

# A Recommendation Procedure for Group Users in Online Communities

오희영<sup>\*</sup>, 김혜경<sup>\*</sup>, 김재경<sup>\*</sup>

<sup>\*</sup>경희대학교 경영대학

서울 동대문 구 회기동 1, 130-701

Tel: +82-2-961-9355, Fax: +82-2-967-0788, E-mail: {oohy, kimhk, jaek}@khu.ac.kr

## Abstract

Nowadays many people participate in online communities for information sharing. But most recommender systems are designed for personalization of individual user, so it is necessary to develop a recommendation procedure for group users, such as participants in online communities. This paper proposes a group recommender system to recommend books for group users in online communities. For such a purpose, we suggest a group recommendation procedure consisting of two phases. The first phase is to generate recommendation list for 'big user' using collaborative filtering, and the second phase is to remove irrelevant books among previous list reflecting the preference of each individual user. The procedure is explained step by step with an illustrative example. And this procedure can potentially be applied to other domains, such as music, movies and etc.

## Keywords:

Group recommender system, Collaborative Filtering, TF-IDF

## 1. 서론

현재 인터넷의 전반적인 확대로 2005년에는 국내에서 약 3천만 명의 인구가 인터넷을 사용하고 있으며 한 개 이상의 온라인 커뮤니티에 가입되어 있거나 직접 커뮤니티를 운영하고 있는 사람이 인터넷 이용자의 50% 정도이다. 커뮤니티의 경우 개인 홈페이지나 블로그의 영향으로 개인 중심의 커뮤니티가 많이 발전하고 있지만 사용자가 정보를 생산해야 하는 개인 중심 커뮤니티에 비해 공통된 관심과 요구를 갖는 구성원들이 정보공유 및 획득을 위해 활동하는 대형 커뮤니티도 지속적으로 증가하고 있다.

온라인 커뮤니티의 경우 정보교류와 지식교환은 커뮤니티의 지속 및 회원 유지에 필수적인 요소이다. 양질의 정보가 발생하지 않는 커뮤니티는 인터넷 사용자들의 관심을 받을 수 없기 때문에 지속적으로

커뮤니티가 정보와 지식을 생성할 수 있도록 지원할 수 있는 시스템이 필요하다.

추천시스템은 사용자들의 성향이나 선호에 적합한 상품을 다양한 기법을 이용해 추천해주는 시스템으로써 Amazon.com이나 CDNow와 같은 전자상거래 사이트에서 성공적으로 사용되었으며 모바일과 P2P 등 다양한 환경에 적합한 추천시스템이 연구되고 있다.

그러나 지금까지 대부분의 추천시스템은 추천의 대상을 개인에게 맞추어 연구되었으며 두 명 이상의 사람들이 모인 그룹에 대한 추천은 고려되지 않고 있다. Usenet News와 같이 개별적인 활동이 주로 이루어지는 분야에서는 그룹을 대상으로 하는 추천에 제약이 있다. 그러나 책이나 음악과 같이 개별적인 활동과 그룹으로 이루어지는 활동이 모두 이루어지는 분야뿐 아니라 레스토랑, 보드게임 및 영화처럼 그룹으로 이루어지는 활동이 보편적인 분야에 대해서도 그룹을 대상으로 하는 추천시스템을 연구한 사례를 찾는 것은 쉽지 않다. 또한 웹 기반 개인 추천시스템에 대한 연구가 많이 이루어지고 있는 반면 온라인 상에서 대표적인 그룹이라고 할 수 있는 커뮤니티에 대한 추천시스템 연구는 상대적으로 미비한 실정이다.

본 연구에서는 구성원들의 관심사에 따라 책을 읽고 토론하며 지식 및 정보를 교환하는 'Online Book Club' 커뮤니티들을 대상으로 구성원들의 개별적인 성향을 고려하면서도 커뮤니티 구성원 모두가 함께 읽고 토론할 수 있는 책을 찾을 수 있도록 지원하는 추천시스템을 제시한다. 온라인 커뮤니티를 위한 책 추천시스템은 대표적인 추천기법인 협업필터링 (Collaborative Filtering: CF)을 이용하여 온라인 커뮤니티에 추천을 제공하며 커뮤니티 내의 개별 사용자들의 선호를 반영하기 위해서 Conjunctive나 Disjunctive 방법과 같이 임계치를 정해 선호하는 아이템을 선정하거나 선호하지 않는 아이템을 제거하는 적합성 기반 필터링을 이용한다. 이와 같은 2단계의 추천절차를 통해 커뮤니티의 개별 구성원들의 만족도를 향상시킬 수 있는 추천을 제공한다.

이 논문의 구성은 2장에서 협업필터링과 그룹추천시스템에 대한 기존 연구를 살펴보고 3장에서는 본 연구에서 제시하는 추천시스템의 전반적인 개요 및 구조도를 설명하며 4장에서 2단계로 이루어진 추천절차를 보여주도록 한다. 5장에서는 본 연구에서 제시된 추천시스템을 적용할 수 있는 간단한 사례를 제시하고 끝으로 6장에서는 본 연구의 시사점 및 한계점과 향후 연구 방향을 논의한다.

## 2. 관련연구

### 2.1 그룹 추천시스템

그룹 추천시스템은 의도적이거나 비의도적으로 형성된 집단에 대해서 모든 구성원들이 함께 이용할 수 있는 아이템에 대해 모두가 만족할만한 추천을 제공하는 시스템으로 기존의 그룹 추천시스템들은 주로 온라인을 이용한 음악, 영화 등의 멀티미디어 콘텐츠에 대해서 연구되었다.

MusicFX[3]는 체육관에서 같은 시간에 운동을 하는 사람들을 대상으로 모두가 선호하는 장르의 음악을 들려주기 위해서 체육관 가입 시에 각 장르에 대한 점수를 받아서 체육관에 있는 사람들의 평균 점수가 높은 음악장르를 방송한다. 이 시스템은 사용자 선호의 변화를 반영할 수 있고 그룹의 만족도를 최대화 할 수 있는 장점이 있으나 초기에 사용자들이 모든 음악장르에 대해 직접 평가를 해야 하는 단점이 있다.

PolyLens[6]는 함께 영화를 보고자 하는 소수의 사람들로 구성된 그룹에 대해 영화를 추천한다. 이 시스템은 그룹 내의 구성원에게 협업필터링 알고리즘을 이용한 추천목록을 생성하고 이를 그룹추천목록으로 결합한다. MusicFX가 그룹의 평균적인 만족도 향상에 초점을 두고 있다면 PolyLens는 그룹 내에서 가장 만족하지 않는 구성원을 기준으로 추천여부를 결정한다. 그러므로 추천되는 영화는 그룹의 모든 구성원들이 선호하는 것으로 결정된다. 이와 같은 만족도 결정 방법은 PolyLens와 같이 소수의 구성원으로 이루어진 그룹일 때만 가능하다.

