

How to improve the diversity on collaborative filtering using tags

Jin-Hyeon Joo*, Geun-Duk Park**

Abstract

In this paper, we propose how to improve the lack of diversity in collaborative filtering, using tag scores contained in items rather than ratings of items. Collaborative filtering has excellent performance among recommendation system, but it is evaluated as lacking diversity. In order to solve this problem, this paper proposes a method for supplementing diversity lacking in collaborative filtering by using tags. By using tags that can be used universally without using the characteristics of specific articles in a recommendation system, The proposed method can be used.

▶ Keyword: Recommendation System, Collaborative Filtering, Tag, Diversity

I. Introduction

정보통신의 발달과 함께 많은 온라인 쇼핑몰이 생겨나면서 소비자들의 소비 패턴이 오프라인 구매에서 온/오프라인 구매로 변화하였다. 이에 소비자들이 접하게 되는 물품들의 수가 급격하게 증가하였고, 다양한 물품들 중에 소비자들은 자신이 원하는 물품을 찾기 힘들게 되면서 개인화 추천 시스템의 중요성이 대두되었다.

협업 필터링(Collaborative Filtering)은 추천 시스템들 중에서 가장 우수한 성능을 지닌 추천 시스템으로써, Goldbug et al.(1992)에 의해 처음 소개되었다. 기존의 콘텐츠 기반 접근 기법의 경우 사용자와 물품 정보만을 사용하여 선호도를 예측하지만, 협업 필터링은 사용자가 아이템에 평가한 정보를 바탕으로 선호도를 예측하기에 콘텐츠 기반 접근 기법보다 높은 정확성을 보장한다[1].

최근 연구에 의하면 추천시스템이 등장한 초반에 비해 연구 영역이나 대상 등이 다양해지고 세분화되었으며, 개인화 추천, 머신러닝, 데이터 마이닝, 빅데이터[2]을 활용하는 등 추천시스템은 다양하게 활용이 되고 있었다[3].

기존의 추천 시스템들은 추천 정확도의 향상을 최우선 과제로 삼아 추천 정확도를 향상 시키는 방향으로 많은 연구가 이루어져 왔다. 하지만 협업 필터링의 과정을 시물레이션한 결과으로써 소비자의 다양성을 해결 수 있다고 Fleder et al.(2009)가 언급한 바가 있으며, 최근 연구에서 실험한 결과

실제로 다양성이 줄어드는 것이 확인되었다[4].

본 논문에서는 앞서 언급된 다양성이 저해되는 문제를 해결하기 위해 물품에 등록되는 태그(Tag)를 사용하여 기존에 협업 필터링보다 높은 다양성을 가지는 태그 기반 협업 필터링(Collaborative Filtering based Tags)를 제안한다.

본 논문의 구성은 다음과 같다. 2장에서는 관련 연구로 기존의 협업 필터링의 단점과 추천 다양성이 왜 중요하고, 이에 대한 연구들을 소개한다. 3장에서는 태그 기반 협업 필터링에서 태그 별 점수(Score)를 구하는 방법과 시스템의 구성 그리고 태그 별 점수를 이용하여 유사 이웃 탐색하는 방법에 대하여 설명한다. 4장에서는 3장에서 설명한 내용을 바탕으로 구현된 태그 기반 협업 필터링을 사용하여 실제 실험한 내용을 정리하였으며, 5장에서는 결론 및 향후 연구에 대하여 언급한다.

II. Preliminaries

1. Related works

1.1 Content based Filtering

내용 기반 필터링의 경우 물품을 분석하여 물품과 물품 혹은

• First Author: Jin-Hyeon Joo, Corresponding Author: Geun-Duk Park
*Jin-Hyeon Joo (joojin4381@gmail.com), Dept. of Computer Science, Hoseo University
**Geun-Duk Park (gdpark@hoseo.edu), Dept. of Computer Science, Hoseo University
• Received: 2018. 05. 28, Revised: 2018. 06. 20, Accepted: 2018. 07. 11.

은 물품과 소비자 선호도 간의 유사성을 분석하여 소비자에게 물품을 추천하는 방식이다.

내용 기반 필터링은 다음과 같은 특징들을 가지고 있다.

