

# 추천시스템과 소셜 네트워크를 융합한 지능형 연구자연결망 구축

## A Study of the Intelligent Researcher Connection Network Build-up that Merges the Recommendation System and Social Network

이 충 무\* · 이 상 기\*\* · 이 병 섭\*\*\*

Choong-Moo Lee · Sang-Gi Lee · Byeong-Seop Lee

### 차 례

- |              |                |
|--------------|----------------|
| 1. 서 론       | 4. 연구자연결망 구축현황 |
| 2. 관련 연구     | 5. 결론 및 발전방향   |
| 3. 연구범위 및 구성 | • 참고문헌         |

### 초 록

개방, 참여, 공유에 기반한 웹 2.0 사상이 다양한 분야로 빠르게 확산되고 있으며, 개인화 기능인 추천시스템과 소셜 네트워크에 관한 연구가 활발하다. 추천시스템과 소셜 네트워크의 경우 각기 다른 영역에서 독자적으로 발전하여왔으며 이들을 융합한 새로운 형태의 서비스모델에 대한 연구는 미미한 실정이다. 본 논문에서는 이용로그, 개인프로파일, 북마크 등을 분석하여 자신과 유사한 분야를 연구 하는 연구자를 지능적으로 추천하여 효율적으로 소셜 네트워크를 구성할 수 있는 연구자연결망을 구축 하였다. 이를 통해 기존 소셜 네트워크 문제인 초기 네트워크 구축의 어려움과 이용자들의 소극적인 참여로 인한 네트워크 확장성 문제를 해결하였다.

### 키 워 드

연구자연결망, 소셜 네트워크, 추천시스템, 웹 2.0 패러다임, 연계 및 융합

\* 과학기술연합대학원 응용정보과학 석사과정  
(M.S. Dept. of Practical Information Science, UST, coolerian@kisti.re.kr)  
 \*\* 한국과학기술정보연구원 정보서비스실 선임연구원  
(Senior Researcher, Dept. of Information Service, KISTI, sklee@kisti.re.kr)  
 \*\*\* 아르고넷 개발팀장  
(Development Team Leader, ARGONET, bslee73@gmail.com)  
 • 논문접수일자: 2009년 2월 18일  
 • 게재확정일자: 2009년 3월 24일

## ABSTRACT

The web 2.0 concept rapidly spreads to the various field which is based on an opening, the participation, and a share. And the research about the recommendation system, that is the personalize feature, and social network is very active. In the case of the recommendation system and social network, it had been developing in the respectively different area and the new research toward the service model of a form that it fuses these is insignificant. In this paper, I'm going to introduce efficient social network which is called the researcher connection network. It is possible to recommend the researcher intellectually who studies the similar field by analyzing the usage log and user profile. Through this study, we could solved the network expandability problem which is due to the user passive participation and the difficulty of the initial network construction that is the conventional social network problem.

## KEYWORDS

Researcher Connection Network, Social Network, Recommendation System, Web 2.0 Paradigm, Linking and Merging

## 1. 서론

정보기술과 인터넷의 발달로 정보가 폭발적으로 증가하고 있으며, 수많은 정보 속에서 정작 자신에게 필요한 지식정보는 쉽게 찾을 수 없는 정보 홍수시대에 살고 있다. 최근 웹 2.0에서는 개인화서비스를 통해 수많은 정보 중에서 개인에게 맞는 맞춤형 정보를 필터링하여 서비스하기 위한 다양한 방법들을 연구하고 있다. Amazon.com과 같은 전자상거래 사이트에서 살 물건을 쇼핑하다 보면, 자신의 취향에 맞는 추천 상품 목록이 나타나는 것을 볼 수 있다. Amazon.com 외에도 디지털 TV 녹

화기인 Tivo나 DVD 대여 서비스인 Netflix 등도 비슷한 추천서비스를 하고 있다. 1990년대부터 본격적으로 연구되기 시작한 자동 추천시스템의 알고리즘은 협력적 여과 방식이다. 이 방식은 개인의 프로파일을 분석하여 비슷한 성향을 가진 다른 사용자들이 구입한 품목을 추천하는 방식이다. 또한, 개인들이 인터넷에 본격적으로 참여하면서 자신과 관심 분야가 같거나 유사한 분야를 연구하는 사람들과 관계를 형성하여 정보를 공유하고 협력하는 모델인 소셜 네트워크가 빠르게 확산되고 있다. 하지만 기존 소셜 네트워크는 대부분 이용자들이 직접 관계를 맺고 싶은 특정 대상을

선택하여 이용자 스스로 소셜 네트워크를 확장해 나가는 방식이다. 이러한 방식은 이용자들에게 자율성을 부여하면서 형식에 얽매이지 않고 개성이 강한 네트워크를 구성할 수 있는 장점도 있지만, 초기에 네트워크를 구축하기 매우 어렵고 활발히 참여하지 않는 사용자들의 경우 네트워크를 구성할 수 없는 한계가 있다. 이런 단점을 극복하기 위해 본 논문에서는 이용자들이 직접 관계를 맺고 싶은 대상을 찾아 네트워크를 구성할 수 있을 뿐만 아니라 같은 논문이나 특허정보를 본 사람이나 관심 분야가 같은 사람을 자동 분석하여 추천하고 소셜 네트워크를 형성할 수 있는 지능형 연구자 연결망을 개발하였다.

## 2. 관련 연구

### 2.1 소셜 네트워크 1.0

소셜은 사회라는 영어단어에서 온 것으로 사람들이 모여 있다는 의미를 가진다. 네트워크는 사람들이 일련의 관계에 의해 모인 관계망을 의미한다. 이 둘의 조합인 소셜 네트워크는 사람들이 연결되어 있는 관계망으로 표현할 수 있다. 사회네트워크란 행위자가 선택하는 행위의 결과이기도 하며, 동시에 행위자의 차후 행위 선택을 제약하는 역할을 한다. 말하자면, 관계망으로서의 사회네트워크는 행위자의 선택과 그 선택이 주는 제약을 담고 있는

구체적인 공간인 것이다. 특히 네트워크가 그 안에 속한 행위자들의 행위를 제약하는 구조로서 역할을 한다는 점에 주목하면 구조라는 실체가 불분명한 개념을 네트워크라는 구체적 실체로 옮겨놓은 것이다(Barnes, Susan B. 2001).

