

자동 추출된 사용자 그룹을 이용한 콘텐츠 및 사용자 히스토리 기반의

사용자 별 콘텐츠 추천 방법

신사임, 양창모, 장세진, 이석필
전자부품연구원 디지털미디어 연구센터
mirror, cmyang, sjjang, lspbio@keti.re.kr

Contents prediction method applying automatically extracted user groups based on users' consuming logs about contents

Saim Shin, Se-Jin Jang, Seok-Pil Lee
Korea Electronics Technology Institute

요 약

본 논문은 사용자의 각종 멀티미디어 콘텐츠 소비 히스토리를 수집하여 체계화 및 패턴 분석을 수행하고, 이를 바탕으로 사용자가 선호할 것으로 예측되는 멀티미디어 콘텐츠들을 추출하여 제공하는 콘텐츠 추천 시스템에 관한 연구이다. 본 논문에서는 콘텐츠 소비와 연관된 사용자 로그와 엔진에서 자동 추출한 사용자 그룹을 통하여 콘텐츠 추천을 수행한다. 각 사용자들의 선호정보 데이터를 분석하여 선호정보 패턴이 유사한 사용자들을 사용자 그룹으로 정의하고, 각 사용자들이 속한 사용자 그룹의 사용자 로그를 활용하여 상호 사용자 별 선호 콘텐츠를 예측한다. 본 시스템은 웹 또는 모바일 환경에서 음악, 방송, 광고, 기사 등의 방대하고 다양한 콘텐츠를 복합적으로 사용자들에게 선별하여 제공해주며, 이들의 연관성과 사용자의 콘텐츠 선호패턴을 반영한 개인 맞춤형 콘텐츠 추천 엔진은 사용자가 선호할만한 콘텐츠들을 추천하여 사용자의 콘텐츠 소비 시의 만족도를 높여줄 수 있다.

1. 서론

현재의 웹 서비스들 상에는 사용자들이 즐길 만 한 다양한 종류와 형식의 콘텐츠들이 산재되어 있다. 이에 따라 콘텐츠 추천 엔진도 사용자들의 요구가 다양해지고 복잡해지고 있으며, 이러한 현상에 발맞추어 더 복잡한 사용자 별 패턴들을 인식하여 정교한 알고리즘을 통한 콘텐츠 추천 엔진으로 진화해야 한다. 기존의 콘텐츠 추천 방식은 사용자 별 콘텐츠 소비 히스토리를 분석하여, 각 사용자들의 기록 분석 결과를 바탕으로 콘텐츠 기반의 추천을 수행한다. 이 방법은 활용하는 데이터의 성격이 단순하고 제한적이어서 데이터 희귀문제가 자주 발생한다. 많은 사용자 로그가 기록된 사용자들은 정확한 분석을 통하여 만족도 높은 콘텐츠 추천을 받을 수 있지만, 사용 빈도가 적은 사용자들의 경우 분석에 활용할 데이터가 부족하여 추천 결과의 양과 충실도가 떨어지게 되고, 결국 이같은 사용자들은 추천 결과의 만족도 저하로 해당 서비스의 이용 빈도가 더 줄어들게 되는 지속적인 악순환이 반복되는 것이다. 이러한 콘텐츠 추천 시스템의 고질적인 문제인 데이터 부족 현상을 극복하기 위한 방법으로 본 연구에서는 협업 필터링 방식 (Collaborative Filtering)의 콘텐츠 추천 알고리즘을 제안하였다. 사용자들의 콘텐츠 선호 패턴을 분석하여 추천엔진이 자동으로 구성해준 사용자 그룹을 기반으로 그룹간 사용자들의 사용자 로그의 이들 사이의 선호패턴 유사도를 가중치로 활용한 콘텐츠 추천 알고리즘을 제안하였다.

본 연구에서 제안한 콘텐츠 추천 엔진을 기반으로 구축된 유무선 개인별 콘텐츠 추천 및 공유 에이전트는 음악, 방송, 광고, 기사 등의 다양한 콘텐츠를 복합적으로 사용자들에게 제공해주어 콘텐츠의 소비 과정에서 많은 사용자들의 서비스 만족도를 높여준다.

