ISSN 1598-0170 (Print)
ISSN 2287-1136 (Online)
http://www.jksii.or.kr

전자 저널 구독 정보 및 웹 이용 로그를 활용한 참고문헌 기반 저널 추천 기법

The Technique of Reference-based Journal Recommendation Using Information of Digital Journal Subscriptions and Usage Logs

이 해 성[1] 김 순 영[1] 김 재 훈[1] 김 정 환[1*]
Hae-sung Lee Soon-young Kim Jay-hoon Kim Jeong-hwan Kim

요 약

전자 학술 정보 유통의 확대에 따라 날로 증가되는 학술 콘텐츠 서비스 수요에 부응하기 위하여 보다 효과적인 학술 콘텐츠 추천 시스템 개발이 요구된다. 학술 콘텐츠 추천 시스템은 정보 소비자의 과거 이용 내역을 기반으로 각 소비자 선호(preference)에 맞는 학술 콘텐츠를 제공함으로써 콘텐츠 이용성을 보다 효과적으로 향상 시킬 수 있다. 본 논문에서는 특정 기관에 소속된 사용자의 선호에 더욱 부합하는 학술 콘텐츠를 제공하기 위하여 기관의 전자 저널 구독 정보 및 웹 이용 로그를 활용한 저널 추천 기법을 제안한다. 제안하는 추천 기법에서는 기관 사용자의 저널 선호도를 효과적으로 예측하기 위하여 기관 유사도(Institution similarity), 그리고 참고문헌의 인용 관계 데이터를 기반으로 저널 유사도(Journal similarity) 및 저널 중요도(Journal importance)를 산출하여 최종적으로 기관 맞춤형 저널 추천 항목을 구성하게 된다. 또한, 제안하는 추천기법이 적용된 기관 맞춤형 저널 추천 시스템 프로토타입을 개발한다. 개발된 저널 추천 시스템은 각 기관의 저널 선호도 예측을 위하여 활용되는 웹 이용로그를 효과적으로 수집하고 이를 추천 기법에 활용하기 용이한 데이터로 가공 처리 하여 별도의 데이터베이스에 저장하여 추천 기법의 저널 선호도 예측을 위한 기반 데이터로 활용한다. 마지막으로 우리는 기존 추천 기법들과의 비교 성능 평가를 통해 제안 기법의 차별성과 우수성을 보인다.

☞ 주제어 : 추천 시스템, 참고문헌, 인용 네트워크, 맞춤형 학술 콘텐츠 서비스, 웹 이용 로그 분석

ABSTRACT

With the exploration of digital academic information, it is certainly required to develop more effective academic contents recommender system in order to accommodate increasing needs for accessing more personalized academic contents. Considering historical usage data, the academic content recommender system recommends personalized academic contents which corresponds with each user's preference. So, the academic content recommender system effectively increases not only the accessibility but also usability of digital academic contents. In this paper, we propose the new journal recommendation technique based on information of journal subscription and web usage logs in order to properly recommend more personalized academic contents. Our proposed recommendation method predicts user's preference with the institution similarity, the journal similarity and journal importance based on citation relationship data of references and finally compose institute-oriented recommendations. Also, we develop a recommender system prototype. Our developed recommender system efficiently collects usage logs from distributed web sites and processes collected data which are proper to be used in proposed recommender technique. We conduct compare performance analysis between existing recommender techniques. Through the performance analysis, we know that our proposed technique is superior to existing recommender methods.

☞ keyword : recommender system, recommendation, citation networks, personalized academic contents, log analysis, usage statics

1. 서 론

인터넷을 기반으로 하는 정보 기술의 눈부신 발전과 스마트폰이나 태블릿 PC와 같은 디지털 기기의 급속한 이용확대는 다양한 정보 콘텐츠들의 소비를 활성화 시키며 인간 생활을 더욱 풍요롭고 윤택하게 만들어 가고 있다. 또한 전자 학술 정보의 폭발적인 생산 및 유통의 확대를 가져오게 되었으며, 다양한 웹 기반의 국내외 학술정보 데이터베이스, 지식포탈 등의 전자 정보원을 활용한 다양한 학술정보서비스들이 제공되고 있다. 이와 같은 변화는 연구자들뿐만 아니라 일반 이용자들의 전자

[1] Dept. of Information Infrastructure, S&T Information Center, Korea institute of Science and Technology Information, 245 Daehak-ro, Daejeon, Korea.
* Corresponding author (kimjh@kisti.re.kr)
[Received 17 June 2016, Reviewed 1 July 2016, Accepted 11 August 2016]

학술 정보에 대한 소비를 더욱 활성화시키고 디지털 콘텐츠 큐레이션(curation)과 같은 개인화된 맞춤형 정보 서비스에 대한 소비 욕구를 더욱 증대시키고 있다[1]. 그러나 학술정보 서비스 요구에 대한 폭발적인 증가와 정보 유통 체제의 발전에도 불구하고 각 기관의 정보 제공 관계자들은 정보 서비스 개발을 위해 요구되는 이용자 수요 분석 등을 위한 기반 데이터를 확보하는데 어려움을 가지고 있다. 때문에 각 기관의 소속된 사용자들의 정보 접근성을 활성화 시키고 개개의 정보 소비요구를 충족시킬 만한 학술 콘텐츠를 제공하는데 많은 한계가 따르게 된다. 또한 각 기관에 소속된 이용자의 정보 소비 요구 및 콘텐츠 확보에 부응하기 위해서는 신규 전자정보원의 지속적인 확대가 매우 중요하지만 이용자들을 만족시킬 만한 신규 전자 정보원을 발굴하는 것 역시 기관 측면에서의 많은 비용을 초래하고 있는 실정이다.

일반적으로 NDSL(National Digital Science Library)과 같은 학술 콘텐츠 서비스 사이트들은 서비스 개선 등의 목적으로 각 기관별로 이용로그를 저장하고 있다[2]. 이러한 학술 콘텐츠 이용로그는 학술 콘텐츠 서비스 제공자뿐만 아니라 각 기관 이용자들에게 보다 개인화된 학술 콘텐츠를 제공하기 위한 중요한 기반요소가 된다. 때문에 학술 콘텐츠 이용 로그 활용의 중요성이 꾸준히 증가하고 있다. 일반적으로 많은 출판사들이 웹 사이트나 각 기관으로 부터의 요청에 따라 이용 통계 정보를 제공하고 있다[3]. 하지만 각 기관은 다수의 출판사로부터 전자 저널을 구독하고 있고 특정 기관에 속한 사용자들 역시 다양한 학술 정보 서비스를 통해 학술 콘텐츠를 이용하고 있기 때문에 효율적으로 이용 정보를 수집하고 이를 통합 관리함으로써 구독 저널 추천 등 맞춤형 정보 서비스에 활용하는데 많은 어려움이 따른다. 때문에 학술 콘텐츠 이용로그는 각 기관의 정보 기반 경영 및 정보 서비스 제공을 위한 중요한 지표로서의 활용 가치가 높음에도 불구하고 효과적으로 활용 되지 못하고 있다.

