

동적 사용자 프로필 및 협업 필터링을 이용한 소셜 네트워크 그룹 추천

Social Network Group Recommendation Using Dynamic User Profiles and Collaborative Filtering

양희태, 차재홍, 안민제, 임중태, 이하, 복경수, 유재수
충북대학교 정보통신공학과

Heetae Yang(yht4534@naver.com), Jaehong Cha(kakasi1527@naver.com),
Minje Ahn(allgrudge86@gmail.com), Jongtae Lim(jtlim@chungbuk.ac.kr),
He Li(lihelol@gmail.com), Kyoungsoo Bok(ksbok@chungbuk.ac.kr),
Jaesoo Yoo(yjs@chungbuk.ac.kr)

요약

최근 SNS(Social Network Service)의 사용이 급격히 증가함에 따라 추천 기법에 대한 연구가 활발히 진행되고 있다. 추천 기법은 사용자들이 좋아하거나 필요할만한 다양한 서비스들을 실시간으로 제공하는 기법이다. 그 중 그룹 추천은 사용자의 성향 정보를 기반으로 적합한 그룹을 제공해 주는 기법이다. 본 논문에서는 소셜 네트워크 환경에서 사용자 프로필 및 협업 필터링을 이용한 그룹 추천 기법을 제안한다. 제안하는 기법은 사용자의 최근 그룹 활동 정보를 수집하여 프로필 정보를 갱신하기 때문에 기존의 정적 프로필 기반의 그룹 추천 기법의 최근 사용자의 성향을 고려하지 못하는 문제점을 해결한다. 또한, 협업 필터링을 통해 그룹 내 자신의 성향과 비슷한 사용자들의 프로필 데이터를 활용하여 그룹을 추천함으로써 사용자에게 좀 더 다양한 그룹을 제공한다. 성능 평가 결과 제안하는 기법이 기존 기법에 비해 사용자의 변화하는 성향이 충분히 반영된 다양한 그룹 추천이 이루어지는 것을 확인 할 수 있었다.

■ 중심어 : | 소셜 네트워크 | 프로필 | 협업필터링 | 그룹추천 | SNS |

Abstract

Recently, as SNS services have been increased, studies on recommendation schemes have been actively done. Recommendation scheme provides various favorable or needed services with users on real time. Group recommendation provides users with suitable groups based on their preference. In this paper, we propose a new group recommendation scheme considering user profiles and collaborative filtering in social networks. The proposed scheme can solve the problems of the static profile based group recommendation scheme because it collects the recent group activities and updates user profiles. It also recommends the more various groups by reflecting the similar tendencies of other users within a group through collaborative filtering. Our experimental results show that the proposed scheme recommends various groups that significantly considers the user's changing preferences compared to the existing scheme.

■ keyword : | Social Network | Profile | Collaborative Filtering | Group Recommendation | SNS |

* 이 논문은 미래창조과학부 및 정보통신산업진흥원의 대학 IT연구센터 지원사업(NIPA-2013-H0301-13-4009)과 2013년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No.2013R1A2A2A01015710)

접수일자 : 2013년 09월 23일

심사완료일 : 2013년 11월 01일

수정일자 : 2013년 11월 01일

교신저자 : 유재수, e-mail : yjs@chungbuk.ac.kr

I. 서론

인맥 관리, 정보 공유 등에 대한 사용자 요구가 급증함에 따라 SNS(Social Network Service)에 대한 사용이 활성화되고 있다. 사용자의 다양한 요구 사항을 빠르게 반영하기 위해 사용자의 성향을 고려한 추천 기법들에 대한 연구들 활발하게 진행되고 있다[1-4]. 이러한 추천 기법 중 그룹 추천은 사용자의 활동이나 성향을 분석하여 사용자의 관심 대상이 되는 추가적인 그룹을 제공하거나 그룹 내에 특정 아이템을 제공하는 것이다[2-4]. 소셜 네트워크에서 그룹은 정적 그룹과 동적 그룹으로 분류된다. 정적 그룹은 특수한 목적을 위해 모인 집단으로써 표면적으로 그룹에서 다루는 콘텐츠 유형이 결정되어 있다. 예를 들어, 테니스 동호회, 영화 모임 등은 특수 목적으로 사용자들이 가입한 집단으로 그룹 내에서 상호 공유되는 정보의 유형이 정해져 있다. 동적 그룹은 정적 그룹과 상반되는 집단으로 특정 이벤트 또는 상황에 따라 일시적으로 생성되었다가 소멸되는 집단을 의미한다. 동일한 TV 채널을 시청하는 사용자, 특정 지역 내에 모인 사용자들이 이에 속한다.

이러한 그룹에 대한 추천은 사용자의 성향에 맞는 그룹을 추천하거나 그룹 구성원에 적절한 아이템을 추천함으로써 서비스 만족도를 향상시키고 향상된 정보를 공유할 수 있게 한다[5-9]. 일반적인 그룹 추천 기법은 사용자의 프로필을 기반으로 관심 대상이 되는 그룹이나 해당 그룹에 적절한 아이템을 제공한다. [2]에서는 TKI(Thomas-Kilmann Conflict Mode Instrument)를 사용하여 설문 조사를 수행하고 사용자의 성향을 파악한다. 이러한 사용자 성향을 이용하여 유사한 성향을 갖는 사용자들이 포함된 그룹을 추천한다. [3]에서는 유사한 성향을 갖는 사용자들을 그룹화하고 그 그룹에 그룹 구성원들에게 신뢰성 있는 아이템을 추천하는 기법을 제안하였다. [4][5]에서는 사회적인 특성을 이용하여 그룹에 아이템을 제공하는 추천 기법을 제안하였다. 사회적인 특성으로 그룹 구성원과의 관계, 구성원의 전문성, 구성원의 성향 등을 고려하여 그룹 구성원 전체의 만족을 최대로 높일 수 있는 추천 방식을 제공한다. [8]에서는 추천 효율을 높이기 위해 해당 분야의 전문적