본 논문에서는 2 장에서 제안하는 사용자 그룹 기반의 콘텐츠 추천 시스템에 대한 설명을 하고, 3 장에서 개발한 콘텐츠 추천을 위한 프로토타입 시스템을 소개하며, 4 장과 5 장에서는 관련연구와 결론 및 향후 연구를 논의한다.

2. 사용자 그룹 기반 콘텐츠 추천 시스템

추천 엔진의 결과를 기반으로 추천 서비스를 제공하도록 구축된 유무선 공유 에이전트는 음악, 방송, 광고, 기사 등의 다양한 콘텐츠를 복합적으로 사용자들에게 제공해주며, 이들의 연관성과 사용자의 콘텐츠 선호 패턴을 반영한 개인 맞춤형 추천엔진은 사용자가 선호할 만한 콘텐츠들을 추천하여 사용자의 만족도를 높여준다.

콘텐츠 추천 알고리즘은 사용자의 콘텐츠 소비 성향을 파악하여 사용자 별로 선호할 만한 콘텐츠들을 예측하는 알고리즘이다. 사용자의 콘텐츠 소비 성향을 파악하는 기존의 대표적인 방법으로는 각 사용자의 콘텐츠 관련 히스토리 (usage history)를 분석하여 이를 통한 사용자 선호 정보 (user preferences)를 파악하는 로그 기반의 콘텐츠 추천 방식이다. 이 방식에서 드러나는 데이터 희귀 문제 (data sparseness problems)는 통계적 처리에 적합하지 않은 양의 데이터를 통계적으로 처리하여 학습 및 예측을 함으로써 소수의 잘못된 선택이 전혀 다른 예측 결과를 만들어 낸다거나, 실제로 의미 있는 결과가 통계 기반 데이터에 의해 무시되어 예측에 반영되지 못하는 현상을 보인다. 이의 극복을 위해서는 양질의 다량의 데이터에 대한 통계적 분석을 통한 결과 도출이 어느 정도 도움은 되지만, 데이터량이 늘어난다고 해도 희귀한 데이터의 비율은 여전히 극소량이기 때문에 근본적인 해결책이라고는 할 수 없다. 그러므로, 새로운 통계적 방법의

제안이 필요하였고, 본 연구에서는 협업 필터링 방식의 협업 사용자 그룹 (Collaborative User Group) 기반의 사용자 그룹 기반의 콘텐츠 추천 방법을 제안함으로써 이러한 문제를 극복해 보고자 시도하였다.

사용자 그룹 (CUG) 기반 콘텐츠 추천 엔진의 전체적인 시스템 구성은 아래 그림과 같다. 콘텐츠 추천 엔진은 크게 CUG 업데이터, 사용자 선호정보 업데이터와 멀티미디어 콘텐츠 순위화 모듈로 구성된다. CUG 업데이터는 콘텐츠 추천 과정에서 유용하게 사용될 수 있는 대상 사용자와 유사한 패턴의 사용자 선호정보를 보이는 사용자들의 클러스터를 그룹화하여 추출한다. 사용자 선호정보 업데이터는 각각의 사용자 별 콘텐츠 타입으로 사용자 선호정보를 주기적으로 업데이트하여 관리하는 모듈이다. 사용자 콘텐츠 순위화 모듈은 각각의 사용자 별 각 콘텐츠 타입 별로 추천할 콘텐츠 목록을 도출하기 위하여 사용자 별 사용자 선호정보를 기반으로 각각의 콘텐츠들을 순위화하는 모듈이다.

