

# 사용자 선호도 변화에 따른 추천시스템의 다양성 적용

나혜연

울산과학기술원 융합경영대학원  
(hyna@unist.ac.kr)

남기환

동국대학교 경영대학 경영정보학과  
(namkh@dongguk.edu)

추천시스템은 시간이 흐를수록 사용자와 기업에게 점점 더 큰 영향을 주고 있다. 최근 코로나(COVID-19) 팬데믹 현상이 전 세계적으로 일어나면서 세대를 뛰어넘어 E-Commerce의 중요성이 증대되었고 추천시스템은 E-Commerce 활성화의 최중심에 있다. 추천시스템이 개발된 이래로 다수의 알고리즘이 추천시스템의 정확도를 올리는 것에 집중되어 있었고, 추천시스템의 희귀성, 다양성, 우연성 등과 같은 다른 가치들이 간과되고 있다. 본 논문에서는 사용자의 만족도는 추천시스템의 정확도에만 달려있지 않고 다양한 성능을 견비했을 때 고객에게 만족스러운 추천서비스 경험을 제공할 것이라 생각하여 다양성을 위한 그래프 기반의 추천시스템을 개발하였다. 사용자 네트워크를 구성한 뒤 카테고리별 활용한 무게중심 변화를 통해 유사도가 낮은 이질적인 사용자를 찾아 추천상품의 유사성을 낮추는 방식으로 다양성을 도모하였다. 또한, 추천의 다양성은 사용자의 다양성 선호 수준에 따라 상이할 것이라는 가정에 따라 사용자의 다양성 선호 수준을 구별하였고 다양성 모델 성능이 사용자 특성별로 차이를 확인할 수 있었다.

전체 연구 결과, 추천시스템의 정확성과 다양성이 트레이드 오프 관계에 놓여있다는 것을 확인할 수 있었지만 본 연구 모델을 통해 근소한 정확도 손실 대비 높은 다양성을 얻을 수 있었다. 본 연구는 그래프 기반의 추천시스템을 통해 사용자의 만족도를 향상시키는 다양성을 실현하였다는 연구적 의의와 사용자 수준을 고려한 추천의 다양성을 적용 결과를 통해 기업의 장기적 이윤을 상승시킬 수 있는 모델 개발이라는 실무적 의의를 꼽을 수 있다.

**주제어** : 추천시스템, 의사결정지원시스템, 다양성, e-커머스

논문접수일 : 2020년 11월 13일    논문수정일 : 2020년 12월 10일    게재확정일 : 2020년 12월 16일  
원고유형 : 학술대회 Fast-track    교신저자 : 남기환

## 1. 서론

지난 20년간, 추천시스템은 사용자의 정보검색을 돕기 위해 많은 연구가 있었고, 아마존이나 넷플릭스와 같은 기업은 이를 적용하여 큰 수익을 올렸다(Adomavicius and Kwon, 2011). 추천시스템의 목적은 개인의 만족도를 상승시키기 위해 개인화된 추천서비스를 제공하는 것이다(Mouzhi et al. 2010). 구체적으로, 추천시스템은 고객이 인터넷 환경에서 편하고 빠르게 구매를

할 수 있도록 유도하여, 기업은 이윤을 높이고 고객은 구매를 통한 만족감을 얻게 하는 것이다. 실제로, 최상위 E-commerce 책임자들은 추천시스템이 고객의 선택에 매우 중요한 영향력을 가진다고 언급했는데, 그 예로 넷플릭스 미디어 선택의 60%, 아마존 매출의 약 35%가 추천시스템에서 나왔다는 사실이다(Lee and Kartik 2019). 다양한 종류의 추천기법이 있지만, 협력필터링은 추천시스템 중에서 가장 우수한 성능을 나타낸다고 알려진 기법으로써, (Goldberg et al.,

1992)에 의해 그 개념이 처음으로 소개되었다 (Jieun et al., 2015). 협력 필터링은 사용자와 비슷한 상품을 구매하는 다른 유저를 찾거나 혹은 비슷하게 구매한 유저의 상품을 추천하는 방식이다 (Schafer et al., 1999).