본 논문에서는 학술 콘텐츠의 정보 접근성을 보다 효과적으로 활성화 시키고 개개인의 학술 정보 소비요구를 충족시키는 것을 목표로 기관 유사도와 저널 유사도를 고려한 정보 소비자 선호도 예측 기법을 제안한다. 또한 정보 소비자의 선호도를 보다 정확하게 예측해내기 위하여 우리는 다수의 정보원 및 다수의 학술 콘텐츠 서비스 사이트들로부터 발생된 학술 이용로그와 각 기관의 전자저널 구독 정보를 수집하고 기관별로 관리하는 학술 콘텐츠 이용 정보 통합 관리 시스템을 설계 및 개발한다. 본 논문의 기여도를 나열하면 다음과 같다. 첫째, 다수의 정보원 및 다양한 학술 정보 서비스 사이트들로부터 제공되는 학술 콘텐츠 이용로그 수집의 어려움을 해결하고 맞춤형 콘텐츠 서비스가 가능하도록 통합 관리함으로써 기관 맞춤형 정보 기반 경영이나 정보서비스 기획 및 개발 효율화에 기여한다. 뿐만 아니라 본 논문을 통해 개발된 추천 시스템의 도움으로 각 기관으로 하여금 기관 소속의 사용자 요구가 반영된 보다 폭넓은 학술 콘텐츠를 효과적으로 확보하는데 유용하게 활용할 수 있다. 둘째, 기존과 차별화된 기관 맞춤형 학술 콘텐츠 추천 서비스를 통해 학술 콘텐츠에 대한 소비자의 요구를 신속히 수용 가능하도록 하여 학술 콘텐츠 자원의 활용범위를 넓히고 고부가가치를 창출할 수 있는 기회를 확대한다.

본 논문의 구성은 다음과 같다. 2장은 관련 연구로 사례조사를 통해 학술 콘텐츠 이용로그 활용의 중요성과 기존의 이용로그 기반 학술 정보 서비스들의 한계점을 살펴보고, 학술 콘텐츠 분야에서 추천 시스템의 필요성에 대하여 논한다. 3장은 학술 이용로그와 각 기관의 전자 저널 구독 정보를 수집하고 기관별로 관리하는 학술 콘텐츠 이용 정보 통합 관리 프로세스 및 시스템 아키텍쳐(Architecture)에 대하여 논하고, 4장에서는 저널에 대한 사용자 선호도를 예측하여 저널 추천 항목을 구성하는 저널 추천 기법을 제안한다. 5장에서는 본 논문에서 제안한 기관 맞춤형 학술 콘텐츠 추천 시스템을 구현하고 시스템의 성능 평가를 실시한다. 마지막으로 6장에서는 결론에 대해 기술하고, 향후 연구에 대해 논한다.

2. 관련 연구

2.1 맞춤형 학술 콘텐츠 추천 시스템

정보기술의 발달과 급속한 변화와 맞물려 전자 학술 자원의 양적 팽창이 빠르게 진행되면서 사용자의 정보 선호(preference)에 맞는 맞춤형 학술 콘텐츠 추천 서비스 제공의 필요성이 더욱 대두되었다. 추천 시스템은 일종의 정보 필터링(information filtering) 기술로서 사용자로 하여금 자신이 선호하는 콘텐츠를 보다 빠르고 정확하게 찾을 수 있도록 한다[4][5]. 추천 시스템에서 개인의 선호를 예측하기 위해 일반적으로 사용되는 협업필터링(collaborative filtering)은 큰 무리의 사용자 집단으로부터 특정 사용자와 유사한 취향을 가진 소규모의 유사 사용자 집단을 발견하는 방식으로 동작한다[6]. 따라서 협업필터링에서는 특정 사용자와 유사한 취향을 가지는 유사 사용자 집단을 추출하기

위하여 특정 사용자의 과거 콘텐츠 이용 데이터뿐만 아니라 전체 사용자의 과거 콘텐츠 이용 데이터에 대한 수집과 관리가 반드시 요구된다. 마찬가지로 학술 콘텐츠 추천 시스템 역시 사용자의 과거 학술 콘텐츠 이용 데이터를 기반으로 개인이 선호하는 연구 주제나 과학기술 분야를 예측하여 이를 토대로 사용자가 관심 있을 법한 학술 콘텐츠를 추천하게 된다. 학술 콘텐츠 이용 히스토리 정보에 기반을 둔 개인 사용자의 선호도 예측기법에 대한 연구는 최근 까지도 꾸준히 수행되어 왔으며, 그중 Kwanghee Hong et al.은 사용자의 프로파일 정보 분석과 논문의 키워드 추론을 통해 논문에 대한 사용자의 선호도를 예측하여 유사 사용자 집단을 구성하는 추천 기법을 제안한바있다[7].

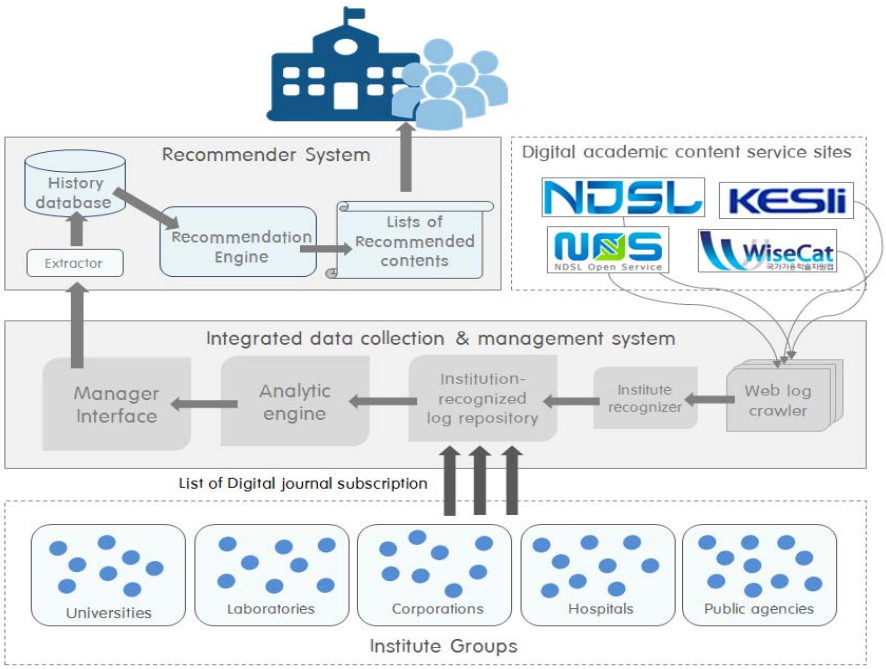
맞춤형 학술 콘텐츠 서비스와 관련하여 최근에는 '콘텐츠 큐레이션(Contents curation)' 역시 정보 과잉의 시대에 의미 있고 가치 있는 정보의 획득과 공유가 중요해짐에 따라 주목 받기 시작했다. '콘텐츠 큐레이션 서비스'는 온라인상의 수많은 콘텐츠들 중 개인의 주관이나 관점에 따라 콘텐츠들을 수집, 정리하고 편집하여 사용자와 관련이 있거나 좋아할 만한 콘텐츠를 제공하는 서비스라 할 수 있다[1]. 미국의 경제지 포브스(Forbes)는 콘텐츠 큐레이션 서비스가 개인화를 통한 인간적 요소에 초점을 맞춘 서비스라고 정의내리고 수많은 콘텐츠에서 사용자의 선호에 맞는 의미 있는 콘텐츠를 발굴하고 가공해 서비스하는 큐레이터(Curator) 역할의 중요성을 언급하였다[8]. Akisato et al.은 미국 SNS업계의 대표적인 콘텐츠 큐레이션 서비스인 'Pinterest'에서 사용자가 업로드한 이미지(Image) 콘텐츠와 해당 콘텐츠에 기술된 글로부터 문맥(Context)를 파악해 냄으로써 이를 기반으로 이미지 콘텐츠 큐레이션을 위한 유사 이미지 콘텐츠 집합을 구성하는 기법을 제안하였다[9][10]. 정보 과잉에 따른 맞춤형 콘텐츠 서비스의 중요성이 날로 증대됨에 따라 '콘텐츠 큐레이션'은 컴퓨터학이나 문헌 정보학은 물론 경영학등 여러 학문 분야에서 다양한 관점으로 연구들이 수행되고 있다[11]-[13]. 그러나 사용자 선호도 예측과 같은 맞춤형 콘텐츠 큐레이션을 위한 개인화 기법에 대한 연구가 아직까지 미미하다. 뿐만 아니라 다양한 형식과 엄청난 규모로 생산되고 있는 전자 학술 정보의 체계적 수집과 서비스뿐만 아니라 기술적인 측면에서 사용자가 원하는 정보 자원을 찾아 분석하고 제공하는 방법에 대한 논의는 현재까지 미흡한 실정이다.

