

협업 필터링 기반 맞춤형 도서 추천 시스템

장민혜*, 정운해**, 박두순*

*순천향대학교 컴퓨터소프트웨어공학과

**순천향대학교 컴퓨터학과

e-mail:jangmh312@naver.com

A Personalized Book Recommendation System Based on the Collaborative Filtering

Min-Hye Jang*, Woon-Hae Jeong**, Doo-Soon Park*

*Dept of Computer Software Engineering, Soonchunhyang University

**Dept of Computer Science and Engineering, Soonchunhyang University

요 약

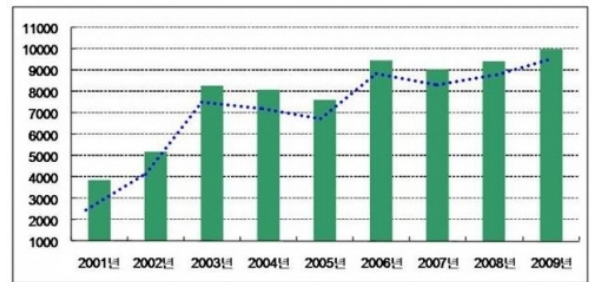
전자상거래 시장의 급격한 성장에 따라 고객이 원하는 정보를 얻기 위해 소요되는 시간과 노력을 절약하기 위한 방안으로 추천 시스템의 필요성이 강조되고 있다. 추천 시스템에 일반적으로 가장 많이 쓰이는 것이 협업필터링 기법이다. 협업 필터링은 추천시스템 분야에서 가장 성공적인 기법으로 전자상거래 포털에서 가장 널리 이용되고 있다. 그러나 희박성, 확장성, 투명성 등의 문제점을 가진다. 본 논문에서는 프로파일링 기법을 사용해 협업필터링의 희박성 문제 해소 방안으로 개인성향을 이용하여, 보다 정확한 추천을 하여 온라인 서점에 적용할 수 있는 추천 시스템이다.

1. 서론

정보화시대가 도래하면서 전자상거래 시장의 변화 또한 빠른 속도로 변화를 맞이했다. 기존 오프라인의 고객층이 시간과 장소에 제약이 없는 온라인으로 이동하면서 더욱 많은 소비를 이끌어 낼 수 있게 되었지만, 수많은 정보 안에서 고객이 진정으로 원하는 정보를 얻기 위한 시간과 노력이 오히려 많이 필요하게 되었다. 따라서 고객이 원하는 정보를 얻기까지 소요되는 시간과 노력을 절약하기 위한 고객 지원 시스템의 필요성이 대두되었다. 따라서 고객 지원 시스템의 일원으로 일대일 맞춤 서비스 제공을 위한 추천 시스템을 통해 최적의 구매의사결정을 할 수 있는 추천 시스템이 필요하다.

도서 시장은 매년 증가 추세를 보이며 온라인 서점의 이용 또한 해마다 늘어나고 있다. <그림 1>에서 볼 수 있듯이 온라인 서점 방문자수가 급격히 성장하며 1인당 체류시간도 크게 늘어 수요에 따른 이익창출의 기회도 많지만 온라인 서점들을 편의성, 접근성, 가격, 구색을 비교해보면 별다른 차별성을 보이지 못한다. 기존의 포털들은 배송 서비스와 가격 면에서 경쟁 우위를 확보하고 있었지만, 모든 온라인 서점들이 점차 보편화 된 서비스를 제공하게 된 시점에서 뚜렷한 경쟁력을 갖춘 포털을 찾아 볼 수 없다[1,2].

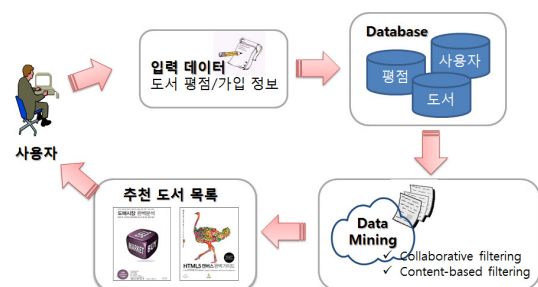
본 논문에서는 전자상거래 포털 중에서도 온라인서점의 니즈에 따른 차별화와 경쟁력을 갖추기 위해 데이터마닝 기법을 적용한 도서 추천 시스템을 제안한다.



<그림 1> 온라인 서점 방문자 추이

2. 협업필터링 기반도서 추천 시스템 설계 및 구현

<그림 2>와 같은 추천 방법으로 도서 추천 시스템을 구축하고 설계하였다.

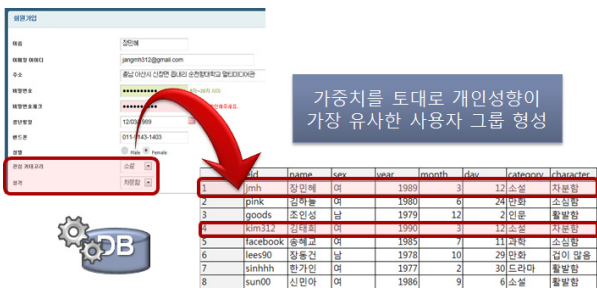


<그림 2> 시스템 구조도

시스템에서 협업 필터링을 이용하는데 있어 데이터 희박성 문제가 있기 때문에 내용기반(content-based filtering) 프로파일링 기법과 협업 필터링(collaborative filtering) 기법 두 가지로 구현하였다. 이를 구분 짓는 기준은 도서 평점 데이터 수에 따라서 방식을 나누도록 했다.