하지만, 추천시스템의 평가지표가 다각화되면서 정확도만이 아니라 다른 특성들에 대한 연구가 진행되고 있다 (Adomavicius and Kwon, 2011). 왜냐하면, 협력 필터링을 기반한 추천시스템이 기업과 고객을 지속적으로 만족시키기 힘든 부분도 있음을 확인하고 있기 때문이다. (Fleder et al., 2009)에서는 분석모델과 시뮬레이션을 통해 협력 필터링은 평균적으로 추천이 한 방향으로 치우치게 한다는 것을 발견했다. 또한, 충분한 평점 확보가 되지 않은 경우 역시 협력 필터링으로부터 편향된 추천의 결과가 나온다는 문제도 지적되고 있다 (Mooney and Roy, 2000). 추천시스템이 거대한 하나의 취향을 만들고 사람들을 그 쪽으로 유인한다고 볼 수 있는 지점이다. 장기적으로 기업 측면에서, 고객이 일부 상품만을 선택하는 양극화 현상이 일어나게 되고 기업의 매출은 감소하게 된다. 고객 측면에서도 고객의 취향이 고정되어 있지 않기 때문에 편향된 추천은 만족스러운 서비스 경험을 만들지 못하게 된다. 추천시스템의 품질은 다양성, 우연성, 희귀성, 불예측성 그리고 커버리지와 같은 가치를 포함하며 이는 사용자의 만족도와 직접적으로 연관되어 있다. 따라서 대부분의 추천시스템과 정보검색 분야는 사용자의 만족도를 향상시키기 위해 다양성의 원리를 연구해야 한다 (Adamopoulos et al. 2014).

추천시스템의 다양성을 실현하는 방법 중 가장 많이 쓰이는 연구 방법은 재순위(reranking)이다 (Marius, 2016). 다양성 초기 연구 (Barry,

2001; Ziegler et al. 2005; Kelly and Bridge 2006)에서 사용자의 선호와 다양성을 조화시키기 위해 탐욕적 재순위(greedy reranking) 방식을 사용했는데, 사용자의 선호와 비선호의 관계를 선형 결합으로 풀어내는 것이 핵심이다. 최근에는 재순위 방식을 더욱 정교한 모델로 업데이트 하거나 최적화를 풀어내는 방식에서 다양성을 시도한다. (Vargas et al., 2011)은 정보검색 연구방식을 적용한 확률 모델로 재순위를 통한 다양성 추천시스템을 시도하였고, (Andrea et al. 2015)은 사용자의 프로파일 아이템들끼리 차이를 계산한 뒤 차이값을 강화시키고 정확도의 안정성을 떨어뜨리는 방식으로 다양성을 모델을 제시하였다.

본 연구에서는 추천시스템이 기업과 고객의 만족도를 더 높이기 위해, 고객의 성향을 상품 카테고리로 상세하게 분류하고 취향이 다른 고객과 유사도를 높게 하여 추천의 범위를 새롭고 다양하게 확장시켰다. 이는 기존 연구에서 사용자의 다양성 성향에 대한 고찰없이 추천시스템 알고리즘의 변형만으로 다양성을 시도한 것과 차별을 두고 있다. 연구에서 가정한 고객의 객관적인 취향 변화는 다양한 측면에서 파악할 수 있는데, 일정 부분은 사용자의 특성이고 또 다른 부분은 환경적 요인 (Yehuda Koren, 2009)으로 반드시 일어나는 일이다. 따라서, 본 연구에서는 사용자의 다양성 수준을 알기 위해 시도하였고 그에 따라 다양성 추천시스템의 효과가 어떻게 다를 수 있는지 결과를 얻어냈다. 또한, 기존 추천시스템 연구에서 많이 사용되는 협력 필터링이 아니라 그래프 기반의 추천시스템을 사용하였다. 그래프 기반의 추천시스템은 여러 방식의 추천시스템 모형을 반영시키며 발전시킬 수 있고 (Huang et al., 2002), 다양한 성격을 반영시킬 수