2.2 학술 콘텐츠 이용로그 처리 기술

맞춤형 학술정보서비스 요구의 증가와 정보 유통 체제의 발전에 따라 학술콘텐츠 이용 데이터는 새로운 전자정보 자원 및 서비스를 개발하고 운영할 수 있는 근거로서 그 관심이 꾸준히 증폭되었다. 이에 따라 학술정보 이용 데이터 활용에 대한 노력이 활발히 이루어지고 있다. 특히, 온라인을 통한 학술 자원의 활용이 급속히 증가하면서, 정보자원 이용량에 대한 일관적이고 표준화된 이용 데이터 수집 시스템 구축의 필요성이 국제적으로 증대되었다. 이용정보의 일관성, 신뢰성의 문제를 해결하기 위한 시도로 COUNTER(Counting Online Usage of Network Electronic Resources)가 2002년 3월에 공식적으로 수립되어 학술 콘텐츠 이용데이터 수집 표준화가 이루어졌다. 표준화를 통해 전자 정보원들 간 비교분석등과 같이 학술 콘텐츠 이용 정보 활용 범위가 보다 넓어지게 되었다[14]. 그러나 이 역시, 출판사나 벤더들의 사이트들을 일일이 접근하여 해당 기관의 이용 정보를 수집하여 처리하는 데 따른 어려움이 발생하게 되었다. 따라서 이용정보 수집에 따른 시간과 비용이 낭비되는 문제를 해결하기 위해 SUSHI(Standardized Usage Statistics Harvesting Initiative)가 개발 되었다. SUSHI는 분산 되어있는 전자정보원들의 이용 데이터를 취합하기 위한 자동요청응답 모델로서 전자 저널의 COUNTER 기반 이용정보를 자동적으로 로컬시스템에 저장할 수 있는 프로토콜이라 할 수 있다[15]. 그러나 SUSHI 기반으로 개발된 학술 콘텐츠 이용정보 수집 시스템은 출판사별 저널 중심의 이용정보만을 수집하여 관리하기 때문에 정보 소비자의 선호도를 예측하는데 어려움이 따르며 맞춤형 저널 추천 서비스를 개발하는데 한계가 따른다.

3. 기관 전자저널 구독 정보 및 학술 콘텐츠 이용로그 통합 수집 및 관리

출판사 별로 수집되던 기존의 학술 콘텐츠 이용로그들이 갖는 한계점을 해결하고 각 기관의 특성을 보다 효과적으로 반영한 맞춤형 저널 추천 시스템 개발이 가능하고자 본 논문에서는 기관 전자저널 구독 정보 및 학술 콘텐츠 이용로그 통합 수집·관리시스템을 제안한다. 그림 1은 제안하는 시스템의 개념적 아키텍처를 보인다. 제안하는 시스템에서는 대학, 연구소, 기업, 의료 기관 그리고 공공기관의 5개 유형으로 구성된 기관 그룹에 속한 각 기관으로부터 전자저널 구독 정보를 수집하여 저장 관리한다.



(그림 1) 기관 전자저널 구독 정보 및 이용로그 통합 수집·관리 시스템의 개념적 아키텍처
(Figure 1) Conceptual architecture of institute's digital journal subscription information and usage log collection & management system

각 기관으로부터 수집된 전자저널 구독 정보는 제안시스템의 레퍼지토리에 기관 식별자가 부여되어 기관별로 저장된다. 또한, 그림 1에서와 같이 한 개 이상의 웹 로그 크롤러(Web log crawler)들은 NDSL(National digital science library)과 전자정보 국가 컨소시엄인 KESLI(Korean electronic site license initiative)가 운영하는 웹 사이트와 같은 국내 전자 학술 콘텐츠 서비스 사이트들로부터 웹 로그를 수집한다[16]. NOS(NDSL open service)는 학술 정보를 오픈 API를 통해 제공하는 서비스이며 WiseCat은 국가가용학술 자원의 이용통계를 제공해주는 서비스이다[17][18]. 수집된 로그는 기관 식별기(Institute recognizer)를 통해 기관 식별이 완료된 형태로 이용 로그를 레퍼지토리에 저장한다. 분석 엔진(Analytic engine)은 레퍼지토리에 저장된 이용 로그를 기반으로 기관별로 표 1과 같은 이용 통계 항목을 생성하여 추천 시스템과 같은 응용 시스템에서 활용할 수 있도록 한다. 그림 1의 추천 시스템은 추출기(Extractor)를 통해 통합 수집·관리 시스템의 분석 엔진으로부터 생성된 이용 통계 데이터를 추출한다. 추출한 이용 통계 데이터는 추천 시스템의 히스토리 데이터베이스 저장되고 이를 기반으로 추천 엔진은 최종적으로 맞춤형 추천 항목을 구성하게 된다.

(표 1) 기관 별 학술 콘텐츠 이용 통계 생성
(Table 1) The creation of academic usage statics by each institute

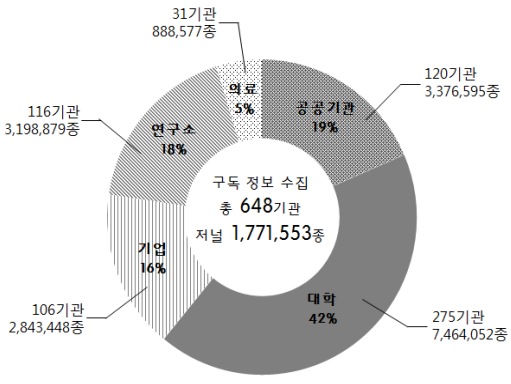
이용 통계 항목	내 용
저널 이용 통계	구독 저널 이용 횟수 미구독 저널 이용 횟수
구독 형태에 따른 저널 이용 통계	OA(Open access) 저널 이용 횟수 유료/무료 저널 이용 횟수
주제 분야별 이용 통계	자연, 응용과학, 사회과학, 인문과학, 기타분야 저널 이용 횟수
기관 유형별 이용 통계	대학, 연구소, 기업, 의료기관, 공공기관 별 이용 횟수

제안 시스템의 관리자 인터페이스(Manager interface)는 그림 2에서 보이는 것과 같이 이용로그 수집과 수집된 로그의 관리 및 이용 통계생성 프로세스를 관리자로 하여금 보다 효율적으로 수행 할 수 있도록 한다.