1. **독립적 정보 활용:** 내용 기반 필터링은 소비자의 선호도를 유추하기 위해서 소비자의 과거 구매 이력이나 프로파일 등의 정보가 필요하다. 따라서 과거 구매 이력이 부족하거나 소비자의 정보가 입력되어 있지 않은 경우 추천 성능이 하향되거나 불가능해진다.

2. **새로운 물품:** 내용 기반 필터링에 새로운 물품이 추가될 경우 협업 필터링과는 달리 소비자의 정보와 새로운 물품 간의 연관성만을 비교 분석하기 때문에 다른 사용자들의 물품 구매와는 상관없이 추천이 가능하다.

3. **과도한 특수화:** 내용 기반 필터링은 앞서 언급한 바와 같이 소비자의 과거 구매 이력과 물품 간의 연관성만을 비교 분석하기 때문에 비슷한 물품들만 추천될 확률이 높다.

내용 기반 필터링의 경우 소비자가 과거에 구매한 이력과 소비자 자신이 입력한 정보에 의해 추천 성능이 영향 받는다. 때문에 내용 기반 필터링은 소비자의 선호도를 더 정확하게 판단하기 위해 최근 혹은 자주 구입한 물품들에 대하여 기계 학습을 사용하여 추천 성능을 향상시키려는 다양한 연구가 제안되었다.

1.2 Collaborative Filtering

협업 필터링은 물품이나 사용자의 정보만을 사용하여 선호도를 예측하는 내용 기반 필터링과는 달리 사용자들이 물품에 내린 평가를 기반으로 선호도를 예측하는 추천 시스템으로 가장 높은 성능을 가진다고 평가된다.

협업 필터링은 우수한 추천 정확성을 가지는 반면에 아래의 문제들이 존재한다.

1. **데이터 희소성:** 앞서 언급한 바와 같이 협업 필터링은 사용자들이 물품에 내린 평가를 기반으로 선호도를 예측한다. 하지만 시스템 초기에는 물품에 대하여 내려진 평가가 없기 때문에 추천을 진행할 수 없는 문제가 발생한다. 이를 Cold-start 문제라 부른다[5].

2. **확장성:** 온라인 쇼핑물을 사용하는 사용자 수 혹은 등록된 물품의 수가 증가할수록 추천 시스템에서 처리해야 하는 데이터 량이 커지기 때문에 유사 사용자 그룹을 찾는 데 소모되는 시간이 길어진다.

3. **Grey sheep:** 협업 필터링의 경우 사용자들이 물품에 내린 평가를 기반으로 유사 사용자 그룹을 찾아 추천할 물품의 목록을 생성하지만, 간혹 일관성 없이 물품을 구매하거나 물품에 평가를 내리는 사용자가 존재한다. 이러한 사용자들로 인해 올바른 추천 물품 목록을 생성하는데 문제된다.

4. **Shilling attack:** 앞서 Grey sheep 문제와 비슷하게 Shilling attack 문제도 추천에 필요한 데이터를 오염시키는 문제이다. 해당 문제는 물품을 판매하는 판매자가 자신의 물품에 대한 평가를 긍정적으로 평가하거나 경쟁 업체의 물품에 부정적인 평가를 줌으로써 올바른 추천을 못하도록 하는 문제이다.

일반적인 협업 필터링을 비롯한 추천 시스템들은 추천 정확성을 중요시하기 때문에 추천 정확성을 높이는 방향으로 많은 연구가 이루어져왔지만 최근 추천 정확성만큼이나 추천 다양성의 중요성이 대두되었다[10].

협업 필터링의 경우 사용자의 물품에 대한 평가 데이터를 기준으로 유사 사용자 그룹을 탐색하여 추천 물품 목록을 생성하기 때문에 시스템 내 사용자들이 자주 구매하는 물품들은 추천 물품 목록에 자주 등장하는 반면에 사용자들에게 생소한 물품들은 추천 물품 목록에 등장하지 않는 문제가 발생한다.

