미 분석
7	중주제별 분석 - 중주제별 대출빈도 분석 - 중주제별 상관계수 산출(동시대출빈도 이용) - NodeXL을 이용한 군집분석 및 시각화 - 각 군집별 의미 분석

본 연구에서 주요하게 사용된 용어에 대한 조작적 정의는 다음과 같다. 첫째, 대출빈도란 특정 장서가 대출된 횟수를 칭한다. 이에 따라 도서관별 대출빈도, 성별대출빈도, 주제별 대출빈도 등을 구하였다. 둘째, 평균 대출빈도는 특정 구분에 따른 평균 대출빈도를 의미한다. 예를 들면, 성별 평균대출빈도에서 남성의 평균 대출빈도는 남성 전체 대출빈도를 남성회원수로 나눈 값을 의미하며, 장서 한편당 평균 대출빈도수는 전체 대출빈도를 장서수로 나눈 값이다. 또한 동시대출이란 도서관 이용자 한 명이 서로 다른 주제 분야의 장서를 함께 대출하는 것을 의미하며, 동시대출빈도란 한 명의 이용

자가 서로 다른 주제 분야의 장서를 함께 대출한 횟수를 의미한다.

4. 데이터 수집 및 전처리

데이터 분석을 위해 수집한 정보는 다음과 같다. 첫째, 장서관련 데이터는 장서등록번호, 장서등록일, 서명, ISBN, 청구기호, 소장 도서관명이다. 둘째, 이용자 관련 데이터는 대출자번호, 성별, 출생연도, 거주지 우편번호이다. 대출자번호는 「개인정보보호법」, 「정보통신망법」에서 정한 바대로 특정 개인을 식별할 수 없도록 하기 위하여 임의의 코드값으로 변환된 값을 수집하였다. 다만 각 회원별 출생연도와 거주지 우편번호의 경우, 정보 보유 비율이 낮아 실제 분석에서는 사용하지 않았다. 셋째, 대출관련 데이터는 대출고유번호, 대출일, 반납일을 수집하였다.

데이터의 정확한 분석을 위하여 수행한 전처리 작업은 다음과 같다. 첫째, 대출과 반납일이 동일한 기록은 전체 대출기록의 약 5% 가량이었으나, 대출 착오 등의 문제를 가지고 있을 가능성이 존재하여 이를 분석에서 제외하였다. 둘째, 9개 도서관이 통합 서지를 사용하지 않고 도서관마다 개별적으로 서지데이터를 보유, 관리하고 있어 분석에 필요한 서명에 대한 표준화 작업을 수행하였다. 예를 들면, 단행본의 경우 서명에 포함되어 있는 특수문자의 표기나 띄어쓰기 등이 각 도서관마다 상이하거나, 서명 뒤에 특별한 의미 없이 존재하는 마침표(.), 슬래쉬(/), 콜론(:)과 같은 기호를 삭제하여 서명을 표준화하였다. 셋째, 주제 분석을 위해

각 도서관에서 사용하고 있는 KDC 분류 번호 앞에 존재하는 별치기호를 삭제하였다. 넷째, 9개 도서관에서 사용하고 있는 주제분류기호 가운데 소숫점 이하는 제외하고 000부터 999까지 존재하는 1,000개의 주제 분류기호를 총류(000)부터 역사(900)까지 10개 대분류와 총류(000)부터 전기(990)까지 100개 중분류로 나누어 변경하였다(〈표 2〉 참조).

〈표 2〉 단계별 데이터 전처리 과정

단계	데이터 전처리
1	대출일과 반납일이 동일한 대출기록 삭제
2	서명 표준화 실시 - 서명에 포함되어 있는 특수문자의 표기나 띄어쓰기 등을 표준화 - 서명 뒤에 존재하는 마침표(.), 슬래쉬(/), 콜론(:)과 같은 기호 삭제
3	KDC 분류 번호 앞에 존재하는 별치기호 삭제
4	KDC 분류 번호를 10개 대분류와 100개 중분류로 변경 - 대분류 변경 예) 019.35 → 000 - 중분류 변경 예) 019.35 → 010

5. 데이터 분석 결과

서울특별시 A구 산하 9개 도서관의 데이터 분석은 앞에서 언급한 데이터 전처리 과정을 거쳐 도출된 9개 도서관의 현황, 각 도서관별 대출빈도, 성별 대출빈도를 살펴보았다. 주제별 대출빈도 분석과 군집분석은 대주제와 중주제로 나누어 각각 실시하였다.