있는 유연성이 있어 (Lei, 2013) 여러 연구에서 활용되고 있다. 추천의 다양성을 증가시킨다는 의미는 추천된 제품 간의 유사성을 감소시킨다는 것이므로, 정확도 손실비용 발생한다(Bradley et al., 2001). 즉, 정확성과 다양성 사이에는 트레이드 오프가 존재하며(Adomavicius et al, 2011), 근소한 손실 혹은 동일한 정확도를 유지하면서 다양성을 극대화시키는 어려운 작업 중 하나이다. 본 연구에서는 추천의 다양성을 넓히기 위해 유사도가 낮은 사용자로부터 추천 목록을 얻지만, 고객의 선호를 낮추지 않는 지점을 찾아내어 매우 근소한 손실의 정확도와 높은 다양성을 얻었다.

## 2. 관련연구

추천시스템은 단순히 높은 정확도를 가지고 있다는 것만으로는 사용자에게 만족을 줄 수 없고, 다양성, 희귀성, 우연성과 같은 여러 요소가 함께 증진되어야 한다(McNee et al. 2006). 모든 요소를 갖춘 추천시스템을 구현하는 데 있어 알고리즘만이 아니라 ‘인간’ 대한 이해가 필요한데, 심리학, 경제학에서는 이러한 인간 심리에 관한 심층적인 연구가 꾸준히 진행되었다. 인간의 욕구를 5가지 계층으로 구분한 매슬로우의 사람은 원하는 욕구가 충족되면 다시 새로운 욕구를 찾는다고 말했다(Maslow, 1943). 예를 들어, 배가 매우 고프 사람은 빵을 찾지만, 충분한 빵을 먹고 나면 욕구상태가 멈추는 것이 아니라 새로운 욕구를 드러내는 것이 인간의 본성이라는 의미이다. 인간은 본능적으로 새로운 것을 추구한다. 이러한 특성은 쇼핑 행동에서 잘 드러난다. 쇼핑에서 개인이 다양성을 찾는 행위는 효용성

에 의해 브랜드를 바꾸는 자체를 의미한다. 무언가를 바꾸는 이유 중 하나는 바꾸는 것 자체로부터의 만족감이 생기기 때문이다(Moshe, 1984). 고센은 재화 혹은 용역이 증가하면 동일한 즐거움의 양은 줄어든다는 한계 효용 법칙을 발표했다. (Itamar, 1990)은 또한 다양한 소비를 하는 소비자는 그들 스스로 자신의 선호를 잘 모르기 때문에 연속적으로 더욱 다양한 제품을 산다고 말하였다. 이는 고객이 가장 선호하는 것이 불확실할 때 다양성 선호로 표출되고, 선호를 바뀌는 것이 고정된 구매에 대한 리스크를 감소시켜 만족감을 높일 수 있음을 의미하고 있다.

추천시스템의 시작은 사용자가 최소한의 검색으로 자신이 좋아하는 제품을 찾도록 도와주거나(Will et al., 1995) 흥미로운 기사들에 대한 집중을 돕는 것이었다(Paul et al., 1994). 이러한 시작과 함께 추천시스템의 연구는 사용자가 앞으로 선택할 제품에 줄 평점을 정확하게 예측하는 것에 집중되어졌다(Kaminskas et al., 2016). 많은 추천시스템이 희귀하고, 우연적이며, 다양하고 흥미로운 추천보다는 더 높은 정확도를 제공하는 것에 관심을 두고 있다(Adamopoulos et al., 2014). 하지만, 추천시스템의 품질은 다양성, 우연성, 희귀성, 불예측성 그리고 커버리지와 같은 가치를 포함하며 이는 사용자의 만족도와 직접적으로 연관되어 있다. 따라서 대부분의 추천시스템과 정보검색 분야는 사용자의 만족도를 향상시키기 위해 다양성의 원리를 연구해야 한다(Adamopoulos et al., 2014). 정보검색 과 관련하여 다양성의 역할은 전형적으로 사용자의 쿼리에서 불명확함과 연관되어 있다(Marius et al., 2016). 여기에서 사용자가 원하는 것이 불명확할 때, 치우친 명백함을 제거하는 것은 다양성이 정확성을 표현할 수 있음을 의미한다. 예를 들어,