Pocket RestaurantFinder[4]는 레스토랑에 함께 가고자 하는 사람들에게 그들의 위치정보와 선호도를 이용하여 적합한 레스토랑을 추천하는 시스템이다. 각 구성원은 레스토랑이 가지고 있는 몇 가지 속성인 위치, 가격, 주요 요리, 시설 및 분위기에 대한 자신의 선호를 휴대용 단말기에 저장하고 있으며 Pocket Restaurant Finder는 각 구성원의 선호와 레스토랑이 가지고 있는 속성값을 계산한 뒤 모두 더해서 레스토랑에 대한 그룹의 선호도를 파악한다. 이 방법은 그룹의 선호도를 파악하는 데에 유용하고 멀티미디어 콘텐츠가 아닌 아이템을 대상으로 하였다라는 특징이 있지만 실제

구현 방법에 대한 고민이 필요하다.

Adaptive Radio[2]는 같은 공간에서 음악을 듣고 있는 사람들을 대상으로 하는 추천시스템이다. 설문지나 사용자 프로파일을 이용해서 사용자의 선호를 파악하는 여타의 추천시스템과 달리 Adaptive Radio는 그룹의 각 구성원들이 현재 방송되는 음악에 대해 선호/비선호를 나타냄으로써 사용자의 선호도를 파악한다. 이 시스템은 추천을 결정하기 위해서 한 명의 구성원이라도 부정적인 반응을 보이는 음악은 방송하지 않으며 대신에 평가가 없는 음악에 대해서는 일단 긍정적인 선호도를 가지고 있다고 판단한다. 그러므로 Adaptive Radio는 개별 구성원들이 들어볼 수 있는 음악의 다양성을 확대하면서도 그룹의 구성원들이 함께 즐길 수 있는 음악을 추천한다는 장점이 있다.

위와 같은 그룹 추천시스템은 그룹의 선호를 결정하는 방법에 대해 두 가지 접근법이 있다. MusicFX와 Pocket RestaurantFinder는 개별 구성원 만족도의 평균을 이용함으로써 만족하는 다수의 의견에 초점을 맞추는 반면 PolyLens와 Adaptive Radio는 추천이 모두에게 수용될 것인지를 확인하기 위해 가장 만족하지 않는 소수의 의견에 초점을 맞추고 있다.

그룹의 선호도를 결정하는 위의 두 가지 방법은 모두 장, 단점을 가지고 있다. 본 연구에서는 그룹의 규모가 다양한 온라인 커뮤니티를 대상으로 추천에 대한 만족도를 향상시키기 위해 2단계 추천 절차를 도입하여 두 가지 방법을 각 단계에 적용한다.

### 2.2 협업필터링 기반 추천시스템

다양한 추천 기법 중 가장 성공적인 상품 추천 기법으로 알려진 협업 필터링은 고객들의 상품에 대한 평가를 이용하는 정보 필터링 기법의 하나로, 일상생활에서 가족, 친구, 동료들의 경험을 통한 구전 효과를 자동화하여 이용하는 것이다. 즉, 추천 대상 고객과 선호도가 유사한 고객들이 과거에 좋아했던 상품을 추천한다. 일반적으로 협업필터링 기반 추천은 크게 (1) 입력 데이터 구성, (2) 유사 선호도 집단 형성, (3) 추천목록생성 단계로 나누어진다.

- (1) 입력 데이터 구성: 입력데이터는 일반적으로  $n$ 개의 아이템에 대한  $m$ 명의 사용자 선호도로 구성되며 이는  $m \times n$ 의 사용자-아이템 행렬  $P$ 로 표현할 수 있다. 즉,  $i$ 번째 사용자가  $j$ 번째 아이템에 대한 선호가 있으면 선호도  $p_{ij}$ 는 1 혹은 평가값을 가지고 선호가 없을 경우 0의 값을 가지게 된다. 아이템에 대한 사용자의 선호도는 직접 입력을 받거나 구매데이터 등으로 부터 추정하여 얻을 수 있다.
- (2) 유사 선호도 집단 형성: 유사 선호도 집단인

네이버를 형성하는 것은 협업필터링 기반 추천에서 가장 중요한 단계이다. 앞에서 구성한 행렬  $P$ 를 이용하여 사용자들 간의 유사도가 높은 네이버를 찾는 과정이다. 유사도는 일반적으로 피어슨 상관관계수 (pearson correlation)와 코사인(cosine)으로 측정되며 네이버의 규모는 임계치 이상의 유사도를 가진 사용자들을 모두 선택하는 방법과 가장 유사도가 높은  $l$ 명의 사용자를 선택하는 방법이 있다.

- (3) 추천 목록 생성: 추천 목록 생성은 네이버의 선호도 정보를 이용하여 추천대상이 구매할 가능성이 높은 아이টে를 결정하는 것이다. 구매할 가능성을 순위화 하는 방법은 선호도를 입력하는 척도에 따라 달라진다. 예를 들어 선호도를 구매횟수라고 하면 추천 대상이 구매하지 않은 아이টে를 중 네이버들이 가장 빈번하게 구매한 아이টে를 추천하게 된다.

그러나 협업필터링은 입력데이터의 희박성, 신상품 추천 문제 그리고 추천 알고리즘의 확장성 문제가 있어서 많은 연구자들이 추천 대상 고객이 선호하는 아이টে와 유사한 특성을 가진 아이টে를 추천하는 내용 기반 필터링(Content-based Filtering:CB)을 협업필터링과 결합하여 사용하였으며 이와 같은 혼합 추천(hybrid recommender) 방식은 두 기법의 영역에서도 향상된 결과를 보여주고 있다.

기존의 그룹 추천시스템은 PolyLens와 같이 협업필터링을 사용하기는 하였으나 해당 시스템은 개별 사용자에게 대한 추천목록을 생성하고 이 목록을 결합하였으므로 협업필터링을 수행하는 것은 개별 사용자에게 대한 것과 다른 것이 없고 그룹내의 사용자와 아이টে의 수가 많을 경우 유용성이 떨어진다고 보고되어 있다. 그러므로 본 연구에서는 커뮤니티 사용자들을 대표하는 Big-user를 커뮤니티의 기본 단위로 하여 Big-user들 간의 유사도를 계산하고 유사한 선호도를 가진 네이버 Big-user를 구성하는 방법을 통해 추천을 수행한다.

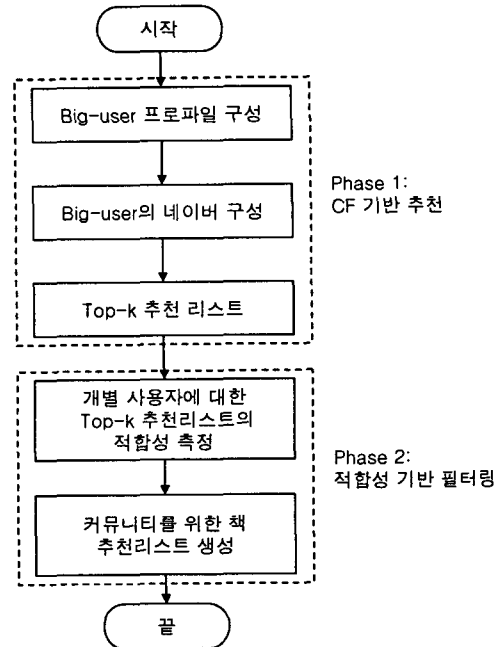
### 3. 온라인 커뮤니티를 위한 추천시스템

#### 3.1 그룹추천시스템의 개요

본 연구에서 제안하는 온라인 커뮤니티를 위한 책 추천시스템은 커뮤니티를 대상으로 하면서도 커뮤니티에 속해 있는 사용자들의 만족도까지 고려하여 추천을 한다.