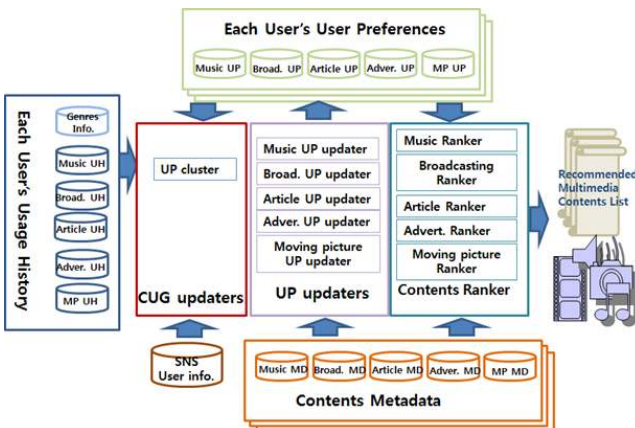


그림 1 CUG 기반 콘텐츠 추천 엔진 구조도

사용자 그룹은 특정 사용자의 콘텐츠 추천을 위하여 참조할 사용자들의 집합을 말한다. 이 모듈에서 추출된 사용자 그룹은 콘텐츠 추천 과정에서 집합적 추천 알고리즘의 기반 데이터로 사용된다. 본 시스템에서는 각 사용자들의 사용 히스토리 (usage history)를 분석하여 추출한 패턴들을 기반으로 해당 사용자의 선호정보를 유추하여 추출한다. 이 과정에서 본 연구는 사용자 선호정보의 추출에 사용하는 사용 히스토리 외에도 해당 사용자에서 기존의 사용자와 소셜 네트워크 상에서 관계를 맺고 있는 온라인/오프라인 상의 유관 사용자들의 선호정보를 함께 분석하여 이들의 사용 히스토리들도 선호정보의 예측에 사용하는 집합적 (Collaborative filtering) 콘텐츠 추천 방식을 제안한다. 또한, 모든 사용자들의 사용자 선호 정보 (User Preferences)들 사이의 유사도를 분석하여, 사용자 선호정보가 유사하여 그룹화 된 사용자들의 사용 히스토리 정보들 또한 콘텐츠 추천에 반영하여 추천 알고리즘을 개선하였다. 이와 같이 참조하는 사용자 그룹을 각각 소셜 사용자 클러스터와 사용자 선호정보 클러스터라고 정의하였다.

2.1. 사용자 그룹 추출

사용자 선호정보 그룹은 대상 사용자와 그 밖의 다른 사용자들의 사용자 선호정보 데이터의 유사도 비교를 통하여 추출하였다. 통계적으로 충분한 사용 히스토리 정보가 보장된 추천 환경이 아닌 경우라면, 사용자 선호정보 데이터는 각 사용자들의 콘텐츠 소비에 관해 사용 히스토리에 드러난 일부의 선호 패턴만을 명시적으로 보여주기 때문에 대상 사용자와 유사한 선호 정보 패턴을 보이는 다른 사용자들의 선호 패턴은 대상 사용자의 내재된 선호 성향을 발견하는 것을 도와줄 수 있다. 그 당시 대중의 콘텐츠 소비 경향 또한 각 사용자들의 콘텐츠 소비 패턴에 직 / 간접적으로 영향을 줄 수 있는데, 이 과정에서 대상 사용자와 비슷한 소비 성향의 다른 사용자들의 사용 히스토리나 선호정보를 참조하여 통계적으로 강건한 콘텐츠 추천을 통하여 추천한다면, 개인화된 정보 외에도 대중적으로 선호되는 콘텐츠들의 추천도 가능하여 사용자의 추천 결과 만족도를 높일 수 있을 것이다. 그러므로, 본 연구에서는 협업 필터링을 위한 유사도 기반 사용자 그룹들을 추천하여 활용하였다.

대상 사용자 u_i 을 위한 이전 선호 정보 벡터 Z_{u_i} 는 식 (1)과 같이 정의될 수 있다.

$$Z_{u_i} = \{g_{a1}^{u_i}, \dots, g_{aN_a}^{u_i}, g_{v1}^{u_i}, \dots, g_{vN_v}^{u_i}, g_{b1}^{u_i}, \dots, g_{bN_b}^{u_i}, g_{m1}^{u_i}, \dots, g_{mN_m}^{u_i}, g_{n1}^{u_i}, \dots, g_{nN_n}^{u_i}\} \quad (1)$$

$$g_{ij} \in \{0,1\}, \forall i \in \{a, v, b, m, n\}, \forall j \in \{N_a, N_v, N_b, N_m, N_n\}$$