추천 다양성을 향상시키는 연구로는 사용자에게 추천되는 물품 목록의 다양성을 향상시키는 방향과 시스템이 추천하는 물품 전체의 다양성을 높이는 방향이 있으며, 사용자에게 추천되는 물품 목록의 다양성을 향상시키는 방향으로의 연구가 더 활발히 이루어지고 있으며, 본 논문에서도 사용자에게 추천되는 물품 목록의 다양성을 향상시키는 방향으로 연구를 진행하였다.

1.3 Diversity and Novelty

앞서 언급한바와 같이 추천 시스템의 성능 향상 연구는 정확성을 향상시키는 방향으로 진행되었으나, 최근에는 롱테일 아이템(Longtail Item), 추천 다양성과 참신성에 관심이 모아지고 있다.

이전부터 추천 시스템이 사용자들이 선호하는 물품들을 추천함으로써 시스템 전체 물품이 다양하게 소비되는지 혹은 특정 물품에 대해서만 소비가 활발히 이루어지는지 많은 논의가 이루어져 왔다. 이에 Fleder et al.(2009)는 협업 필터링의 과정을 시뮬레이션을 진행하여 결과적으로 소비의 다양성을 감소시킬 위험이 있음을 확인하였다.

추천 시스템의 다양성을 높이는 연구로 개인에게 추천되는 물품 목록의 다양성을 높이는 방법과 시스템이 추천하는 물품 목록의 다양성을 높이는 방법이 있으며, 일반적으로 전자의 연구가 활발히 이루어지고 있다.

개인의 추천 물품 목록에 대하여 다양성을 증대시키는 방안으로 Ziegler et al.(2005)는 물품 간 유사도가 크도록 아이템의 순서를 정함으로써 사용자들에게 비슷한 종류의 물품들이 등장하지 않도록 하여 사용자의 관심 영역을 최대한 포함하는 추천 물품 목록을 생성하는 탐욕 재랭킹(Greedy Re-ranking) 알고리즘을 제안하였다[11]. Zhang et al.(2008)은 각 물품 간의 거리 행렬을 계산하여 추천 물품 목록에 포함되어 있는 물품 간 거리를 최대화하는 방향으로 다양성을 증가시켰다. Adomavicius et al.(2012)는 추천 물품 목록을 생성 후 사전에 설정한 임계값에 미달하는 물품들에 대하여 유명도가 낮은 물품들을 선택하여 사용자에게 추천하는 방식을 사용하여 사용자에게 다양한 물품들을 소개하는 방법을 제안하였다.

시스템이 추천하는 물품 목록의 다양성을 높이기 위한 방안으로 Park et al.(2008)는 롱테일에 위치한 물품들을 군집화한

후, 군집 내에서 리뷰 혹은 평점이 부족한 물품들의 정보를 추정함으로써 인기가 없는 물품들이 추천 물품 목록에 들어갈 수 있도록 하였다. Schein et al.(2002)는 최신 영화의 경우 출연하는 배우들의 추천 점수를 계산한 후 이들의 합을 추천 점수로 사용하는 콘텐츠기반 추천 방법을 제안하였다.

본 논문에서 제안하는 방식과 유사한 것으로는 Schein et al이 제안한 방식이 있지만, 출연하는 배우를 사용하는 방식은 영화에만 해당하는 반면에 태그를 사용하는 방식은 영화뿐만 아니라 음식, 노래 등의 다양한 분야에 활용할 수 있다는 장점과 콘텐츠 기반 추천 방법이 아닌 협업 필터링과 사용함으로써 사용자의 선호도를 보다 더 정확하게 반영하고 있다.

III. The Proposed Scheme

앞서 언급한 바와 같이 기존의 협업 필터링은 높은 정확성을 가지는 반면 낮은 다양성을 갖는다. 이와 같은 문제를 해결하기 위해 본 논문에서는 태그를 사용하여 다양성을 향상시킨 태그 기반 협업 필터링을 제안한다.

구매한 모든 물품에 대하여 사용자들이 평점을 매겼다고 가정할 때, 기존 협업 필터링에서 유사 사용자 그룹을 찾는 방법의 경우 현재 사용자(A)가 구매했던 물품들(I_a)이 비교하려는 사용자(B)가 구매했던 물품들(I_b) 내에 포함($I_a \in I_b$)되어 있어야 유사도 계산이 가능하다. 하지만 A 사용자가 물품을 구매하면서 집합의 크기가 커지게 된다. 집합의 크기가 커지면서 조건에 맞는 B 사용자를 찾을 수 없어지면서 유사 사용자 그룹의 크기가 줄어든다.