5.1 A구 산하 공공도서관 현황

서울특별시 A구는 2006년 6월 최초의 구립

〈표 3〉 A구 산하 공공도서관의 장서 및 이용자 현황

구분	현황	비고
서지수	140,681건	ISBN 정보누락 11,343건 포함
장서수	219,686권	서지 한 건당 평균 1.56권의 복본 보유
회원수	83,759명	여성 31,952명, 남성 19,836명, 정보누락 3,676명

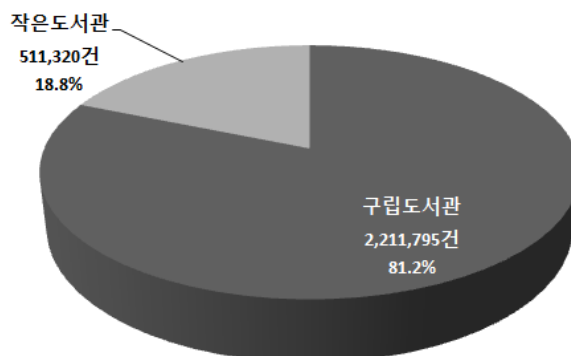
공공도서관이 설립된 이래 2014년 7월을 기준으로 A구청과 A구시설관리공단에서 2개의 공공도서관과 7개의 작은도서관을 운영하고 있다. A구 산하 9개 공공도서관에서 소장하고 있는 서지 레코드수는 ISBN이 존재하지 않는 11,343건을 포함하여 140,681건이었으며, 장서수는 219,686권으로 서지 레코드 한 건당 평균 1.56권의 복본을 보유하고 있다. 또한 9개 도서관에서 통합 관리하고 있는 전체 회원수는 83,759명이었다(〈표 3〉 참조).

5.2 도서관별 대출빈도 분석

서울특별시 A구 산하 9개 도서관은 2006년 6월 '가' 도서관이 가장 먼저 설립된 이후 2012년 3개, 2013년 1개, 2014년 4개 도서관이 각각

개관하였으며, 지난 8년 2개월 동안 하루 평균 929건 가량의 장서 대출이 발생하여 누적 대출빈도는 2,723,115건을 기록했다. 2개 구립도서관의 대출빈도는 2,211,795건(81.2%), 7개 작은도서관의 대출빈도는 511,320건(18.8%)이었다. 9개 도서관 가운데 22.2%에 해당하는 2곳의 구립도서관이 전체 대출빈도의 81.2%를 차지하는 영향력을 보유하고 있는 것으로 나타나 파레토의 80:20 법칙과 매우 유사한 결과를 보였다(〈그림 1〉 참조).

각 도서관별 하루 평균 대출빈도를 살펴보면, 작은도서관에 비해 규모가 큰 구립도서관인 '가' 도서관과 '나' 도서관을 합친 하루 평균 대출빈도는 730.9권이었으며, '가' 도서관과 '나' 도서관은 각각 하루 평균 743.4권과 353.4권씩을 대출하는 것으로 나타났다. '나' 도서관의 하



〈그림 1〉 구립도서관과 작은도서관의 대출빈도 및 비율

〈표 4〉 서울특별시 A구 산하 공공도서관의 대출 현황

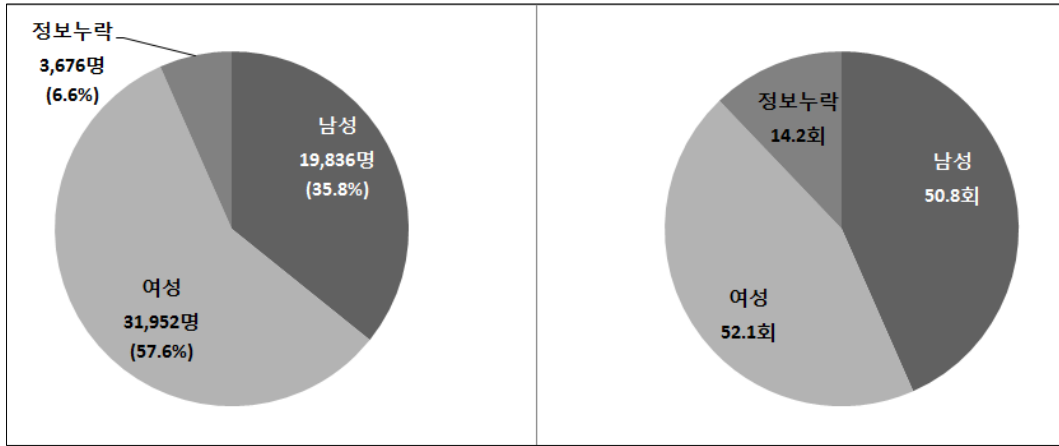
구분	최초 대출	운영일수 ¹⁾	장서수	대출빈도	평균 대출빈도		
					일 평균	장서1권당	
구립도서관	'가'	'06년 06월	2,929	123,726	2,177,519	743.4	17.6
	'나'	'14년 04월	97	23,224	34,276	353.4	1.5
	소계	—	2,929	123,726	2,211,795	730.9	15.1
작은도서관	'다'	'12년 02월	866	16,846	227,370	262.6	13.5
	'라'	'12년 03월	846	19,274	182,416	215.6	9.5
	'마'	'12년 05월	797	18,373	85,969	107.9	4.7
	'바'	'13년 12월	216	5,350	5,571	25.8	1.0
	'사'	'14년 03월	125	3,099	4,323	34.6	1.4
	'아'	'14년 03월	106	2,505	3,435	32.4	1.4
	'자'	'14년 03월	103	2,410	2,236	21.7	0.9
	소계	—	3,059	189,173	511,320	167.1	7.5
전체	'06년 06월	6,085	214,807	2,723,115	447.5	7.5	