식 (1)에서 g_{ij} 은 각 형태 별 콘텐츠 - 광고 콘텐츠 (a), 비디오 콘텐츠 (v), 방송 콘텐츠 (b), 음악 콘텐츠 (m), 신문 기사 콘텐츠 (n) - 에 포함되는 콘텐츠들의 모든 장르를 의미한다. N_k 는 k 를 위한 장르의 개수이다. 사용자 그룹 클러스터는 대상 사용자 벡터 Z_{u_i} 와 대상 사용자 u 를 제외한 다른 사용자들 사이의 dice coefficient 유사도를 계산하여 그룹화를 진행하며, 그 유사도 계산식은 식 (2)와 같다.

$$D_{u_i, u_j} = \frac{2 \sum_k g_k^{u_i} g_k^{u_j}}{\sum_k (g_k^{u_i})^2 + \sum_k (g_k^{u_j})^2}$$

$$N = N_a + N_v + N_b + N_m + N_n$$

$$D_{avg}^{u_i} = \frac{\sum_j D_{u_i, u_j}}{N_u - 1} \quad \text{----- (2)}$$

식 (3)은 u_i 를 위한 사용자 클러스터를 보여주고 있다. D 는 dice coefficient 계산값을 의미하고, N_u 는 본 시스템에서 대상으로 하는 모든 콘텐츠 추천 대상 사용자들의 수이다. 유사도 비교를 통해 추출한 최종 사용자 선호정보 클러스터는 PC 로 정의되었다.

$$PC_{u_i} = \{u_x \mid u_x \text{ is } D_{u_i, u_x} > D_{avg}^{u_i}\} \text{----- (3)}$$

사용자 선호정보 클러스터 (PC_{u_i})는 u_i 와의 dice coefficient 값이 전체 사용자와 u_i 와의 평균 dice coefficient 값보다 큰 사용자들의 집합으로 정의 된다. PC_{u_i} 에 속한 사용자들은 대상 사용자와 선호 콘텐츠에 관해 유사한 패턴을 보여주기 때문에, 그들의 사용 히스토리 데이터는 대상 사용자의 최근 선호정보 예측에 활용 가능하다. 대상 사용자의 PC 기반 사용자

선호정보의 예측은 더욱 견고하고 대중의 경향을 반영한 사용자 선호정보의 예측을 도와줄 것이다.

2.2. 사용자 선호 정보 자동 추출

사용자 선호정보 업데이트 모듈은 주기적으로 각각의 사용자들에 대한 각 콘텐츠 타입 별 사용자 선호정보를 새롭게 업데이트된 사용자 히스토리를 기반으로 예측한다. 이 결과를 반영하여 각각의 콘텐츠 타입 별 콘텐츠 순위화 모듈은 사용자들의 추천 콘텐츠 목록을 추출한다. 콘텐츠 추천 엔진의 선호 정보 업데이트와 추천 목록 순위화를 설명하기 위하여 식 (4)의 정의가 필요하다.

$$\begin{aligned}
 UH_{u_i} &= \{uh_1, uh_2, \dots, uh_n\} \\
 P_{u_i} &= \{p_x \mid \forall p_x = (uh_k, uh_{k+1}, \dots, uh_{k+k}), k' < 5\} \\
 F_{u_i} &= \{f_{p_x}^{u_i} \mid \text{the frequency of } p_x, p_x \in P_{u_i}\} \quad (4)
 \end{aligned}$$

UH_{u_i} 는 u_i 를 위한 사용 히스토리 (Usage History) 집합이다. P_{u_i} 는 UH_{u_i} 에 나타나는 연속적인 사용 히스토리 행동 패턴 시퀀스를, F_{u_i} 는 P_{u_i} 각각의 사용 히스토리 패턴의 대상 사용자의 UH_{u_i} 에서 나타나는 빈도수를 말한다.