먼저 추천 서비스는 사이트에 가입한 사용자에게 한하며, 회원가입을 하는 절차에 있어 고객의 성별, 나이, 성격, 관심 카테고리 정보를 입력 받는다[3]. 이는 내용기반 프로파일링을 하기 위해 받는 데이터로써, 협업 필터링의 희박성 문제를 해결하기 위한 방안으로 평점 데이터가 적을 시 발생하는 추천의 정확성을 확보하기 위함이다. 서비스를 이용하면서 고객은 도서 평점을 입력하게 되고, 이 데이터가 희박성 문제를 해결할 수 있는 최소 기준치에 일치하게 되면 시스템은 자동으로 프로파일링 기법에서 협업필터링 기법으로 서비스를 전환한다. 이는 평점 데이터 개수가 10개 일 때를 기준으로 하며, 고객은 자신과 유사한 그룹을 찾아 그 그룹에서 평점이 높은 도서를 추천 받는다.

데이터가 없는 신규 회원의 경우 유사도를 이용한 그룹화를 할 수 없는 희박성의 문제가 발생한다. 따라서 회원가입 시 입력 받는 성별, 나이 성격, 관심 카테고리의 데이터를 기반으로 사용자와 가장 유사한 사용자를 찾아 그 사용자의 평점을 이용한다. <그림 3>은 유사 사용자 선정에 대한 개념도를 보여준다.



<그림 3> 유사그룹 형성과정

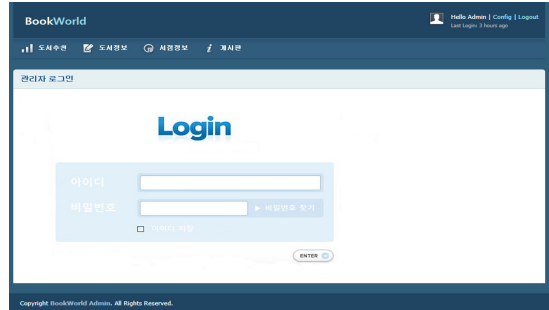
[표 1]은 전체 시스템의 구현 환경을 나타낸 것이다.

[표 1] 시스템 구현 환경

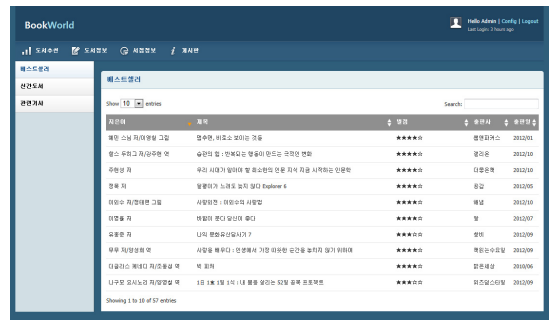
시스템 구성요소	시스템 세부 내용
운영체제	Windows7
개발언어	JAVA
개발환경	Eclipse - Indigo
JDK	JDK version 1.6
Framework	Spring 3.0
Web Server	Apache Tomcat 6.0
Web Language	JSP, jQuery
DBMS	MySQL

<그림 4>는 관리자 로그인 페이지이다. 관리자는 도

서를 등록, 수정, 삭제 등을 할 수 있으며, 회원 및 도서 DB를 관리하고 게시판 이용 또한 관리한다. <그림 5>는 도서정보 메뉴에 있는 베스트셀러 페이지이다. 지은이, 제목, 별점, 출판사, 출판일의 각 항목을 오른쪽 화살표를 선택하여 내림차순과 오름차순으로 정렬할 수 있다.

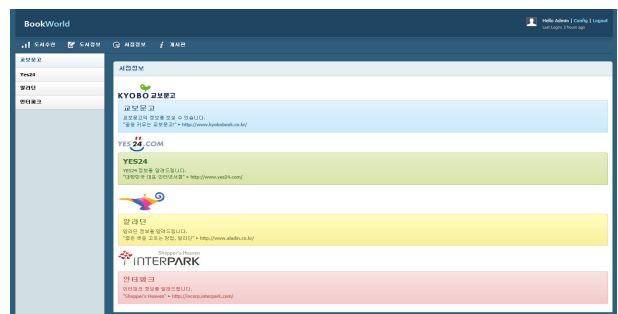


<그림 4> 관리자 로그인 페이지



<그림 5> 베스트셀러 화면

<그림 6>은 서점 정보 메뉴 화면이다. 교보문고, 알라딘, 인터파크 등의 온라인 서점 정보로 구성되며, 선택하게 되면 해당 서점으로 바로가는 기능을 구현하였다.



<그림 6> 서점 정보 메뉴 화면

3. 결론

본 논문에서는 데이터마이닝 기법을 적용한 도서 추천 시스템을 구현하였다. 단순한 서점의 역할을 벗어나서 사용자가 원하는 도서를 미리 추천해 줌으로써 사용자의 만족도를 높일 수 있고, 서비스를 제공하는 측면에서는 마케팅 측면에서도 사용할 수 있다.

본 시스템의 특징은 사용자가 웹 브라우저를 통해 개인 정보를 입력하고 이를 통해 개인에게 맞는 도서를 추천 받을 수 있다는 점이다. 또한 사용자가 늘어나고 데이터가 쌓임으로 인해 더욱더 정확한 영화 정보를 추천하는 시스템으로 확장 될 수 있다.

참고문헌

- [1] Amazon.com Recommendations : Item-to-Item Collaborative Filtering, IEEE Internet Computing, Jan/Feb 2003.
- [2] 이용구, "필터링 기법을 이용한 대출 도서 추천 시스템 구축에 관한 실험적 연구", 연세대학교 대학원 석사학위논문, 2002.
- [3] 장슬기, 박두순, 정영식 "개인 성향과 협업필터링 기반 영화추천 시스템 성능 향상" 한국정보처리학회, 춘계학술 발표대회 논문집 제 17권 1호, 2010.4