

추천 다양화 방법을 적용한 콜드 아이템 추천 정확도 향상

한정규[†], 천세진^{**}

Improved Cold Item Recommendation Accuracy by Applying an Recommendation Diversification Method

Jungkyu Han[†], Sejin Chun^{**}

ABSTRACT

When recommending cold items that do not have user-item interactions to users, even we adopt state-of-the-arts algorithms, the predicted information of cold items tends to have lower accuracy compared to warm items which have enough user-item interactions. The lack of information makes for recommender systems to recommend monotonic items which have a few top popular contents matched to user preferences. As a result, under-diversified items have a negative impact on not only recommendation diversity but also on recommendation accuracy when recommending cold items. To address the problem, we adopt a diversification algorithm which tries to make distributions of accumulated contents embedding of the two items groups, recommended items and the items in the target user's already interacted items, similar. Evaluation on a real world data set CiteULike shows that the proposed method improves not only the diversity but also the accuracy of cold item recommendation.

Key words: Item Cold Start, Recommender System, Recommendation, Diversification, Maximum Marginal Relevance, Proportionality.

1. 서 론

클릭, 구매 등 사용자와 아이템 사이의 상호작용이 충분히 존재하지 않는 콜드 아이템(Cold item)을 추천할 경우 추천시스템의 주요 알고리즘인 협업필터(Collaborative Filter)[1,2]를 사용하여 사용자의 취향에 맞는 아이템을 정확하게 추천하기 어려운 아이템 콜드 스타트(Item Cold-Start) 문제가 발생한다. 콜드 스타트 문제를 개선하기 위해, 콜드 아이템

이 시스템에 등록될 때부터 보유하고 있는 콘텐츠 정보를 인공지능망의 입력으로 받아, 협업필터를 사용하여 계산되는 아이템의 상호작용자질(Interaction feature)을 예측하여 협업필터에 적용하는 방법이 최근에 연구되어 기존 방법 대비 개선된 성능을 보이고 있다[3,4].

인공지능망을 사용하는 방법[3,4]은 실제 상호작용 정보를 사용하지 않고 콜드 아이템의 콘텐츠로부터 상호작용자질을 “예측”하는데, 실제 상호작용 정

※ Corresponding Author : Sejin Chun, Address: RS-625, Dong-a University, 37 Nakdong-daero 550(obaegosip)be, Hadan 2(i)-dong, Busan, Republic of Korea, TEL : +82-51-200-7923, FAX : +82-, E-mail : sjchun@dau.ac.kr

Receipt date : Aug. 11, 2022, Approval date : Aug. 26, 2022

[†] Division of Computer Engineering and AI, Dong-A University (E-mail : jkhan@dau.ac.kr)

^{**} Division of Computer Engineering and AI, Dong-A University (E-mail: sjchun@dau.ac.kr)

※ This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. 2021R1G1A1092248) and also supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government (MSIT) (No. 2021R1F1A1050937).

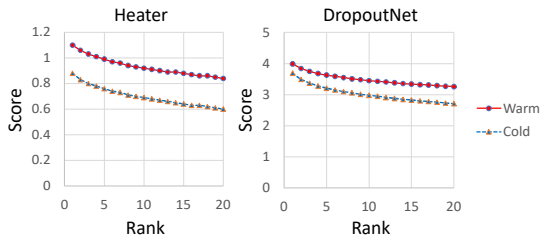


Fig. 1. Average recommendation score w.r.t. item rank.

보를 충분히 사용할 수 있는 워밍 아이템(Warm Item)에 비해 예측한 사용자 취향에 대한 알고리즘의 확신도가 낮아지는 경향이 관측된다. Fig. 1은 본 논문에서 제안 방법 평가에 사용한 CiteULike¹⁾ 데이터를 사용해 학습한 Heater[4], DropoutNet[3] 모델이 사용자에게 추천한 상위 20개의 아이템에 부여한 추천점수 평균을 워밍 아이템과 콜드 아이템에 대해 각각 나타낸 그래프이다. 동일 순위에서 워밍 아이템이 콜드 아이템에 비해 추천점수가 높음을 알 수 있다.

사용할 수 있는 정보량의 차이로 알고리즘이 계산하는 추천점수의 크기뿐만 아니라 사용자 취향에 부합하는 아이템을 추천하는 성능도 콜드 아이템이 워밍 아이템에 비해 낮은 경향이 관측된다. Table 1은 Heater와 DropoutNet 각각이 사용자에게 추천한 상위 10개, 20개 아이템의 콘텐츠 임베딩벡터 평균과 추천대상 사용자가 소비한 아이템의 콘텐츠 임베딩벡터 평균의 코사인 유사도[5]를 측정된 결과를 각각 나타내었다.