‘사과’를 검색했을 때, 먹는 사과, 기업 애플이 동시에 나오는 일은 다양성을 통한 정확성 확보이다. 추천시스템은 모든 제품들 사이에서 유사함을 최소화하기 위해 독특한 제품들로 대체하거나 혹은 추천제품들 사이에서 너무 명백한 제품은 제거할 수 있다(Billsus and Pazzani, 2000). 이는 비슷하고 반복되는 아이템을 제거하여 지루함을 덜고 다양성을 확보하는 방향이다. 반복되는 정확도는 지루함을 만들어낼 수 있다. (Ziegler et al., 2005)는 사용자의 관심 영역을 최대한 포함하는 추천 아이템리스트를 만들기 위하여 내부 리스트 유사도 개념을 제안하였는데 이는 아이템 간의 유사도가 크도록 아이템의 순서를 정함으로써 사용자들에게 비슷한 종류의 아이템들을 등장시키지 않도록 하여 다양성을 확보했다. (Zhang et al., 2008)은 이러한 문제를 각 아이템 간 행렬을 계산하고 추천 리스트에 포함된 아이템의 거리를 최대화하는 최적화 문제로 정리하였다. (Zhang et al., 2012)은 정확성 외에 다양한 가치를 포함한 알고리즘을 제시했고 (Zhou et al., 2010)은 특정 의미 단어 혹은 일부 정보에 의존하지 않고 추천의 정확성과 다양성 모두를 포함한 하이브리드 형식의 추천시스템을 제안했다.

### 3. 다양성을 위한 추천시스템 모델

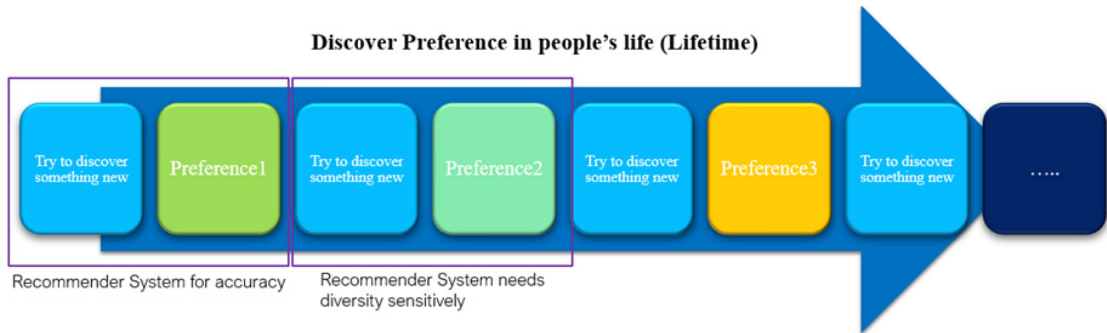
#### 1.1 추천시스템의 다양성 정의

고객이 구매 활동에서 다양한 상품을 쇼핑하는 것은 서비스 혹은 제품을 선택할 때 개인의 다양성 선호에 대한 성향에 달려있다(Barbara, 1995). 지난 연구에서 살펴본 것처럼, 인간은 본

능적으로 전체 생애 주기 동안 자신의 선호를 끊임없이 바꾸며 살아간다. 선호를 탐구하는 과정에서 여러가지 경험이 필요하며, 만들어진 취향은 고정되지 않고 새로운 취향으로 변화한다. 예를 들어, 드라마/멜로 영화를 즐기던 사용자가 어떠한 계기를 통해 액션, 코미디, 공포 등 다른 장르의 영화로 취향이 변화하거나 확장될 수 있다. 추천시스템은 고객의 구매 변화에 대한 즐거운 경험을 제공할 수 있어야 하고, 적절한 다양성 제공은 고객에게 취향 변화의 도움과 추천시스템 사용에 대한 만족감을 높일 수 있다. 추천시스템이 처음 개발되던 시점에서는 사용자의 첫 취향을 찾아내는 것에 초점이 맞춰져 있을 수밖에 없다. 하지만, 사용자의 구매가 반복적으로 이루어졌다면, 추천시스템은 사용자의 과거 기록을 바탕으로 사용자의 취향변화에 민감하게 반응하며 다양성을 시도해야 한다.