(그림 2) 관리자 인터페이스 개발
(Figure 2) Development of the manager interface

관리자 인터페이스를 통해 관리자는 표 1에서 나열된 통계 생성 항목과 날짜별로 이용 로그에 대한 수집 요청이 가능하고, 통계 생성과 이미 생성된 통계에 대한 삭제를 수행할 수 있다. 그림 3은 본 연구를 통해 개발된 전자저널 구독 정보 및 이용로그 통합 수집·관리 시스템에서 수집 및 구축되어 있는 기관 유형에 따른 전자 저널 구독 현황을 보인다.

표 2는 제안 시스템에서 전자저널 구독 정보를 수집하고 있는 전체 648 기관 중 임의의 10개 기관에 대한 학술 콘텐츠 이용 통계 일부를 보인다. 각 기관별 이용 통계는 그림 1의 제안 시스템을 구성하고 있는 분석 엔진을 통해 산출된다.



(그림 3) 전자저널 구독 정보 수집 기관 현황
(Figure 3) The current state of institute's subscription for digital journals

(표 2) 기관 별 학술 콘텐츠 이용 통계
(Table 2) Institute's usage statics of academic contents

기관	기관 유형	이용자 수	NDSL		NOS	DDS
			상세 보기	원문 보기		
A	연구소	91	156	288	257	47
B	공공	21	51	59	25	5
C	대학	146	568	1,036	639	0
D	대학	306	293	282	232	1
E	의료	29	54	78	64	0
F	대학	781	1,776	1,796	1,715	9
G	연구소	31	70	166	181	10
H	기업	11	18	7	0	21
I	공공	272	337	379	6,395	171
J	기업	35	5	2	0	10

표 2에서 각 기관의 이용자 수는 일정기간 동안 학술 콘텐츠를 이용한 각 기관에 소속된 이용자 수를 의미한다. 상세보기와 원문보기는 각각 국가 과학기술 전자 도서관인 NDSL을 통하여 특정 기관 이용자가 학술 콘텐츠 메타(Meta) 정보에 접근한 횟수와 논문이나 보고서등 학술 콘텐츠 원문을 이용한 횟수이다. NOS와 DDS(Document delivery service)는 Open API를 이용하여 학술 콘텐츠 정보를 활용한 횟수와 원문복사 서비스를 이용한 횟수를 각각 의미한다. 결과적으로 기관의 전자저널 구독 정보 및 이용 로그 통합 수집·관리 시스템 개발을 통해 기관 별 이용 통계 데이터를 생산해냄으로써 이를 기반으로 보다 효과적인 맞춤형 저널 추천 서비스가 가능하다.

4. 기관 유사도 및 저널 유사도를 고려한 기관 맞춤형 저널 추천 기법

본 연구에서 제안하는 저널 추천기법은 각 기관 소속 연구자들의 선호에 보다 적합한 저널 목록을 기관 단위로 구성하여 제공함으로써 기관 맞춤형 정보 서비스가 가능하도록 한다. 이를 위해 우리는 3장에서 제안한 전자저널 구독 정보 및 이용 로그 통합 수집·관리 시스템을 통해 생성된 표 1에 해당하는 기관 별 이용 통계 데이터를 기반으로 기관 유사도 및 저널 유사도를 고려하여 각 기관 사용자들의 학술 콘텐츠에 대한 선호를 예측한다.

본 논문에서 제안하는 추천 기법은 기관 유사도 및 저널 유사도를 기반으로 기관 단위 정보 소비자의 저널 선호도를 예측하여 저널 추천 항목을 구성한다. 표 3은 본 연구에서 제안하는 저널 선호도 예측 기법을 통해 추천 항목을 구성하는 알고리즘을 보인다.

(표 3) 기관 I_i의 저널 선호도 예측 기반 추천 알고리즘

(Table 3) Recommendation algorithm based on the prediction of Institute I_i's journal preferences

알고리즘 1.
01. P ← 기관 I_i에서 발표한 논문이 실린 저널 집합
02. S ← 기관 I_i에서 구독/미구독 저널별 이용 통계 집합
03. R ← 기관 I_i에서 발표한 논문들의 참고문헌 집합
04. **FOR EACH** (기관 집합 I)
05. **BEGIN**
06. **IF** (I_i의 기관 유형 == I_j의 기관 유형) **THEN**
07. $Type\,W_{ij}$ ← 1
08. **ELSE** $Type\,W_{ij}$ ← 0 **END**
09. **FOR EACH** (저널 집합 P_j)
10. **BEGIN**
11. $Pub\,W_{ij}$ ← 기관 i와 j 의 발표 논문의 저널 주제 분야 간 유사도
12. **END**
13. **FOR EACH** (구독 저널 집합 S_j)
14. **BEGIN**
15. $Suc\,W_{ij}$ ← 기관 i와 j의 구독 저널의 주제 분야 유사도
16. **END**
17. $ISim_{ij}$ ← $Pub\,W_{ij}$ + $Suc\,W_{ij}$ + $Type\,W_{ij}$
18. **FOR EACH (** 상위 k의 $ISim$값을 갖는 유사 기관 집합 SI **)**
19. **BEGIN**
20. **FOR EACH** (SI_j의 구독 저널 집합 J)
21. **BEGIN**
22. **IF (**저널 J_k == 기관 I_i의 미구독 저널**) THEN**
23. **FOR EACH(**기관 I_i에서 발표한 논문이 실린 저널 집합 P**)**
24. **BEGIN**
25. $Ref\,W_{lk}$ ← 저널 P_k와 J_k의 인용 유사도
26. **END**
27. $Imp\,W_{ik}$ ← 기관 I_i의 저널 k에 대한 중요도
28. JP_{ik} ← $Ref\,W_{lk}$ • $Imp\,W_{ik}$
29. **END**
30. 상위 k의 $JSim$을 갖는 저널들을 추천 항목으로 구성
31. **END**

본 연구에서 제안하는 표 3의 추천 알고리즘에서는 먼저, 기관 유사도 $ISim$를 통해 기관 i에 대한 유사 기관 집단을 추출한다. 다음으로 추출된 유사 기관들에서 이용된 저널들을 대상으로 저널 선호도 JP를 산출하고 상위의 저널 선호도 값을 갖는 저널 목록을 구성하여 이를 추천 항목으로 제공하게 된다.