코사인 유사도의 값이 1.0에 가까울수록 사용자가 추천시점까지 소비한 아이템의 콘텐츠 경향과 추천시스템이 사용자에게 추천한 아이템의 콘텐츠 경향이 유사하다고 생각할 수 있다. Table 1은 워밍 아이템 추천이 콜드 아이템 추천에 비해 “훨씬” 더 정확하게 사용자의 콘텐츠 취향에 맞는 아이템을 추천함을 보인다.

콜드 아이템 추천성능 개선을 위해, 본 논문에서

Table 1. Cosine similarity of contents embedding vectors between user preference and recommended items.

k	Heater [4]		DropoutNet [3]	
	cold	warm	cold	warm
10	0.55	0.76	0.48	0.62
20	0.57	0.78	0.52	0.66

1) <https://github.com/js05212/citeulike-a>

는 기존 추천 알고리즘의 추천 결과에 “추천 다양화 방법”을 적용하는 방법을 제안한다. 구체적으로는 기존의 최신 콜드 스타트 아이템 추천 기법을 사용하여 사용자에게 추천된 콜드 아이템 추천 리스트를 입력으로, 사용자에게 추천되는 아이템 집합의 콘텐츠 특징 분포가 추천대상 사용자의 콘텐츠 취향 분포와 유사해지도록 재배열하는 후처리(Post processing) 다양화 방법을 적용한다. 다양화 방법을 적용하지 않아도 정확한 추천을 위한 사용자의 취향은 충분히 파악할 수 있어, 다양화 방법을 적용할 경우 추천 결과의 다양성은 증가하지만 추천 정확성이 어느 정도 감소하는 워밍 아이템추천과 달리 콜드 아이템은 추천 다양화 방법의 적용으로 다양성과 정확성이 모두 증가할 수 있음을 보인다. 또한 제안 방법은 후처리 방법이기 때문에, 적용에 있어서 서비스 중인 기존 추천시스템의 구조 변경을 최소화할 수 있다.

본 논문은 다음과 같이 구성된다. 2장에서 관련 연구, 3장에서 제안 방법의 구조를 설명하고, 4장에서는 실제 데이터를 사용하여 수행한 실험 결과를 바탕으로 성능을 논의한다. 마지막 5장에서는 본 논문을 정리한다.

2. 관련 연구

본 장에서는 제안 방법의 적용대상이 되는 콜드 스타트 추천 알고리즘에 대해 살펴본 다음, 제안 방법이 속하는 추천 결과 다양화 연구에 대해 정리한다.

2.1 인공신경망을 이용한 상호작용 정보 예측

콜드 스타트 문제를 개선하기 위한 연구로, 인공신경망을 사용하여 아이템이 생성될 때부터 보유하고 있는, 예를 들어 영화 줄거리 요약, 장르 등 콘텐츠 정보로부터 아이템에 대한 사용자의 상호작용 패턴을 예측하여 콜드 아이템의 추천성능을 개선하기 위한 연구가 이루어지고 있다[3,4,6,20]. 이들은 먼저 BPR[7] 등 행렬분해 알고리즘[1]을 사용하여 사용자, 아이템 각각에 대한 상호작용 자질을 계산한 다음, 사용자나 아이템 각각의 콘텐츠 정보를 인공신경망의 입력으로 하여 미리 계산한 상호작용 정보를 직접[4,6] 혹은 간접적[3]으로 예측한다. 이들 방식은 예측된 상호작용 자질이 아닌 실제로 계산된 상호작용 자질도 사용할 수 있기 때문에 콜드 아이템뿐만이 아닌 워밍 아이템에 대한 추천도 가능하여 활용성이

높으나, 서론에서 언급한 바와 같이 사용자 취향에 부합하는 콜드 아이템 추천성능은 워밍 아이템 추천에 비해 크게 낮은 경향이 있다.

2.2 추천 다양화 방법

평가의 객관성 및 용이성으로 인해 일반적인 추천 시스템 연구에서는 사용자가 실제 상호작용한 아이템 중 일부를 은닉(Hide-out)하고 은닉한 아이템을 정확히 추천했는지를 평가하는 추천 정확도가 주된 평가관점으로 사용되고 있다[8].

추천 정확도를 개선하기 위한 알고리즘은 PRP (Probability Ranking Principle)[9]에 근거하고 있어 정확도를 높이기 위한 최선의 해를 구할 수 있으나, 개개의 아이템이 사용자 취향에 가장 부합하는지만을 판단 근거로 삼기 때문에 동일한 그룹에 속한 서로 유사한 아이템을 다수 추천하는 문제점이 발생한다.