추천 물품 목록을 생성할 때 유사 사용자 그룹에서 구매한 물품들에 대하여 점수를 예측하여 높은 점수를 가진 물품들을 추천 물품 목록에 넣기 때문에 유사 사용자 그룹의 크기가 작다는 것은 같은 물품이 등장할 확률이 증가하기 때문에 결과적으로 다양성이 감소한다.

태그 기반 협업 필터링은 앞서 언급한 문제를 해결하기 위해 유사 사용자 그룹을 찾을 때 물품 평점이 아닌 태그 점수를 기준으로 유사 사용자 그룹을 찾는다. 단, 물품과 태그 간의 관계는 N:1 이므로 물품에 대한 평점을 해당 물품에 해당하는 태그들에 대한 점수로 변환해야 한다. 아래는 물품 평점을 태그 점수로 변환할 때의 점수 표를 나타낸다.

Table 1. Score translation table

Item Rating	Tag Score
4.0 ~ 5.0	5
3.0 ~ 3.9	3
2.0 ~ 2.9	1
1.0 ~ 1.9	-3
0.0 ~ 0.9	-5

위의 <표1>을 기준으로 변환된 점수는 누적 합계로써 사용자 객체에 저장하여 추후 추천 물품의 점수를 예측할 때 사용된다. 하나의 물품에 대한 점수를 예측할 때, 해당 물품의 태그들에 관련된 사용자 객체 내에 저장된 점수들을 가져와 총합 점수를 계산한다.

3.1 System Design

본 논문에서 기존 협업 필터링과 제안된 태그 기반 협업 필터링 간의 평가를 진행하기 위하여 다음과 같이 프로그램을 설계하였다.

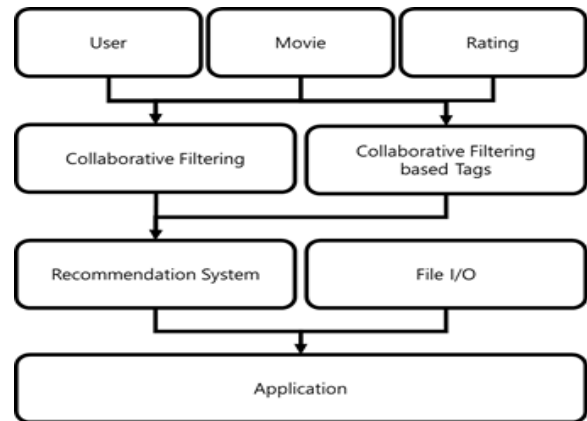


Fig. 1. System structure

프로그램 내에서 기존 협업 필터링과 태그 기반 협업 필터링 간의 설계상 차이점은 존재하지 않는다. 다만 기존 협업 필터링과는 달리 태그 기반 협업 필터링은 사용자 객체에 대하여 평가(Rating) 점수 데이터를 불러올 때 앞서 언급했던 점수 표 <표 1>을 참고하여 태그 별 점수(Score) 데이터로 변환하여 저장한다.

Recommendation System 클래스에서는 추천 목록을 생성을 위한 함수가 존재하며(createRecommendList), 해당 함수 내에서는 다음과 같은 작업이 발생한다.

- 1) 유사 사용자 그룹 검색
- 2) 유사 사용자 그룹 내 유사도가 0.5 미만인 사용자를 그룹에서 제외
- 3) 유사 사용자 그룹 내 사용자들이 구매한 물품에 대한 평가(혹은 점수)들의 합계 계산하여 추천 물품 목록 생성
- 4) 추천 물품 목록을 평가(혹은 점수)를 기준으로 내림차순 정렬

위 과정 중 유사 사용자 그룹을 검색하는데 있어서 User 클래스에서는 가상 함수를 제공함으로써 비교할 기존 협업 필터링 클래스(Collaborative Filtering) 및 태그 기반 협업 필터링 클래스(Collaborative Filtering based Tags) 각각이 다른 방식으로 검색을 할 수 있도록 하였다.