루 평균 대출빈도는 '가' 도서관의 하루 평균 대출빈도의 절반에 미치지 못하는 47.5%에 불과했다. 이는 '나' 도서관이 개관한지 3개월 정도 지난 시점이기 때문에 지역주민들의 '나' 도서관의 개관에 대한 인지부족 등의 원인으로 인해 발생한 결과라 추측할 수 있다. 7개 작은도서관의 경우 개관 시기가 2012년 2월과 3월로 가장 빠른 '다' 도서관과 '라' 도서관의 대출비율이 작은 도서관 전체 대출빈도(511,320건)의 80.1%로 나타났으며, 기타 5개 도서관의 대출비율이 19.9%로 나타났다. 7개 작은도서관은 모두 하루 평균 300권 미만의 대출이 발생했다. 또한 A구 산하 2개 구립도서관과 7개 작은도서관 개관 시기가 빠를수록 누적대출빈도와 하루 평균 대출빈도 모두 큰 경향을 보이고 있다(〈표 4〉 참조).

5.3 성별 대출빈도 분석

앞에서도 언급한 바와 같이 A구 산하 9개 도서관은 83,759명의 회원을 보유하고 있다. 하지만 2006년 6월 '가' 도서관이 A구에서 가장 먼저 개관된 이후 약 8년 동안 9개 도서관의 2,723,115건의 장서를 대출한 전체 회원수는 약 66%인 55,464명이었으며, 여성 회원이 31,952명(57.6%)으로 19,836명(35.8%)인 남성 회원보다 12,116명(21.8%) 더 많았다. 또한 성별 정보가 존재하지 않는 회원도 3,676명(6.6%)이 존재했다. 성별 1인당 평균 대출빈도는 평균 52.1회를 대출한 여성회원이 50.8회를 기록한 남성 회원보다 0.3회 더 많이 대출한 것으로 나타나 근소한 차이를 보였다. 성별 정보가 존재하지 않는 회원도 평균 14.2회를 대출하였다(〈그림 2〉 참조).

1) 도서관 운영일수는 휴관일을 포함하였음.



〈그림 2〉 성별 대출회원수(좌) 및 1인당 평균 대출빈도(우)

5.4 대주제별 분석

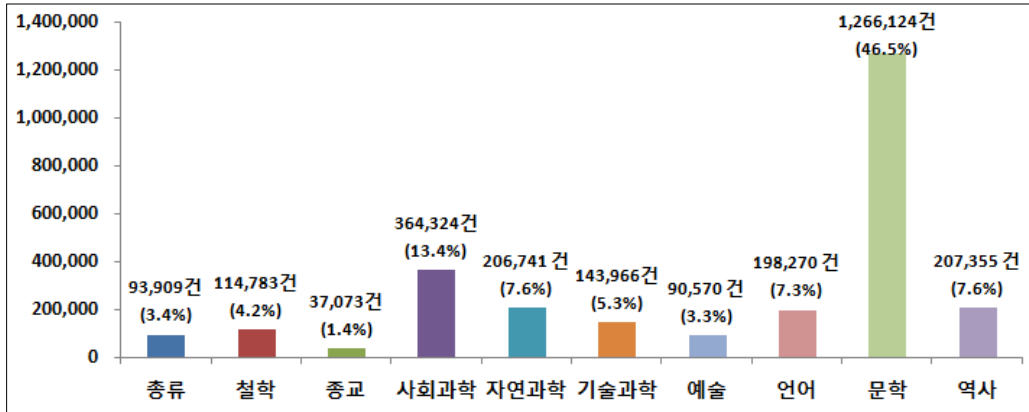
A구 산하 9개 도서관의 주제별 대출현황을 파악하기 위하여 KDC(한국십진분류) 기준에 따라 총류(000)부터 역사(900)까지 10개의 대분류와 총류(000)부터 990(전기)까지 100개 중분류로 나누어 주제별 대출빈도 분석을 실시하였다.

먼저 10개 대분류별 대출빈도를 분석한 결과, 문학(800)이 1,266,124건(46.5%)으로 가장 높은 비율을 차지했다. 다음으로 사회과학(300) 364,324건(13.4%), 역사(900) 207,355건(7.6%), 자연과학(400) 206,741건(7.6%), 언어(700) 198,270건(7.3%), 기술과학(500) 143,966건(5.3%), 철학(100) 114,783건(4.2%), 총류(000) 93,909건(3.4%), 예술(600) 90,570건(3.3%), 종교(200) 37,073건(1.4%)의 순으로 나타났다. A구 산하 9개 도서관의 전체 대출빈도의 82.4%(2,242,814건)가 문학, 사회과학, 역사, 자연과학, 언어 5개 분야에 집중되고 있는 것으로 나타났다(〈그림 3〉 참조).

림 3〉 참조).

A구 산하 9개 도서관에서 발생한 대출기록에 대해 어떠한 대주제들이 함께 대출되고 있는지를 분석하기 위하여 대주제 분야 간 네트워크 분석을 수행하였다. 먼저 PASW 18을 통해 주제별 상관계수를 산출하였으며, NodeXL을 사용하여 대주제별 군집화 및 시각화를 표현하였다. 네트워크에서의 각 노드 간의 연결은 주제별 상관계수 값을 이용하여 패스파인더 네트워크 지수를 생성하여 표현하였다. 대주제별 군집화 및 시각화 결과는 다음과 같다.