정보를 가지고 있는 전문가의 정보를 반영하여 추천 시스템에 활용하였다. [10]에서는 SNS를 비롯한 온라인이나 디지털 환경에서 활동한 정보들이 아이템이나 그룹 및 다양한 콘텐츠를 추천받거나 구매함에 미치는 영향력에 대해 분석하였다.

기존 그룹 추천 기법들은 정적인 프로필에 기반 하여 선호도를 생성하고 그에 해당하는 아이템이나 관련된 그룹을 추천한다. 그러나 대부분의 사용자들은 그룹 가입 시에 입력된 프로필을 거의 변경하지 않기 때문에 사용자의 성향이 변경된 내용을 추천에 활용할 수 없다. 이로 인해 추천의 정확성이 저하되고 정보의 최신성이 반영되지 못하는 문제점이 있다. 또한, 추천 과정에서 사용자 프로필 정보만을 반영하기 때문에 다른 사용자들의 최근 성향을 반영하지 못하는 문제점이 있다.

본 논문에서는 사용자 동적 프로필과 소셜 네트워크 사용자들의 성향을 이용한 정적 그룹 추천 기법을 제안한다. 본 논문에서 정적 그룹은 사용자들이 특수한 목적으로 모인 온라인 커뮤니티로 정의한다. 제안하는 기법은 사용자의 최근 활동 정보를 분석하여 동적 프로필을 생성한다. 동적 프로필은 그룹 활동 정보를 통해서 지속적으로 갱신되기 때문에 사용자의 최근 성향을 반영할 수 있다. 또한 그룹 추천 과정에서 그룹 내의 존재하는 사용자들의 활동 정보를 수집하여 그룹 내에서 최근 사용자들의 성향 정보를 반영한다. 추천 과정에서 다른 사용자와 자신의 성향을 동시에 고려하기 위해 협업 필터링을 사용하여 그룹을 추천한다. 본 논문에서는 정적 프로필 기반에서 변화하는 사용자의 성향의 최신성을 반영하지 못한다는 점을 해결하며, 해당 분야의 자신과 비슷한 성향을 가진 다른 사용자의 프로필 정보 또한 고려하기 때문에 기존의 기법보다 사용자에게 질 높고 정확한 그룹 추천 서비스를 제공한다.

본 논문의 나머지 구성은 다음과 같다. II장에서는 그룹 추천 기법에 관한 기존 연구에 대해 기술하고 III장에서는 제안하는 그룹 추천 기법에 대하여 기술한다. IV장에서는 제안하는 기법의 우수성을 입증하기 위해 성능 평가를 수행한 결과를 기술한다. 마지막 V장에서는 본 논문의 결론 및 향후 연구에 대해 기술한다.

II. 관련 연구

프로필 기반 그룹 추천은 사용자가 서비스에 가입할 때 입력한 정적인 프로필 정보를 활용하여 그룹을 추천하는 기법이다. 이러한 기법은 사용자 그룹 가입 시 입력한 정적 프로필 정보를 활용하여 사용자의 성향 정보를 도출하고 사용자의 성향에 적합한 그룹을 추천할 수 있다. [2]에서는 사용자의 프로필 정보를 생성하기 위해 TKI를 사용하였다. [그림 1]은 TKI의 5가지 성향을 나타낸다. 사용자들의 성향을 적극성 및 협조성에 따라 Competing, Collaborating, Avoiding, Accomodating, Compromising 으로 구분한다. 추천된 그룹 내 자신과 친분이 있는 사용자와의 관계성을 10가지 속성 비교를 통해 추천된 그룹에 대한 신뢰도를 향상시켰다.

[3]에서는 4단계를 거쳐 사용자의 프로필 정보를 기반으로 그룹을 생성하고 해당 그룹에 아이템을 추천하는 기법을 제안하였다. 첫 번째 단계는 특정 아이템 평가 정보를 이용하여 사용자 성향을 벡터화 하여 사용자들의 유사성을 판별하고 두 번째 단계는 사용자들을 그룹화 한다. 세 번째 단계는 그룹 내의 사용자들이 기존에 아이템 평가 정보를 이용하여 그룹 성향을 생성한다. 마지막 단계에서는 기존에 평가된 아이템과 평가되지 않는 아이템의 유사도를 비교하여 평가되지 않은 아이템의 평가 점수를 예측한다.

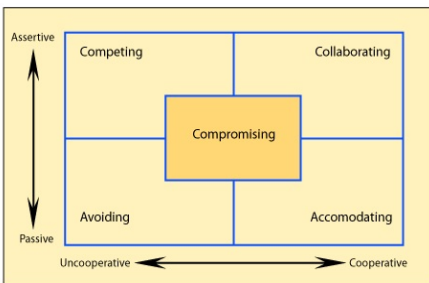


그림 1. TKI 구성도

다수 사용자들로부터 획득한 선호도 정보에 따라 사용자들의 관심사들을 자동적으로 추천하기 위해 협업 필터링 기법들이 연구되었다[11-13]. 협업 필터링은 사용자들의 선호도를 기반으로 유사한 성향을 가지는 다른 사용자들을 식별하고 식별된 다른 사용자가 가지는

정보 중 사용자가 없는 정보를 추천해 주는 기법이다. [11]에서는 소셜 환경에서 협업 필터링을 사용할 경우 추천 결과에 미치는 다양한 속성을 분석하고 이를 평가 하였다. [12]에서는 사용자 성향과 위치 정보를 협업 필터링에 적용하여 웹 서비스를 추천하는 기법을 제안하였다.