프로그램의 Main 내에서는 다음과 같은 목록을 유지하고 있다. 이전 추천 물품 목록의 경우 다양성(Diversity)를 측정하기 위하여 유지하며, 노출된 물품 목록은 참신성(Novelty)를 측정하기 위하여 유지된다.

- 이전 추천 물품 목록(prevRecommendList)
- 노출된 물품 목록(exposeMovies)

본 논문에서는 기존 협업 필터링과 태그 기반 협업 필터링 간의 정확성(Precision), 다양성, 참신성 방면에서 측정하였으며, 이에 대한 측정은 다음의 수식을 기반으로 진행하였다.

3.2 Search for Similarity User Group based Tag Score

본 논문에서 제안하는 태그 기반 협업 필터링에서는 사용자 간의 유사도를 측정할 때 사용자가 물품에 내린 평가 기준으로 유사도를 측정하는 기존 협업 필터링과는 달리 사용자가 태그를 평가한 점수를 기반으로 측정한다.

태그 점수의 경우 앞서 언급한 바와 같이 물품과 태그 간의 관계가 1:N 이기 때문에 물품에 대한 평점을 태그 점수로 변환하는 작업이 필요하다.

프로그램 시작 시 File I/O 클래스가 평점 데이터 및 영화 데이터를 읽어오는데, 이 시점에서 평점 데이터와 영화 데이터 간 매칭을 통해 사용자 데이터를 생성한다. 이 때, 영화 데이터 내에 있는 태그를 읽어서 평점 데이터를 태그 별 점수로 변환하는 작업을 수행한다.

변화된 태그 별 점수는 사용자 객체에 저장되며, 해당 점수는 태그 별로 누적된다. 즉, 물품에 대한 평가가 좋지 않을 경우 해당 물품에 속하는 모든 태그들의 점수들이 감점된다. 점수를 감점 시킴으로써 평점이 좋지 않은 물품들이 추천 물품 목록 내에 잘 보이지 않게 됨으로 사용자에 대한 선호도 평가가 이루어진다.

본 논문에서 진행한 실험에서는 추천된 물품들 목록 중에서도 하나의 물품을 선택하여 구매한다. 이 때, 구매한 물품에 대하여 내리진 평가를 앞서 언급한 바와 같이 태그 별 점수로 변환하여 사용자의 태그 별 점수를 수정하도록 하여 추천이 될 때마다 사용자에 대한 선호도가 지속적으로 수정이 될 수 있도록 한다.

3.3 Data Structure

본 논문에서 실험에 사용한 데이터는 MovieLens 사이트로부터 제공하는 데이터를 기준으로 진행하였다. MovieLens에서 제공하는 데이터는 영화(movie), 평점(rating)을 제공한다.

영화 데이터는 movie id, title, genes 항목으로 나뉘어 있고, 평점 데이터는 user id, movie id, rating, timestamp로 이루어져 있다.

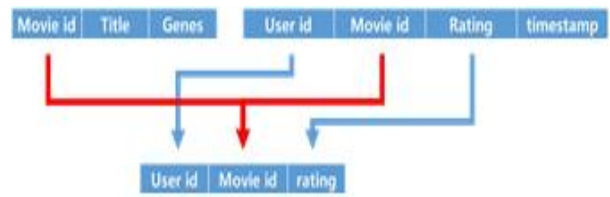


Fig. 2. Data Mapping

사용자(user)에 대한 정보가 별도로 제공되지 않기 때문에 본 논문에서 구현한 프로그램 내에서는 평점 데이터를 기반으로 사용자 데이터를 구축하였다. 사용자 데이터는 user id, movie id, rating로 이루어져 있으며, 태그 기반 협업 필터링은 편의성을 위하여 score 항목이 추가되었다.

IV. Experiments

본 논문에서 진행된 실험은 MovieLens에서 제공하는 데이터 셋을 사용하였다. 해당 데이터 셋은 3.3. 데이터 구조에서 언급한 바와 같이 영화 데이터와 평점 데이터로 구분되어 있으며, 본 논문에서는 해당 데이터들을 Movie ID 기준으로 맵핑하였다. 결과적으로 영화 27,278개, 사용자 7,120명 그리고 태그 20개로 이루어진 데이터 셋을 얻을 수 있었으며, 이를 실험에 사용하였다. 평가 기준은 총 3가지 방향으로 진행되었으며 각각 다양성(Diversity), 참신성(Novelty) 그리고 정확도(Precision)에 대하여 평가를 진행하였다.