추천 다양화 기법은 이를 개선하기 위해, 사용자에게 제공된 추천 아이템 리스트에 속하는 개개의 아이템 상호 간의 차이를 최대화하는 기법[10], 사용자가 소비한 아이템의 콘텐츠 특징의 누적 분포를 계산하여 사용자 콘텐츠 취향 분포로 삼고, 추천하는 아이템 리스트에 포함된 아이템의 콘텐츠 특징의 누적 분포가 사용자 콘텐츠 취향 분포와 유사하도록 아이템을 추천하는 콘텐츠 취향 기반 다양화 방법 [11,12] 등이 연구되었다. 이들 연구는 추천 정확도가 약간 감소하더라도 추천 아이템의 다양성 증가가 사용자 경험에 긍정적으로 작용함을 보였다. 또한 Zhu 등은 콜드 아이템 추천에 있어서 비인기 아이템 그룹에 속하여 사용자에게 노출 기회가 적은 아이템을 취향에 부합하는 사용자에게 더 많은 추천 기회를 주는 방법[13]을 연구하였다.

[10,11,12]의 연구는 워밍 아이템 추천환경에서 추천되는 아이템의 콘텐츠 다양화 방법 및 그 효과에 대한 연구이고, [13]의 연구는 콜드 아이템 추천 정확도를 어느 정도 희생하여 아이템 노출 기회 평등을 개선하기 위한 연구인데 비해, 제안 방법은 콜드 아이템의 추천 정확도 향상을 목적으로 추천 다양화 방법을 적용한 연구라는 점이 차이점이다.

3. 비율기반 추천 다양화 방법

본 장에서는 제안 방법의 개요를 먼저 설명하고,

이어 알고리즘의 핵심인 목적함수에 대하여 설명한다.

3.1 개요

콜드 아이템 추천은 워밍 아이템 추천에 비해 사용자 콘텐츠 취향에 부합하는 아이템을 추천하는 능력이 크게 떨어지는 경향이 있다(Table 1). 이는 기존 알고리즘[3,4]이 콜드 아이템의 콘텐츠 정보로부터 상호작용자질을 “예측”하는데, 예측 과정에서 하나의 아이템에 포함되어 있는 여러 콘텐츠 요소 중 제일 강하게 나타나는 특정 콘텐츠 요소가 상호작용자질 예측에 강하게 반영되어, 결과적으로 제일 강한 콘텐츠 요소를 기반으로 사용자에게 아이템을 추천하게 되어 추천 정확도 저하로 이어진다고 생각된다.

예를 들어 Fig. 2에 나타난 사용자는 IR(정보검색)에 40%, DB(데이터베이스)에 30%, DC(분산컴퓨팅)에 20%, GT(그래프이론)에 10%의 흥미를 가지고 있는데 논문1이 IR:DB가 6:4인 내용(콘텐츠 요소), 논문2가 IR:GT가 9:1인 내용일 경우 기존 방법은 사용자에게 논문1보다 논문2를 추천할 수 있다.

이를 완화시키기 위해 제안 방법은 (1) 사용자의 취향에 부합하는 아이템을 추천하면서, (2) 추천된 아이템의 콘텐츠 분포가 사용자의 콘텐츠 취향 분포를 따르는 아이템을 추천하도록 한다. 구체적으로는 Fig. 3처럼 추천대상 사용자 u 를 위해 기존 콜드 스타트 추천 방법을 사용하여 알고리즘이 계산한, 아이템이 사용자 취향에 부합하는 정도(추천 점수)를 사용하여 계산한 추천 아이템 리스트 I_u^c 를 입력으로, 추천 다양화 기준에 따라 I_u^c 에 포함된 아이템의 순서를 재배치한 최종 추천 아이템 리스트 I_u 을 생성하는 MMR[14] 기법의 후처리 알고리즘을 사용한다.

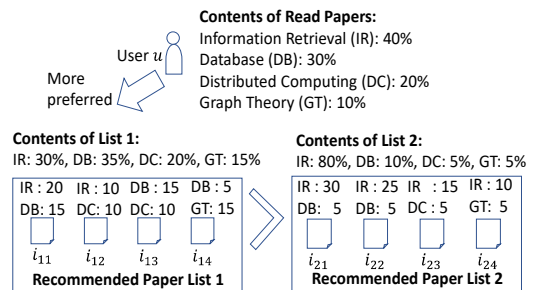


Fig. 2. Contents Distribution based Diversification.

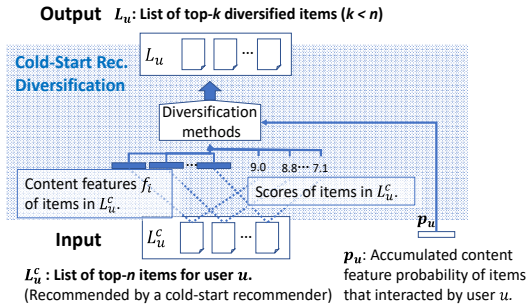


Fig. 3. Overview of Diversification.