4.1 유사 기관 집단 추출

본 연구에서 제안하는 저널 추천 기법에서는 추천 항목 구성을 위한 대상 저널들을 우선적으로 고려하기 위하여 유사 기관 집단을 찾고 유사 기관들에서 이용된 저널들을 대상으로 선호도를 예측한다. 알고리즘 1의 $ISim_{ij}$은 저널 k에 대한 기관 i와 j간의 유사도를 의미하며, 산출식은 식(2)와 같다.

$$ISim_{ij} = Pub\,W_{ij} + Suc\,W_{ij} + Type\,W_{ij} \quad (2)$$

식 (2)에서 $Type\,W_{ij}$는 기관 유형에 따른 유사 가중치를 의미한다. $Type\,W_{ij}$는 0과 1중 하나의 값이 할당되며, 기관 i와 j의 기관 유형이 같으면 1, 다르면 0값을 가지게 된다. $Pub\,W_{ij}$는 기관 소속 연구자들이 발표한 논문들이 발행된 저널의 주제 분야를 고려한 유사 가중치이다. 즉, 기관 i와 j에 각각 소속된 연구자들이 동일한 주제 분야를 갖는 저널에 논문을 많이 발표할수록 비슷한 연구를 수행하는 유사 기관으로 판단되는 것이다.

$$Pub\,W_{ij} = \frac{\sum_{k=1}^{n}(p_{ik}-\overline{p_i})(p_{jk}-\overline{p_j})}{\sqrt{\sum_{k=1}^{n}(p_{ik}-\overline{p_i})^2 \cdot \sum_{k=1}^{n}(p_{jk}-\overline{p_j})^2}} \quad (2.1)$$

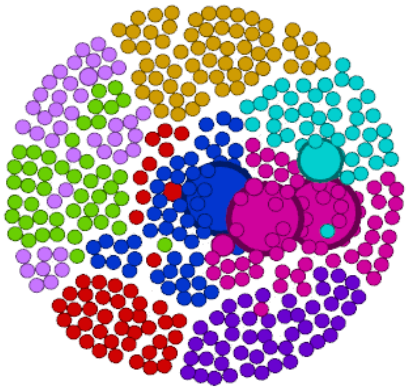
논문이 발행된 저널의 주제 분야를 고려한 유사 가중치 $Pub\,W_{ij}$는 식 (2.1)와 같이 피어슨 상관계수(Pearson correlation coefficient)를 적용하여 산출 된다[19]. 식 (2.1)의 $\overline{p_i}$와 $\overline{p_j}$는 각각 기관 i와 j의 소속 연구자들이 발표한 논문이 발행된 전체 저널 수이며, p_{ik}와 p_{jk}는 기관 i와 j의 주제 분야 k에 해당하는 저널의 논문 발행 횟수를 각각 의미한다.

다음으로 식 (2)의 $Suc\,W_{ij}$는 기관에서 구독 중인 저

널의 주제 분야를 고려한 유사 가중치이다. 예를 들어, 기관 i와 j에서 구독중인 저널의 주제 분야가 동일한 횟수가 많을수록 유사 기관으로 파악 된다. 식 (2.2)는 $Suc\,W_{ij}$의 산출식을 보인다. 여기서, $\overline{s_i}$와 $\overline{s_j}$는 각각 기관 i와 j에서 구독중인 전체 저널 수이며, s_{ik}와 s_{jk}는 기관 i와 j의 주제 분야 k에 해당하는 저널 구독 횟수를 각각 의미한다.

$$Suc\,W_{ij} = \frac{\sum_{k=1}^{n}(s_{ik}-\overline{s_i})(s_{jk}-\overline{s_j})}{\sqrt{\sum_{k=1}^{n}(s_{ik}-\overline{s_i})^2 \cdot \sum_{k=1}^{n}(s_{jk}-\overline{s_j})^2}} \quad (2.2)$$

그림 4는 식 (2)의 기관 유사도를 기반으로 400여개 기관에 대한 K-평균 군집화(K-Means Clustering)를 수행하고 이에 대한 결과를 가시화 한 것 이다[20].



(그림 4) 기관 유사도 기반 K-평균 군집화
(Figure 4) K-Means clustering based on institute similarity

각 군집의 중심 값은 식(2)의 $Pub\,W_{ij}$와 $Suc\,W_{ij}$의 값으로 구성된다. 색상별로 총 8개의 군집이 생성되었으며 같은 색상을 가지는 기관들은 상대적으로 유사한 기관으로 고려될 수 있다.

4.2 참고문헌 저널 유사도 기반 선호도 예측

저널 선호도 JP_{ik}는 참고문헌 저널 간 유사도로 기관 i에서 과거 이용된 저널 l과 저널 k의 유사한 정도를 나타내며 산출식은 식(3)과 같다.

$$JP_{ik} = Ref\,W_{lk} \bullet Imp\,W_{ik} \quad (3)$$

식(3)에서 $Ref\,W_{lk}$는 저널간 인용 유사성(Citation similarity)을 고려한 가중치로 코사인 거리(Cosine distance) 계산법이 적용된 산출식은 식(3.1)과 같다[21].

$$Ref\,W_{lk} = \frac{\sum_{q=1}^{n} l_q \times k_q}{\sqrt{\sum_{q=1}^{n}(l_q)^2} \times \sqrt{\sum_{q=1}^{n}(k_q)^2}} \quad (3.1)$$

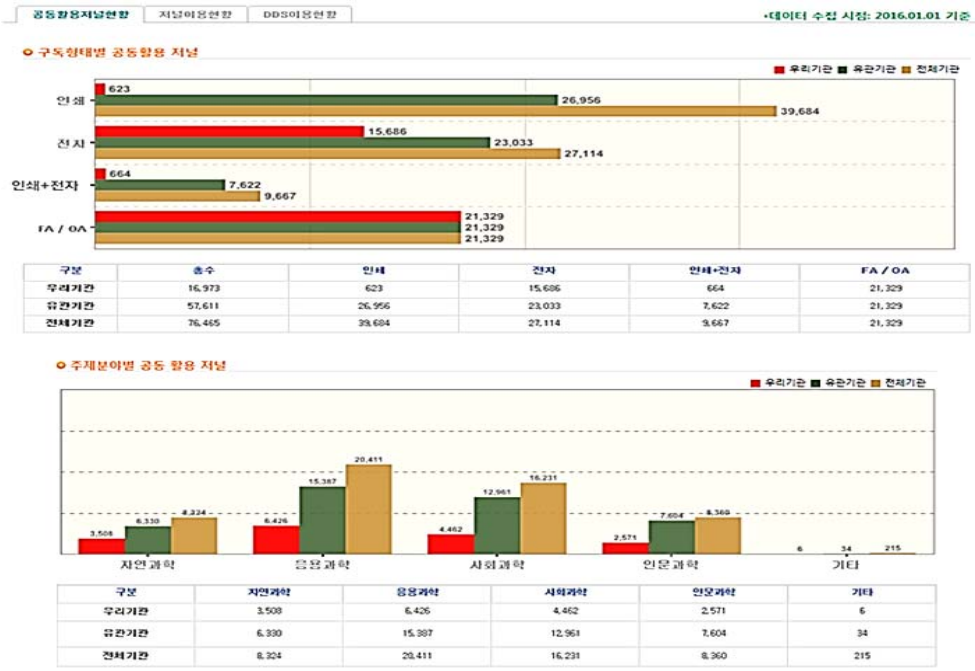
식(3.1)에서 l_q와 k_q는 저널 l과 k에서 발행된 논문들이 저널 q를 통해 발행된 논문을 인용한 횟수를 각각 의미한다. 즉, 저널 l과 k에서 각각 인용된 논문이 발행된 저널이 동일할수록 높은 $Ref\,W_{lk}$ 값이 산출된다. 또한, 식(3)의 $Imp\,W_{ik}$은 기관 i의 저널 k에 대한 중요도로 식(3.2)를 통해 산출되며, i기관의 소속 연구자들이 발표한 논문에서 인용한 참고문헌을 많이 발행한 저널 일수록 중요도가 높게 산출된다.

$$Imp_{ik} = \frac{rj_{ik}}{\overline{ri}} \quad (3.2)$$

식(3.2)에서 $\overline{r_i}$는 기관 i의 소속 연구자들이 인용한 저널의 평균인용 횟수를 의미한다. 예를 들어, 기관 i의 소속 연구자들이 발표한 논문에서의 전체 인용수가 10,000건이고, 해당 논문들을 발행한 저널 수가 총 100종이라면 $\overline{r_i}$는 100으로 계산된다. rj_{ik}는 저널 k에 대한 기관 i의 소속 연구자들이 발표한 논문에서의 인용 횟수이다. 즉, 기관 i의 소속 연구자가 발표한 논문이 저널 k를 통해 발표된 논문을 많이 이용할수록 저널 k에 대한 기관 i의 저널 중요도 $Imp\,W_{ik}$는 높게 산출된다.