하지만, 추천시스템의 다양성을 위해서는 정확성에 대한 비용을 필요로 하는데, 이는 정확성과 다양성 사이에 트레이드 오프가 발생한다는 의미이다(Adomavicius et al., 2011). 추천의 다양성이 증가한다는 의미는 추천된 제품 간의 유사성이 감소했다는 의미이기 때문이다(Bradley et al., 2001). 따라서, 많은 연구에서 근소한 정확도 손실비용 혹은 동일한 정확도를 유지하면서 다양성을 극대화하는 방향은 어려운 작업 중 하나이다.

본 논문에서 고객의 다양성을 확인할 수 있는 데이터로 제품의 카테고리를 사용하였다. 카테고리는 상품에 대한 분류와 상세한 특징을 표현한 것이다. Amazon Grocery에서는 한 개의 상품이 4개 이상의 카테고리로 분류되어 있다. 예를 들어, A음료에 대해 'Beverages', 'Coffee, Tea & Cocoa', 'Coffee', 'Roasted Coffee Beans' 로 카테고리



〈Figure 1〉 User's preference change and recommender system

리가 분류되어 있다. 유저 B가 A음료를 자주 구매했다면, 단순히 ‘Beverages’ 단계만이 아니라 ‘Coffee’ 중에서도 ‘Roasted Coffee’를 선호한다는 것을 알 수 있다. 유저 B는 A음료에 대한 만족도가 높지만, Roasted Coffee 대신에 Non-Roasted Coffee를 원할 수도 있고 혹은 decaffeination 커피로 취향을 달리 바꿀 수도 있다. 또한 Coffee 대신에 Black Tea, Green Tea 등 새로운 차를 시도해보고자 할 수도 있다. 다양성 확보의 전제는 사용자 구매이력이 어느정도 쌓여 취향을 알 수 있는 상태이고, 과거 구매 카테고리를 활용하여 새로운 제품을 다양하게 추천하는 것이 가능하다.

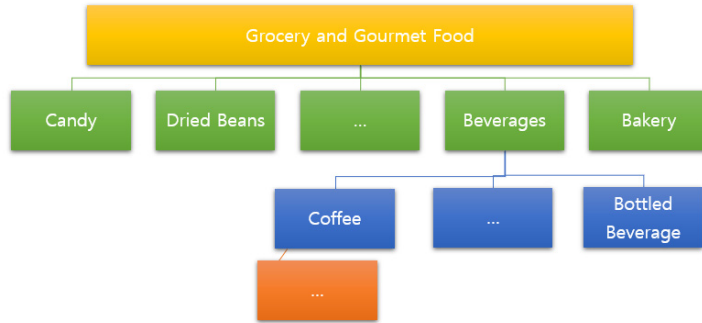
추천의 다양성을 반드시 카테고리만으로 실현할 수 있지는 않다. 연구자가 데이터를 얻을 수 있다면 사용자가 구매한 제품의 텍스트 설명, 사용자의 인구학적 정보 (성별, 거주도시, 수입소득, 결혼여부, 자녀여부, SNS사용 등) 역시 다른 유사도를 갖는 사용자를 찾는 데 도움이 될 것으로 본다. 다양성을 구현하기 위한 기준을 정하는데 중요한 점은 얼마나 상세하게 특징을 설명했는지 여부이다. 왜냐하면, 사용자 혹은 상품끼리의 이질성을 알기 위해서는 거시적인 차이점만

으로는 구분이 어렵기 때문이다. 예를 들어, 수입소득이라는 인구학적 정보로 다양성을 구현하려고 할 때, 단순히 수입이 많고 적음과 같은 이분적인 구분을 통해 이질성을 찾아내고자 한다면 정확한 결과를 얻기 어려울 것이다. 상세한 기준은 풍부한 다양성을 보장하여 보다 나은 결과를 얻게 한다.