Fig. 4에 제안 방법의 의사코드를 나타내었다. 제안 방법은 사용자 u 를 위해 기존 콜드 스타트 추천 방법을 사용하여 생성한 n 개의 아이템으로 구성된 추천 리스트 L_u^c 을 입력으로 사용한다. L_u^c 에 포함된 아이템 각각에 대해, 최종적으로 u 에게 추천되는 아이템의 리스트인 L_u 에 해당 아이템이 포함되었을 때의 다양화 점수를 계산하여, 이를 최대화하는 아이템 (i)을 L_u 에 추가시키는 아이템 선정 작업(Line 03~07)을 수행한다. 한 번의 선정작업에 하나의 아이템이 L_u 에 추가되므로 이를 k 회 반복 (Line 02) 하여 아이템 k 개로 구성된 최종 추천 리스트 L_u 을 출력한다.

3.2 목적함수

다양화 알고리즘은 목적함수 (Fig. 4, Line 04)의 내용에 따라 결과가 달라진다. 제안 방법은 [11,12]과

Algorithm CS_Div:

Input:
 L_u^c : an ordered list of n cold items for user u calculated by a cold start recommender.
 k : the number of items added to output list L_u .
 λ : a diversification weight.
 p_u : the content feature preference probability distribution of user u .

Output:
 L_u : an ordered list of k recommended cold items for user u .

Algorithm:

```

01:  $L_u = \phi$ 
02: for 1 to  $k$ :
    //Select the item which maximizes accuracy,
    // and diversity (content preference similarity)
03:   for each  $i \in L_u^c$ :
04:      $i^{add} = \operatorname{argmax}_{i \in L_u^c} \operatorname{obj}(L_u, i, \lambda, p_u)$ 
05:    $L_u = L_u \cup \{i^{add}\}$ 
06:    $L_u^c = L_u^c - \{i^{add}\}$ 
07: endfor
08: endfor
09: return  $L_u$ 
    
```

Fig. 4. Diversification Algorithm.

유사하게 추천된 아이템의 콘텐츠 분포가 사용자의 콘텐츠 취향 분포를 따르도록 목적함수를 정의한다.

예를 들어 Fig. 2와 같이 사용자가 IR, DB, DC, GT 콘텐츠를 포함한 논문을 40-30-20-10의 비율로 읽었다면, 추천 리스트에 포함된 논문의 콘텐츠 비율이 80-10-5-5이 되는 논문보다, 콘텐츠 비율이 30-35-20-15이 되는 논문을 추천하게 되어 사용자 콘텐츠 취향 분포와 더 유사한 콘텐츠로 구성된 아이템을 추천한다.

$$\begin{aligned}
 & \operatorname{obj}(L_u, i, \lambda, p_u) \\
 &= \sum_{j \in L_u \cup \{i\}} s_{u,j} - \lambda \cdot \operatorname{KLdiv}(p_u, p_{L_u \cup \{i\}})
 \end{aligned} \tag{1}$$

제안 방법에서 사용한 목적함수는 식(1)과 같다. 사용자 u 를 위한 최종 추천 리스트 L_u 에 추가시킬지를 평가하는 대상이 되는 아이템을 i 라 할 때, 추천 리스트 $L_u \cup \{i\}$ (i 를 기존 L_u 에 추가할 경우)에 포함된 각 아이템 j 추천점수 $s_{u,j}$ 의 합을 계산하고, $L_u \cup \{i\}$ 에 포함된 아이템 전체를 사용하여 계산한 아이템 콘텐츠 요소의 누적 분포 $p_{L_u \cup \{i\}}$ 와 사용자의 콘텐츠 취향 분포 p_u 사이의 유사도를 계산한 다음, 추천점수와 콘텐츠 유사도의 가중합을 목적함수 값으로 한다.

분포 유사도를 측정하기 위해 [12]를 따라 두 분포의 거리를 측정하는 KL divergence [15] (식 (1)의 KLdiv)를 사용하였다. 두 분포의 형태가 완전히 일치하면 KL divergences는 0, 차이가 크면 클수록 높은 양수 값을 가진다. 다양성 가중치 λ 는 0.0에서 1.0 사이의 범위를 가지는 다양성 분포 가중치이며 0.0에 가까울수록 기존 알고리즘을 사용하여 계산한 추천 점수를 우선한다.

p_u 는 식(2)를 사용하여, $p_{L_u \cup \{i\}}$ 은 식(3)의 아이템 리스트 L 의 내용을 $L_u \cup \{i\}$ 로 하여 계산한다.

$$p_u = \operatorname{softmax}\left(\frac{1}{|L_u|} \sum_{i \in L_u} f_i\right) \tag{2}$$

$$p_L = \operatorname{softmax}\left(\frac{1}{|L|} \sum_{j \in L} f_j\right) \tag{3}$$

p_u 는 사용자 u 가 상호작용한 아이템 집합 L_u 에 속하는 각각의 아이템 i 의 콘텐츠에 대한 임베딩 벡터인 콘텐츠자질 f_i 의 평균값에 대한 softmax[16]로 계산된다. 유사하게 $p_{L_u \cup \{i\}}$ 은 추천 아이템 리스트 $L_u \cup \{i\}$ 에 속한 각 아이템 j 의 콘텐츠자질 f_j 의 평균

값에 대한 softmax로 계산된다. 콘텐츠자질은 아이템 시스템에 등록될 때부터 보유하고 있는 콘텐츠 정보를 사용하여 계산되므로 콜드 아이템 여부와 상관없이 계산 가능하다. 임베딩 벡터의 원소값은 음의 값을 가질 수 있으므로 확률분포로 변환하기 위해 softmax를 사용한다.