사회학에서 연구하고 있는 사회연결망이론(소셜 네트워크 이론)은 인간이 여러 사람들과 관계를 맺어가는 것으로 인간관계의 행위(action)와 구조(structure)의 상호역동성을 설명하는 이론이다. 이는 행위자들의 관계를 중심으로 무원칙하고 혼란스러우며 복잡해 보이는 이면에 존재하는 패턴을 파악하려는 시도로 이해할 수 있으며, 따라서 사회네트워크 분석(SNA : Social Network Analysis)은 일정한 사람들 사이의 특정한 연계 전체의 특성으로 연계에 포함된 사람들의 사회적 행위를 설명하려는 것이다. 또한 최근 온라인을 중심으로 등장하고 있는 소셜 네트워크는 인맥구축, 사회연결망, 지인네트워크 등으로 소개되고 있지만 이는 앞서 설명한 사회연결망 이론에 뿌리를 두고 있는 것이다.

소셜 네트워크 즉, 인맥 서비스는 오프라인 상에서 과거부터 존재해왔으며 아직 남아있는 것도 있다. 사람들의 인맥은 지연, 학연, 혈연에 기반한다. 이런 사회 환경으로 인해 지연, 학연, 혈연에 대한 소셜 네트워크는 오래 전부터 있어왔으며, 향우회(지연), 동창회(학연), 종친회(혈연)가 바로 그것이다. 결혼과 관련해서는 중매인을 예로 들 수 있다. 중매인은 결혼

적령기의 남녀에 대한 정보를 가지고 이를 통해 남녀의 만남을 주선하는 서비스를 제공한다. 이와 비슷한 형태의 서비스를 제공하는 결혼정보업체도 대표적인 오프라인 소셜 네트워크다. 다만, 결혼을 원하는 사람들이 자발적으로 가입하여 이뤄지는 것이 다르다. 펜팔도 소셜 네트워크의 또 다른 사례다. 편지를 주고받길 원하는 사람들을 중개해주는 역할을 하는 것이 펜팔 업체의 임무다. 이러한 오프라인 소셜 네트워크는 지금도 여전히 존재하며, 온라인으로 전환하는 경우도 많다. 하지만 사회 변화에 따라 사라져가는 사례도 있다. 하지만 이러한 모델들이 응용되어 본격적인 소셜 네트워크로 발전하는 경우도 더러 있다.

## 2.2 소셜 네트워크 2.0

지금과 같은 소셜 네트워크 형태는 웹이 활성화되기 시작한 2000년 초에 등장하기 시작하였다. 이들은 사용자들이 오락적, 혹은 비즈니스적인 인적 네트워크를 구성하는 데 도움을 주었다. 이들은 유즈넷이나 펜팔사이트 등 소셜 네트워크 형성 그 자체를 목표로 한 오프라인 소셜 네트워크와 달리 소셜 네트워크 형성을 다른 비즈니스 모델을 구현하는 데 필요한 수단으로 활용하기 시작하였다. 이 소셜 네트워크 2.0에 해당하는 서비스는 Friendster(www.friendster.com), MySpace(www.myspace.com), Facebook(www.facebook.com), del.icio.us(delicious.com), Freechal(www.

freechal.com), iloveschool(www.ilove school.co.kr), cyworld(www.cyworld.com) 등이다. 이런 다양한 소셜 네트워크 사이트는 niche형, matching형, 블로그형, 폐쇄형, 영상 중심형으로 분류할 수 있다.

Niche형 소셜 네트워크는 특정그룹 전용의 커뮤니티로 활용되는 소규모 특화형 사이트이다. 애완견과 고양이를 다루는 Dogster(www.dogster.com), Catster(www.catster.com), 가족을 주제로 하는 Farmster(www.famster.com), 유아부모를 대상으로 하는 TotJot, 다이어트를 주제로 하는 Traineo(www.traineo.com), 그리고 여러 가지 소 관심사를 주제로 사람들이 모이는 Daum 및 Naver cafe 등은 niche형 소셜 네트워크로 볼 수 있다.

Matching형 소셜 네트워크는 미국의 LinkedIn(www.linkedin.com), 영국의 Ecademy(www.ecademy.com) 등이 있는데 이들은 비즈니스에 특화된 사교 및 교류장소를 제공하여 비즈니스 매칭기회를 창출하고, 인재소개 서비스, 비즈니스 교류, 스쿨 동창회, 기업 OB 활용 등을 제공한다.

블로그형 소셜 네트워크는 2002년에 생긴 FriendSter, MySpace, Facebook 등이 대표적이다. Friendster는 초창기에 엄청난 인기를 끌었지만, 사용자들이 인적 네트워크를 구축한 다음 뚜렷한 할 거리를 제공하지 않고 그대로 방치했기 때문에, 무엇인가 할 거리를 제공한 MySpace에 자리를 내주고 말았다. 그리고 2004년 오픈한 Facebook은 처음에는 Havard

〈표 1〉 온라인 소셜 네트워크 분류

대분류(수)	중분류(특성)	서비스명	내용
개인 중심 커뮤니티	관계 중심	MSN 메신저 등	채팅, 메신저 서비스
		미니홈피, MSN 홈페이지	인맥중심 개인연결망
	관심 중심	블로그	정보중심 개인연결망
대중·그룹 커뮤니티	관계 중심	아이러브스쿨	학연중심 그룹연결망
	관심 중심	Daum 카페, Naver 카페 등	정보중심 그룹연결망

에서만 서비스하였으나, 점차 Ivy league, 고등학생 네트워크, 기업 네트워크로 확대하다가 2006년 9월 전면 개방하였다. 특히, Facebook은 소셜 OS를 표방하며, Third party를 통해서 신규 서비스를 창출하고 있으며 캐나다, 노르웨이, 이집트, 요르단, 레바논, 파나마, 남아프리카 등에 진출하였다.