기존 연구에서는 정적 프로필을 기반으로 하여 사용자의 성향을 파악하고 그 성향에 맞는 그룹을 제공한다. 본 논문에서는 기존 정적 프로필 기반의 그룹추천 방식에서 사용자의 최근 성향을 반영하지 못하는 문제점을 동적 프로필을 활용하여 해결 하였으며, 자신의 성향과 비슷한 다른 사용자의 정보 또한 고려함으로써 기존 정적 프로필 기반 추천 방식보다 다양하고 정확한 그룹 추천을 제공한다.

III. 제안하는 그룹 추천 기법

1. 제안하는 시스템의 구조

제안하는 그룹 추천 기법은 사용자의 최근 성향과 그룹 내 사용자들의 성향을 동시에 고려하여 그룹 추천의 신뢰성 및 만족도를 향상시킨다. 사용자의 활동 내역을 통해 동적으로 프로필을 생성하고 그룹 내 다른 사용자의 프로필을 고려하기 때문에 신뢰도가 향상된다. 또한, 그룹 추천 기법에 협업 필터링을 적용하여 그룹 내에서 자신의 성향과 유사한 다른 사용자가 가입한 그룹을 추천하기 때문에 만족도를 향상시킨다. [그림 2]는 제안하는 기법의 전체 구성을 나타낸 것이다.

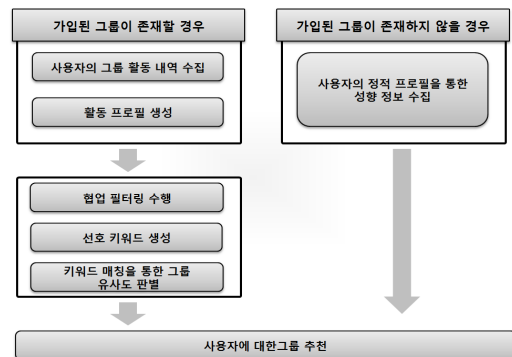


그림 2. 시스템 전체 구성도

그룹 추천을 위해 우선 그룹에 가입된 여부를 판별한다. 만약 사용자가 그룹에 가입되어 있지 않다면 사용자가 기술한 정적 프로필 정보를 이용하여 그룹을 추천한다. 만약 사용자가 그룹에 가입되어 있다면 그룹 내에서 사용자가 최근에 수행한 활동을 기반으로 사용자의 선호도를 생성하고 동적 프로필을 생성한다. 그룹 내 다른 사용자들의 성향을 반영하기 위해 사용자가 가입한 그룹 내 사용자들의 최근 활동 정보를 분석하여 협업 필터링을 수행한다. 이후 협업 필터링을 수행하여 도출한 정보를 통해 선호도를 생성한다. 선호도는 사용자의 특성을 반영한 키워드로 도출되며 이를 선호 키워드라 정의한다. 생성된 선호 키워드를 통해 그룹 내 태그와 키워드에 대한 유사도를 비교하여 그룹 추천을 수행한다.

2. 활동 프로필

정적 프로필은 사용자가 관심 있는 항목을 직접 기술한 것으로 대부분의 사용자는 해당 프로필을 변경하지 않는다. 따라서 정적 프로필은 사용자의 최근 성향을 파악하기 위한 목적으로 사용하는 것은 문제가 있다. 제안하는 기법에서는 사용자의 최근 성향을 표현하기 위해 활동 프로필을 사용한다. 활동 프로필은 사용자의 그룹 활동 정보를 기반으로 최근 사용자의 관심 항목을 추출하여 생성한 것으로 주기적으로 갱신되는 동적 프로필과 같다. 그룹에 가입하여 활동을 시작할 때에는 사용자가 정적 프로필을 입력하지만 그룹 활동을 수행할수록 최신 성향을 반영한 활동 프로필이 갱신된다. 활동 프로필 갱신은 동일한 카테고리에 있는 그룹 활동 정보를 수집하여 사용자의 성향 정보를 생성한다. 예를 들어, '음악'이란 카테고리가 존재하고 사용자가 3개의 '음악' 관련된 그룹에서 활동한다고 가정하자. 활동 프로필은 사용자가 가입한 음악 관련 3개의 그룹 정보에서 수행한 활동을 종합하여 최근 성향을 갱신한다.