먼저 각 대주제 분야 간 상관관계를 알아보기 위하여 Pearson 상관분석을 실시한 결과, 모든 대주제 분야 간에는 양의 상관관계가 있는 것으로 나타났다. 계수값의 크기에 따라 철학과 사회과학($r=.573$), 자연과학과 문학($r=.572$), 문학과 역사($r=.536$), 자연과학과 역사($r=.533$), 기술과학과 예술($r=.524$), 사회과학과 역사($r=.514$) 등의 순서로 주제 분야 간 상관관계가 높았다(〈표 5〉 참조).



〈그림 3〉 KDC 대주제별 대출빈도

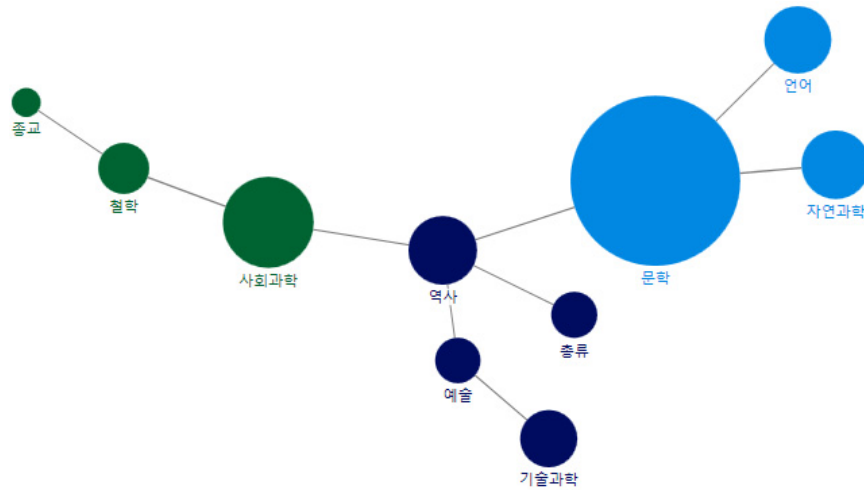
〈표 5〉 대주제별 상관계수

	총류	철학	종교	사회과학	자연과학	기술과학	예술	언어	문학	역사
총류	1	.308***	.211***	.446***	.450***	.363***	.441***	.249***	.397***	.459***
철학		1	.409***	.573***	.208***	.353***	.302***	.113***	.284***	.315***
종교			1	.343***	.182***	.223***	.210***	.140***	.209***	.275***
사회과학				1	.469***	.430***	.388***	.220***	.482***	.514***
자연과학					1	.269***	.274***	.247***	.572***	.533***
기술과학						1	.524***	.159***	.313***	.417***
예술							1	.143***	.356***	.443***
언어								1	.346***	.236***
문학									1	.536***
역사										1

*p< .05, **p< .01, ***p< .001

〈표 3〉의 상관계수 값을 기준으로 NodeXL을 이용하여 대주제별 군집을 나누어 본 결과, 3개로 나누어졌다. 제1군집은 총류, 기술과학, 예술, 역사의 4개 주제가, 제2군집은 자연과학, 언어, 문학의 3개 주제가, 제3군집은 철학, 종교, 사회과학의 3개 주제가 포함되었다. 10개의 대주제 가운데 역사가 연결정도(=4), 매개중심성(=29), 인접중심성(=0.066667)로 모든 네트워크 관련 지수가 가장 높았다. 다음으로

문학이 연결정도(=3), 매개중심성(=15), 인접중심성(=0.052632)로 나타났으며, 사회과학이 연결정도(=2), 매개중심성(=14), 인접중심성(=0.052632) 등의 순으로 나타났다. 분석 결과, 역사 분야가 다른 주제와 함께 대출이 많이 되고 있어 역사 분야를 중심으로 도서관 신규 서비스를 모색한다거나 역사 분야를 연계한 장서구성이나 서가 구성 등을 활용할 수 있을 것으로 나타났다(〈그림 4〉, 〈표 6〉 참조).



〈그림 4〉 대주제별 군집분석 결과

〈표 6〉 대주제별 네트워크 지수

구분	연결정도	매개중심성	근접중심성
총류	1	0	0.043478
철학	2	8	0.040000
종교	1	0	0.030303
사회과학	2	14	0.052632
자연과학	1	0	0.037037
기술과학	1	0	0.034483
예술	2	8	0.047619
언어	1	0	0.037037
문학	3	15	0.052632
역사	4	29	0.066667

5.5 중주제별 분석

대주제별 분석에 이어 100개 중주제에 대한 분야별 대출 빈도를 분석하였다. 한국문학(810)이 가장 많은 434,346건(16.0%)의 대출을 기록하였으며, 영미문학(840) 329,408건(12.1%), 문학(800) 194,851건(7.2%), 영어(740) 173,697건(6.4%), 자연과학(400) 123,316건(4.5%), 일본문학 및 기타 아시아 제 문학(830) 117,747건

(4.3%), 경제학(320) 117,558건(4.3%), 교육학(370) 95,369건(3.5%), 생활과학(590) 82,361건(3.0%), 아시아 역사(910건) 73,532건(2.7%) 등의 순서로 대출을 많이 하고 있는 것으로 나타났다. 또한 한국문학(810)을 비롯한 상위 20개의 주제 분야의 대출빈도는 전체 대출빈도의 81.2% (2,210,406건)를 차지하는 것으로 나타나 파레토의 80:20 법칙과 거의 유사한 결과를 보였다. 상위 20개 중주제 분야 가운데 문학 분야가 7개

포함되어 있었다. 이들의 대출빈도는 1,232,938건으로 문학 분야 전체 대출기록(1,266,124건) 가운데 97.4%를 기록했으며, A구 산하 공공도서관 전체 대출기록(2,723,115건) 가운데 45.3%를 차지하는 것으로 나타났다(〈표 7〉 참조).