폐쇄형 소셜 네트워크로는 일본의 Mixi (<http://mixi.jp>), YUCASEE(<http://yucasee.jp>) 등이 있다. Mixi는 미리 형성되어 있던 인간 관계나 커뮤니케이션을 기반으로 초대를 통해 가입한다. 일정한 신상정보를 가지고 있는 사람끼리 연결되는 장소를 제공하고 있다. 그리고 YUCASEE는 순 금융자산 1억엔 이상의 부유층을 위한 폐쇄형 소셜 네트워크로서 Abraham Holdings가 개설하여 그들이 제휴한 회사에서 추천 받은 사람만 회원가입이 가능하다. YUCASEE의 역할은 소셜 네트워크, 회원 간 질의응답, 입소문, 최신기사를 제공하고 있다.

영상 중심형 소셜 네트워크는 웹에서 제공하는 영상을 등록, 공유할 수 있는 서비스로서

YouTube가 대표적이다(김명숙 2007).

온라인 소셜 네트워크는 연결망의 형태 혹은 크기에 따라 개인중심 커뮤니티와 대중그룹 커뮤니티로 크게 분류할 수 있으며, 이는 다시 개인 간 인맥의 형성 및 유지, 강화를 목적으로 하는 관계 중심 커뮤니티와 정보를 중심으로 연결되어 있는 관심 중심 커뮤니티로 나눌 수 있다(〈표 1〉 참조). 현재 제공되고 있는 서비스들 가운데 개인 중심 커뮤니티는 개인간 실시간 커뮤니케이션을 지원하는 메신저, 개인 홈페이지이며, 관심 중심의 서비스는 블로그를 들 수 있다. 물론 블로그는 사용하는 용도에 따라 관계 중심으로 분류할 수도 있으나, 대부분의 블로그 제공 사이트들이 제공하고 있는 RSS, 트랙백 등의 기능은 개인의 관계보다는 정보서비스에 중점을 두고 있다.

### 2.3 소셜 네트워크 도서관 적용사례

OCLC(Online Computer Library Center)의 설문조사 내용에 따르면 도서관은 소셜 네

트위크를 운영할 필요가 없다는 결과가 나왔다. 단지 전체의 14%만이 도서관이 이용자를 위하여 소셜 네트워크를 운영할 필요가 있다고 대답했으며, 도서관이 소셜 네트워크 사이트를 운영한다면 13%정도가 참여하겠다고 응답하였다(OCLC 2007a). 그러나 OCLC는 도서관이 주도하여 소셜 네트워크를 결부시키지 못했기 때문에 도서관의 기능과 소셜 네트워크의 결합이 이루어지지 않았다고 하였다(OCLC 2007b). KERIS의 라이브러리 2.0에 관련된 설문 조사를 보면, 소셜 네트워크의 개념을 알고 있는 응답자는 전체의 6% 밖에 없었다. 하지만 커뮤니티 서비스를 비슷한 관심 분야를 지닌 이용자들에게 제공해야한다는 의견이 21%를 차지하였다(이지연, 주수형, 민지연 2007).

이 수치는 이용자들 간 네트워크가 정보 수요가 적지 않음을 나타낸다.

도서관은 소셜 네트워크 기능을 통해 정보 수요자에게 조금 더 심도 있는 경험을 제공할 수 있으며 도서에 대한 리뷰에 메타아키텍처를 결합하면 그 책에 대한 정보의 교류가 가능하다고 Fred Stutzman은 말하였다. 많지는 않지만 몇 군데 도서관과 정보서비스 기관에서 정보수요자들과 비슷한 분야에 관심을 두고 있는 다른 정보수요자들을 연결해주는 서비스를 진행하고 있다(OCLC 2007b).

최근에는 게시판, 채팅, 블로그 등을 활용하여 도서관이 적극적으로 이용자와 대화를 시도하고 있다. <표 2>에서와 같이 켄터키 대학 도서관은 MySpace를 거의 홍보 수단으로

<표 2> 상용 소셜 네트워크를 활용한 도서관 서비스 사례(조재인 2008)

도서관명	서비스이름	소셜 네트워크	주요내용
켄터키 대학	오픈 라이브러리	MySpace	<ul style="list-style-type: none"> <li>- 온라인 계정에 도서관 연락처, 사진, 장서정보 등 등록</li> <li>- 학생들에게 가시적으로 도서관 서비스를 어필할 수 있는 홍보 수단으로 평가됨.</li> </ul>
텐버 공공도서관	틴 스페이스	MySpace	<ul style="list-style-type: none"> <li>- 10대들과 함께 할 수 있는 아이디어를 소셜 네트워크를 통해 실천</li> <li>- 도서관서비스 뿐 아니라, YouTube 동영상 콘테스트, 이용자 리뷰, 북마크 디자인 콘테스트, 사진 콘테스트 등 각종 이벤트를 통해 10대 유인</li> <li>- 마이스페이스에 계정 생성 후, 10대 이용자 등록률 41% 증가</li> </ul>
OCLC	Facebook Worldcat	Facebook	<ul style="list-style-type: none"> <li>- OCLC 검색 API를 설치하여 WorldCat을 검색할 수 있도록 하고 신규도서 추천 등 각종 부가기능 제공</li> <li>- 이용자가 WorldCat 홈페이지에 설정한 "Favorite World Cat List"를 Track할 수 있는 기능 제공</li> <li>- 다른 이용자에게 WorldCat API 설치를 추천하는 기능 등을 제공</li> </ul>

사용하고 있고, 덴버 공공도서관도 MySpace에서 '틴 스페이스'라는 서비스를 통하여 각종 콘텐츠 등으로 이벤트를 열어서 10대 이용자들이 많이 찾도록 하였다. 하지만 이는 홍보수단에 가까운 정도이다. 한편 OCLC는 Facebook에 WorldCat을 검색할 수 있도록 검색 API를 설치하였다. 앞의 두 도서관처럼 단순한 홍보수단은 넘어서는 서비스이긴 하지만 상용 소셜 네트워크 서비스의 연결망을 이용하고 있는 점에서 아직 미미한 수준이라고 볼 수 있다. 위와 같이 상용 소셜 네트워크 서비스를 이용한 도서관과 다르게 독자적인 소셜 네트워크 서비스를 구축하는 도서관도 조금씩 늘어나고 있다. 한 예로 OCLC는 WorldCat Community 서비스를 들 수 있다. WorldCat Community에서는 서지아이템의 상세정보, 평점과 리뷰 등을 이용자들이 작성할 수 있다(조재인 2008).