활동 프로필을 갱신하기 위해 우선 사용자가 가입한 모든 그룹에 대해 각각의 활동량을 계산한다. 활동량은 사용자가 최근에 동일한 카테고리로 분류된 그룹에서 활동한 내역을 수치로 나타낸 값이다. 활동량은 사용자가 그룹 내에서 수행한 활동에 영향을 미칠 수 있는 속

성 값을 일정 기간 동안 수집하여 계산한다. 이때, 속성 값은 작성한 글, 댓글, 체크인 수 등이 해당될 수 있다. 사용자 각각의 활동량은 (식 1)을 이용하여 계산한다. 그룹 i 내에서 사용자의 활동량 PV_i^j (Profile Value)은 (식 1)과 같이 그룹에서 사용자 j 가 수행한 활동량을 나타낸다. (식 1)에서 N_{ij} 는 그룹 i 내에서 j 사용자가 정의한 임계 시간 내 수행한 활동량을 계산하기 위한 속성, α_{ij} 는 N_{ij} 에 활동에 영향을 줄 수 있는 속성 값에 부여된 가중치, t 는 현재 시각, m 은 사용자의 활동량 정보가 수집되는 최초시점, n 은 수집되는 최초시점을 조절하는 변수이다. p 와 k 는 과거부터 현재까지 활동량을 종합할 때 각각의 시점에서 부여될 PV_i^j 값의 가중치를 조정하기 위한 상수이다. 예를 들어, p 와 k 에 의해서 수식의 기울기 증가율이 높아지면 현재 시점에 가까운 과거 활동정보의 가중치는 높아지며 기울기의 증가율이 낮다면 현재 시점에서 가까운 과거 활동정보의 가중치는 낮아진다.

$$PV_i^j = \sum_{m=t-n}^t (10e^{-k/p}) N_{ij} \alpha_{ij} \quad (1)$$

동일한 그룹 카테고리에서 수행한 모든 활동량을 결합하여 사용자의 활동 프로필을 갱신한다. 계산된 PV_i^j 값의 종합에서 높은 값을 가지는 k 개의 그룹의 키워드 정보를 사용자의 활동 프로필을 생성하기 위해 사용한다. 그룹 키워드는 그룹의 특성을 설명하기 위해 부여한 값으로 그룹 관리자에 의해 부여되거나 텍스트 마이닝을 통해 자동으로 생성된다. 활동 프로필은 사용자가 활동한 그룹 카테고리별로 생성된다. 동일한 그룹 카테고리에 포함된 각 그룹 키워드들에서 활동 프로필로 사용하기 위한 키워드를 선출하기 위해 각 그룹 키워드에 KV_{ik}^j 를 부여한다. KV_{ik}^j 는 그룹 i 에 포함된 키워드 k 에 부여된 사용자 j 의 PV_i^j 값으로 각 키워드에 부여된 가중치를 의미한다. 동일한 그룹 카테고리에 포함된 모든 그룹에서 KV_{ik}^j 가 부여되면 (식 2)을 이용하여 사용자 j 에 대한 키워드 k 의 선호도를 생성한다. 키워드에 대한 선호도가 생성되면 임계치 이상의 선호도를 갖는 키워드를 활동 프로필로 생성한다. 즉, 사용자가 많이

활동한 그룹 내에 포함된 키워드들을 사용자의 활동 프로파일로 사용한다.

$$KW_k^j = \sum_{i=1}^n KV_{ik}^j \quad (2)$$

[그림 3]은 사용자 활동량을 계산하는 과정을 나타낸 것이다. 만약 사용자가 특정 그룹에서 [그림 3(a)]에 해당하는 속성 값에 대한 활동을 했다고 가정하자. 또한 활동량 계산 범위를 최근 4달 내에 활동한 활동량을 기준으로 정한다고 가정하자. 활동량을 계산하기 위한 각 속성에 부여된 가중치는 [그림 3(b)]와 같고 각 시점별 가중치는 [그림 3(c)]와 같다고 하자. 1월의 경우 $((50 * 0.5) + (90 * 0.3) + (200 * 0.2)) * 0.1$ 이 되어서 최종적으로 1월의 활동량은 9가 도출된다. 2월, 3월, 4월의 계산도 동일하게 이루어지며 최종적으로 최근 4달간에 이 그룹에서 활동한 활동량은 53으로 결정된다.

	1월	2월	3월	4월	Total
작성한 글 (M_{ij})	50	40	20	40	200
댓글 수 (N_{ij})	90	30	30	90	240
체크인 수(M_{ij})	200	150	100	50	600

(a) 사용자 활동 정보

	작성글	댓글	체크인	Total
가중치(α_{ij})	50%	30%	20%	100%

(b) 속성 별 가중치

	1월	2월	3월	4월	Total
가중치($10e^{-kt}$)	10%	20%	30%	40%	100%

(c) 시점 별 가중치

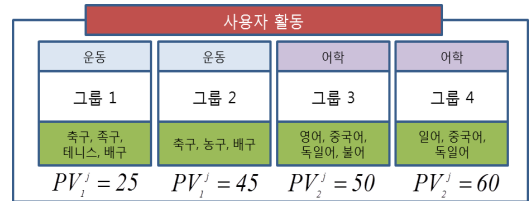
	1월	2월	3월	4월	Total
활동량(PV_i^j)	9	12	9	23	53

(d) 최종 활동량

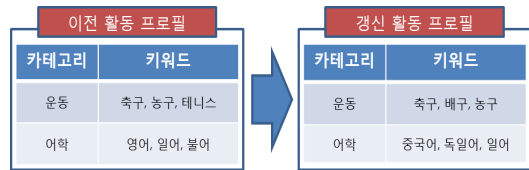
그림 3. 활동량 계산 과정

[그림 4]는 사용자 활동 프로ファイルを 갱신하는 과정을 나타낸 것이다. 만약 사용자 j 가 운동과 어학에 관련된 2개의 카테고리에 활동하고 사용자의 활동량은 (식 1)

을 이용하여 [그림 4(a)]와 같이 생성되었다고 가정하자. [그림 4(b)]와 같이 운동 카테고리의 활동 프로파일은 축구, 농구, 테니스이었으나 (식 2)를 이용하여 활동 프로 파일을 갱신할 경우 축구, 배구, 농구가 된다. 이와 유사하게 어학 카테고리의 활동 프로파일도 중국어, 독일어, 일어로 변경된다.