실험 내 물품 추천은 [유사 사용자 그룹 검색 → 유사 사용자 그룹 내 구매한 물품 목록 생성 → 구매한 물품 목록 내 물품들의 점수 예측 → 추천 물품 목록 생성 → 점수가 높은 물품 구매 → 유사 사용자 그룹 검색부터 다시 시작] 과정으로 이루어져 있고, 해당 과정을 100번 진행하였다.

기존 협업 필터링(Collaborative Filtering, CF)와 본 논문에서 제안한 태그 기반 협업 필터링(Collaborative Filtering based Tags, CFT) 간에 실험 상 차이점은 유사도 계산과 물품에 대한 점수 예측 외에는 차이가 없다.

실험에서 사용된 사용자 선정은 가장 많은 유사 사용자를 갖는 사용자를 선택하였으며, 이 후 추천 정확성을 측정할 때 정확하게 추천되었는지 평가하기 위한 지표로써 선택된 사용자의 구매 목록에서 무작위로 구매 목록의 1/2을 삭제한다.

다양성(Diversity)와 참신성(Novelty)의 경우 아래에 보이는 수식을 기반으로 측정되었으며, 식 (1)에서의 L_1 은 이전에 추천된 목록을 의미하고 L_2 는 현재 추천된 목록을 의미한다. 식 (2)에서의 A_t 는 t시점까지 추천된 전체 물품 목록을 나타내고 식 (1)과 식 (2)에서 공통적으로 들어가는 N 은 추천된 목록의 개수를 의미한다.

$$Diversity = \frac{|L_2 - L_1|}{N} \quad (1)$$

$$Novelty = \frac{|L - A_i|}{N} \quad (2)$$

다양성(Diversity) 측면에서는 CF의 경우 초반에는 CFT와 비슷하거나 높은 값을 가지는 것을 확인할 수 있었다. 하지만 후반에는 값이 0으로 고정되는 것을 확인하였다. 이와 같은 결과는 CF의 유사 사용자 그룹의 수가 크게 줄면서 유사 사용자가 1명만 남게 되어 추천 물품 목록 내에 있는 물품들이 중복되어 추천되는 것으로 판단된다. 반면에 CFT는 비록 초반에는 높은 값을 가지지 못했지만 CF와는 달리 후반에도 값이 나오는 것을 확인하였다.

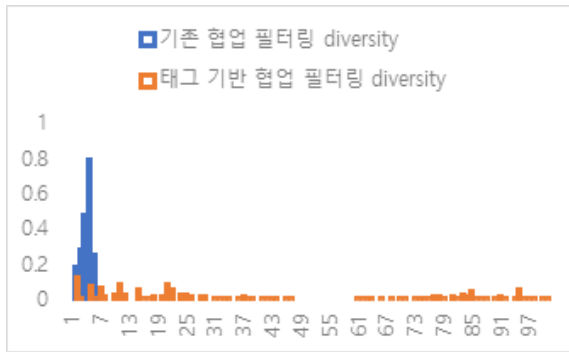


Fig. 3. Comparison graph for diversity

참신성(Novelty)도 다양성과 비슷하게 초반에는 CF의 값이 높지만 후반에는 CFT의 값이 높았다. 이와 같은 결과는 앞서 언급한 바와 같이 유사 사용자 그룹의 크기 때문에 발생한 것으로 보인다. 아래의 <Figure 4>에 보이는 바와 같이 CF의 유사 사용자 그룹 크기는 폭포수와 같이 급격하게 감소하는 반면에 CFT의 유사 사용자 그룹은 천천히 감소하는 것을 확인할 수 있다.