4. 평가 및 고찰

본 장에서는 제안 방법의 성능 평가에 사용한 데이터 및 실험 환경에 대해 설명한 다음, 평가 결과를 보고하고 제안 방법의 성능을 고찰한다.

4.1 실험 데이터

콜드 아이템 관련 추천 연구 [3,4,6]에서 사용한 CiteULike 데이터를 사용하여 제안 방법을 평가하였다. CiteULike는 사용자가 본인의 관심 논문 리스트에 등록한 학술논문에 관한 데이터이다. 총 16,980개의 논문에 대한 사용자 5,551명이 생성한 204,986회의 등록에 대한(등록: 1 비등록: 0) 데이터로 구성된다. 콘텐츠 데이터로는 논문 제목과 초록이 제공된다. 먼저 [3,4]의 연구와 동일한 처리를 통해 콘텐츠 자질을 계산하였다. 구체적으로는 tf-idf [17]를 사용하여 전체 논문 집합(Corpus)에서 나타나는 단어에 대한 중요도를 계산한다. 다음에 tf-idf 점수가 높은 8,000단어를 선별하여, 각 논문에 대해 선별한 8,000개의 단어가 나타나는지의 여부를 벡터로 나타내고 이들을 특이값분해[18]를 사용하여 논문마다 300차원의 벡터로 표현되는 콘텐츠자질 f_i 를 계산하였다.

CiteULike 논문 중 1/5의 논문을 임의로 추출하여 콜드 아이템으로 선정하고, 이들 논문에 대한 사용자의 등록 데이터를 콜드 아이템 평가를 위한 데이터로 사용하였다. 구체적으로는 콜드 아이템에 대한 등록 데이터 30%를 기존 방법[3,4] 학습 결과의 검증(Validation), 나머지 70% 중 1/5은 제안 방법 성능 검증, 나머지 4/5를 제안 방법 성능 테스트(Test)에 사용하였다.

콜드 아이템으로 선정하지 않은 나머지 4/5의 논문은 워밍 아이템으로 선정하고 이들에 대한 사용자의 등록 데이터 중 75%를 기존 방법[3,4] 학습(Training)에, 나머지 25%중 1/5은 제안 방법 성능 검증, 나머지 4/5를 제안 방법 성능 테스트에 사용하였다. 테스트

Table 2. Test data statistics.

	Cold	Warm
# of Users	4,886	5,272
# of Candidate items	2,371	11,358
# of Ground Truths	22,173	32,787

Table 3. Evaluated Configurations.

Method	Description
HC/ HW	사용자 u 에게 Heater[4]을 사용하여 k (=20)개의 콜드(HC)/워밍(HW) 아이템을 추천.
DivHC/ DivHW	사용자 u 에게 Heater를 사용하여 선택한 n (=60)개의 콜드(DivHC)/워밍(DivHW) 아이템(R_u^c)을 제안 방법의 입력으로 하여 k (=20)개의 콜드(DivHC)/워밍(DivHW) 아이템(R_u)을 추천.
DC/ DW	사용자 u 에게 DropoutNet[3]을 사용하여 k (=20)개의 콜드(DC)/워밍(DW) 아이템(R_u)을 추천.
DivDC/ DivDW	DivHC/DivHW의 Heater를 DropoutNet으로 변경.

데이터 내역은 Table 2에 나타내었다.

4.2 실험 환경

제안 방법을 Table 3에 나타난 8종류의 구성에 대해 평가하였다. 제안 방법이 개요에서 설명한 가정을 개선하여 콜드 아이템 추천성능 정확도가 향상되는지를 확인하기 위해 기존 추천 방법 및 제안 방법을 콜드 아이템 추천과 워밍 아이템 추천에 대해 각각 평가하였고, 제안 방법이 개선하는 문제가 특정 알고리즘에 한정된 문제가 아니며 콘텐츠 정보로부터 상호작용 정보를 예측하는 방법에 공통적인 문제인지를 검증하기 위해 Heater[4] 및 DropoutNet[3]을 사용하여 생성한 추천 리스트를 제안 방법의 입력으로 사용하였다.