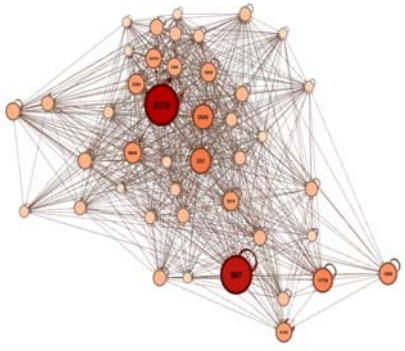
그림 5는 표 1의 기관 D에서 발표된 논문의 저널과 해당 논문에서 인용한 문헌이 발행된 저널 간 인용 관계 네트워크(Relationship network)를 보인다.

그림 5에서 적색의 노드들은 기관 D에서 발표된 논문에서 상대적으로 많이 인용된 논문들이 발행된 저널들로 $Imp\,W_{ik}$값이 다른 저널들 보다 클 가능성이 높다. 즉, 그림 5의 저널 식별자 '8370'과 '987'에 해당하는 노드들은 다른 저널들보다 상대적으로 중요도가 높은 저널을 의미하며 본 연구에서 제안하는 추천 기법을 통해 기관 D에 대한 저널 추천 항목으로 구성 될 가능성이 높다.



(그림 6) 기관의 전자 저널 구독 정보 수집 현황
(Figure 6) The current status of the collected institute's digital journal subscription information

결과적으로 본 연구에서 제안한 저널 추천 기법은 각 기관의 전자 저널 구독 정보 및 이용 통계를 기반으로 기관 유사도뿐만 아니라 참고문헌 기반 저널 유사도를 고려하여 보다 정확하게 저널에 대한 각 기관 단위의 선호도를 예측한다. 따라서 기관에 소속된 이용자나 연구자들이 필요로 하는 학술 콘텐츠를 보다 효과적으로 제공할 수 있다.



(그림 5) 기관 D의 발표 논문
참고문헌 기반 저널 간 인용 관계도
(Figure 5) Journal's citation network based on references of institute D's published papers

5. 저널 추천 시스템 개발

본 절에서는 제안된 기관 맞춤형 저널 추천 기법을 기반으로 하는 저널 추천 시스템 프로토타입을 구현하여 제안 기법의 유용성을 보인다.

5.1 데이터 구성

구현한 시스템은 각 기관의 전자저널 구독 정보와 이용로그의 수집 및 관리를 통해 저널 선호도 예측을 위한 기반 데이터를 구축 및 제공한다.

또한, 해당 데이터를 기반으로 본 논문의 4장에서 제안한 기관 유사도와 저널 유사도를 고려한 저널 추천 기법을 적용하여 기관 단위의 선호에 부합하는 기관 맞춤의 저널 추천 항목을 최종 구성하게 된다. 그림 6은 본 연구를 통해 개발된 기관 전자저널 구독 정보 및 이용로그 통합 수집·관리 시스템의 사용자 인터페이스를 통해 제공되는 전체 기관 및 '행복 대학교'의 구독 현황을 보인다. 전체 400여개 기관을 통해 수집된 전자 저널 수는 총 76,465종으로 이중, '행복 대학교'와 유사한 기관들에

공동활용저널현황 | 저널이용현황 | DDS이용현황 ›데이터 수집 시점: 2016.01.01 기준

우리기관 이용자가 활용한 학술 자원 현황

구분	이용자수	NDSL 플랫폼에서 이용					출판사 이용건수
		상세보기	원문보기	NOS	DDS신청	KISTI 호스팅 원문	
우리 기관	82	316	403	489	26	11	593
유관기관총계	2,967	8,073	9,592	16,247	1,648	42	73,243
전체기관총계	33,315	89,295	93,243	83,350	3,153	109	123,243

(그림 7) 기관 이용자들의 학술 자원 이용 현황
(Figure 7) The current status of Institute's use of academic contents

서 구독하고 있는 저널 수는 57,611종이다. 저널의 주제 분야에 따른 구독 현황은 '사회과학' 분야가 총 16,231종으로 가장 많은 기관에서 '사회과학'분야의 저널을 구독하고 있는 것으로 파악할 수 있다. 그림 7은 기관 이용자들이 수집된 웹 이용 로그의 기관 별 분석을 통해 파악된 기관 이용자들의 학술 자원 이용 현황을 보인다. 기관 별로 파악된 전체 이용자 수는 33,315명으로 이중, '행복 대학교' 소속 이용자 수와 '행복 대학교'와 유사한 기관들에 속한 이용자 수는 각각 82명과 2,967명으로 집계되었다. 또한, 이용 유형 중 '원문보기'가 전체 93,243건으로 집계되어 학술 콘텐츠 서비스들 중 원문에 대한 이용률이 가장 높은 것으로 파악될 수 있다.

따라서 본 연구에서는 그림 6과 그림 7에서 보이는 바와 같은 기관의 구독 정보와 이용 통계 데이터를 기반으로 하여 각 기관의 저널 선호도를 예측하고, 이를 통해 기관 맞춤형 저널 추천 항목을 구성하게 된다.

5.2 성능 평가

본 논문에서는 제안한 저널 추천 기법의 유용성과 효과성을 검증하기 위해 추천 시스템 프로토타입을 개발하고 이를 통한 성능평가를 실시한다. 표 4는 본 논문의 4장에서 제안한 식(2)의 기관 유사도 산출식을 통해 추출된 '행복 대학교'의 유사 기관들 중 상위 5개의 기관들을 나열한 것이다.

(표 4) 기관 '행복 대학교'의 유사 기관 목록
(Table 4) The similar institute list of 'Happy university'

	기관 A	기관 B	기관 C	기관 D	기관 E
유사도	0.743	0.6262	0.5952	0.5835	0.4913

표 4에서는 기관 A가 가장 큰 유사도 값을 가지며, 이는 다른 기관들 보다 '행복 대학교'와 유사하다는 것을 의미한다.