A구 산하 9개 도서관에서 발생한 대출기록에 대해 어떠한 중주제들이 함께 대출되고 있는지를 분석하기 위하여 중주제 분야 간 네트워크 분석을 수행하였다. 분석 방법은 대주제별 분석과정과 동일한 방법을 사용하였다. 중주제별 군집화 및 시각화 결과는 다음과 같다.

중주제별 군집은 총 10개로 형성되었다. 먼저 제1군집은 10개의 군집 가운데 가장 많은 15

개의 주제 분야를 포함하고 있다. 제1군집은 경제학(320) 117,558건, 사회학, 사회문제(330) 36,818건, 기독교(230) 20,020건, 정치학(340) 16,600건, 사회과학(300) 16,129건 등의 순서로 대출이 많이 되었다. 제1군집에는 사회과학 분야 7개, 철학과 종교 분야가 각 3개씩 포함되어 있었으며, 기타 주제로는 공연예술 및 매체 예술(680)과 신문, 저널리즘(070)도 포함되어 있었다. 제1군집은 정치, 경제, 종교, 철학, 예술 분야가 다양하게 존재하고 있으며, 철학과 종교 분야의 경우 서양 철학과 기독교가 같은 군집으로 포함되어 있는 특징을 보이고 있다.

제2군집은 13개 분야를 포함하고 있다. 제2군

〈표 7〉 KDC 100개 중주제별 대출빈도

대출순위	주제	대출		누적 대출	
		빈도	비율	빈도	비율
1	한국문학(810)	434,346	16.0%	434,346	16.0%
2	영미문학(840)	329,408	12.1%	763,754	28.0%
3	문학(800)	194,851	7.2%	958,605	35.2%
4	영어(740)	173,697	6.4%	1,132,302	41.6%
5	자연과학(400)	123,316	4.5%	1,255,618	46.1%
6	일본문학 및 기타 아시아 제 문학(830)	117,747	4.3%	1,373,365	50.4%
7	경제학(320)	117,558	4.3%	1,490,923	54.8%
8	교육학(370)	95,369	3.5%	1,586,292	58.3%
9	생활과학(590)	82,361	3.0%	1,668,653	61.3%
10	아시아 역사(910)	73,532	2.7%	1,742,185	64.0%
11	풍속, 예절, 민속학(380)	70,966	2.6%	1,813,151	66.6%
12	프랑스문학(860)	59,829	2.2%	1,872,980	68.8%
13	독일문학(850)	55,464	2.0%	1,928,444	70.8%
14	지리(980)	49,593	1.8%	1,978,037	72.6%
15	심리학(180)	43,481	1.6%	2,021,518	74.2%
16	중국문학(820)	41,338	1.5%	2,062,856	75.8%
17	윤리학, 도덕철학(190)	40,359	1.5%	2,103,215	77.2%
18	사회학, 사회문제(330)	36,818	1.4%	2,140,033	78.6%
19	전기(990)	35,216	1.3%	2,175,249	79.9%
20	의학(510)	35,157	1.3%	2,210,406	81.2%
—	기타	512,709	18.8%	2,723,115	100%
합계		2,723,115	100%	2,723,115	100%

집은 자연과학(400) 123,316건, 아시아(아세아) 역사(910) 73,532, 전기(990) 35,216건) 역사(900) 31,726건, 수학(410) 29,427건 등의 순서로 대출빈도가 많았다. 제2군집은 자연과학(400)을 포함한 자연과학 분야가 5개 분야로 가장 많았고, 다음으로 역사 4개, 기술과학 2개, 어학과 종교가 각 1개씩 포함되어 있다. 이 13개 주제 분야 가운데 자연과학과 기술과학 분야가 7개 주제로 절반 이상을 차지했으며, 역사 분야가 함께 포함되어 있는 것이 특징적이라 할 수 있다.

제3군집은 1,273,213건의 대출을 기록하여 10개 군집 가운데 가장 많은 대출이 발생하였다. 12개 주제 분야로 구성된 제3군집은 단일 중주제 분야로는 가장 많은 434,346건의 대출빈도를 기록한 한국문학(810) 이외에 영미문학(840) 329,408건, 영어(740) 173,697건, 일본문학 및 기타 아시아 제 문학(830) 117,747건 등의 순으로 나타났다. 이 제3군집은 문학 분야 9개, 영어(740)와 중국어(720)가 포함된 언어 분야 2개, 그리고 일반전집, 총서(080)가 포함된 총류가 1개 포함되어 있다. 영미문학(840)과 영어(740), 중국문학(820)과 중국어(720)가 같은 군집에 포함되어 있는 특징을 보이고 있다.