<그림 1>과 같이 온라인 커뮤니티를 위한 책 추천은 크게 두 단계로 나누어져 있다. 우선 1단계는 CF(collaborative Filtering)기반 추천으로 Big-user 프로파일을 구성하고 타겟 Big-user에 대한 네이버를

구성한 뒤 추천리스트를 생성하는 과정이고 2단계는 Big-user에 대한 추천후보목록과 커뮤니티의 개별 사용자들 간의 적합성을 판단하여 적합하지 않은 책은 추천하지 않는 적합성 기반 필터링 과정이다.



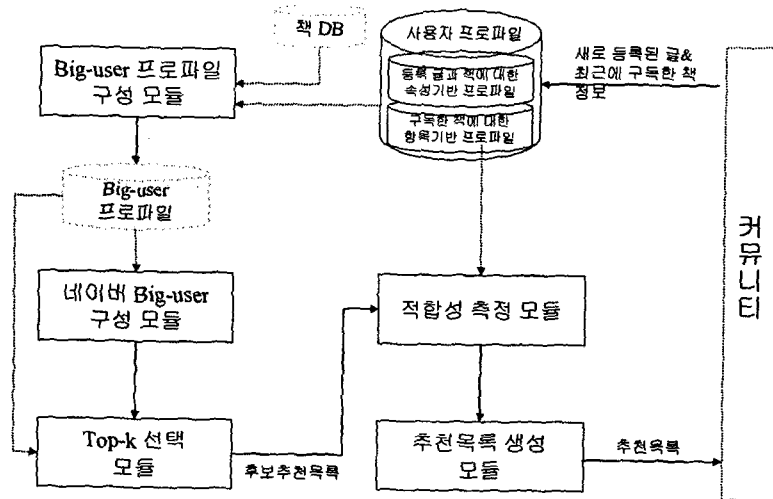
<그림 1> 온라인 커뮤니티를 위한 책 추천 과정

개별 사용자가 아닌 커뮤니티에 책을 추천하기 위해서는 우선 커뮤니티에 가입되어 있는 사용자들의 개별적인 선호를 분석하여 커뮤니티를 대표하는 가상의 프로파일로 구성해야 한다. Big-user는 커뮤니티의 특성을 대표하는 가상의 사용자로 커뮤니티에 속한 개별 사용자 프로파일을 이용해서 생성된다.

모든 커뮤니티에 대한 Big-user가 생성된 후에는 각 Big-user들 사이의 유사도를 계산하여 추천 대상인 타겟 Big-user와 높은 유사도를 가진 Big-user들로 네이버를 형성한다.

네이버가 결정된 후에는 네이버 Big-user들이 구독한 책들에 대해 타겟 Big-user가 구독할 가능성을 계산하여 가능성이 큰 순서대로 타겟 Big-user에 대한 추천후보목록을 만든다. 추천후보목록이란 두번째 단계에서 이루어지는 필터링 과정을 거치기 전의 big-user에 대한 초기 추천목록을 의미한다.

타겟 Big-user에 대한 추천후보목록이 생성되면 타겟 커뮤니티의 개별 사용자들과 추천후보목록의 책들이 가지는 적합성을 계산한다. 적합성이란 사용자들이 추천후보목록의 각 책에 대해 가지고 있는 유사도를 바탕으로 커뮤니티에 속한 사용자들이 추천된 책에 대해 느끼는 만족도를 의미한다. 그리고 임계치 이하의 적합도를 가진 책들을 추천후보목록에서 제거하여 커뮤니티 내의 사용자들과 적합성이 높은 책으로 구성된 최종 추천목록이 생성된다.



<그림 2> 온라인 커뮤니티를 위한 책 추천시스템의 구조도

이와 같이 본 연구에서 제안하는 그룹추천시스템은 CF를 기반으로 하는 1단계와 적합성 기반 필터링의 2단계로 구성되며 협업필터링을 이용하여 타겟 Big-user와 유사한 선호도를 가진 네이버 Big-user들의 구독 정보를 통해 구독 가능성이 높은 책 목록을 생성할 수 있다. 또한 2단계를 통해 타겟 Big-user의 선호에 맞추어진 추천목록을 개별 사용자의 선호와 비교하는 과정을 통해 개인 프로필을 정규화하지 않음으로써 활발한 활동을 하는 사용자들의 의견이 상대적으로 많이 반영된 Big-user 프로필로 인해 커뮤니티 내의 적극적인 사용자만 만족할 수 있는 추천목록이 제공되는 것을 방지할 수 있다.

### 3.2 그룹추천시스템의 구조도

<그림 2>는 온라인 커뮤니티를 위한 책 추천시스템의 구조도를 표현한 것이다. 이 추천시스템은 보는 바와 같이 Big-user 프로필 구성 모듈, 네이버 구성 모듈, Top-k 선택 모듈, 적합성 측정 모듈, 추천리스트 생성 모듈로 구성되어 있다. 5가지 모듈은 각각 커뮤니티에 등록된 글과 구독한 책에 대한 사용자 프로필, Big-user 프로필 및 책 데이터베이스와 연계되어 책을 추천한다.

Big-user 프로필 구성 모듈은 커뮤니티 내의 등록된 글과 구독한 책에 대한 사용자 프로필, 그리고 책의 속성정보가 저장되어 있는 데이터베이스를 이용해서 Big-user 프로필을 구성한다. Big-user 프로필을 구성하기 위해서는 모든 커뮤니티 사용자들의 선호를 평등하게 반영해야 하지만 커뮤니티 내의 활발한 활동을 하는 사용자에 대한 보상이 사용자들에게 동기를 부여함으로써 커뮤니티의 활성화를 가능하게 하기 때문에 개별 사용자들의 프로필을 정규화하지 않도록 한다. Big-user 프로필은 각 키워드에 대한 Big-user의 선호도를 나타내는 행벡터로 이루어져

있다.

네이버 Big-user 구성 모듈은 Big-user 프로필을 이용해서 Big-user간의 유사도를 계산하고 높은 유사도를 가진 Big-user들로 네이버를 형성한다. 이때 유사도는 각 Big-user들이 가진 행벡터의 코사인 각도를 이용해서 계산된다.

Top-k 선택 모듈은 타겟 Big-user에 대한 추천후보목록을 생성하기 위해서 네이버 Big-user가 구독한 책들에 대한 타겟 Big-user의 구독 가능성을 측정하는 것이다. 측정된 PLS가 높은 순서대로 k개의 책을 선택하여 추천후보목록을 생성한다.