(표 5) 기관 '행복 대학교'의 저널 추천 목록
(Table 5) The recommended journal list of 'Happy university'

	8370	9648	231	8180	10791
선호도	0.9431	0.8823	0.7723	0.7712	0.7634

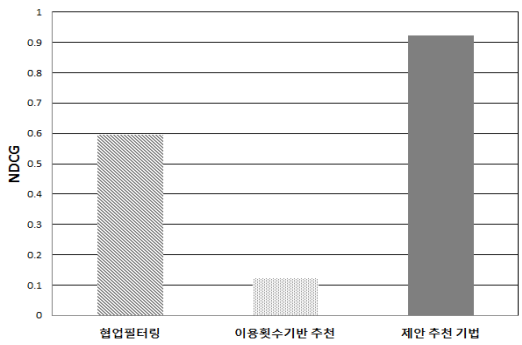
표 5는 표 4와 같은 유사 기관들에서 이용된 저널들 중 식(3)의 저널 선호도 산출식을 통해 각 저널에 대한 '행복 대학교'의 선도를 산출하여 상위 값을 갖는 5개의 저널 목록을 보인다. 즉, 표 5와 같이 저널 선호도에 따라 상위 K개의 저널을 구성하여 최종 추천 항목으로 제공할 수 있다. 표 5에서는 저널 식별자 '8370'을 갖는 저널이 가장 높은 선호도 값을 가지며, '행복 대학교' 소속 이용자들이 가장 선호할만한 저널로 간주된다. 그림 8은 '행복 대학교'의 저널 추천 서비스 화면을 보인다. 총 20종의 저널이 추천 항목으로 구성 되었으며, 각 저널이 갖는 선호도에 따라 정렬되어 제공된다. 그림 8에서의 이용 건수는 해당 저널의 전체 이용 건수를 의미하며, 제공된 추천 목록에서는 이용 건수와 상관없이 저널 선호도에 따라 정렬된 것을 알 수 있다. 따라서 본 연구에서 제안된 저널 추천 기법을 통해 각 기관의 저널 선호에 부합하여 추천 저널 목록이 구성되었음을 알 수 있다. 또한, 추천 항목에 '행복 대학교'에서 미구독 중인 저널들이 추천 항목으로 구성된 것을 확인할 수 있으며 이를 토대로 저널 구독 의사결정을 위한 근거 자료로서 활용될 수 있다.

우리기관 추천 저널 목록

NO	서명	이용건수	구독여부
1	Diamond and related materials (Elsevier / 0925-9635)	106	미구독
2	Thin solid films (Elsevier Sequoia / 0040-6090)	139	미구독
3	Surface & coatings technology (Elsevier Sequoia / 0257-8972)	149	미구독
4	Journal of the European Ceramic Society (Elsevier Science Publishers / 0955-2219)	45	미구독
5	Applied surface science (New York] / 0169-4332)	38	미구독
6	Ceramics international (Ceramurgica / 0272-8842)	22	미구독
7	Infection and immunity (American Society for Microbiology / 0019-9567)	58	미구독
8	Ceramic engineering and science proceedings (American Ceramic Society] / 0196-6219)	17	미구독
9	Materials science & engineering. properties, microstructure and processing. A, Structural materials (Elsevier Sequoia / 0921-5093)	19	미구독
10	Nuclear instruments & methods in physics research. Section B, Beam interactions with materials and atoms (North-Holland Physics Pub / 0168-583x)	17	미구독
11	Materials science forum (Trans Tech Publications / 0255-5476)	17	미구독
12	Japanese journal of applied physics. Part 1, Regular papers, short notes and review papers (Publication Board, Japanese Journal of Applied Phy / 0021-4922)	16	미구독
13	Journal of nuclear materials = Journal des matériaux nucléaires (North Holland Pub. Co / 0022-3115)	14	미구독
14	Journal of materials processing technology (Elsevier / 0924-0136)	11	미구독
15	Journal of crystal growth (North-Holland Pub. Co / 0022-0248)	24	미구독
16	Chemical vapor deposition : CVD (VCH Verlagsgesellschaft ; / 0948-1907)	11	미구독
17	Key engineering materials (Trans Tech Publications / 1013-9826)	10	미구독
18	Scientific American (Scientific American, inc., etc.] / 0036-8733)	10	미구독
19	Materials letters (North-Holland / 0167-577x)	9	미구독
20	International journal of refractory metals & hard materials (MPR Pub. Services / 0263-4368)	14	미구독

(그림 8) '행복 대학교' 저널 추천 목록
(Figure 8) The recommended journal list of 'Korea university'

제안된 추천 기법을 통해 구성된 저널 추천 항목의 정확도에 대한 비교 성능 평가를 위해 본 연r에서는 추천 기법의 정확도 평가 방법으로 널리 사용되고 있는 NDCG(Nomalized dicounted cumulative gain)을 이용 한다 [22]. NDCG는 이상적인 결과 항목과 추천 알고리즘에 의해 구성된 항목이 유사할수록 큰 값을 가지게 된다. 즉, NDCG 값이 높을수록 추천 정확도가 높은 것을 의미한다. 효율적인 비교 성능 평가를 위해 우리는 제안 기법이외에도 많은 추천 시스템들에서 일반적으로 활용되고 있는 협업 필터링 기법과 이용 횟수에 따른 추천 기법의 3가지 기법들 간 추천 항목 정확도 평가를 실시한다. 여기서, 이용 횟수에 따른 추천 기법은 단순히 이용 횟수가 많은 저널들을 추천 항목으로 구성하는 것이다. 우리는 총 5번의 걸쳐 각 기법들에 추천 정확도 평가를 수행하였으며, 이들 평가 값들의 평균을 취해 최종 NDCG값을 산출하였다. 그림 9는 NDCG 성능 비교 결과 그래프를 보인다. 그림 9에서 보이는 바와 같이, 단순 이용횟수만을 고려한 추천 항목의 정확성 보다는 널리 사용되는 협업필터링 기반의 추천 정확성이 높은 것으로 나타났다.



(그림 9) NDCG 성능 비교
(Figure 9) Performance comparison of NDCG

그러나 본 연구에서 제안하는 기관 유사도와 저널 선호도를 고려한 추천 기법이 기존의 두 기법에 보다 가장 큰 NDCG 값을 가지는 것을 알 수 있다. 이는 K종의 저널 목록을 구성함에 있어 기존의 추천 기법들보다 본 연구에서 제안하는 기법이 기관 이용자가 갖는 저널에 대한 선호에 보다 부합하여 추천 항목을 정확하게 구성한다는 것을 의미한다.

6. 결론 및 향후 연구

학술정보 서비스 요구에 대한 폭발적인 증가와 정보 유통 체제의 발전에도 불구하고 각 기관의 정보 제공 관계자들은 정보 서비스 개발을 위해 요구되는 이용자 수요 분석 등을 위한 기반 데이터를 확보하는데 어려움을 가지고 있다. 때문에 각 기관의 소속된 사용자들의 정보 접근성을 활성화 시키고 개개의 정보 소비요구를 충족시킬 만한 학술 콘텐츠를 제공하는데 많은 한계가 따른다. 또한 각 기관에 소속된 이용자의 정보 소비 요구 및 콘텐츠 확보에 부응하기 위해서는 신규 전자정보원의 지속적인 확대가 매우 중요하지만 이용자들을 만족시킬 만한 신규 전자 정보원을 발굴하는 것 역시 기관 측면에서의 많은 비용을 초래하고 있는 실정이다.

본 논문에서는 학술 콘텐츠의 정보 접근성을 보다 효과적으로 활성화 시키고 개개인의 학술 정보 소비요구를 충족시키는 것을 목표로 기관 유사도와 저널 유사도를 고려한 정보 소비자 선호도 예측 기법을 제안하였다. 또한 정보 소비자의 선호도를 보다 정확하게 예측해내기 위하여 우리는 다수의 정보원 및 다수의 학술 콘텐츠 서비스 사이트들로부터 발생된 학술 이용로그와 각 기관의 전자 저널 구독 정보를 수집하고 기관별로 관리하는 학술 콘텐츠 이용 정보 통합 관리 시스템을 설계 및 개발하였다. 구현한 시스템은 각 기관의 전자저널 구독 정보와 이용로그의 수집 및 관리를 통해 저널 선호도 예측을 위한 기반 데이터를 구축 및 제공하며, 제안한 기관 유사도와 저널 유사도를 고려한 저널 추천 기법을 통해 기관 단위의 선호에 부합하는 기관 맞춤의 저널 추천 항목을 최종 구성한다. 우리는 기존 추천 기법들과의 비교 성능 평가를 실시하여 제안 기법을 통해 기관 이용자가 갖는 저널에 대한 선호에 보다 부합하여 추천 항목을 구성함으로써 맞춤형 저널 추천 서비스에 있어 기존 추천 기법들과의 차별성과 우수성을 보였다.