Heater, DropoutNet, 제안 방법은 Tensorflow 2.6²⁾ 및 Python 3.8³⁾을 사용하여 구현하였다. Heater, DropoutNet의 하이퍼 파라미터(Hyper parameter) 값은 원논문의 설정을 따랐으며 사용자와 논문의 상

2) <https://www.tensorflow.org/>

3) <https://www.python.org/>

호작용자질은 BPR[7]을 사용하여 크기 200차원의 벡터를 계산하였다.

4.3 평가지표

추천 정확도 및 추천되는 아이템의 콘텐츠 분포가 어느 정도 사용자 콘텐츠 취향 분포와 유사한지를 평가하기 위한 취향 분포 유사도 지표를 사용하여 평가하였다.

추천 정확도를 측정하는 지표로 *Precision@k* ($P@k$) [19]와 *Recall@k* ($R@k$)[19]을 사용하였다. $P@k$ 는 추천대상 사용자를 위해 추천된 상위 k 개의 아이템 중 몇 개가 해당 사용자의 정답 데이터에 존재하는지의 비율을 계산하고, $R@k$ 는 추천대상 사용자의 정답 데이터에 존재하는 아이템 중 몇 개가 사용자를 위해 추천된 상위 k 개의 아이템에 속하는지의 비율을 계산한다. 두 지표 모두 1.0이 최고 성능, 0.0이 최저 성능을 나타낸다.

사용자 콘텐츠 취향 분포에 대한 추천 아이템의 콘텐츠 분포 유사성을 측정하는 지표로는, 사용자 u 의 콘텐츠 취향 분포 p_u (식(2))와 u 에게 추천된 상위 k 개의 아이템 리스트 L_u^k 를 식(3)의 L 의 로 두고 식(3)을 사용하여 계산한 아이템 콘텐츠 분포 $p_{L_u^k}$ 사이의 KL-Divergence의 평균($KL@k$, 식(4))과 코사인 유사도의 평균($Cos@k$, 식(5))을 사용하였다. 식(4), (5)에서 U 는 추천 대상 사용자 집합을 의미한다.

$$KL@k = \frac{1}{|U|} \sum_{u \in U} KLdiv(p_u, p_{L_u^k}) \quad (4)$$

$$Cos@k = \frac{1}{|U|} \sum_{u \in U} Cosine.Sim(p_u, p_{L_u^k}) \quad (5)$$

$KL@k$ 은 최소값인 0에 가까울수록 두 분포 p_u 와

$p_{L_u^k}$ 의 형태가 유사함을 뜻하고, $Cos@k$ 는 1에서 -1의 범위의 값을 가지며 1에 가까울수록 p_u 와 $p_{L_u^k}$ 가 유사함을 뜻한다.

$P@k$, $R@k$, $KL@k$ 및 $Cos@k$ 의 k 로는 10과 20을 사용하였다.

4.4 성능 평가 및 고찰

Table 4에 4.2장에서 설명한 8종류의 구성에 대한 성능 측정 결과를 나타내었다. 검증 데이터를 사용한 평가 결과에서 $R@20$ 이 가장 높았던 다양성 가중치 λ 값을 사용하여 테스트한 결과를 나타내었다.

Heater에 제안 방법을 적용하여 콜드 아이템을 추천하는 DivHC는 목적함수를 통해 명시적으로 향상시키려 의도한 콘텐츠 취향 유사도 ($KL@k$, $Cos@k$)가 Heater만을 사용하는 HC에 비해 향상되었을 뿐 아니라 정확도 지표($P@k$, $R@k$)도 HC에 비해 향상되었음을 확인할 수 있다.

반면에 워م 아이템을 추천하는 DivHW는 목적함수를 통해 명시적으로 향상시키려 의도한 콘텐츠 취향 유사도는 HW보다 향상되었으나 정확도 지표는 HW에 비해 감소하였음을 확인할 수 있다. 이는 [10,11, 12]와 같은 추천 다양화 기법 적용연구에서 보고되는 일반적인 정확도-비정확도 성능의 트레이드오프 현상이다. DropoutNet에 제안 방법을 적용하여 아이템을 추천하였을 때도 Heater의 경우와 유사한 경향을 보인다.

Fig. 5, Fig. 6에 다양성 가중치 λ 값에 따른, 기존 추천 방법(Heater, DropoutNet)에 대한 제안 방법의 $R@20$, $KL@20$ 성능 향상 비율을 각각 나타내었다. $R@20$ 은 콜드 아이템 추천에 제안 방법을 적용할 경우 λ 값이 커짐에 따라 점차 증가하다 성능향상 비율

Table 4. Recommendation Performance of the Diversification method.