시청 중의 사용자 행동양식을 관찰하기 위하여 패턴 풀 (pattern pool)을 제안한다. 패턴 풀에는 사용자 행동기록에서 추출 가능한 모든 연속적인 사용자 행동패턴들을 수집한다.

패턴 풀에 수집된 행동패턴들은 특성추출 과정에서 평가를 통하여 선호정보 추출에 사용할 선호패턴 및 비선호 패턴 목록을 결정하게 된다. 특성추출과정은 패턴 풀에 수집한 패턴들에 대하여 사용자 행동기록에서 그 빈도수를 기반으로 상호 정보값 (Mutual Information)을 계산한다. 특정 패턴 p_x 를 위한 상호 정보값의 계산식은 식 (5)와 같다.

$$MI_{p_x^{u_i}} = \log\left(\frac{p(f_{p_x}^{u_i} \mid pos)}{p(f_{p_x}^{u_i})}\right) \quad \text{----- (4)}$$

식 (5) 서 $p(f_{p_x}^{u_i} \mid pos)$ 는 u_i 가 선호하는 콘텐츠들 시정하는 동안 p_x 의 등장 확률이다. $f_{p_x}^{u_i}$ 는 p_x 의 빈도수로 정의되고 $p(f_{p_x}^{u_i})$ 는 전체 UH_{u_i} 에서 p_x 의 등장 확률을 의미한다. MI 는 해당 패턴과 대상 사용자가 선호하는 콘텐츠와의 등장 확률의 독립성 정도를 통계적으로 설명해 준다. 만약 해당 패턴의 MI 값이 0 보다 크다면, 대상 사용자에게 해당 패턴의 등장은 사용자가 선호할 만한 콘텐츠일 가능성이 높다. 이러한 경우의 패턴들을 선호 패턴 (positive pattern)으로 규정한다. MI 가 0 인 패턴은 사용자의 선호도와 완전히 독립적인 무관한 패턴이다. 또한, MI 가 0 보다 작은 경우는 사용자 선호도와 패턴의 빈도가 반비례하는 부정적인 패턴이라고 할 수 있다.

본 연구에서는 특성추출 과정에서 사용자의 선호도를 점수화하기 위한 선호패턴과 비선호도를 반영하기 위한 비선호 패턴을 식 (1)의 MI 값을 기준으로 관리한다.

패턴 풀의 전체 패턴의 평균 상호 정보값인 MI_{avg} 값을기준으로 MI_{avg} 와 0 보다 큰 MI 를 가지는 패턴을 선호패턴으로, 0 보다 작은 MI 의 패턴을 비선호 패턴으로 결정하였다.

이러한 패턴수집과 특성추출 과정은 사용자가 방송을 시청하지 않는 시간에 주기적으로 새로운 패턴을 수집하여 최적의 선호 / 비선호 패턴으로 관리한다. 자동적인 사용자 별 행동패턴 관찰과정은 사용자의 최신의 선호정보를 관리하는데 유리하다.

또한 본 연구의 특성 추출 방법은 사용자 별로 유의미한 패턴의 선정에 대한 수작업 없이도 사용자 선호 정보와 연관된 특성 패턴들을 자동으로 수집하고 관리할 수 있는 방법이다. 또한, 사용자의 빈번한 사용자 선호정보 수작업 업데이트 없이도 정확한 사용자의 선호 패턴을 인식하여 콘텐츠 추천에 반영할 수 있다.

사용자 선호정보에 포함될 선호정보 관련 아이템들은 앞서 결정된 선호 / 비선호 패턴의 빈도수들을 기반으로 점수화한다. 점수화 방법은 아래의 수식 (6)과 같이 특성 추출기로부터 추출한 선호 / 비선호 패턴들을 사용하여 진행한다. 수식 (6)은 데이터 희귀 문제를 극복하고 대상 사용자가 선호할 것 같은 더 많고 다양한 콘텐츠들을 추천하기 위하여 협업 필터링 방식으로 사용자 클러스터에 속한 사용자들의 사용 히스토리를 함께 적용하여 추천을 진행한다.