몇몇 도서관은 이용자들이 단순히 도서관과 이용자 간의 단순한 연결로 정보를 얻는 방법이 아닌 다른 이용자들과도 정보를 주고받으며 그 안에서 더욱 양질의 정보를 제공할 수 있는 연결망 구축을 진행하고 있다. 상용 소셜 네트워크 서비스도 계속해서 변화하고 있으며 도서관에서 소셜 네트워크를 접목하는 서비스도 이제 시작 단계이다. 아직 도서관에서 소셜 도구를 사용하는 것이 단순 홍보용에 그치기 때문에 서비스 제공 방법에 대하여 앞으로 많은 고민과 실험으로 더욱 양질의 정보를 제공할 수 있기를 기대 한다.

## 2.4 학술연구자를 위한 소셜 태그 및 북마크

연구자는 연구 시작단계에서부터 연구와 관련된 다양한 참고정보원을 수집하여 관리한다. 다양한 리소스를 통해 필요한 자료를 수집하고, 북마크한 콘텐츠에는 각종 주석과 리뷰를 생성한다. 그리고 수집된 자료는 참고문헌 형식으로 반출하여 관리한다. 최근 소셜 태그, 북마크는 인터넷 정보자원을 대상으로 다양한 참고정보원을 수집·관리하는 데 활용되고 있으며, 더 나아가 동일한 관심사를 가지는 연구자가 수집한 정보를 공유하는 훌륭한 도구가 되고 있다.

Connotea, CiteULike는 태그를 중심으로 한 소셜북마크 사이트이지만, 공통된 관심사를 공유한 타인과 메시지를 교환하거나 직접적인 커뮤니케이션을 시도할 수 있으며, 더 나아가 온라인 커뮤니티나 그룹을 생성할 수 있다. <표 3>과 같이 기본 기능은 같지만 Connotea와 CiteULike는 기본 기능은 같지만 조금의 차이점들이 있다. Connotea가 운영하는 연구자 커뮤니티는 범람하는 인터넷 콘텐츠에서 의미 있는 자원을 선별하여 동료 연구자 간에 공유하기 위하여 운영하고 있다. 위키를 기반으로 커뮤니티 참여자가 그룹 안내 페이지를 작성·갱신하고 있으며, 커뮤니티 페이지에서는 관련된 태그 리스트와 소속된 멤버에 관한 정보가 제공된다. CiteULike는 열린 접근 기반의 참고문헌 관리도구로서 2004년 11월 Richard Cameron이 개발하여 개인적으로 운영하다가 현재 Springer의 후원을 받고 있다.

〈표 3〉 Connotea, CiteULike의 주요특징(황혜경, 이재운 2006)

구분	Connotea	CiteULike
운영주체	네이처 출판국	리처드 카멜론 개인운영
기본 기능	<ul style="list-style-type: none"> <li>- 개인별 북마크 관리 및 소셜태깅</li> <li>- 메타데이터 자동추출 및 참고문헌 관리</li> <li>- 연구자 커뮤니티 서비스</li> </ul>	
검색방식	포털형 검색방식	서지 검색방식
콘텐츠	공식출판물과 블로그 등 비공식 출판물 망라	검증받은 공식 출판물

CiteULike의 기본적인 취지는 모든 저자들이 한 장소에서 그들이 인용하고자 하는 모든 논문을 자유롭게 얻을 수 있기를 바라는 마음에서 설계되었다. CiteULike의 연구자 커뮤니티는 Connotea 서비스와 유사하지만, 회원정보를 더 많이 노출하고 있으며, CiteULike는 학술지 및 학술 회의록에 발표된 연구논문, 보고서, 전문도서 혹은 석·박사 학위논문과 같이 논문심사를 거친 검증받은 자원을 대상으로 한다. 연구자가 검증받지 않은 기타 정보자원을 북마크하여 관리를 할 수 없는 것이 아니지만, 이런 자원들은 CiteULike의 서포트 리스트에는 공개되지 않고 단지 연구자 개인의 라이브러리에서만 관리될 수 있도록 처리하고 있다. 커뮤니티 멤버의 프로필, 블로그 뿐만 아니라 이 연구자가 어떠한 자원을 어떠한 태그로 북마크하고 있는가에 대한 정보를 시간대별로 표시하고 있다(황혜경, 이재운 2006).

## 2.5 Netflix와 Amazon.com의 추천시스템

Netflix([www.netflixprize.com](http://www.netflixprize.com))는 DVD 대여와 온라인 영화감상을 주된 사업으로 하는

미국회사로 1997년에 설립되었다. 1999년에 월정액 모델로 변경한 이후 2000년부터는 완전한 월정액 서비스만을 제공하고 있으며 가입자 820만 명에 약 5,500만 개의 DVD를 보유하고 있다. Netflix는 자사의 성공에 자체 개발한 영화추천 서비스 Cinematch가 중요한 역할을 했다고 주장한다. 대체로 처음 서비스를 이용하는 사람들은 자신이 예전부터 보고 싶었던 영화들을 임대한다. 그 후에는 주로 신작, 특히 화제작이나 인기작을 중심으로 임대 신청을 하는데, 이렇게 되면 짧은 기간 동안 대량의 DVD를 구매해야 하고 이들이 실제로 유통되는 기간도 매우 짧아지는 문제가 있다. 따라서 사업의 핵심은 발매된 지 오래된 DVD이지만 이전 임대 기록과 영화에 대한 평가를 바탕으로 해서 이 오래된 영화 타이틀을 임대하도록 만드는 데 있다. Netflix는 이런 측면에서 매우 성공적이었고 현재 매일 발송하는 190만 장의 DVD는 수만 개의 타이틀로 이뤄져 있다고 한다. Netflix에 방문하면 4GB가 넘는 방대한 데이터를 다운로드할 수 있는데, 여기에는 1만7,770개의 영화 타이틀에 대한

정보와 48만189명의 사용자가 등록한 1억48만507개의 평가 자료가 들어 있다. 자체 개발한 추천 시스템인 Cinematch는 이 데이터베이스를 기초로 학습되었고 시스템의 예측 오류를 평가하는 데 많이 사용되는 RMS라는 평가지표를 사용했을 때 약 0.9525의 오류를 보인다고 한다. 다시 말해 이 시스템이 어떤 사용자가 어떤 영화에 대해 평점 3.7을 줄 것 같다고 예측하면 실제 사용자는 거의 대부분 3점이나 4점을 주고, 2점이나 5점 같은 터무니없는 경우는 거의 발생하지 않는다는 것이다. 이 정도면 상당히 정확한 예측시스템이다(김종규 2008).