Rec.	Alg. (Div. W)	P@10	P@20	R@10	R@20	KL@10	KL@20	Cos@10	Cos@20
Cold	HC	0.103	0.077	0.263	0.382	0.115	0.084	0.55	0.57
	DivHC($\lambda=0.95$)	0.102	0.079	0.271	0.402	0.064	0.064	0.62	0.62
Warm	HW	0.037	0.039	0.093	0.185	0.067	0.053	0.76	0.78
	DivHW($\lambda=0.9$)	0.030	0.036	0.072	0.179	0.036	0.037	0.83	0.83
Cold	DC	0.076	0.063	0.194	0.311	0.137	0.098	0.48	0.52
	DivDC($\lambda=0.99$)	0.086	0.067	0.232	0.345	0.063	0.065	0.61	0.61
Warm	DW	0.049	0.043	0.108	0.181	0.093	0.069	0.62	0.66
	DivDW($\lambda=0.99$)	0.039	0.041	0.086	0.180	0.038	0.040	0.77	0.77

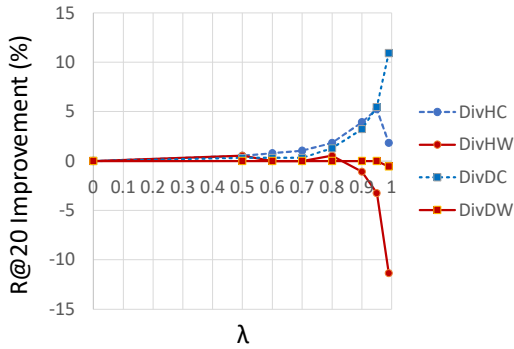


Fig. 5. R@20 improvement w.r.t. λ .

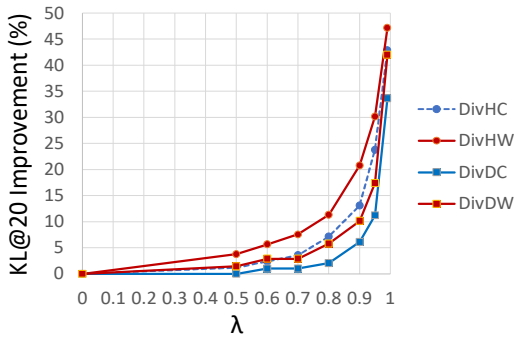


Fig. 6. KL@20 improvement w.r.t. λ .

이 일정 이상의 값에서 증가가 감소하는데 비해 워밍업 아이템 추천에 적용할 경우 λ 값이 커지면 감소한다. 반면 KL@20 성능 향상 비율은 추천 아이템에 관계없이 λ 값이 커질수록 증가한다.

실험 결과로부터, 3.1절에서 가정한 “아이템에서 제일 강하게 나타나는 콘텐츠 요소가 콜드 아이템의 상호작용 자질 예측에 강하게 반영”되는 점을 개선하면 추천 정확도를 향상시킬 수 있다고 생각된다. 반면 워밍업 아이템의 경우 이미 충분한 상호작용 데이터를 사용해서 추천 정확도를 향상시키기 위한 정보를 충분히 학습 하였기 때문에, 제안 방법을 적용하여도 추천 정확도 향상은 어려우며 기존 추천 다양성 연구와 유사하게 정확도와 비정확도 성능의 트레이드오프 관계를 보인다고 판단된다.

5. 결 론

본 논문에서는 최신 콜드 아이템 추천 방법의 개선점을 추천 정확도 측면에서 살펴보고, 추천 정확도를 개선하기 위해, 비정확도 성능을 개선하기 위해

사용되는 추천 다양화 기법을 적용하는 방법을 제안하였다.

실제 CiteULike 데이터를 사용한 성능 측정에서 제안 방법을 적용할 경우 콜드 아이템에 대한 추천 정확도가 기존 알고리즘 대비 약 5-10% 증가함을 보여, 추천 다양화 기법의 적용이 콜드 아이템 추천 정확도 향상에 기여함을 보였다. 결과적으로 이미 구현된 기존 콜드 아이템 추천시스템을 크게 변경하지 않고 추가적인 추천성능 향상을 이룰 수 있는 방향을 제시하였다.

다만 제안 방법은 그 구조상 아이템이 충분한 양의 콘텐츠 데이터를 보유하고 있을 경우에 좋은 성능을 보일 것이라 예상된다. 콘텐츠 데이터의 정보량에 따른 제안 방법의 성능 변화 특성 및 주어진 콘텐츠 데이터에서 제안 방법을 적용하기에 유용한 정보를 추출하는 방법에 관한 연구가 후속 연구과제이다.

REFERENCE

[1] Y. Koren, R. Bell, and C. Volinsky, “Matrix Factorization Techniques for Recommender Systems,” *Computer*, Vol. 42, No. 8, pp. 30-37, 2009.

[2] E. Im and H. Yong, “PARAFAC Tensor Reconstruction for Recommender System based on Apache Spark,” *Journal of Korea Multimedia Society*, Vol. 22, No. 4, pp. 443-454, 2019.