$$\begin{aligned}
 UHM_Score_{g_k}^{u_i} &= \frac{\sum_{u_j} f_{p_x}^{u_j} - \sum_{u_j} f_{p_x}^{u_i}}{N_p} \\
 N_p &= N_{pos} + N_{neg} \\
 N_c &= N_{pc} + N_{sc} \quad (6) \\
 UPM_Score_{g_k}^{u_i} &= \sum_{u_j} D_{u_i, u_j} \cdot UHM_Score_{g_k}^{u_j}
 \end{aligned}$$

사용자 u_i 에 대한 선호 정보 아이템 g_k 에 대한 UHM_Score 값은 식 (6)에서 선호 패턴의 합과 비선호 패턴의 합의 차이를 통하여 계산된다. N_{pos} 와 N_{neg} 는 전체 선호 패턴과 비선호 패턴의 개수이고, D_{u_i, u_j} 는 UPM_Score 에서 UHM_Score 의 중요성에 대한 가중치를 표현한다. D_{u_i, u_j} 가 u_i 와 대상 사용자와의 선호 정보 유사도를 가지고 가중치로 적용하였다. 제안하는 시스템은 대상 사용자와 대상 사용자의 유사 사용자 그룹에 속한 사용자들의 선호 정보와 이들의 유사도 값을 적용하여 다양한 형태의 콘텐츠 추천에 활용하였다. 선호정보 예측모듈에서 대상 사용자 u_i 에 대하여 가장 큰 UPM_Score 값을 보이는 선호 정보 아이템들을 선택하여 대상 사용자의 선호 정보를 업데이트하게 된다.

3. 프로토타입 시스템

그림 2는 앞서 설명한 사용자 그룹 기반의 추천 엔진의 결과를 활용하여 구축한 멀티미디어 콘텐츠 추천 서비스 프로토타입을 보여주고 있다. 이 시스템은 음악, 뉴스기사, 방송, 광고, 유튜브 동영상 콘텐츠의 다섯가지 종류의 웹 상의 다양한 콘텐츠에 대한 추천을 제공하고 있다.

본 시스템에서는 콘텐츠 타입 별 추천 목록, 콘텐츠 통합 추

천 목록 등 사용자의 편의와 필요에 따라 다양한 형태로 추천 서비스를 이용할 수 있도록 사용자 인터페이스를 제공하고 있다. 또한, 각자의 사용자 선호정보와 사용자 그룹 정보를 확인하고 선호정보의 경우 사용자가 수정할 수 있도록 하여, 사용자의 니즈를 직접 반영하여 콘텐츠 추천 시스템의 만족도를 향상시킬 수 있도록 구성하였다. 또한, 본 시스템 전 부분에서 사용자의 콘텐츠 및 다양한 기능에 대한 사용 히스토리는 이 시스템 상에서 사용자 별로 기록되어, 지속적으로 다음 콘텐츠 추천의 기반 데이터로 활용되도록 구성되었다.



그림 2 사용자 그룹 기반 콘텐츠 추천 서비스 프로토타입 시스템

4. 관련 연구

추천 시스템을 구분하는 가장 일반적인 방법은 추천 방식에 따른 분류이며, 추정하는 방법에 따라 콘텐츠 기반 추천과 협업 필터링 기반 추천, 그리고 두 가지 방법을 결합한 하이브리드 추천 등으로 구분된다.

기존의 많은 콘텐츠 기반 추천 방법은 텍스트 기반 아이템에 대한 추천을 위해 연구되었고, 대상 아이템은 일반적으로 키워드를 기반으로 표시되는 경우가 많았다. 따라서 기존 콘텐츠 기반 추천 시스템에 텍스트는 가장 중요한 특성 정보 가운데 하나로 인식되었다. 텍스트 기반 콘텐츠 추천 시스템에서는 특성 정보의 가중치 결정을 위해 특정한 단어가 문서 내에 얼마나 자주 등장하는지 나타내는 값인 단어 빈도수 (TF, Term Frequency)와 특정 단어가 나타난 문서 빈도수의 역수인 IDF (Inverse Document Frequency)를 사용한 TF-IDF 알고리즘이 대표적으로 적용되었다. 웹 페이지의 콘텐츠에 포함되어 있는 중요한 100 개의 키워드를 이용하여 사용자에게 웹 페이지를 추천하는 Fab 시스템과 128 개의 의미 있는 단어의 집합을 이용하는 콘텐츠 기반 추천 시스템인 Syskill & Webert 시스템이 기존 대표적인 콘텐츠 기반 추천 시스템 연구이다 [1].