본 연구에서는 아마존의 디테일한 카테고리의 분류를 사용하여 사용자의 취향을 분류하고, 취향이 다른 유저와 가깝게 연결되는 알고리즘을 통해 유저에게 익숙하지 않은 새로운 상품을 추천해주고자 시도하였다.

### 3.2 데이터

마케팅 측면에서 저관여제품은 고관여제품에 비해 선택의 주관성이 낮다(Radder et al., 2008). 저관여제품은 습관성 제품, 음식, 가정에서 사용하는 제품 등을 말한다. Amazon에서는 다양한 데이터를 오픈하고 있는데, Grocery and Gourmet Food는 저관여제품 데이터로서 추천시스템의 다양성을 적용하는데 적절하다고 판단하였다. 연구에서 사용한 데이터는 Amazon Grocery and Gourmet Food의 Review와 Meta Data이다. 전체



〈Figure 2〉 Category taxonomy

사용자의 수는 2,695,974명이며, 283,507개의 상품이 있다. Amazon Grocery and Gourmet Food는 [그림 2]과 같이 카테고리가 분류되어 있다. 가장 상위의 분류는 26가지, 두 번째는 127가지 등 밑으로 내려갈수록 상세하게 제품이 분류된다.

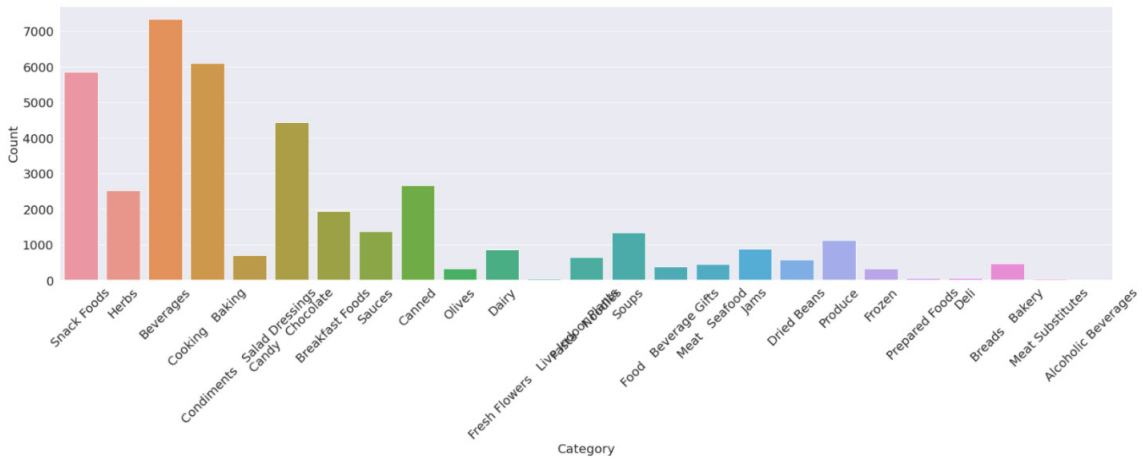
다양성 알고리즘을 적용하기 위해서는 사용자의 기존 구매가 어느정도 확보되어야 선취향계산이 가능하여 다양성 알고리즘에 합리적이다. 하지만 Amazon Grocery and Gourmet Food 데이터의 1%의 사람들이 10개의 제품을 구매하는 데 그쳤다. 10개의 제품 역시 동일한 제품 10개를 포함했기 때문에 구매 제품의 종류가 최소한 5개 이상 되는 사용자 0.4%를 대상으로 모델의 알고리즘을 적용하였다.