Fig. 4. Comparison graph for novelty

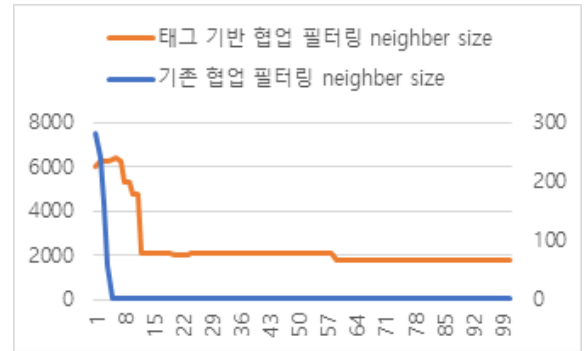


Fig. 5. Comparison graph for size of similarity user group

추천 정확성(Precision)의 경우 아래의 식 (3)에 보이는 바와 같이 앞서 언급했던 사용자의 구매 물품에서 삭제되었던 물품이 추천 목록 내에 등장하면 옳게 추천한 것으로 판단하였다. 결과적으로 보면 추천이 지속될수록 CF가 CFT보다 높은 정확성을 갖는 것을 확인할 수 있었다. 이와 같은 결과는 추천이 지속될 때 CF는 다양한 물품이 추천되지 않는 반면, CFT는 새로운 물품이 지속적으로 노출됨으로써 정확하게 추천된 물품의 개수가 줄어들어 발생하는 것으로 보인다.

$$Precision = \frac{\text{정확하게 추천한 아이 템 개수}}{\text{추천한 전체 아이 템 개수}} \quad (3)$$

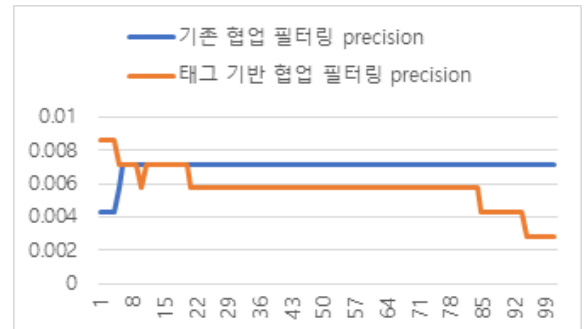


Fig. 6. Comparison graph for precision

V. Conclusions

본 논문에서는 기존 협업 필터링의 추천 다양성이 부족한 문제점을 개선한 태그 기반 협업 필터링을 제안 및 구현하였다. 기존의 추천 다양성 증가를 위해 제안된 방법으로는 앞서 관련 연구에서 언급했다시피 Ziegler et al이 제안한 물품 간의 유사도를 활용하는 방안과 Zhang et al이 제안한 물품 간 거리 행렬을 구하는 방법 그리고 Adomavicius et al이 제안한 임계값 미달의 물품을 추천하는 방법들이 있으나, 이러한 방

법들은 사용자의 선호도를 반영하지 못하고 기존 사용자가 구매한 물품과 유사한 물품들 중 노출도가 낮은 물품을 추천 물품 목록에 포함함으로써 추천 다양성을 증가시켰다. 하지만 본 논문에서 제안한 태그 기반 협업 필터링은 기존 사용자가 구매한 물품에 대한 평점을 기준으로 태그 별 점수를 매김으로써 사용자의 태그 별 선호도를 반영하여 추천 다양성이 높은 추천 물품 목록을 생성하고 있다.

실험의 결과에 보이다시피 본 논문에서 제안한 태그 기반 협업 필터링의 추천 정확성의 경우 추천이 지속될수록 점차적으로 낮아지는 현상이 발생하는데, 이는 추천 정확성과 추천 다양성 간의 관계가 반비례를 이루고 있기에 발생하는 현상이다.

본 논문에서 제안된 태그 기반 협업 필터링의 경우 장기적인 측면에서 볼 때, 새로운 물품들을 추천하는 성능이 기존 협업 필터링보다 높은 반면에 단기적인 측면에서는 기존 협업 필터링보다 다소 떨어지는 성능을 가지는 문제가 있다. 향후 연구에서는 신경망(NN, Neural Network) 등의 AI(Artificial Intelligence)을 도입하여 성능 향상을 추구할 예정이다.