Jeffrey P. Bezos 는 요미우리 신문과의 인터뷰를 통해서 “고객이 Amazon.com을 이용할 때마다 좋은 경험으로 인식시키는 게 중요하다. 성장을 위해서는 두 가지 방법이 있다. 하나는 열심히 광고를 해서 자신들의 서비스가 지닌 장점을 호소하는 것이고, 다른 하나는 입소문을 통해 사용자의 수를 늘리는 것이다”라고 언급하였다(모리켄 2008). 이 두 가지 실천방법으로는 추천시스템과 지원프로그램을 들 수 있다. 상품 라인업의 확대, 신속한 배송 체제의 확립이 충실한 고객서비스를 의미한다면, 웹에서의 기술개발은 사용자를 끌어들이기 위한 이윤으로 연결시키는 시스템이라 할 수 있다.

추천시스템 기능은 어떤 상품을 구입하려는 사용자에게 “이 물건을 구입한 사람은 이런 상품들도 구입하였습니다.”라는 말과 함께 다른 상품을 추천하는 기능이다. Amazon.com

은 상품 페이지를 관람하기만 해도 반드시 그와 관련된 다른 상품을 추천한다. 예를 들면, 뇌는 만남을 통해 성장한다라는 책을 구입하려는 경우에는 뇌를 키운다 등 6권의 책이 소개된다. 이런 식으로 자동적으로 표시된 상품들은 선택한 상품과 연관성이 많은 것들이어서 사용자들의 구매 욕구를 자연스럽게 불러 일으킨다. 이러한 추천시스템은 데이터베이스의 조합을 통해 가능해진다. 추천을 실현하는 알고리즘에는 크게 세 가지가 있다. 협조 필터링(collaborative filtering), 클러스터 모델(Cluster Models), 검색 베이스 방식(Search-Based Methods)이다. Amazon.com 개발 담당자에 의하면 Amazon.com의 기능은 이 세 가지 기능과는 약간의 차이가 있다고 한다. 즉, 상품 간 협조 필터링(Item to Item Collaborate Filtering)이라고 지칭하였다. 하지만 이 세 방법은 몇 가지 문제점이 있다.

우선 협조 필터링은 같은 상품을 구매한 사용자와 관련성이 높은 다른 상품을 추천하는 방식이다. 상품 수가 100만 종에 달하고 사용자 수가 4,700만 명에 달하는 Amazon.com 입장에서는 이 조합을 0.5초 이내에 하기란 사실상 불가능하다. 사용자의 속성에 따라 유사한 상품별로 구분하는 클러스터 모델은 다양한 상품을 구매한 사용자의 속성을 정의하는 것 자체가 힘들며, 사용자 입장에서 유익한 추천 상품을 끄집어내기가 쉽지 않다. 검색 베이스 방식은 검색된 상품에 관련된 키워드 혹은 상품의 속성에 따라 다른 추천 상품을 표시할

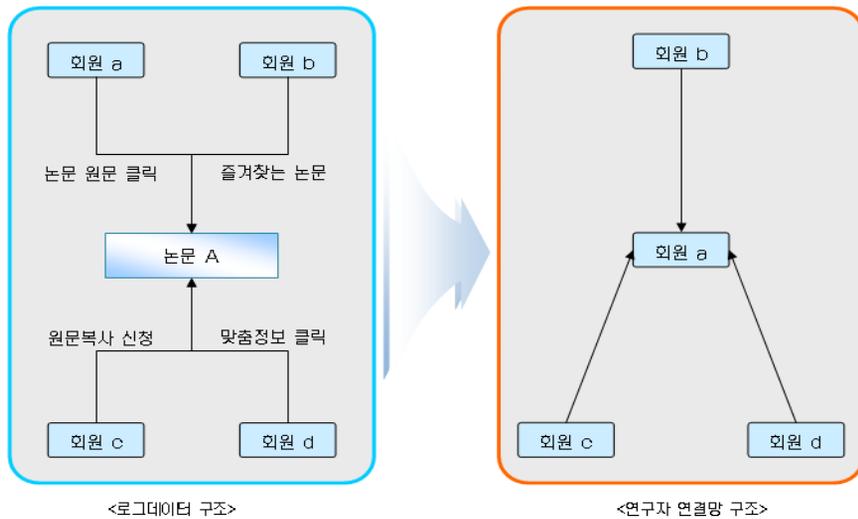
수 있지만, 상품의 개요를 데이터베이스로 정리해야 하는 번거로움이 있다. 게다가 추천 상품 은 베스트셀러 상품이거나 완전히 똑같은 키워 드를 가진 상품으로 한정되는 경향이 있다.

이를 극복하기 위해 Amazon.com의 추천 시스템은 2단계 과정을 거친다. 첫 번째는 사용 자가 구입하려고하는 상품들 간의 matching, 유사성의 수치화 작업이다. 하지만 사용자가 구입한 여러 가지 상품에 반드시 커다란 연관 성이 있지는 않다. 실용서를 검색한 다음에 소 설이나 잡지를 살펴보기도 하고, 과학서적을 찾아본 후 건강기구를 알아보기도 하기 때문 이다. 두 번째는 구입대상 상품과 관련성 있는 모든 상품들의 유사점에 관한 데이터베이스를 참조하는 것이다. 여기서 관련성이란 그 책이 속한 분야, 키워드에 포함된 단어, 저자 이름 을 비롯한 복수의 관련성을 말한다. Ama-

zon.com에서는 이러한 것들과 관련성이 높은 것들을 계산하여 수치가 높은 순서대로 추천 한다(모리켄 2008).