적합성 측정 모듈은 추천후보목록의 책들과 타겟 커뮤니티의 개별 사용자들과의 적합도를 측정한다. 이는 Big-user 프로필 구성에 이용되는 사용자 프로필이 개별적으로 정규화되지 않음으로써 활동을 많이 하는 사용자의 의견이 더 많이 반영되어 커뮤니티를 대표하는 특성이 왜곡되는 것을 방지하기 위한 것으로 추천리스트 생성 모듈에서는 측정된 적합도를 바탕으로 개별 사용자들이 추천후보목록에서 임계치 이하의 적합도를 가지는 책들을 제거한다. 이 과정을 통해서 커뮤니티에서 활발히 활동하는 사용자에 대한 보상뿐만 아니라 개별 사용자의 만족도까지 고려함으로써 추천에 대한 전체적인 만족도를 향상시키고 기존 그룹추천시스템의 취약점을 해결할 수 있다.

그리고 커뮤니티에 새로운 글이 등록되거나 사용자들이 새로운 책을 구독하게 되면 사용자 프로필이 업데이트 되고 Big-user 프로필에도 변화를 주게 되어 네이버 구성 및 추천리스트가 달라지게 된다. 커뮤니티를 대상으로 하는 추천이므로 실시간으로 추천목록을 제공하지는 않아도 되지만 정기적으로 각 모듈이 추천절차를 수행해서 변화된 선호를 만족하는 추천목록을 제공한다.

#### 4. 온라인 커뮤니티를 위한 책 추천 절차

기존의 그룹 추천시스템은 Big-user를 만들거나 각각의 사용자들을 대상으로 생성된 추천목록을 통합하는 방법이 일반적이었다. 그러나 Big-user의 경우 의견을 많이 제시하는 사용자에게 의해 그룹의 선호도가 왜곡될 수 있고 개별 추천리스트 통합은 그룹 내의 사용자 수가 많아질수록 시간과 비용이 많이 소모된다는 단점이 있다.

본 연구에서는 많은 사용자들이 존재하는 큰 규모의 커뮤니티를 대상으로 하기 때문에 추천목록 구성의 효율성을 위해 Big-user를 생성하고 각 사용자들의 선호를 최대한 반영하기 위하여 커뮤니티 사용자들의 특성과 후보추천목록의 적합성을 판별하는 2단계 추천 방법을 제시한다.

온라인 커뮤니티를 위한 책 추천시스템은 커뮤니티에 등록된 글과 사용자들의 구독 정보가 저장된 사용자 프로파일과 책 데이터베이스를 이용해서 Big-user 프로파일을 구성하고 Big-user 간의 유사도를 계산하여 네이버와 후보추천목록을 생성한 뒤 적합하지 않은 후보추천 목록을 필터링한다.

이 추천시스템은 CF 기반 추천과 적합성 기반 필터링의 두 단계로 구성된다.

<표 1> 사용자 프로파일의 구성

속성 기반 사용자 프로파일 $M_u=(P_{ki})$						
	Community	CRM	Reco-mmend	Data Mining	Social Network	P2P
kim	MIS Study	6	0	10	3	8
	Best CRM	2	4	6	3	18
Oh	MIS Study	18	0	18	0	0
	AI	14	0	24	2	1
Chae	MIS Study	7	3	4	3	8
:	:	:	:	:	:	:
user $h$	community $l$					
항목 기반 사용자 프로파일 $M_h=(P_{hk})$						
	Book1	Book2	Book3	Book4	Book5	...
Kim	0	0	1	0	0	...
Oh	1	0	1	0	0	...
Chae	0	0	1	0	0	...
:	:	:	:	:	:	:

##### 4.1 단계 1: CF 기반 추천

Big-user에 대한 후보추천목록을 생성하기 위해 수행되는 CF 기반 추천절차는 우선 각 커뮤니티를 대표하는 Big-user를 구성하고 추천 대상인 커뮤니티의 Big-user와 다른 커뮤니티의 Big-user간의 유사도를 계산하여 네이버를 형성한다. 그리고 구독 가능성 점수를 나타내는 PLS 계산을 통해 타겟

Big-user가 구독할 가능성이 높은 책들로 후보추천목록을 결정한다.

##### 4.1.1 Big-user 프로파일 생성

개별 사용자를 위한 추천 알고리즘은 다양한 방법으로 연구되었지만 이런 추천방법들은 그룹을 대상으로 하는 추천에 바로 적용할 수 없다. 그룹에 속해 있는 사용자들은 서로 같지 않은 성향을 가지고 있기 때문에 사용자들의 개별적인 프로파일들을 그룹을 대표할 수 있는 하나의 프로파일로 구성하는 단계가 필요하다.

그룹의 선호도를 학습하기 위해서 일반적으로 두 가지 접근법이 있다. [1][2][3][4][6][8] 첫 번째 방법은 그룹의 성향을 대표하는 'Big-user'를 만들어서 'Big-user'에 대한 추천리스트를 생성하는 것이다. 그리고 두 번째는 그룹에 속해있는 사용자마다 개별적인 추천목록을 생성하고 이를 결합하여 그룹에 대한 추천목록으로 재구성하는 방법이다. 본 연구에서는 CF에 기반하여 추천 대상 커뮤니티가 구독할 가능성이 높은 책 목록을 구성하므로 네이버를 이용할 수 있도록 그룹의 선호도를 나타내는 방법으로 'Big-user'를 이용한다.

사용자 프로파일은 사용자들이 가입한 커뮤니티와 각 커뮤니티에 등록한 글, 그리고 사용자들이 구독한 책에 대한 정보를 포함하고 있다. Big-user 프로파일은 <표 1>에서와 같은 사용자 프로파일을 이용해서 구성된다.

사용자 프로파일은 사용자의 선호도가 있는 책(구독한 책)에 대한 항목 기반 프로파일과 사용자가 가입한 커뮤니티, 해당 커뮤니티에 등록한 글 및 구독한 책의 특성을 나타내는 속성정보에 대한 프로파일로 구성되어 있다. Big-user 프로파일을 생성하기 위해서는 각 커뮤니티에 속해있는 사용자 프로파일을 이용한다.

프로파일에 포함된 속성 정보는 사용자들이 커뮤니티에 등록한 글과 사용자들이 구독한 책에서 추출된 키워드로 나타나며, 따라서 Big-user 프로파일은 Big-user  $i$ 의 키워드  $j$ 에 대한 가중치로 구성된 행렬  $M_Q=(P_{ij})$ 로 정의할 수 있다.