REFERENCES

- [1] Jieun Son, Seoung Bum Kim, Hyunjoong Kim, Sungzoon Cho. "Review and Analysis of Recommender Systems." *Journal of the Korean Institute of Industrial Engineers*, 41.2 pp. 185-208, April 2015
- [2] Jong-Chan Lee, Moon-Ho Lee. "Big data-based information recommendation system." *Journal of the Korea Institute of Information and Communication Engineering*, 22.3 pp. 443-450. March 2018.
- [3] Eun-Young Bae, Seok-Jong Yu. "Keyword-based Recommender System Dataset Construction and Analysis". *The Journal of Korean Institute of Information Technology*, 16(6), pp. 91-99. June 2018
- [4] Seok-Jong Yu. "A Study of Improvement of Individual Item Diversity in Collaborative Filtering-based Recommendation." *The Journal of Korean Institute of Information Technology*, 14.8 pp. 89-94. October 2016
- [5] Hyunwoo Kim, Hyoung-Joo Kim. "Item Recommendation using Tag Expansion and Temporal Information." *Journal of KIISE : Computing Practices and Letters*, 18(7), pp. 521-527. July 2012.
- [6] Min Sehee, Yeom Hongseon. "Recipe Recommendation System based on Ingredient By using Hybrid Filtering." *Journal of the Korean Information Science Association*, 1575-1577. December 2014.
- [7] Han Yong Eom, Chang Hyeon Park. "Recommender System based on Hidden Markov Model." *Korea Electronic Engineering Conference*, 78-81. November 2001.
- [8] Nachiketa Sahoo, Param Vir Singh, and Tridas Mukhopadhyay. "A hidden Markov model for collaborative filtering". *MIS Q.* 36, 4 1329-1356. December 2012.
- [9] Jong-Hun Kim, Kyung-Yong Chung, Joong-Kyung Ryu, Kee-Wook Rim, Jung-Hyun Lee. 2008. A Recommendation Agent System Using HMM-Based Collaborative Filtering in Smart Home Environment. 11-13. Nov 2008.
- [10] N. Lathia, S. Hailes, L. Capra, and X. Amatriain, "Temporal Diversity in Recommendation Systems", *SIGIR*, Geneva, pp. 210-217, July. 2010.
- [11] C-N. Ziegler, S. M. McNee, J. A. Konstan, and G. Lausen, "Improving recommendation lists through topic diversification", *WWW 2005*, Chiba, Japan, pp. 22-32, Jan. 2005.
- [12] M. Zhang and N. Hurley, "Avoiding Monotony: Improving the diversity of Recommendation Lists", *RecSys 2008*, Lausanne, Switzerland, pp. 123-130, Oct. 2008.
- [13] T. Zhou, Z. Kuscsik, J-G. Liu, M. Medo, J. R. Wakeling, and Y-C. Zhang, "Solving the apparent diversity-accuracy dilemma of recommender systems", *Proc. of the National Academy of Sciences of the United States of America* 107, 10, pp. 4511-4515, Feb. 2010.
- [14] O. Celma and P. Herrera, "A New Approach to Evaluating Novel Recommendations", *RecSys 2008*, Lausanne, Switzerland, pp. 179-186, Oct. 2008.
- [15] G. Adomavicius and Y. Kwon, "Improving Aggregate Recommendation Diversity Using Ranking-based Techniques", *IEEE Transactions on Knowledge and Data Engineering*. Vol. 24, No. 5, pp. 896-911, May 2012.
- [16] Seok-Jong Yu. "Frequency-sensitive Diversification in Collaborative Filtering". *Journal of the Korea Information Technology Association*, 13(7), 93-98. July 2015.

Authors



JinHyeon Joo received B.S and M.S degrees in Computer Engineering from Asan Hoseo University in 2011 and 2013 respectively. He is currently doing a Ph.D degree in Computer Engineering at Hoseo University. His current research interests include

Recommendation System, Artificial Intelligence, Internet of Things technology.



Geunduk Park received M.S. and Ph.D. degrees in Computer Engineering from Seoul National University in 1997 and 2005 respectively, and B.S. degree in Computer Science and Statistics from Seoul National University in 1993. He is currently a

professor of the department of computer engineering at Hoseo university. His current research interests include embedded software engineering, service oriented architecture, and semantic web technology.