(a) 사용자 활동량



(b) 활동 프로파일 갱신

그림 4. 활동 프로파일 갱신 과정

3. 그룹 추천

사용자와 동일한 그룹 카테고리 내에서 활동하는 다른 사용자의 성향을 반영하기 위해 제안하는 기법에서는 활동 프로파일에 대한 협업 필터링을 수행한다. 즉, 그룹 추천을 위해 사용자의 성향을 나타내는 활동 프로파일과 그룹 내에 존재하는 자신과 유사한 성향을 가지는 사용자들의 성향을 고려하여 선호 키워드를 생성한다. 협업 필터링을 수행하기 위해 우선 사용자와 동일한 그룹 카테고리 내에서 활동하는 사용자에 대해 유사한 성향을 가지는 사용자를 선별한다. 유사한 성향을 가지는 사용자가 도출되면 도출된 협업 필터링을 수행하여 유사한 사용자의 활동 프로파일 키워드 항목 중에서 자신의 활동 프로파일 키워드 항목에 존재하지 않는 키워드를 추출한다. 추출된 키워드에 대해 우선 순위를 부여하기 위해 사용자와의 유사도에 따라 각각의 키워드에 점수를 부여한다. 이렇게 부여한 점수가 임계치 이상일 경우 선호 키워드로 선정하고 이를 통해 그룹을 추천한다.

사용자와 유사한 성향을 가진 사용자를 선별하기 위해 사용자와 동일한 그룹 카테고리 내에서 활동하는 사용자를 검색하고 사용자들의 활동 프로필의 유사성을 비교한다. 유사성은 동일한 그룹 카테고리에 활동한 사용자의 활동 프로필에 포함된 키워드를 비교하여 계산한다. (식 3)는 사용자들의 유사성을 계산하는 식이다. 이때, SM_{ij}^k 은 그룹 카테고리 k 에서 사용자 i, j 의 활동 프로필 유사성, UK_i^k 는 그룹 카테고리 k 에서 사용자 i 의 활동 프로필 키워드 수, SK_{ij}^k 는 그룹 카테고리 k 에서 사용자 i, j 에 대한 공동 활동 프로필 키워드 수이다.

$$SM_{ij}^k = \frac{SK_{ij}^k}{UK_i^k} \times \frac{SK_{ij}^k}{UK_j^k} \quad (3)$$

[그림 5]는 유사한 성향을 가지는 사용자를 판별하는 과정을 나타낸 것이다. 사용자 A의 활동 프로필을 중심으로 (식 3)를 이용하여 사용자 B, 사용자 C, 사용자 D, 사용자 E와의 유사성을 계산한다. 예를 들어, [그림 4](a)와 같이 사용자 B는 사용자 A와 3개의 동일 키워드를 포함하고 있고 사용자 C, D, E는 각각 2, 1, 2의 동일 키워드를 포함하고 있다. 사용자 유사성을 계산할 경우 [그림 4](b)와 같이 사용자 B의 유사성은 약 0.56이고 사용자 C, D, E는 약 0.33~0.06이다. 만약 유사성 임계치가 0.5이라고 가정하면 유사한 사용자라고 판단되는 사용자는 사용자 B가 된다.

키워드	키워드	키워드	키워드	키워드
축구	축구	축구	축구	축구
족구	족구	족구	골프	족구
농구	농구	테니스	탁구	하키
야구	테니스	탁구	하키	테니스
사용자A 활동 프로필	사용자B 활동 프로필	사용자C 활동 프로필	사용자D 활동 프로필	사용자E 활동 프로필

(a) 사용자 활동 프로필의 키워드

사용자	유사도
사용자B	0.56
사용자C	0.33
사용자D	0.06
사용자E	0.33

(b) 사용자 유사도

그림 5. 사용자 유사성 추출 과정

협업 필터링은 특정 사용자와 유사한 성향을 가지는 다른 사용자의 성향을 반영하여 그룹을 추천하기 위해 사용한다. 즉, 협업 필터링을 통해 유사한 성향을 가지는 사용자 활동 프로필을 추출하여 그룹 추천에 활용한다. 그룹 추천을 위한 대상 그룹을 검색하기 위해 선호 키워드를 사용한다. 선호 키워드는 동일한 그룹 카테고리 내에서 유사한 성향을 가지는 사용자의 활동 프로필 중 특정 사용자가 관심을 가질 것으로 판단되는 키워드를 추출한 것이다. 추출한 키워드에 대해 우선 순위를 부여하여 선호 키워드로 정의한다. (식 4)를 통해 추출된 사용자의 활동 프로필과 그룹을 추천받은 사용자의 활동 프로필을 협업 필터링하여 선호 키워드를 추출한다. 이때, KP_{ip}^k 는 사용자 i 에 대해 그룹 카테고리 k 에서 협업 필터링을 통해 추출한 키워드 p 의 점수를 나타내며 n 은 유사 사용자 수를 나타낸다. 추출된 키워드에 우선 순위를 부여하기 위하여 (식 3)을 통해 도출된 유사성 값을 이용하여 최종 키워드 점수를 부여한다. 점수가 부여된 키워드 중 임계치 이상의 값을 가지는 키워드를 선호 키워드로 선정한다.

$$KP_{ip}^k = \sum_{j=1}^n SM_{ij}^k \quad (4)$$

[그림 6]은 KP_{ip}^k 를 계산하는 과정을 나타낸 것이다. 운동에 관련된 그룹에서 협업 필터링을 통해 추출된 키워드가 ‘테니스’, ‘탁구’, ‘하키’라고 가정하자. 테니스의 경우 사용자 B의 유사도 점수가 0.56, 사용자 C의 경우 0.33, 사용자 E의 경우 0.33이기 때문에 최종적으로 테니스 키워드의 KP값은 1.22이다. 임계치 값이 1이라고 가정할 경우 최종적으로 ‘테니스’ 키워드가 선호 키워드로 선정된다.