세 번째 군집과 동일한 수인 12개 주제 분야로 구성된 제4군집은 심리학(180) 43,481건, 윤리학, 도덕철학(190) 40,359건, 불교(220) 6,501건, 동양 철학, 사상(150) 5,261건 등의 순으로 대출빈도가 나타났다. 제4군집은 심리학(180)을 포함한 철학 분야와 불교(220)를 포함한 종교 분야가 각 6개 분야씩 포함되어 있는데, 주로 동양권의 철학 분야와 종교 분야가 포함되어 있는 특징을 보였다.

11개 주제 분야를 포함하고 있는 제5군집의

대출빈도는 총류(000) 29,397건, 유럽(920) 6,953건, 천문학(440) 6,299건, 일본어 및 기타 아시아 제어(730) 5,042건 등의 순서로 나타났다. 제5군집은 자연과학 분야 3개, 역사 분야 3개, 기술과학 분야 2개, 그리고 총류(000), 국방, 군사학(390), 일본어 및 기타 아시아 제어(730)가 각 하나씩 포함되어 있었다.

다섯 번째 주제와 동일한 수인 11개 주제 분야를 포함하고 있는 제6군집은 6개의 예술 분야가 포함되어 있었으며, 기술과학 분야 2개, 그리고 강연집, 수필집, 연설문집(040), 광물학(460), 언어(700)가 각 하나씩 포함되어 있다. 제6군집의 전체적인 특징으로는 예술 분야의 주제 이외에도 이와 관련된 다른 주제 분야들이 함께 포함되어 있다는 것이다.

9개의 주제 분야가 포함되어 있는 제7군집은 생활과학(590) 82,361건, 지리(980) 49,593건, 의학(510) 35,157건, 공예, 장식미술(630) 10,428 등의 순서로 대출빈도를 보였다. 제7군집은 4개의 기술과학 분야와 2개의 역사 분야, 그리고 식물학(480), 공예, 장식미술(630), 독일어(750) 등이 각각 포함되어 있다.

제8군집은 7개의 주제 분야로 구성되어 있는데, 이들의 대출빈도를 살펴보면 문학(800) 194,851건, 교육학(370) 95,369건, 풍속, 예절, 민속학(380) 70,966건 등의 순서로 나타났다. 제8군집의 경우 총류 2개, 사회과학 2개, 철학, 문학, 역사 분야가 각 1개씩 포함되어 있었다. 제9군집은 6개의 주제 분야로 구성되어 있으며, 음악(670) 11,221건과 오락, 스포츠(690) 7,738건의 대출빈도를 가진 예술 분야가 2개 포함되어 있었으며, 나머지 4개 주제 분야는 프랑스어를 비롯한 유럽 언어들이 포함되어 있는 것을 확인할 수 있다.

〈표 8〉 주제별 군집 결과

구분	주제분야수	주제 분야	대출빈도
군집1	15	신문, 저널리즘(070), 철학의 체계(130), 서양철학(160), 논리학(170), 종교(200), 기독교(230), 신도(260), 사회과학(300), 통계학(310), 경제학(320), 사회학, 사회문제(330), 정치학(340), 행정학(350), 법학(360), 공연예술 및 매체예술(680)	233,499
군집2	13	비교종교(210), 자연과학(400), 수학(410), 지학(450), 생명과학(470), 동물학(490), 기술과학(500), 공학, 공업일반, 토목공학, 환경공학(530), 한국어(710), 역사(900), 아시아(아세아) 역사(910), 아프리카(930), 전기(990)	363,334
군집3	12	일반전집, 총서(080), 중국어(720), 영어(740), 한국문학(810), 중국문학(820), 일본문학 및 기타 아시아 제 문학(830), 영미문학(840), 독일문학(850), 프랑스문학(860), 스페인 및 포르투갈 문학(870), 이탈리아문학(880), 기타 제문학(890)	1,273,213
군집4	12	형이상학(110), 인식론, 인과론, 인간학(120), 경학(140), 동양 철학, 사상(150), 심리학(180), 윤리학, 도덕철학(190), 불교(220), 도교(240), 천도교(250), 힌두교, 브라만교(바라문교)(270), 이슬람교(회교)(280), 기타 제종교(290)	100,285
군집5	11	총류(000), 국방, 군사학(390), 물리학(420), 화학(430), 천문학(440), 기계공학(550), 전기공학, 전자공학(560), 일본어 및 기타 아시아 제어(730), 유럽(920), 북아메리카(940), 남아메리카(남미)(950)	62,835
군집6	11	강연집, 수필집, 연설문집(040), 광물학(460), 건축공학(540), 제조업(580), 예술(600), 건축술(610), 조각 및 조형미술(620), 서예(640), 회화, 도화(650), 사진예술(660), 언어(700)	56,763
군집7	9	식물학(480), 의학(510), 농업, 농학(520), 화학공학(570), 생활과학(590), 공예, 장식미술(630), 독일어(750), 오세아니아(960), 지리(980)	199,467
군집8	7	백과사전(030), 일반 연속간행물(050), 철학(100), 교육학(370), 풍속, 예절, 민속학(380), 문학(800), 양극지방(970)	397,726
군집9	6	음악(670), 오락, 스포츠(690), 프랑스어(760), 스페인어 및 포르투갈어(770), 이탈리아어(780), 기타 제어(790)	20,873
군집10	4	도서학, 서지학(010), 문헌정보학(020), 일반 학회, 단체, 협회, 기관(060), 향토자료(090)	15,120