Big-user 프로파일 구성 단계에서 커뮤니티를 대표하는 Big-user 고유한 특성을 파악하기 위해 TF-IDF로 키워드의 가중치를 계산한다. 다수의 사용자를 포함하고 있는 커뮤니티의 Big-user 프로파일을 구성하는 것이므로 다수의 문서를 포함한 폴더형 프로파일의 TF-IDF 가중치를 계산할 수 있는 SwiftFile[7]을 이용한다.

$$rf(Q_i, d_j) = \frac{freq(Q_i, d_j)}{freq(Q_i, d_{all})} \quad (1)$$

식 (1)은 특정 Big-user가 가지고 있는 모든

키워드에 대해 특정 키워드가 가지는 상대적인 빈도수를 의미한다.  $rf(Q_i, d_j)$  는 Big-user  $Q_i$ 가 포함하는 특정 키워드  $d_j$ 의 빈도수를  $Q_i$ 가 가지고 있는 전체 키워드의 빈도수로 나눈 값이다. 그러므로 특정한 Big-user가 특정 키워드를 많이 가지고 있을수록  $rf(Q_i, d_j)$ 는 더 높은 값을 갖는다.

$$rf(Q_{all}, d_j) = \frac{freq(Q_{all}, d_j)}{freq(Q_{all}, d_{all})} \quad (2)$$

식 (2)는 전체 Big-user에 대한 특정 키워드의 상대적인 빈도수를 나타낸다. 즉, 전체 Big-user가 특정 키워드를 많이 가지고 있을수록  $rf(Q_{all}, d_j)$ 는 높은 값을 갖는다.

식 (1), (2)에서 얻어진 Big-user가 가지는 키워드의 상대적인 빈도수를 가지고 Big-user의 키워드에 대한  $TF$ (Term Frequency)를 계산한다.

$$TF(Q_i, d_j) = \frac{rf(Q_i, d_j)}{rf(Q_{all}, d_j)} \quad (3)$$

식 (3)은 키워드  $d_j$ 가 각각의 Big-user에게서 가지는 가중치를 전체 Big-user에게서 가지는 가중치로 나누어 줌으로써 정규화한 것이다.  $TF$ 가 높은 키워드는 해당 Big-user의 관심도가 높은 것을 의미한다.

$TF$ 가 키워드에 대한 Big-user의 상대적인 가중치를 의미한다면  $IDF$ 는 해당 키워드가 Big-user들을 구별 짓는 특징을 가지고 있는지를 나타낸다.

$$IDF(d_j) = \frac{1}{df(d_j)} \quad (4)$$

식 (4)에서  $df(d_j)$ 란 전체 Big-user 중에서 특정 키워드가 한 번이라도 나타나는 Big-user의 비율을 의미한다.  $IDF$  가중치가 높은 키워드는 소수의 Big-user에게만 나타나는 것이므로 Big-user 간의 특징을 구별하는 역할을 할 수 있다.

Big-user  $Q_i$ 가 가지고 있는 키워드  $d_j$ 의  $TF-IDF$  가중치  $c(Q_i, d_j)$ 는 다음 식 (5)와 같이  $TF$ 와  $IDF$ 를 곱해서 계산할 수 있다. 이 값은 특정 키워드에 대한 해당 Big-user의 특징을 반영한다.

$$c(Q_i, d_j) = TF(Q_i, d_j) \times IDF(Q_i, d_j) \quad (5)$$

<표 2> 속성 기반 Big-user 프로파일  $M_Q=(P_{ij})$

키워드( $d_j$ ) Big-user( $Q_i$ )	CRM	Recommend	Data Mining	social network	P2P
MIS Study	2.82	0.95	0.45	4.35	0.32
Case Review	1.43	3.13	2.73	0.00	1.63
Best CRM	3.39	0.00	0.00	1.59	0.00
AI World	0.00	1.51	4.26	22.96	1.36
Next e-Biz	3.46	1.87	0.22	0.00	0.00

#### 4.1.2 네이버 형성

Big-user 프로파일이 생성되면 유사한 선호도를 가지고 있는 Big-user를 찾기 위해 Big-user들 사이의 유사도를 계산한다.

식 (6)과 같이 타겟 Big-user  $T$ 와 다른 Big-user  $Q_i$ 들간의 유사도  $sim(T, Q_i)$  는 4.1.1 에서 계산된  $c(Q_i, d_j)$ 를 나타내는 행벡터의 코사인 각도를 이용해서 계산할 수 있다.

$$sim(T, Q_i) = \frac{\vec{T} \cdot \vec{Q}_i}{|\vec{T}| \times |\vec{Q}_i|} = \frac{\sum_{j=1}^l c_{Tj} \times c_{ij}}{\sqrt{\sum_{j=1}^l (c_{Tj})^2} \times \sqrt{\sum_{j=1}^l (c_{ij})^2}} \quad (6)$$

두 Big-user의 유사도가 가장 높을 때는 '1', 가장 낮을 때는 '-1'이며 계산된 유사도가 큰 순서대로 상위  $f$  개의 Big-user로 네이버를 형성한다. 이 때, 타겟 Big-user 자신도 유사도 1을 가진 네이버로 선택된다. 이와 같은 방법은 일반적인 협업필터링에서는 사용되지 않지만 본 연구에서는 그룹에 대한 추천이 목적이므로 타겟 커뮤니티에 속해 있는 개별 사용자들은 공통의 관심사나 목적을 가지고 있기 때문에 타겟 Big-user의 네이버라고 볼 수 있다. 그러므로 타겟 커뮤니티 내의 사용자들이 개별적으로 읽은 책은 커뮤니티 차원에서 함께 읽고 토론할 수 있는 추천 대상이 될 수 있다.

예를 들어 타겟 Big-user인 MIS Study와 다른 커뮤니티의 Big-user 간의 유사도는 <표 3>과 같으며  $f = 3$ 일 때, 이 중 유사도가 높은 MIS Study=1, Best CRM=0.83, AI World=0.83이 네이버  $N=\{MIS Study, Best CRM, AI World\}$ 가 된다.

<표 3> 타겟 Big-user와의 유사도 측정  $sim(T, Q_i)$

$sim(MIS Study, MIS Study)$	1.00
$sim(MIS Study, Case Review)$	0.35
$sim(MIS Study, Best CRM)$	0.83
$sim(MIS Study, AI World)$	0.83
$sim(MIS Study, Next e-Biz)$	0.56

### 4.1.3 추천후보목록 생성

추천후보목록은 네이버 Big-user들이 구독한 책 중에서 타겟 Big-user의 구독 가능성이 높은 책들로 구성되며 타겟 Big-user  $T$ 가 책  $B_k$ 를 구독할 가능성은 다음 <식 7>과 같은  $PLS$ (Purchase likeliness score)를 통해 계산된다.  $PLS$ 를 계산하기 위해서는 항목기반 사용자 프로파일에 저장되어 있는 네이버 Big-user들이 구독한 책에 관한 데이터를 이용해야 한다. <표 4>는 네이버 Big-user가 구독한 책의 목록과 구독한 개별 사용자의 수를 정리한 것이다.

$$pls(T, B_k) = \frac{\sum_{n_i \in N} p_{n,k} \times sim(T, n_i)}{\sum_{n_i \in N} sim(T, n_i)} \quad (7)$$

$$(p_{n,k} = \frac{\sum_{h \in n_i} V_k}{U} (V_k = \begin{cases} 1 & \text{if } B_k \in u_{h \in n_i} \\ 0 & \text{if } B_k \notin u_{h \in n_i} \end{cases}))$$

식 (7)에서  $sim(T, n_i)$ 는 타겟 Big-user와 네이버 Big-user $_i$  간의 유사도이며  $p_{n,k}$ 는 네이버 Big-user $_i$ 의 전체 사용자의 수( $U$ ) 중 책  $B_k$ 를 읽은 사용자의 수를 나타낸다.  $V_k$ 는 사용자가 책에 대해 가지는 값으로 해당 책을 읽었을 때는 '1', 그렇지 않을 때는 '0'을 가지게 되므로 커뮤니티 내의  $V_k$ 를 모두 더하면 책  $B_k$ 를 읽은 사용자의 수를 알 수 있다.