	사용자 B 프로필 KP	사용자 C 프로필 KP	사용자 E 프로필 KP	Total KP
테니스	0.56	0.33	0.33	1.22
탁구	0	0.33	0	0.33
하키	0	0	0.06	0.06

그림 6. 선호 키워드 추출 과정

선호 키워드가 선정되면 이를 이용하여 사용자에게

추천 그룹을 검색한다. 그룹에 부여된 키워드와 선호 키워드를 비교하여 일치되는 키워드의 수에 따라 추천 순위가 결정된다. [그림 7]은 추천 키워드를 기반으로 그룹이 추천되는 과정을 설명한다. (식 4)를 통해 선정된 선호 키워드가 ‘테니스’, ‘탁구’, ‘하키’ 라 가정하자. 선호 키워드를 통해 그룹을 검색한 결과 [그림 6]의 그룹A~D가 결과 값으로 도출 되었다고 하자. 추천되는 그룹의 우선순위는 선호 키워드와 그룹 키워드의 일치 수를 기준으로 판단한다. 결과적으로 그룹 추천 우선순위는 그룹 B, C, A, D 순으로 결정된다.

그룹	키워드
그룹A	축구, 농구, 테니스, 골프, 하이킹, 족구
그룹B	테니스, 탁구, 하키, 축구, 농구, 야구, 배구
그룹C	하키, 탁구, 축구, 당구, 배구, 하키, 골프
그룹D	축구, 농구, 테니스, 축구, 당구, 골프

그림 7. 추천 우선순위 추출 과정

[그림 8]은 제안하는 그룹 추천 알고리즘을 나타낸 것이다. 입력은 프로필에 대한 정보인 Profile_SET이고 출력은 추천 그룹에 대한 정보인 Group_SET으로 구성된다. 그룹 활동을 처음 시작하는 사용자는 정적 프로필의 정보를 활용하여 그룹 추천이 수행된다(Line2-4). 활동 그룹이 존재하는 사용자의 경우 그룹 활동 내역을 통해 활동 프로필을 추출하고 사용자가 가입한 그룹을 검색하여 그룹을 카테고리별로 분류한다.(Line6-8). (식 2)를 통해 그룹 내에서 유사 성향을 가지는 사용자를 검색하고 유사 성향을 갖는 사용자에게 존재하는 키워드를 도출하여 점수를 부여한다(Line9-17). 점수가 부여된 키워드 중 임계치 이상의 점수를 가지는 선호 키워드를 추출하고 그룹 정보와 비교하여 가장 많은 수의 키워드가 매칭되는 그룹을 사용자에게 추천한다 (Line18-21).

```

그룹 추천
입력 : Profile_SET      출력 : Group_SET
저장리스트 : List_Key_P (추출 키워드 점수 저장 리스트)
              List_Key_D (선호 키워드 저장 리스트)

1:  if(그룹이 존재 하지 않을 경우)
2:      정적 프로필(Profile_SET) 수집;
3:      선호 키워드(Keyward_SET) 수집;
4:      그룹 검색;
5:  else if(그룹이 존재할 경우)
6:      활동 프로필(Profile_SET) 추출;
7:      사용자가 가입한 그룹검색;
8:      카테고리별 그룹 분류;
9:      성향이 유사한 사용자를 SMij을 통해 계산;
10:     for(SM에 의해 추출된 사용자)
11:         KPij 계산;
12:         List_key_P[] += KP;
13:     end for
14:     if(임계치와 List_Key_P 비교)
15:         상위n개 List_Key_D에 저장;
16:     end if
17:     선호키워드 후보 키워드에 KP를 이용한 점수부여;
18:     선호키워드 추출;
19:     선호키워드를 통한 그룹 유사도 비교;
20:     Group_SET 생성;
21:     그룹 추천;
22: end if
    
```

그림 8. 그룹 추천 알고리즘

IV. 성능평가

제안하는 기법의 우수성을 입증하기 위해 성능평가를 수행하였다. 이전에 제안된 기법인 프로필의 갱신이 이루어 지지 않는 정적 프로필 기반의 그룹 추천방식과 제안하는 기법의 그룹추천 방식의 그룹 추천 정확도에 대하여 비교 평가를 수행하였다. 임의로 생성한 100,000명의 성향을 데이터 집합으로 사용하였으며 RAM 4GB, CPU Intel(R) Core(TM) i3-2100, Window 7 Ultimate K OS 컴퓨터 환경에서 C# 언어로 구현하였다. [표 1]은 실험 평가 환경을 나타내며 [표 2]는 실험에서 초기에 입력된 프로필의 키워드와 제안하는 기법을 사용하였을 때 갱신된 프로필 키워드를 비교하여 나타낸 것이다.

