

하이브리드 협업필터링을 통한 개인화 여행지 추천 기법

Personalized Travel Destination Recommendation Scheme through Hybrid Collaborative Filtering

신중훈, 송지현, 북경수, 유재수+
충북대학교 정보통신공학부

Jonghoon Shin, Jihyeon Song, Kyoungsoo Bok,
Jaesoo Yoo

Department of Information & Communication
Engineering, ChungBuk National University

요약

최근 주말의 개념이 확장되고 일상보다 여유를 우위를 두는 사람들이 많아짐에 따라 여행 산업이 발전하고 있다. 본 논문에서는 사용자 성향을 기반으로 하이브리드 협업 필터링을 이용한 여행지 추천 기법을 제안한다. 사용자별 여행지 선호도를 생성하고 사용자 기반 협업 필터링을 통해 후보 여행지를 생성하고 아이템 기반 협업 필터링을 수행하여 여행지 성향 점수를 생성한다. 여행지 성향 점수와 여행지별 성향을 고려하여 최종 여행지를 추천한다.

I. 서론

한국문화관광연구원에 따르면 2010년부터 2016년까지 관광 횟수가 약 27%가 증가하였다[1]. 이와 함께, 사용자에게 여행지 정보를 제공하기 위한 다양한 서비스들이 개발되고 있다. 대표적으로 트립 어드바이저, 익스피디아, 대한민국 구석구석 등이 있다. 그러나 대부분의 여행지 정보 서비스들은 대중적인 사용자의 관심사 또는 평판을 기반으로 여행지 정보를 제공한다. 즉, 개인의 성향에 적합한 개인화된 여행지 정보를 제공하지 못하고 있다.

사용자가 경험하지 못한 장소를 제공하기 위해 협업 필터링(CF : Collaborative Filtering) 기법이 활용되고 있다 [2]. 협업 필터링은 사용자 기반 협업 필터링과 아이템 기반 협업 필터링으로 구분된다. 사용자 기반 협업 필터링은 유사한 사용자들만을 고려하고 상품간의 유사도를 고려하지 않기 때문에 예측의 정확성이 떨어질 수 있다. 아이템 기반 협업 필터링은 사용자와 선호가 전혀 다른 사용자도 고려될 수 있기 때문에 정확성이 떨어진다. 이러한 문제점을 보완하기 위하여 본 논문에서는 두 가지의 기법을 결합한 하이브리드 필터링(HCF : Hybrid CF)

기반의 여행지 추천 기법을 제안한다. 사용자별 장소 방문 횟수와 평점을 이용하여 사용자를 선호도를 생성하고 이를 이용하여 사용자 기반 협업 필터링을 통해 선호도 기반 여행지를 선출한다. 선출된 선호도 기반 여행지에 대해 아이템 협업 필터링을 수행하여 여행지 성향을 생성한다. 또한, 여행지별 성향을 고려하여 최종 여행지를 추천한다.

II. 제안하는 개인화 여행지 추천 기법

1. 시스템 구조

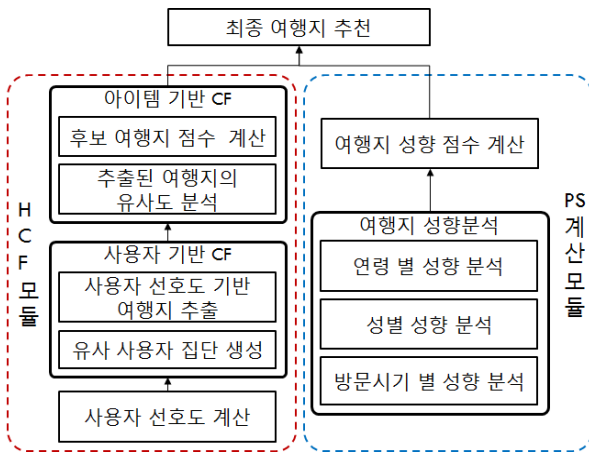
본 논문에서 기존 협업 필터링을 통한 여행지 추천 기법의 문제점을 해결하기 위해 사용자 기반 협업 필터링과 아이템 기반 협업 필터링을 결합한 하이브리드 협업 필터링 기반 여행지 추천 기법을 제안한다. 성별, 연령, 여행 기간을 이용하여 여행지 별 성향을 분석하고 하이브리드 협업 필터링의 결과와 결합하여 최종 여행지를 추천한다.

그림 1은 제안하는 여행지 추천 기법의 시스템 구조를 나타낸다. 개인 맞춤 여행지 추천을 위해 우선 여행지 방문 횟수와 평점을 이용하여 사용자의 선호도를 계산한다. HCF 모듈에서는 계산된 사용자의 선호도를 기반으로 사용자 기반 협업 필터링을 통해 선호도 기반 여행지를 생성한다. 다음으로 아이템 기반 협업 필터링을 통해 생성된 여행지의 유사도를 분석하여 후보 여행지 점수 TS 를 계산한다. PS 계산 모듈에서는 후보 여행지를 방문한 사용자들의 연령, 성별, 여행 기간 정보를 이용하여 여행지 성향 점수 PS 를 계산한다. HCF 모듈을 통해 생성된 후보 여행지와 PS 계산 모듈을 통해 분석된 여행지

+ 교신 저자 : yjs@cbnu.ac.kr

본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 대학ICT연구센터육성 지원사업(ITP-2018-2013-1-00881)과 2016년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원(No. 2016R1A2B3007527)과 2017년도 정부(과학기술정보통신부)의 재원으로 한국연구재단-차세대정보·컴퓨팅기술개발사업의 지원을 받아 수행된 연구임(No. NRF-2017M3C4A7069432)

성향을 고려하여 최종 여행지를 추천한다.