네이버 Big-user가 구독한 책의 목록은 커뮤니티의 개별 사용자들 중 한 명이라도 읽은 책은 모두 포함되며 커뮤니티 내에서 책을 읽은 사용자의 수를 커뮤니티 전체 사용자의 수로 나누어 정규화 한다. 또한 타겟 Big-user 자신이 유사도 1을 가진 네이버로 구성되므로 타겟 Big-user의 책 목록도  $PLS$ 를 계산한다. 그러나 타겟 커뮤니티의 모든 유저가 읽은 책  $B_k$ 는 제외하도록 한다.

Top-k 선택 모듈에서 타겟 Big-user가 가지는  $PLS$ 가 측정되었다면 가장 높은 점수를 가진 k개의 책들로 추천후보목록을 구성한다. <표 4>는 네이버

big-user의 커뮤니티에서 구독한 책의 목록 및 구독횟수와 함께 타겟 Big-user가 각 책에 대해 가지는  $PLS$ 를 보여주고 있다. Book<sub>3</sub>은 타겟 커뮤니티의 모든 사용자가 구독하였으므로 추천에서 제외하고  $k = 4$ 일 때, 추천후보목록은 Book<sub>1</sub>, Book<sub>8</sub>, Book<sub>12</sub>, Book<sub>14</sub>로 구성된다.

### 4.2 단계 2: 적합성 기반 필터링

적합성 기반 필터링은 CF 기반 추천과정에서 생성된 추천후보목록의 책들 중에서 커뮤니티를 구성하는 사용자들의 개별적인 선호와 적합하지 않은 것을 제거함으로써 추천에 대한 커뮤니티의 만족도를 향상시킬 수 있다.

적합성 기반 필터링 단계는 우선 커뮤니티 내에 속해있는 사용자들의 특성을 나타낼 수 있도록 구성된 프로파일을 이용해서 적합성 측정 모듈이 추천후보목록과 개별 사용자 프로파일의 유사도를 계산하고 추천목록 생성 모듈에서 임계치를 만족하지 않는 책들을 목록에서 제거하는 과정으로 이루어져 있다.

#### 4.2.1 커뮤니티의 개별 사용자 프로파일

커뮤니티에 속해 있는 개별 사용자들에 대한 프로파일은 Big-user 프로파일을 구성할 때 사용되지만 Big-user에 대해 생성된 추천후보목록의 적합성을 판단하기 위해서도 이용할 수 있다.

커뮤니티 사용자들의 프로파일 구성 또한 사용자들이 구독한 책과 커뮤니티에 등록된 글에서 정보를 얻기 때문에 사용자-키워드 행렬로 나타낼 수 있다.

개별 사용자 프로파일은 커뮤니티 내에서 각 사용자들이 가지는 키워드의 특징을 나타내는 것으로  $TF-IDF$ 가중치를 이용해서 표현할 수 있다.

$$tf(u_{h \in C_i}, d_j) = \frac{freq(u_h, d_j)}{\max freq(u_h, d_j)} \quad (8)$$

<표 4> 네이버 Big-user의 책 구독정보 및 PLS 측정

	U	구독 횟수( $\sum V_k$ )														
		B <sub>1</sub>	B <sub>2</sub>	B <sub>3</sub>	B <sub>4</sub>	B <sub>5</sub>	B <sub>6</sub>	B <sub>7</sub>	B <sub>8</sub>	B <sub>9</sub>	B <sub>10</sub>	B <sub>11</sub>	B <sub>12</sub>	B <sub>13</sub>	B <sub>14</sub>	B <sub>15</sub>
MIS Study	5	2	0	5	0	1	0	0	2	0	3	0	3	0	3	0
Best CRM	4	2	1	1	0	2	0	0	2	2	1	0	0	0	4	0
AI World	3	2	0	2	1	0	0	1	1	1	0	1	3	0	2	0
$pls(T, B_k)$		0.51	0.08	-	0.10	0.23	-	0.10	0.41	0.26	0.30	0.10	0.54	-	0.75	-

<표 5> 커뮤니티 'MIS Study'의 속성기반 개별 사용자 프로파일  $M_C=(P_{hj})$

키워드( $d_j$ ) 사용자( $u_h$ )	CRM	Recommend	Data Mining	Social Network	P2P
Kim	1.25	0.00	1.88	0.00	0.00
Oh	0.00	0.00	0.00	1.11	2.50
Chae	1.25	1.50	0.00	1.00	0.00
Seo	0.83	0.00	0.00	1.67	0.83
Park	1.25	0.83	1.67	0.00	0.00

커뮤니티에 가입한 사용자가 키워드에 대해서 가지는 TF 가중치는 식 (8)에서와 같이 커뮤니티 내에서 사용자가 가지고 있는 키워드의 빈도수를 사용자에게서 가장 많이 출현한 키워드의 빈도수로 나누어 준 값이다. 이는 빈도수를 정규화하여 상대적인 빈도수를 알 수 있게 한다.

IDF 가중치는 식 (9)에서 보여지는 것과 같이 해당 키워드를 전체 사용자 중 몇 명의 사용자들이 가지고 있는지에 대한 비율의 역수로 계산한다.  $V_j$ 는 키워드  $d_j$ 에 대한 개별 사용자들의 값을 나타내는 것으로 사용자  $u_h$ 가 키워드  $d_j$ 를 가지고 있으면 '1', 그렇지 않으면 '0'의 값을 가진다.

$$idf(u_{h \in C_i}, d_j) = \frac{U}{\sum_{h=1}^U V_j} \quad (9)$$

$$(V_j = \begin{cases} 1 & \text{if } d_j \in u_{h \in C_i} \\ 0 & \text{if } d_j \notin u_{h \in C_i} \end{cases})$$

개별 사용자들의 키워드  $d_j$ 에 대한 TF-IDF 가중치  $a(u_h, d_j)$ 도 TF 값과 IDF 값을 곱해서 나타낼 수 있으며 키워드  $d_j$ 가 커뮤니티 내의 소수의 사용자에게서 빈번하게 나타나면 높은 가중치를 얻게 된다.

$$a(u_{h \in C_i}, d_j) = tf \times idf = \frac{freq(u_h, d_j)}{\max freq(u_h, d_j)} \times \frac{U}{\sum_{h=1}^U V_j} \quad (10)$$

<표 5>는 'MIS Study' 커뮤니티내의 사용자들이 가지는 속성기반 개별 사용자 프로파일로써 커뮤니티 내의 다른 사용자들에 대해 각 사용자들이 어떤 속성에 대한 선호도가 더 높은지를 보여준다.