표 1. 실험 환경

파라미터	값
사용자 수	100,000
그룹 수	104
PV_i^k	1 ~ 100

표 2. 활동 프로필 갱신

초기 프로필	갱신 프로필
CN	DV
DX	AF
CO	BJ
BS	CR
AB	AJ

[그림 9]는 갱신된 프로필을 기반으로 추천된 그룹과 협업 필터링을 이용하여 도출된 선호 키워드로 선정된 기존 후보 키워드들의 점수 값을 나타낸다. [표 2]를 이용하여 사용자의 동적 프로필을 갱신하고 선호 키워드는 KP_{ij}^k 값 중 상위 5개를 포함하는 값으로 임계치를 설정하였다. 초기 활동 프로필 정보를 통해 5개의 선호 키워드(CN, DX, CO, BS, AB)를 지정 하였다. 사용자가 활동하는 각 그룹의 PV_i^k 값을 각각 1~100로 랜덤하게 부여함에 따라 초기 활동 프로필의 키워드(CN, DX, CO, BS, AB)에서 (DV, AF, BJ, CR, AJ)로 갱신되는 것을 확인 할 수 있었다.

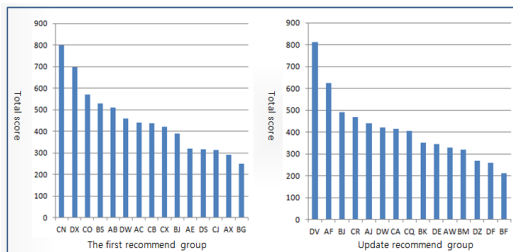


그림 9. 추천 키워드 및 점수 비교

[표 3]은 추천 정확도를 확인하기 위한 성능 평가 환경을 나타낸다. [그림 10]는 기존 기법과 제안하는 기법의 그룹 추천에 대한 정확도를 비교한 것이다. 기존 기법과 제안하는 기법을 이용하여 지속적으로 성향이 변

화하는 사용자에게 그룹을 추천 하였을 때 사용자의 현재 성향과 추천된 그룹의 키워드의 일치율을 비교하여 정확도를 비교하였다. 개인의 성향을 50 가지로 분류 하였으며 개인이 성향이 빠르게 변화하는 그룹(Fast tendency group, F_Group), 일반적으로 변화하는 그룹(Normal tendency group, N_group), 거의 변하지 않는 그룹(Static tendency group, S_group)으로 분류하여 10,000번의 성능 평가 결과에 대한 평균한 값이다. 그룹 추천에 대한 정확도는 기존 기법과 비교하여 평균 50% 성능이 향상된 것을 확인할 수 있었다.

표 3. 실험 환경 2

파라미터	값
그룹 성향	3
개인 성향	50
추천 횟수	10,000

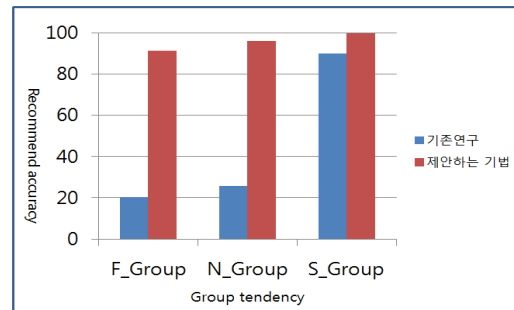


그림 10. 그룹 추천 정확도 비교

V. 결론 및 향후 연구

본 논문에서는 소셜 환경에서 활동 프로필과 협업 필터링을 이용한 그룹 추천 기법을 제안하였다. 기존 그룹 추천 기법은 정적 프로필을 사용하기 때문에 신뢰성 및 만족도가 저하되는 문제점이 있다. 제안하는 기법에서는 활동 프로필의 지속적인 갱신과 그룹 내 사용자들의 성향을 고려하여 기존 기법의 문제점을 해결했다. 성능평가 결과 기존의 기법은 시간의 흐름에 따른 사용자의 성향 변화를 충분히 반영하지 못하여 추천되는 그룹이 사용자의 성향과의 일치율이 낮게 나오는 것을 확

인할 수 있었다. 제안하는 기법은 사용자의 성향 변화에 따라 프로필이 갱신되기 때문에 기존기법에 비해 약 50% 정확도를 향상시켰다. 향후 연구로는 그룹 추천을 수행할 때 다른 카테고리와의 연관성까지 고려하여 좀 더 다양한 그룹을 추천하는 기법을 연구할 예정이다.

참고 문헌

[1] 송정진, 이상철, 김상욱, “사용자의 다양성 성향을 고려하는 추천 기법”, 정보과학회논문지:데이터베이스, 제40권, 제2호, pp.99-105, 2012.

[2] L. Quijano-Sanchez, J. A. Recio-Garcia, and B. Diaz-Agudo, “Personality and Social Trust in Group Recommendations,” Proc. International Conference on Tools with Artificial Intelligence, Vol.3, pp.121-126, 2010.

[3] L. Boratto, S. Carta, A. Chessa, M. Agelli, and M. L. Clemente, “Group Recommendation with Automatic Identification of Users Communities,” Proc. International Workshop on Intelligent Web Interaction, Vol.3, pp.547-550, 2009.

[4] M. Gartrell, X. Xing, Q. Lv, A. Beach, R. Han, S. Mishra, and K. Seada, “Enhancing Group Recommendation by Incorporating Social Relationship Interactions,” Proc. International ACM SIGGROUP Conference on Supporting Group Work, pp.97-106, 2010.