〈표 9〉 중주제별 네트워크 지수

구분	연결정도	매개중심성	근접중심성
사회학, 사회문제(330)	10	3,538	0.002208
독일문학(850)	6	2,552	0.002045
한국문학(810)	4	2,513	0.002141
문학(800)	5	2,154	0.001898
풍속, 예절, 민속학(380)	4	1,947	0.001745
생활과학(590)	6	1,937	0.001815
자연과학(400)	7	1,929	0.001600
의학(510)	2	1,850	0.001996
건축술(610)	4	1,357	0.001621
예술(600)	4	1,172	0.001451

6. 결론 및 제언

본 연구는 서울특별시 A구 산하 9개 도서관 이용자의 대출기록을 모두 조사하여 주제별 장서이용 행태를 분석하고 주제분야간 어떠한 연관성을 지니고 있는가를 알아보는 것을 그 목적으로 하였다. 이를 위하여 A구 산하 9개 도서관에서 2006년 6월부터 2014년 6월까지 발생한 2,723,115건의 대출기록을 수집하여 분석하였다. 한국십진분류(KDC)에 따라 총류(000)부터 역사(900)까지 10개 대분류와 총류(000)부터 전기(990)까지 100개 중분류로 나누어 주제별 대출빈도를 분석하였으며, 도서관 이용자의 대출기록을 중분류에 따른 상관관수에 근거한 군집분석을 통하여 주제별 연관성을 분석하였다.

연구 결과를 요약하면 다음과 같다.

첫째, A구 산하 9개 공공도서관에서 소장하고 있는 서지 레코드수는 140,681건, 장서수는 219,686권으로 서지 레코드 한 건당 평균 1.56권의 복본 장서를 보유하고 있었으며, 9개 도서관을 이용하는 전체 전체 회원수는 83,759명이었다.

둘째, 9개 도서관의 누적 대출빈도는 2,723,115건이었으며, 2개 구립도서관의 대출빈도는 2,211,795건(81.2%)으로 7개 작은도서관의 대출빈도는 511,320건(18.8%)의 4배가 넘어 2개 구립도서관의 영향력이 매우 큰 것으로 나타났다.

셋째, 9개 도서관에 등록되어 있는 83,759명의 회원 가운데 66.2%인 55,464명의 회원이 1회 이상 대출한 것으로 나타났으며, 여성 회원수(31,952명)가 남성(19,836명)보다 12,116명(21.8% 포인트) 더 많았다. 성별 대출빈도는 1,663,107건

(61.1%)을 대출한 여성회원이 1,007,713건(37.0%)을 기록한 남성 회원보다 655,394건(24.1% 포인트) 더 많이 대출한 것으로 나타났다. 또한 성별에 따른 주제별 선호도 역시 통계적으로 유의한 차이를 보였다.

넷째, 10개 대주제별 대출빈도를 분석한 결과, 문학, 사회과학, 역사, 자연과학, 언어 등의 순서로 대출이 많이 되는 것으로 나타났다. 이상위 5개 분야의 대출빈도가 전체 대출빈도의 82.4%를 차지할 정도로 특정 분야의 대출이 집중적으로 발생했다. 10개 대주제는 3개 군집으로 나뉘어졌으며, 사회과학, 자연과학, 기술과학 분야 모두 서로 다른 군집으로 형성되었다. 역사 분야는 10개 대주제 가운데 연결정도(=4), 매개중심성(=29), 인접중심성(=0.066667)이 모두 가장 높았다.

다섯째, 중주제에 대한 분야별 대출 빈도를 분석한 결과, 한국문학, 영미문학, 문학, 영어, 자연과학 등의 순서로 많은 대출이 발생했다. 대출 빈도 상위 20개 분야가 전체 대출빈도의 81.2%를 차지하는 것으로 나타나 중주제 분류에서도 특정 분야의 대출이 집중적으로 발생하는 것을 확인할 수 있었다. 100개의 중주제는 총 10개의 군집을 형성하였으며, 대출빈도 상위 1, 2위인 한국문학과 영국문학이 포함된 제3군집의 대출빈도가 전체 대출빈도의 절반에 가까운 46.8%를 차지하였다. 또한 중주제 분야의 네트워크 관련 지수 분석 결과, 역사 분야의 네트워크 지수가 상대적으로 높았던 대주제 분야와는 전혀 다르게 문학분야와 사회과학 분야의 매개중심성이 높은 것으로 나타났다.

본 연구를 통해 알아본 연구 결과의 의의는 다음과 같다.

첫째, 9개 도서관의 대출빈도와 대주제, 중주제별 대출빈도를 분석한 결과, 소수의 개체(도서관 또는 중주제 분야)가 대다수의 대출빈도를 발생시키는 전형적인 파레토의 80:20 법칙과 매우 유사한 결과를 보인다는 것을 확인할 수 있었다.