#### 4.2.2 적합성 측정

적합성이란 후보 추천리스트에 속한 책들이 사용자들의 개별적인 선호와 어느 정도 유사한 특성을 가지고 있는지를 의미하는 것으로 사용자와

임계치 이하의 유사도를 가진 책은 후보 추천리스트에서 제거된다.

개별 사용자들과 추천후보목록의 책 간의 유사도는 개별 사용자와 책의 속성기반 프로파일을 이용해서 계산할 수 있다. 타겟 커뮤니티 내의 사용자  $u_h$ 와 후보추천리스트의 책  $B_k$  사이의 유사도  $sim(u_{h \in C_i}, B_{k \in R_T})$ 는 식 (11)과 같이 계산할 수 있다.

$$Sim(u_{h \in C_i}, B_{k \in R_T}) = \frac{\vec{u}_h \cdot \vec{B}_k}{|\vec{u}_h| \times |\vec{B}_k|} = \frac{\sum_{j=1}^m a_{hj} \times b_{kj}}{\sqrt{\sum_{j=1}^m (a_{hj})^2} \times \sqrt{\sum_{j=1}^m (b_{kj})^2}} \quad (11)$$

식 (11)에서  $a_{hj}$ 는 사용자  $h$ 가 키워드  $j$ 에 대해 가지는 가중치이며  $b_{kj}$ 는 책  $k$ 가 키워드  $j$ 에 대해서 가지는 가중치이다. 유사도는 Big-user 사이의 유사도와 마찬가지로 가장 높을 때는 '1', 가장 낮을 때는 '0'의 값을 가진다.

<표 6> 후보추천리스트와 'MIS Study' 회원 간 유사도

	Kim	Oh	Chae	Seo	Park
Book1	0.96	0.39	0.48	0.41	0.90
Book8	0.00	0.00	0.64	0.15	0.35
Book12	0.24	0.16	0.39	0.87	0.21
Book14	0.59	0.04	0.74	0.55	0.66

#### 4.2.3 추천후보목록 필터링

적합도 측정 모듈에서 후보추천리스트와 타겟 커뮤니티 사용자들 간의 유사도가 계산되면 개별 사용자들에게 적합하지 않은 책들을 추천후보목록에서 제거하게 된다. 이 과정은 커뮤니티를 대표하는 Big-user에 대한 추천목록이 생성되었어도 Big-user 프로파일이 개별 사용자 프로파일을 정규화해서 이용하지 않았기 때문에



누락될 수 있는 사용자들의 선호를 고려하기 위한 것이다.

실제로 온라인 커뮤니티에서는 몇 명의 사용자들만이 활발한 활동을 하고 대다수의 사용자들은 주로 커뮤니티에 존재하는 정보를 얻는 활동을 하기 때문에 Big-user 프로파일을 구성할 때 활발한 사용자들의 의견이 많이 반영될 수 밖에 없다. 그러므로 Big-user에 대한 추천리스트는 정보를 얻어가는 활동을 주로 하는 사용자들을 소외시키게 되므로 극단적으로 대부분의 커뮤니티 사용자들이 선호하지 않는 책이 추천될 수 있다. 본 연구에서 제안하는 그룹추천시스템은 선호하는 책을 커뮤니티 차원에서 함께 읽고 정보를 교환하는 것 뿐만 아니라 커뮤니티의 개별사용자들 또한 더 다양한 분야에서 자신의 선호를 벗어나지 않는 책을 접할 수 있어야 한다. 그리고 선호하지 않는 책이 추천될 경우 개별 사용자들이 느끼는 추천의 신뢰도가 낮아지게 되면 그룹추천의 의미가 퇴색된다.

그러므로 적합성 판단 과정을 통해 적합하지 않은 책을 추천후보목록에서 제거함으로써 대부분의 커뮤니티 사용자들이 선호하는 책들로 추천을 제공할 수 있도록 한다.

적합성을 판단하는 기준은 다양한 방법으로 정할 수 있다. 우선 임계치를 정해서 커뮤니티 사용자 중 한 명이라도 임계치  $\tau$  이하의 유사도를 가지고 있는 책을 제거하는 것이다. 임계치는 전체 책에 대한 유사도의 평균이나 각 책에 대한 가장 낮은 유사도를 비교하는 등 임의로 정할 수 있다. 두 번째는 유사도가 높은 사용자들이 많은 순서대로 책을 순위화해서 상위의 책을 선택하는 방법이다. 그러나 본 연구에서 적합성 기반 필터링을 제안하는 목적이 좀 더 유사한 책으로 추천목록을 구성하기 위한 것이 아니라 적합하지 않은 책을 추천후보목록에서 제거하는 것이므로 전자의 방법을 이용한다.

## 5. 예제 : 온라인 커뮤니티 'MIS Study'를 위한 그룹 추천

본 연구에서 제안하는 온라인 커뮤니티를 위한 책 추천시스템은 사용자들이 커뮤니티 활동을 하면서 등록한 글과 구독한 책에 대한 프로파일과 책 데이터베이스를 이용하며 CF 추천 절차와 적합성 기반 필터링 절차를 통해 커뮤니티를 대상으로 책을 추천한다.

추천에 사용되는 사용자 프로파일은 사용자가 가입한 커뮤니티와 해당 커뮤니티에 등록된 글에서 추출한 속성에 대한 선호도로 이루어진 속성 기반 사용자 프로파일과 사용자가 어떠한 책을 읽었는지에 대한 항목 기반 프로파일로 이루어져 있다. 그리고 책 데이터베이스에는 책의 속성 정보가 포함되어 사용자가 구독한 책에 대한 정보를

이용할 수 있다.

사용자, 커뮤니티, 책의 프로파일이 위와 같이 저장되어 있다면 온라인 커뮤니티 'MIS Study'를 위한 책 추천 과정은 다음과 같다.

우선 존재하는 모든 커뮤니티에 대한 Big-user 프로파일은 <표 2>와 같이 구성한다. Big-user 프로파일은 각 커뮤니티에 속해 있는 사용자들의 프로파일을 결합하여 구성되며 각 키워드에 대한 가중치로 표현된다. Big-user 프로파일이 구성되었다면 추천 대상인 'MIS Study'와 각 Big-user들 간의 유사도를 계산하고 가장 유사한 선호도를 지닌 Big-user들을( $f=3$ ) 네이버 Big-user로 구성한다. 이 때, 'MIS Study'는 유사도 1을 가진 네이버 Big-user가 되어 커뮤니티의 개별 사용자들이 읽은 책을 추천에 이용할 수 있도록 한다. 커뮤니티 'MIS Study'와의 유사 선호도는 <표 3>과 같이 MIS Study가 1, Best CRM과 AI World가 0.83이며 Next e-Biz=0.56, Case Review=0.35로 유사도가 높은 순인 MIS Study, Best CRM, AI World가 네이버로 구성된다.