[5] S. AmerYahia, S. Basu Roy, A. Chawla, G. Das, C. Yua, “Group Recommendation: Semantics and Efficiency,” Proc. International Conference on Very Large Data Bases, Vol.2, No.2, pp.754-765, 2009.

[6] X. Liu, Y. Tian, M. Ye, and W. Lee, “Exploring Personal Impact for Group Recommendation,” Proc. International conference on Information and Knowledge Management, pp.674-683, 2012.

[7] C. Wan, B. Kao, and D. W. Cheung,

“Location-sensitive Resources Recommendation in Social Tagging Systems,” Proc. international conference on Information and Knowledge Management, pp.1960-1964, 2012.

[8] 정연오, 이성우, 이지형, “개인화된 전문가 그룹을 활용한 추천 시스템”, 한국지능시스템학회, 제23권, 제1호, pp.7-11, 2013.

[9] 김재영, 이석원, “온톨로지 기반 영화 메타데이터 간 연관성을 활용한 영화 추천 기법”, 한국지능정보시스템학회, 제19권, 제3호, pp.25-44, 2013.

[10] 이한석, “SNS 디지털 환경의 특성이 문화콘텐츠 구매의도에 미치는 영향”, 한국콘텐츠학회논문지, 제12권, 제7호, pp.336-345, 2012.

[11] 안성만, 김인환, 최병구, 조윤호, 김은홍, 김명균, “소셜네트워크 분석을 통한 협업필터링 추천 성과의 이해”, 한국전자거래학회지, 제17권, 제2호, pp.129-147, 2012.

[12] M. Tang, Y. Jiang, J. Liu, and X. Liu, “Location Aware Collaborative Filtering for QoS-Based Service Recommendation,” Proc. International Conference on Web Services, pp.202-209, 2012.

[13] Y. Cui, S. Song, L. He, and G. Li, “A Collaborative Filtering Algorithm Based on User Activity Level,” Proc. International Conference on Business Intelligence and Financial Engineering, pp.80-83, 2012.

저자 소개

양희태(Heetae Yang)

준회원



- 2012년 2월 : 충북대학교 정보통신공학과(공학사)
- 2012년 3월 ~ 현재 : 충북대학교 정보통신공학과 석사과정

<관심분야> : 데이터베이스 시스템, 소셜 네트워크 서비스, 빅데이터, 무선 센서 네트워크 등

차 재 홍(Jaehong Cha)

준회원

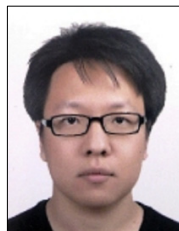


- 2012년 2월 : 충북대학교 정보통신공학과(공학사)
- 2012년 3월 ~ 현재 : 충북대학교 정보통신공학과 석사과정

<관심분야> : 데이터베이스 시스템, 소셜 네트워크 서비스, 빅데이터, 무선 센서 네트워크 등

안 민 제(Minjae Ahn)

준회원



- 2012년 2월 : 충북대학교 정보통신공학과(공학사)
- 2012년 3월 ~ 현재 : 충북대학교 정보통신공학과 석사과정

<관심분야> : 데이터베이스 시스템, 소셜 네트워크 서비스, 빅데이터, 무선 센서 네트워크 등

임 중 태(Jongtae Lim)

정회원



- 2009년 2월 : 충북대학교 정보통신공학과(공학사)
- 2011년 2월 : 충북대학교 정보통신공학과(공학석사)
- 2011년 3월 ~ 현재 : 충북대학교 정보통신공학과 박사과정

<관심분야> : 데이터베이스 시스템, 시공간 데이터베이스, 위치기반 서비스, 모바일 P2P 네트워크, 빅데이터

이 하(He Li)

정회원



- 2006년 7월 : 운남대학교 컴퓨터공학과(공학사)
- 2010년 8월 : 충북대학교 정보통신공학과(공학석사)
- 2010년 8월 ~ 현재 : 충북대학교 정보통신공학과 박사과정

<관심분야> : 데이터베이스 시스템, 모바일 소셜 네트워크 등

복 경 수(Kyungsoo Bok)

종신회원



- 1998년 2월 : 충북대학교 수학과(이학사)
- 2000년 2월 : 충북대학교 정보통신공학과(공학석사)
- 2005년 2월 : 충북대학교 정보통신공학과(공학박사)

• 2005년 3월 ~ 2008년 2월 : 한국과학기술원 정보전자연구소 Postdoc

• 2008년 3월 ~ 2011년 2월 : 가인정보기술 연구소 차장

• 2011년 3월 ~ 현재 : 충북대학교 정보통신공학부 조빙부교수

<관심분야> : 데이터베이스 시스템, 자료저장시스템, 이동객체 데이터베이스, RFID 및 센서네트워크, 모바일 P2P 네트워크, 빅데이터

유 재 수(Jaesoo Yoo)

종신회원



- 1989년 2월 : 전북대학교컴퓨터공학과(공학사)
- 1991년 2월 : 한국과학기술원 전산학과(공학석사)
- 1995년 2월 : 한국과학기술원 전산학과(공학박사)

• 1995년 3월 ~ 1996년 8월 : 목포대학교 전산통계학과 전임강사

• 1996년 8월 ~ 현재 : 충북대학교 정보통신공학부 및 컴퓨터정보통신연구소 교수

<관심분야> : 데이터베이스시스템, 정보검색, 센서네트워크 및 RFID, 멀티미디어 데이터베이스, 분산객체컴퓨팅, 빅데이터