둘째, 중주제 분야 간 동시대출빈도를 고려했을 때, 동일한 대주제 분야에 속한 중주제 분류가 아닌 서로 다른 대주제 분야로 분류된 중주제 분야가 유사성이 더 높은 경우를 확인할 수 있었다. 현재 우리나라의 많은 도서관이 이동자료실이나 별도의 서가 배치가 필요한 경우를 제외하면 거의 대부분 KDC나 DDC 분류기호 순으로 서가를 배열하고 있는 것이 일반적이다. 하지만 본 연구에서 얻어진 결과를 개별 도서관에서 활용한다면 분류기호 순 서가배열과 동시대출빈율이 매우 높은 서로 다른 주제분야를 동일한 서가나 인접한 서가로 배치함으로써 이용자들에게 편의를 제공할 수 있는 방안을 모색할

수 있는 근거를 마련했다 할 수 있다.

다만 본 연구는 대출빈도가 각 군집이나 주제분야별로 비슷하게 분포하지 않고 한쪽으로 치우친 비대칭적인 분포가 나타난 상황을 고려하지 않고, 대출빈도에 상관없이 10개 대주제와 100개 중주제 분야로 나누어 군집분석을 수행하였다. 중주제별 대출빈도 분석에서 나타난 바와 같이 대출순위 1위인 한국문학(810)의 대출빈도(434,346회)는 20위인 의학(510)의 대출빈도(35,157회)의 12배가 넘기 때문에 대출빈도를 보유하고 있기 때문에 한국문학을 의학 분야와 비슷한 대출빈도를 가진 세부 주제영역으로 나누어 분석을 실시하지 못한 것은 본 연구의 제한점이라 할 수 있다. 따라서 향후에는 계층적 군집 분석을 활용하되, 특정 대출빈도를 기준으로 하여 대출빈도가 많은 주제분야는 좀 더 세분화할 수 있는 적정 임계점을 찾아 이를 기준으로 주제별 연관성을 분석하는 후속 연구가 이루어져야 할 것이다.

참 고 문 헌

- 김선애 (2013). 대학도서관의 자료 이용에 관한 연구 - A대학도서관 이용자의 대출통계 분석을 중심으로. 한국비블리아학회지, 24(4), 115-131.
- 김선애, 서혜란 (2009). 공공도서관 소장자료 현황 분석- 장서의 주제별 분포 및 노후화 현황. 한국비블리아학회지, 20(1), 151-164.
- 도서관법 시행령([시행 2010.3.19] [대통령령 제22075호, 2010.3.15., 타법개정]
<http://www.law.go.kr/lsInfoP.do?lsiSeq=103216#0000> [2014년 8월 13일 인용].
- 박영애, 이재운 (2010). 대출 및 상호대차 통계를 활용한 작은도서관 장서 평가에 대한 실험적 연구. 한국문헌정보학회지, 44(2), 333-356. <http://dx.doi.org/10.4275/KSLIS.2010.44.2.333>
- 유경중, 박일중 (2009). 도시지역과 군지역에 위치한 공공도서관의 자료이용 특성에 관한 비교연구:

- 경남지역 4개 공공도서관의 대출기록을 중심으로. 한국비블리아학회지, 20(1), 39-57.
- 이진경, 이지연 (2011). 이용자 장서이용행태 분석을 통한 공공도서관의 장서개발에 관한 연구. 한국비블리아학회지, 22(3), 151-171.
- Ochola, J. N. (2002). Use of circulation statistics and interlibrary loan data in collection management. *Collection Management*, 27(1), 1-13.
- Pareto, V. (1964). *Cours d'Économie politique: Nouvelle édition* par G.-H. Bousquet et G. Busino, Librairie Droz, Geneva, 299-345.
- Trueswell, R. L. (1969). Some behavioral patterns of library users: The 80/20 rule. *Wilson Library Bulletin*, 43, 458-461.

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

- Kim, S. (2013). A study on the collection use of an academic library: Focused on the analysis of circulation statistics of the a-academic library users. *Journal of the Korean Biblia Society for Library and Information Science*, 24(4), 115-131.
- Kim, S., & Suh, H. (2009). Examining the collection of public libraries in terms of subject and currency. *Journal of the Korean Biblia Society for Library and Information Science*, 20(1), 151-164.
- Lee, A., & Lee, J. (2011). A study of collection development by analyzing users' collection-use behaviors in public libraries. *Journal of the Korean Society for Library and Information Science*, 22(3), 151-171.
- Libraries Act enforcement ordinance([enforced 2010.3.19] [Presidential decree No. 22075, 2010.3.15., 타법개정. Retrieved from <http://www.law.go.kr/lsInfoP.do?lsiSeq=103216#0000> [2014년 8월 13일 인용].
- Park, Y., & Lee, J. Y. (2010). An experimental study on small library collection evaluation utilizing circulation statistics and interlibrary loan data. *Journal of Korean Library and Information Science Society*, 44(2), 333-356. <http://dx.doi.org/10.4275/KSLIS.2010.44.2.333>
- Yoo, K., & Park, I. (2009). A comparative study of the use characteristic of public library collection in urban and rural areas: Focused on the circulation data of four libraries in the Gyungangnam-do Province. *Journal of the Korean Society for Library and Information Science*, 20(1), 39-57.