[3] M. Volkovs, G. Yu, and T. Poutanen, “Dropoutnet: Addressing Cold Start in Recommender Systems,” *Proceedings of Advances in Neural Information Processing Systems*, Paper ID 2563, 2017.

[4] Z. Zhu, S. Sefati, P. Saadatpanah, and J. Caverlee, “Recommendation for New Users and New Items via Randomized Training and Mixture-of-Experts Transformation,” *Proceedings of the 43rd International ACM SIGIR Conference*, pp. 1121-1130, 2020.

[5] A.R. Lahitani, A.E. Permanasari, and N.A. Setiawan, “Cosine Similarity to Determine Similarity Measure: Study Case in Online Essay Assessment,” *Proceedings of 4th International Conference on Cyber and IT Service*

- Management*, pp. 1-6, 2016.
- [6] J. Han and S. Chun, "Addressing the Item Cold-Start in Recommendation Using Similar Warm Items," *Journal of Korea Multimedia Society*, Vol. 24, No. 12, pp. 1673-1681, 2021.
- [7] S. Rendle, C. Freudenthaler, Z. Gantner, and L. Schmidt-Thieme, "BPR: Bayesian Personalized Ranking From Implicit Feedback," *Proceedings of the 25th Conference on Uncertainty in Artificial Intelligence*, pp. 452-461, 2009.
- [8] G. Shani and A. Gunawardana, "Evaluating Recommendation Systems," In: F. Ricci, L. Rokach, and B. Shapira, (eds) *Recommender Systems Handbook*. Springer, 2011.
- [9] S.E. Robertson, "The Probability Ranking Principle In IR," *Journal of Documentation*, Vol. 33, No. 4, pp. 294-304, 1977.
- [10] C. Ziegler, S.M. McNee, J.A. Konstan, and G. Lausen, "Improving Recommendation Lists Through Topic Diversification," *Proceedings of the 14th international conference on World Wide Web*, pp. 22-32, 2005.
- [11] V. Dang and W. Bruce Croft, "Diversity by Proportionality: An Election-Based Approach to Search Result Diversification," *Proceedings of the 35th International ACM SIGIR Conference*, pp. 65-74, 2012.
- [12] H. Steck, "Calibrated Recommendations," *Proceedings of the 12th ACM Conference on Recommender Systems*, pp. 154-162, 2018.
- [13] Z. Zhu, J. Kim, T. Nguyen, A. Fenton, and J. Caverlee, "Fairness among New Items in Cold Start Recommender Systems," *Proceedings of the 44th International ACM SIGIR Conference*, pp. 767-776, 2021.
- [14] J. Carbonell and J. Goldstein, "The Use of MMR, Diversity-Based Reranking for Reordering Documents and Producing Summaries," *Proceedings of the 21th Annual International ACM SIGIR Conference*, pp. 335-336, 1988.
- [15] S. Kullback and R.A. Leibler, "On Information and Sufficiency," *The Annals of Mathematical Statistics*, Ann. Math. Statist. Vol. 22, No. 1, pp. 79-86, 1951.
- [16] C.M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.
- [17] A. Rajaraman and J. Ullman, *Data Mining. In Mining of Massive Datasets*, Cambridge University Press, pp. 1-17, 2011.
- [18] S. Brunton and J. Kutz, "Singular Value Decomposition (SVD)," *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*, pp. 3-46, Cambridge University Press, 2019.
- [19] Y. Tamm, R. Damdinov, and Al. Vasilev, "Quality Metrics in Recommender Systems: Do We Calculate Metrics Consistently?," *Proceedings of 5th ACM Conference on Recommender Systems*, pp. 708-713, 2021.
- [20] Y. Wei, X. Wang, Q. Li, L. Nie, Y. Li, X. Li, and T. Chua, "Contrastive Learning for Cold-Start Recommendation," *Proceedings of the 29th ACM International Conference on Multimedia*, pp. 5382-5390, 2021.



한 정 규

2005년 2월 서울대학교 컴퓨터공학부(공학학사)
2007년 2월 서울대학교 전기컴퓨터공학부(공학석사)
2018년 3월 Waseda University, Computer Science and Communications

Engineering (공학박사)

2020년 9월~현재 동아대학교 컴퓨터AI공학부 조교수
2018년 6월~2020년 8월 NAVER AiRS 근무
2007년 7월~2014년 4월 NTT Software Innovation Center 근무
관심분야: 추천 시스템, 정보 검색, 데이터 마이닝



천 세 진

2009년 8월 서울시립대 컴퓨터과학부(공학사)
2018년 8월 연세대 컴퓨터과학과(공학박사)
2021년 3월~현재 동아대학교 컴퓨터AI공학부 조교수

2018년 10월~2020년 8월 미국 표준기술연구원(NIST) ITL 근무

관심분야: 지식그래프 표현 및 추론, 온톨로지 기반 모델링, 스트림 데이터 처리