### 3. 연구범위 및 구성

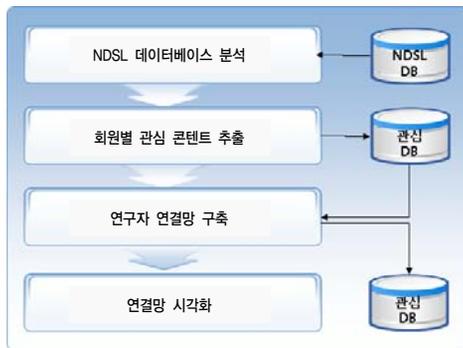
본 연구는 KISTI에서 서비스하는 NDSL 회원에 국한하여 과학기술분야에 종사하는 연구자들을 대상으로 하였으며, 이 연구를 통해 구축되는 소셜 네트워크를 “연구자 연결망”이라 하였다. 연구방법은 NDSL 회원들이 등록된 관심분야와 정보검색 히스토리, 북마크한 콘텐츠 등을 활용하여 같은 주제 분야에 관심을 갖는 다른 회원들을 같은 네트워크로 구성 하였다. <그림 1>은 특정 논문 A를 열람한 회원과 같은 논문을 즐겨찾기에 저장한 회원, 원



<그림 1> 연구자연결망 구조

문복사 신청한 회원, 그리고 메일링 서비스를 통해 상세정보를 본 회원들은 모두 같은 논문에 관심을 갖는다는 점에 착안하여 구성된 연구자연결망 구조이다.

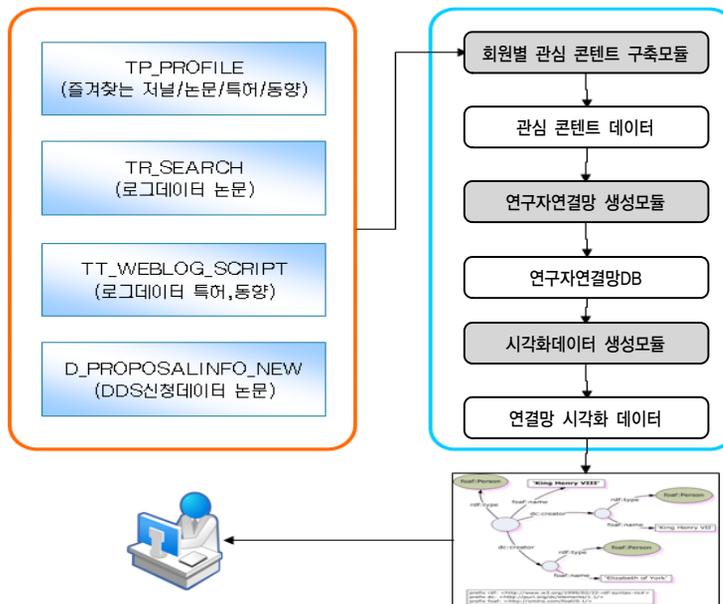
연구자연결망 구축 프로세스를 좀 더 상세하게 기술하면 먼저 회원의 관심분야나 정보를 저장하고 있는 NDSL의 데이터베이스를 분



〈그림 2〉 연구자연결망 구축 프로세스

석하여 데이터를 정제한 후 회원별 관심 DB를 구축한다. 이렇게 구축된 관심 DB에서 각각의 회원별로 같은 콘텐츠에 관심을 갖고 있는 다른 회원들과의 관계도를 분석하여 연구자연결망을 구축한다. 구축된 연구자연결망은 네트워크 다이어그램과 같은 시각화도구를 활용하여 웹으로 표현하였다(〈그림 2〉 참조).

본 논문에서 구현하고자 하는 연구자연결망의 전체적인 구성은 〈그림 3〉과 같다. 〈그림 3〉의 좌측에 있는 4개의 테이블은 연구자연결망을 구축하는 데 기반이 되는 NDSL DB다. 각각의 DB에는 원문복사신청, 로그정보, 즐겨찾기 정보 등이 저장되어 있으며, 이 4개의 DB를 대상으로 관심 콘텐츠를 추출하고, 추출된 콘텐츠를 활용하여 연구자연결망 데이터를 생성한 후 시각화를 위한 XML 데이터를 생성한다.

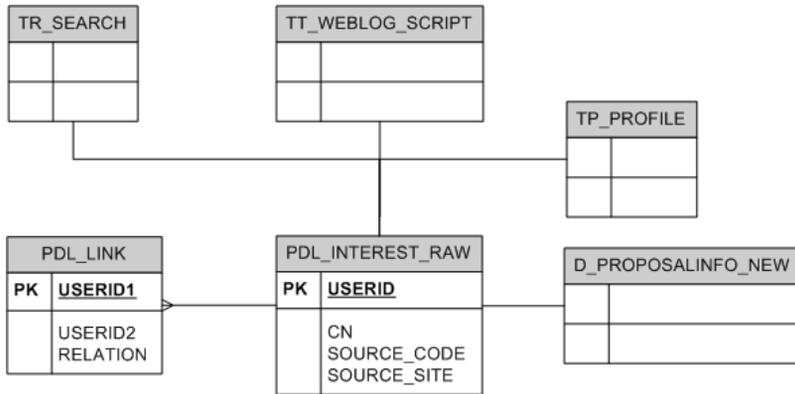


〈그림 3〉 연구자연결망 전체 구성

연구자연결망을 구축하는 데 필요한 데이터를 저장하기 위한 데이터베이스 구조는 <그림 4>, <표 4>, <표 5>과 같으며, 이것은 연구자연결망을 시각적으로 표현하기 위한 중간단계의 데이터를 저장하는 곳이다.