협업 필터링 (CF, Collaborative Filtering)이란, 정보 필터링 기술을 이용한 방법으로 기존에 문서 추천에 활용된 콘텐츠 기반 필터링 (Content-based filtering)과 대비되는 용어로, 문서나 뉴스를 선택할 때 다른 사용자들의 협력적인 도움을 받는 것으로 정의된다. 즉, 협업 필터링 기반 추천 시스템은 다른 사용자가 이전에 평가한 특정 아이템의 정보를 기반으로 목표 사용자에게 그 아이템의 유용성을 예측하는 시스템이다.

이러한 협업 필터링 기반 추천 시스템은 이후, 그룹렌즈 (GroupLens)에 의하여 최근접 이웃 방법 (K-nearest neighbor)을 기반으로 유사 사용자에게 대한 정보를 이용하여 추천을 수행하는 방법으로 발전해 갔으며, 콘텐츠 기반 추천, 인구 통계학적 추천 등 다른 추천 기법에 비하여 추천의 정확도가 높아 영화, 문서, 음악 등 다양한 도메인에 널리 활용되고 있다 [2, 3]

5. 결론 및 향후 연구

본 논문은 자동 추출한 사용자 그룹 정보를 활용한 콘텐츠 추천 알고리즘을 소개하였다. 사용자 그룹은 사용자들의 사용자 선호정보를 분석하여 활용함으로써 콘텐츠 점수와 과정에서 활용하는 사용자 히스토리의 범위를 확장하였다.

이러한 협업기반 콘텐츠 추천 알고리즘은 사용자 별 데이터 부족 문제를 보완해주고, 취향이 비슷한 사용자들의 사용 패턴을 기반으로 사용자의 선호도에도 부합하면서도 광범위한 콘텐츠의 추천을 가능하게 해준다. 또한, 사용자들의 트렌드와 경향이 실시간으로 반영되고 나타나는 최근의 웹 기반 및 소셜 네트워크 환경에서는 다양한 정보 안에서 대중과 연관된 사용자들의 콘텐츠 소비 패턴을 찾아내어 공유하는 행위가 매우 중요하게 부각되고 있으므로, 사용자 그룹을 통한 콘텐츠 추천 방식의 적용은 이러한 사용자들의 콘텐츠 소비 방식의 변화를 반영할 수 있는 매우 적합한 추천 방식이다.

향후 연구로는 각 콘텐츠 고유의 특성을 반영하여 콘텐츠 타입별로 세분화된 특징들을 활용하여 콘텐츠 종류별로 정교한 추천 방식으로 알고리즘을 세분화할 계획이다. 또한, 사용자 그룹 추천 과정에서 적용하는 사용자 선호정보도 좀 더 세분화한 그룹핑 방법과 그룹 추출에 대한 연구를 진행할 것이다.

Acknowledgement

본 연구는 지식경제부 산업융합원천기술개발사업의 지원을 받아 수행된 연구결과임. (No. 10031824)

참고 문헌

- [1] M. Balabanovic and Y. Shoham, "FAB: content-based, collaborative recommendation," *Proceedings of Communications of the ACM*, Vol. 40, No. 3, pp. 66-72, 1997.
- [2] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl, "GroupLens: an open architecture for collaborative filtering of netnews," *Proceedings of ACM Conference on Computer supported Collaborative Work*, pp. 175-186, 1994.
- [3] X. Su and T. Khoshgoftaar, "A survey of collaborative filtering techniques," *Advances in Artificial Intelligence*, Vol. 2009, pp. 1-20, 2009.