향후에는 참고문헌의 인용/피인용 관계에 대한 의미정보를 활용하여 저널 선호도 예측의 정확성을 높이기 위한 개인화 기법과 추천 항목에 대한 사용자 만족도를 보다 높이고자 네트워크 분석기반의 저널 랭킹(Ranking) 기법에 대한 연구를 수행할 것이다.

참 고 문 헌 (Reference)

[1] Fotopoulou, Aristea, and Nick Couldry, "Telling the story of the stories: online content curation and digital engagement," Information, Communication & Society, Vol. 18, no. 2, pp. 235-249, 2015. http://doi.org/10.1080/1369118X.2014.952317

[2] NDSL, http://scholar.ndsl.kr/index.do

[3] Won-Sik Shim, "Strategies for Leveraging Usage Statistics of Electronic Resources," Journal of the Korean Society for Information Management, Vol. 22, no. 2, pp. 5-21, 2005. http://doi.org/10.3743/KOSIM.2005.22.2.005

[4] Chenguang Pan, Wenxin Li, "Research Paper Recommendation with Topic Analysis," International Conference On Computer Design And Applications (ICCDA 2010), Vol. 4, pp. 264-268, 2010. http://dx.doi.org/10.1109/ICCDA.2010.5541170

[5] Masashi Shimbo, Takahiko Ito, Yuji Matsumoto, "Evaluation of Kernel-based Link Analysis Measures on Research Paper Recommendation," JCDL '07 Proceedings of the 7th ACM/IEEE-CS Joint Conference, pp.354-355, 2007. http://dx.doi.org/10.1145/1255175.1255245

[6] Badrul Sarwar, George Karypis, Joseph Konstan, and John Riedl, "Item-Based Collaborative Filtering Recommendation Algorithms," WWW10, pp. 285-295, May 1-5, 2001. http://dx.doi.org/10.1145/371920.372071

[7] Lee, Joonseok, et al., "Personalized Academic Paper Recommendation System," 2015. https://arxiv.org/abs/1304.5457

[8] Steven Rosenbaum, "The Curation Explosion, And Why Humans Still Trump Tech", http://www.forbes.com/sites/stevenrosenbaum/2015/07/26/the-curation-explosion/#6c9f0aed231c.

[9] Gilbert, Eric, et al., "I need to try this?: a statistical overview of pinterest," Proceedings of the SIGCHI conference on human factors in computing systems. ACM, pp. 2427-2436, 2013. http://dx.doi.org/10.1145/2470654.2481336

[10] Kimura, Akisato, et al., "Image context discovery from socially curated contents," Proceedings of the 21st ACM international conference on Multimedia, ACM, pp. 565-568, 2013. http://dx.doi.org/10.1145/2502081.2502149

[11] Yang-Ha Park, "A Study on the Book Recommendation Standards of Book-Curation Service for School Library,"

Journal of Korean Library and Information Science Society, Ver. 47, no. 1, pp. 279-303, 2006. http://www.riss.kr/link?id=A101821207

[12] Hyunho Lee, Wonjin Lee, “A Study on the Design of Curation System of Customized Sport Convergence Contents for Activation of Sport for All,” Journal of Korea Multimedia Society, Vol. 19, no. 2, pp. 396-404, 2006. http://dx.doi.org/10.9717/kmms.2016.19.2.396

[13] Kwon, Jaekwang, Choi, Sungwoo, Yu, Jehyun, Jung, Inyoung, Jung, Byunghee. “Social curation service with broadcasting contents,” proceeding of Korea Broadcast Engineering, pp. 187-190, 2015.

[14] Ashcroft, Linda, ″Developing competencies, critical analysis and personal transferable skills in future information professionals,″ Library review, Vol. 53, no. 2, pp. 82-88, 2004. http://dx.doi.org/10.1108/00242530410522569

[15] Baker, Gayle, and Eleanor J. Read. ″Vendor supplied usage data for electronic resources: a survey of academic libraries,″ Learned Publishing Vol. 21, no. 1, pp. 48-57, 2008. http://dx.doi.org/10.1087/095315108X247276

[16] KESLI, http://kesli.or.kr

[17] NOS, http://nos.ndsl.kr

[18] WiseCat, http://wisecat.ndsl.kr

[19] Benesty, Jacob, et al, ″Pearson correlation coefficient,″ Noise reduction in speech processing, Springer Berlin Heidelberg, pp. 1-4, 2009. http://dx.doi.org/10.1007/978-3-642-00296-0_5

[20] Hartigan, John A., and Manchek A. Wong, ″Algorithm AS 136: A k-means clustering algorithm,″ Journal of the Royal Statistical Society. Series C (Applied Statistics), Vol. 28, no. 1, pp. 100-108, 1979. http://dx.doi.org/10.2307/2346830

[21] Foote, Jonathan T., ″Content-based retrieval of music and audio,″ Voice, Video, and Data Communications. International Society for Optics and Photonics, APA, pp. 138-147, 1997. http://dx.doi.org/10.1117/12.290336

[22] MLA Shani, Guy, and Asela Gunawardana. ″Evaluating recommendation systems,″ Recommender systems handbook. Springer US, pp. 257-297, 2011. http://dx.doi.org/10.1007/978-0-387-85820-3_8

◐ 저 자 소 개 ◑

이 해 성 (Hae-sung Lee)
2003년 경기대학교 전자계산학과 졸업(학사)
2010년 경기대학교 대학원 컴퓨터학과 졸업(석사)
2015년 경기대학교 대학원 컴퓨터학과 졸업(박사)
2015~현재 한국과학기술정보연구원 박사후연구원
관심분야 : 데이터베이스, 빅데이터, 데이터마이닝
E-mail : hae2425@kisti.re.kr

김 순 영 (Soon-young Kim)
1992년 충남대학교 문헌정보학과 졸업(학사)
2006년~현재 한국과학기술정보연구원 선임연구원
관심분야 : 메타데이터 레지스트리, 디지털 정보자원
E-mail : maya@kisti.re.kr

김 재 훈 (Jay-hoon Kim)
1999년 연세대학교 문헌정보학과 졸업(학사)
2008년 성균관대학교 경영대학원 졸업(석사)
2006년~현재 한국과학기술정보연구원 선임연구원
관심분야 : 도서관 컨소시엄, 디지털 정보자원
E-mail : jay.kim@kisti.re.kr

김 정 환 (Jeong-hwan Kim)
2009년 충남대학교 문헌정보학과 졸업(박사)
2006년~현재 한국과학기술정보연구원 책임연구원
관심분야 : 도서관 컨소시엄, 정보자원 공동수집 및 공동활용, 디지털 정보자원 개발,
E-mail : kimjh@kisti.re.kr