관심 콘텐츠 테이블의 경우 PDL\_INTEREST\_RAW가 되는데 각 회원별로 어떤 콘텐츠에 관심을 갖고 있었는지를 저장하는 테이블이다. CN은 NDSL에서 데이터를 식별하기

위해 사용하는 로컬 식별자이다. SOURCE\_CODE에는 향후 활용을 위해 CN형태로 변형하기 이전에 소스가 되는 테이블에 저장되어 있던 원형데이터를 그대로 저장하였다. SOURCE\_SITE는 이용로그, 맞춤정보, 원문복사 신청, 북마크 등 연구자연결망 구성을 위해 관련 정보를 수집한 원천 소스의 출처에 대한 구분이다. PDL\_LINK 테이블에는 각 회원별로 같은 콘텐츠에 관심을 가진 다른 회원들의 정



<그림 4> 연구자연결망 ERD

<표 4> 관심 콘텐츠 테이블구성

항목명	데이터 형식	항목설명	비고
USERID	VARCHAR2(30)	회원ID	PK
CN	VARCHAR2(100)	ControlNo	
SOURCE_CODE	VARCHAR2(100)	소스 테이블 정보	
SOURCE_SITE	VARCHAR2(2)	소스 테이블 구분	

<표 5> 연구자연결망 테이블구성

항목명	데이터 형식	항목설명	비고
USERID1	VARCHAR2(30)	회원ID 1	PK
USERID2	VARCHAR2(30)	회원ID 2	
RELATION	VARCHAR2(100)	관련콘텐츠 수	

보를 보관하고 있으며, 어느 정도 관련이 있는 지는 RELATION에 빈도수로 관리한다.

#### 4. 연구자연결망 구축현황

NDSL 회원 프로파일, 즐겨찾기 논문, DDS, 이용로그, 한국인 저자 DB를 분석하여 <표 7>

과 같이 관심 DB를 구축하였다. 본 연구에서 사용된 소스데이터는 수시로 추가되는 변동 데이터로 2008년 10월 1일 이전까지 축적된 데이터까지 연구에 활용하였다.

<그림 5>는 구축된 데이터를 기반으로 연구자연결망을 시각화하기 위해 생성한 XML 파일 포맷이다. XML 파일을 각 회원별로 생성하여 플래쉬 형태로 개발된 시각화도구에

<표 7> 연구자연결망 구축현황

구분	구축건수	비고
관심 콘텐츠	총 건수 : 4,145,688 회원 : 208,047	관심분야 : 235,883 개인프로파일 : 205,963 원문이용 : 1,906,944 웹 로그 : 1,796,873
연구자연결망	총 건수 : 6,892,295 회원 : 181,191	총 건수는 같은 콘텐츠를 이용하여 관계가 형성된 회원 노드간 연결건수이다

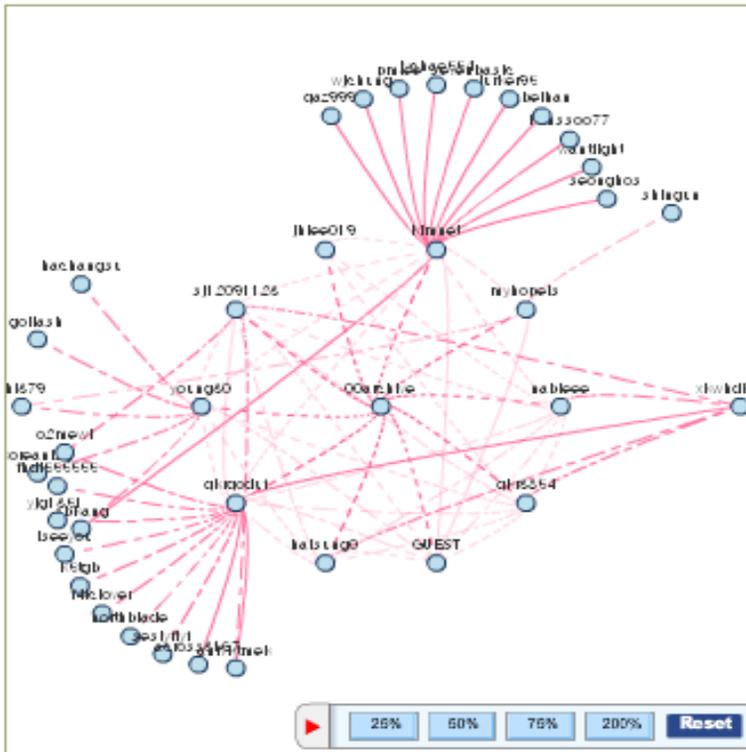
```
<?xml version="1.0" encoding="UTF-8" ?>
- <OntologySet xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
- <ConceptTypeList>
  <ConceptType color="yellow" shape="circle">Concept</ConceptType>
  <ConceptType color="blue" shape="circle">Object</ConceptType>
  <ConceptType color="violet" shape="triangle">DataType</ConceptType>
</ConceptTypeList>
- <RelationTypeList>
  <RelationType color="0xff6699" shape="solid">NT</RelationType>
  <RelationType color="0xff6699" shape="dotted">UF</RelationType>
  <RelationType color="0xff6699" shape="solid-dotted">RT</RelationType>
</RelationTypeList>
- <ConceptFacetList>
  <ConceptFacet type="000001">기기/장치/부속</ConceptFacet>
  <ConceptFacet type="000002">내용</ConceptFacet>
  <ConceptFacet type="000003">물질/재료</ConceptFacet>
  <ConceptFacet type="000004">속성-수량</ConceptFacet>
  <ConceptFacet type="000005">생물</ConceptFacet>
</ConceptFacetList>
- <MenuList>
  <Menu command="REDRAW" target="Background">전체 다시그리기</Menu>
  <Menu command="REDRAW_CONCEPT" target="Concept">컨셉을 중심으로 다시그리기</Menu>
  <Menu command="GETINFO_CONCEPT" target="Concept">컨셉정보 보기</Menu>
  <Menu command="REDRAW_RELATION" target="Relation">관계를 중심으로 다시그리기</Menu>
</MenuList>
- <ConceptList>
- <Concept type="Object">
  <ConceptSID>0196650006</ConceptSID>
  <ConceptFacet>000001</ConceptFacet>
  <ConceptKor>0196650006</ConceptKor>
  <ConceptEng />
  <ConceptLink>0196650006</ConceptLink>
  <ConceptAdditionalInfo />
</Concept>
```

<그림 5> 연구자연결망 시각화를 위한 XML파일

XML파일을 로딩함으로써 비주얼한 연구자연결망으로 시각화가 가능하다.