추천후보목록을 만들기 위해서 네이버 Big-user들이 구독한 책들에 대해서 'MIS Study'의 구독 가능성 점수를 계산한다. 이 때, 'MIS Study'의 모든 사용자들이 구독한 Book<sub>3</sub>와 어떤 네이버 Big-user도 구독하지 않은 Book<sub>6</sub>, Book<sub>13</sub>, Book<sub>15</sub>는 계산에서 제외된다. PLS 계산이 완료되면 점수가 높은 Book<sub>1</sub>, Book<sub>8</sub>, Book<sub>12</sub>, Book<sub>14</sub>로 후보 추천리스트( $k=4$ )를 구성하여 타겟 Big-user에 대한 CF 기반 추천후보목록을 만든다.

추천후보목록이 생성되면 적합성 측정 모듈이 'MIS Study'의 개별 사용자들과 추천목록의 책들과의 적합성 여부를 판단한다. 적합성은 사용자와 책의 유사도로 측정하며 'MIS Study'의 사용자들과 후보추천리스트의 적합도는 <표 6>과 같다. 만약 적합성의 기준이 유사도 0.2 이하의 사용자가 전체 사용자의 20% 이하라면 <표 4>에서 0.2 이하의 유사도를 가진 사용자가 2명 이상인 Book<sub>8</sub>이 추천리스트에서 제거된다.

그리고 제외된 Book<sub>8</sub> 이외의 Book<sub>1</sub>, Book<sub>12</sub>, Book<sub>14</sub>가 최종 추천목록으로 'MIS Study'에 제공된다.

온라인 커뮤니티를 대상으로 하는 추천은 일정 기간을 두고 정기적으로 제공되므로 다음 추천 제공 시에는 그 동안 커뮤니티에 등록된 새로운 글과 선호하는 책의 정보로 업데이트 된 사용자 프로파일을 이용해서 1, 2단계 추천절차를 다시 수행한다.

## 6. 결론

현재 많은 사람들이 온라인 커뮤니티에 참여해서 활동하고 있으나 지금까지 연구된 많은 웹 기반 추천시스템은 대부분이 개인들의 상품구매와 같이

개별 사용자에게만 초점을 맞추고 있다.

본 연구에서는 연구 사례가 미비할 뿐만 아니라 많은 구성원이 있는 그룹에 대한 추천시스템이 효율성과 만족도의 딜레마에 빠져있는 상황에서 다양한 크기의 커뮤니티에도 적용 가능하면서 전체적인 만족도도 충족할 수 있는 2단계로 구성된 그룹 추천시스템을 제안한다.

공통의 관심사에 대한 책을 읽고 토론하는 온라인 커뮤니티에 대해서 1단계는 협업필터링(CF)에 기반한 추천과정으로 각 커뮤니티의 Big-user를 만들어서 Big-user의 네이버를 구성한다. 네이버가 구독한 책 목록 중 타겟 Big-user가 구독할 가능성이 높은 책으로 추천후보목록을 구성하므로 기존의 그룹추천이 협업필터링을 이용한 경우 개별 사용자에게 대해 추천리스트를 생성함으로써 발생하는 비효율성을 규모가 큰 커뮤니티에서도 해결할 수 있다. 그리고 그룹을 대상으로 하는 추천에서 가장 큰 문제가 되는 개별 사용자들의 선호도 반영 또한 2단계의 적합성 기반 필터링을 통해 커뮤니티의 사용자들이 선호하지 않는 책을 추천에서 제외할 수 있다.

이와 같은 2단계 추천 과정은 앞에서 설명한 바와 같이 추천의 효율성 및 고객의 만족도를 제고할 수 있으며 또한 커뮤니티 내의 정보교환을 활발하게 함으로써 커뮤니티의 발전과 회원유지를 가능하게 할 것으로 기대한다.

그러나 협업필터링의 문제점으로 지적되었던 입력 데이터의 희박성과 선호정보가 없는 새로운 커뮤니티에 대한 추천 및 신상품 추천은 앞으로 해결해야 할 한계점이다. 책은 속성 정보를 가지고 있으므로 문제가 되지 않지만 커뮤니티 내의 사용자들이 글을 등록하거나 책을 읽는 활동을 활발히 하지 않으면 커뮤니티의 선호도를 파악하는 것이 어렵기 때문에 데이터가 많지 않은 커뮤니티와 새로 생성된 커뮤니티에 대한 추천방법이 필요하다. 그리고 커뮤니티의 개별 사용자들이 임의로 구독한 책에 신간이 포함될 수 있지만 체계적인 방법을 통해 신간을 우선 추천할 수 있는 연구가 이루어져야 한다.

향후에 본 연구의 추천시스템의 성능을 검증하고 커뮤니티의 특성에 따라 추천 결과가 어떻게 달라지는지를 알아보기 위해 실제 커뮤니티 데이터를 이용해서 실험을 수행하도록 할 것이다.

## [참고문헌]

- [1] A. Jameson.(2004) "More than the sum of its members: Challenges for group recommender systems." Proceedings of the International Working Conference on Advanced Visual Interfaces, pp.48-54
- [2] D. L. Chao, J. Balthrop, and S. Forrest.(2004) "Adaptive Radio: Achieving consensus using negative preferences." Technical Report TR-CS-2004-08, Dept. of Computer Science, The University of New Mexico

- [3] J. F. McCarthy and T. D. Anagnost. (1998) "MusicFX: An arbiter of group preferences for computer supported collaborative workouts." Proceedings of the 1998 Conference on Computer-Supported Cooperative Work, pp.363-372
- [4] J. F. McCarthy (2002) "POCKET RESTUARANTFINDER: A Situated Recommender Systems for Groups.", Workshop on Mobile Ad-Hoc Communication at the 2002 ACM Conference on Human Factors in Computer Systems
- [5] Michael J. Pazzani (1999) "A Framework for Collaborative, Content-Based and Demographic Filtering." Artificial Intelligence Review Vol.13 No.5-6 pp. 393-408
- [6] M. O'Connor, D. Cosley, J. A. Konstan, and J. Riedl (2001) "PolyLens: A Recommender System for Groups of Users." Proceedings of the seventh European Conference on Computer Supported Cooperative Work, pp.199-218
- [7] R. B. Segal and J. O. Kephart(2000) "Swiftfile: An intelligent assistant for organizing e-mail." In AAAI 2000 Spring Symposium on Adaptive User Interfaces
- [8] Y. Zhiwen, Z. Xingshe, and Z. Daqing (2005) "An adaptive in-vehicle multimedia recommender for group users." Proceedings of the Vehicular Technology Conference, 2005. Vol.5, pp.2800- 2804
- [9] 김재경, 조운호, 김승태, 김혜경 (2005) "모바일 전자상거래 환경에 적합한 개인화된 추천시스템" 경영정보학연구, 제15권 제3호, pp.223-241
- [10] 류미, 박주석, 김재경(2004) "사용자 지식을 반영한 메일 폴더 추천 방법론" 한국지능정보시스템학회논문지, 제10권 3호, pp.133-146