▶▶ 그림 1. 여행지 추천 기법 구조도

2. HCF 모듈

HCF 모듈은 사용자 기반 협업 필터링과 아이템 기반 협업 필터링을 통해 사용자가 방문하지 않은 후보 여행지 점수를 계산한다. 사용자 기반 협업 필터링에서는 유사 사용자 집단의 선호도를 기반으로 한 여행지를 추출한다. 아이템 기반 협업 필터링에서는 후보 여행지의 예측 선호 점수 TS 를 계산한다.

우선 수식 1을 사용하여 협업필터링에 쓰이는 사용자 선호도 $SigS_{u,i}$ 를 계산한다.

$$SigS_{u,i} = \alpha * Score_{u,i} + \beta * \frac{Visit_{u,i}}{Total\ Visit_u} \quad (수식 1)$$

$Score_{u,i}$: 사용자 u 가 여행지 i 에 마지막으로 남긴 평점

$Visit_{u,i}$: 사용자 u 가 여행지 i 에 방문한 횟수

$Total\ Visit_u$: 사용자 u 의 총 방문 횟수

아이템 기반 협업 필터링은 여행지간 유사도만을 고려하기 때문에 정확성이 떨어지는 문제점이 있다. 이를 보완하기 위해 먼저 사용자 기반 협업필터링을 적용한다. 보다 정확한 사용자 유사도를 계산하기 위하여 여행지의 계층적 구조를 이용한다[3]. 계산된 사용자 유사도를 기반으로 여행자와 유사한 사용자 집단을 생성하고, 이들이 평가를 남긴 여행지들로 선호도 기반 여행지를 생성한다.

선호도 기반 여행지 중 여행자가 방문하지 않은 여행지의 점수를 예측하기 위해 아이템 기반 협업 필터링을 수행한다. 선호도 기반 여행지에 남겨진 평가를 기반으로 피어슨 상관계수를 이용해 여행지간 유사도를 계산한다. 계산된 유사도를 이용해 여행자 u 가 방문하지 않은 여행지 i 에 대한 예측 선호도 $TS_{u,i}$ 를 계산한다. 수식 2는 이를 나타낸 식이다.

$$TS_{u,i} = \frac{\sum_{all\ evaluated, N} (s_{i,N} * US_{u,N})}{\sum_{all\ evaluated, N} (|s_{i,N}|)} \quad (수식 2)$$

N : 사용자가 평가를 남긴 여행지

$s_{i,N}$: 여행지 i 와 N 의 유사도

$US_{u,N}$: 사용자 u 가 여행지 N 에 남긴 평가

3. PS 계산 모듈

각 여행지는 여행지 특성에 따라 선호도가 다르다. 예를 들어, 이태원 쇼핑거리는 60대 사용자보다 20대 사용자가 많이 방문하고, 눈썰매장은 봄보다 겨울에 많이 방문하는 등 선호도가 상이할 수 있다. 따라서 여행지 추천을 위해서는 여행지 성향을 고려할 필요가 있다. PS 계산 모듈은 여행자의 성별, 연령, 방문 시기에 맞는 여행지의 성향을 분석한다.

여행지 성향분석 단계에서는 여행자의 연령대를 a , 성별을 g , 여행할 시기를 m 이라고 할 때, 관광지 i 에 대해서 a 연령대가 평가한 점수 $AgeS_i$, 성별 g 가 평가한 점수 $GenS_i$, m 월의 평가 점수 $MonS_i$ 를 계산한다. 이 때, 오래된 평점에는 시간 가중치를 적용하여 최신 평점이 더 높게 반영되도록 한다. 분석된 여행지 성향을 이용하여 최종적으로 관광지 i 에 대한 성향점수 PS_i 를 계산한다. 이 때 각 평가점수 별 적절한 반영비율을 두어 점수를 계산한다. 수식 3은 이를 나타낸 것이다.

$$PS_i = \alpha AgeS_i + \beta GenS_i + \gamma MonS_i \quad (수식 3)$$

where $\alpha + \beta + \gamma = 1$

마지막으로 두 모듈의 결과 값 $TS_{u,i}$, PS_i 를 고려하여 최종 여행지 순위를 매기고 높은 순으로 사용자에게 여행지를 추천해준다.

III. 결론

본 논문에서는 하이브리드 협업필터링을 통한 개인 맞춤형 여행지 추천 기법을 제안하였다. 제안하는 기법은 여행자의 성향을 정확하게 반영하기 위하여 하이브리드 협업필터링 기법을 사용했다. 또 여행자의 연령, 성별, 여행 시기에 따른 관광지의 성향을 고려하였다. 최신 정보를 반영하기 위하여 시간가중치를 부여하였다. 향후 연구에서는 제안하는 기법을 실제 구현하여 성능을 평가할 예정이다.

■ 참고 문헌 ■

[1] e-나라지표, http://www.index.go.kr/potal/main/EachDtlPageDetail.do?idx_cd=1652
 [2] 손창환, 김기수, "추천 시스템의 예측 정확도 향상을 위한 하이브리드 협업 필터링", 한국산업경영학회 2006년도 동계 학술대회, pp.561-587, 2006.
 [3] Min-Joong Lee and Chin-Wan Chung, "A User Similarity Calculation Based on the Location for Social Network Services", Proc. International Conference on Database Systems for Advanced Applications, pp.38-52, 2011