〈그림 6〉은 구축된 회원별 관심 DB를 기반으로 개발한 회원별 연구자연결망이다. 특정 회원을 중심으로 이 회원과 같은 콘텐츠에 관심을 갖는 회원들을 연결하여 시각화 한 예시 화면이다. 회원은 원으로 표시하였으며, 회원 간의 관계는 빈도수에 따라 20이상은 실선으로, 5~20관계는 굵은 점선으로, 5이하의 관계는 얇은 점선으로 표현하여 연구자들 간 상관관계 정도를 쉽게 알아 볼 수 있도록 처리하였다.

〈그림 7〉은 본 연구자의 개인화 홈페이지를 통해 관련 분야를 연구하는 연구자들을 연계한 연구자연결망이다. 확장성 있게 소셜 네트워크를 설계하기 위해 관심분야를 변경 또는 추가할 수 있도록 설계하였으며, 특정회원을 삭제하거나 신규로 추가할 수 있도록 개발하였다. 아울러 관련 연구자들이 저작한 논문과 최근 관심분야, 인기콘텐츠 등을 쉽게 볼 수 있도록 구성하였다. 부가적으로 회원들의 관심분야와 특히 IPC분류, 국가과학기술표준분류, DDC분류 등 다양한 분류체계 간 매핑을 하였다. 매핑은 국가과학기술표준분류를



〈그림 6〉 연구자연결망 시각화 예시화면



〈그림 7〉 저자의 연구자연결망

기준으로 타 분류체계를 매핑하는 방식으로 개발하였다.

본 연구자연결망을 통해 다른 회원이 등록한 북마크 정보를 공유할 수 있으며, 맞춤정보, 학술논문 태그, 즐겨찾기, 이용로그 등을 활용하여 내가 검색하고자 하는 분야의 전문가를 쉽게 추천하고 연계할 수 있다. 또한 위젯기반의 개인화 기능과 연계하여 관심저널, 관심특허, 개인저작논문 관리기능 등을 제공할 수 있다.

## 5. 결론 및 발전방향

웹 2.0 패러다임과 개인화서비스 강화에 부응하여 추천시스템과 소셜 네트워크에 대한 연구가 활발하며, 서로 간 융합을 통해 고객 가치를 제고하고 신규서비스를 창출하려는 움직임이 늘고 있다.

본 논문에서는 서로 다른 분야인 추천시스템과 소셜 네트워크를 융합한 신개념의 연구자연결망을 구축하였다. 연구자연결망은 같

은 분야를 연구하는 연구자들끼리 지능적으로 소셜 네트워크를 구축하여 관련 정보를 공유하고 최신 연구동향을 신속하게 파악할 수 있도록 하기 위해 개발한 온라인 휴먼네트워크이다. 기존 소셜 네트워크의 경우 연구자들이 자신들의 소셜 네트워크를 구축하기 위해 수동으로 연구자들을 연결해 나가는 방식인데 비해, 본 방식은 이용로그, 개인프로파일, 북마크, 맞춤정보 등 다양한 정보를 분석하여 자동으로 구축하는 지능화된 방식이다. 기존방식이 연구자들에게 자율성을 부여하면서 형식에 얽매이지 않고 개성이 강한 네트워크를 구성할 수 있는 장점이 있지만, 초기에 네트워크를 구축하기가 매우 어렵고 활발히 참여하지 않는 연구자의 경우 폐쇄적일 수 있는 한계가 있다. 본 방식에서는 연구자들이 직접 관계를 맺고 싶은 대상을 찾아 네트워크를 구성할 수 있을 뿐만 아니라, 같은 논문이나 특허정보를 본 사람이나 관심 분야가 같은 연구자를 지능적으로 분석하여 자동으로 소셜 네트워크를 구축할 수 있다.

글로벌 연구자연결망으로 확장 등을 고려하여 연구자프로파일 서술(description)표준인 FOAF(Friend Of A Friend)에 기반하여 연구자프로파일을 작성할 필요가 있으며, 학술논문, 연구보고서 등의 공저자 관계도 연구자연결망에 활용할 필요가 있다. 공저자의 경우 동명이인에 대한 식별이 필수적이며 이를 해결하기 위해서는 저자전거에 대한 연구가 선행되어야 한다.

## 참고문헌

- 김명숙. 2007. 소셜 네트워크 Service, KT미래 기술연구소.
- 김종규. 2008. Netflix 영화추천 시스템 분석 『마이크로소프트웨어』, 2008(11).
- 모리켄. 2008. 『구글·아마존화하는 사회』. 하연수 역, 서울: 경영정신(작가정신).
- 소셜 네트워크의 비즈니스 활용: 일본 동향. [cited 2009. 02. 01.]  
 <<http://elfinx.springnote.com/pages/98904>>.
- 조재인. 2008. 도서관 정보 수요자를 위한 소셜 네트워크 서비스 도입에 관한 연구. 『한국도서관정보학회지』, 39(2): 169-186.
- 이지연, 주수형, 민지연. 2007. 『라이브러리 2.0 과 도서관서비스 발전 방향』. [서울]: 한국교육학술정보원. 연구보고 KR 2007-11.
- 황혜경, 이재윤. 2006. 이용자의 참여, 협력, 공유를 근간으로 하는 학술정보 포털서비스 : Connotea와 CiteULike를 중심으로. 『한국정보관리학회 학술대회논문집』, 13 (2006): 63-70.
- Barnes, Susan B. 2001. *Online connections: internet interpersonal relationships*. New Jersey: Hampton Press Inc.
- Greg Linden, Brent Smith, Jeremy York. 2003. "Amazon.com Recommendations: Item-to-Item Collaborative Filtering." *IEEE Internet Computing*,

- 7(1): 76-80.
- OCLC. 2007a. "Libraries and Social Net-  
working" Sharing, Privacy and Trust  
in Our Networked World : A Report  
to the OCLC Membership. Dublin:  
OCLC.
- \_\_\_\_\_. 2007b. "Libraries and Social Net-  
working: The thoughts of nine ex-  
perts about our increasingly online  
lives." *NextSpace: The OCLC News  
Letter*, 1(7): 4-10.