### 3.3 다양성 알고리즘

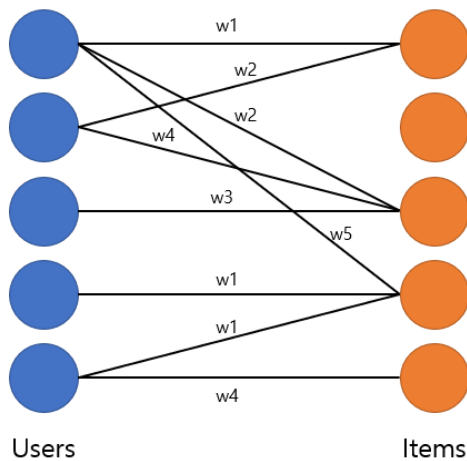
“명백함”에서 벗어나는 것은 우연성, 다양성을 볼 수 있는 하나의 방법이다. 평균적인 유사도를 낮출수록 다양성은 올라가게 된다(Charu, 2016). 협력필터링은 높은 유사도의 사용자 혹은 제품 사이에서 추천을 이룬다. 이러한 방식이 계속 유지되면 추천된 제품들이 다시 추천되는 현

상이 반복되어 강자가 더욱 강하게 되는 현상이 나타나게 된다(Fleder and Hosanagar, 2009). 사용자에게 새로운 제품을 보여주지 못하기 때문에 어느 지점에서 사용자는 진부함을 느끼기 쉽고, 기업 입장에서는 새로운 매출 가능성이 줄어들게 된다. Amazon Grocery and Gourmet Food의 판매 카테고리를 통해 이러한 불균형을 확인할 수 있다. <Figure 3>는 가장 최상위의 26개의 카테고리를 나타낸 것으로 구매편향이 확인된다. Beverages, Cooking Baking, Snack Foods는 Top 3 카테고리이며, Fresh Flowers, Meat Substitutes, Prepared Food는 비중이 거의 없다. 본 연구에서는 사용자 간의 카테고리 유사도를 최소화하여 추천의 다양성과 롱테일 법칙을 구현하였다.

그래프 베이스 추천시스템을 구현하기 위해서는 사용자와 상품 두 가지 집합으로 이뤄진 이분 그래프(Bipartite Graph)를 구성해야 한다. 이를  $B(U, I, E)$ 로 표현하며,  $U$ 는 사용자 집합,  $I$ 는 상품집합 그리고  $E$ 는 두 집합의 연결을 의미한다. 사용자와 상품은 과거 구매를 기반으로 무방향 연결되며 사용자가 준 평점은 두 연결의 연결강도( $w$ )를 결정한다. 이러한 Bipartite Graph는 추천 시스템을 가장 잘 표현하는 방법 중 하나로, 두



<Figure 3> Purchased category's bias



<Figure 4> Users and items graph,  $B(U, I, E)$

집합의 연결은 구매 뿐 아니라 클릭, 평점 등 다양한 요소를 고려하여 만들 수 있다. <Figure 4>는 이분 그래프를 표현한 것이다.

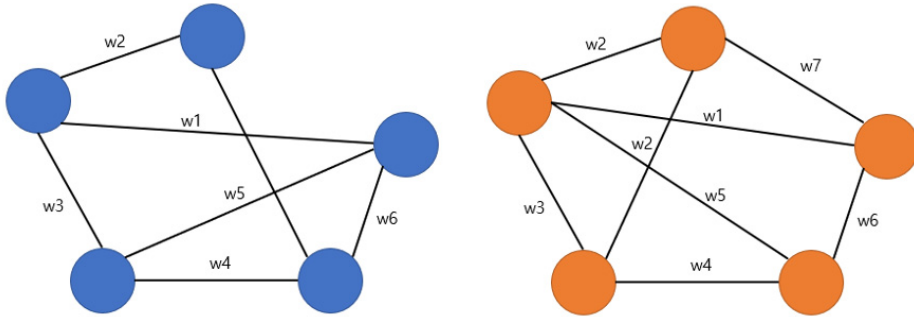
이분 그래프의 추천시스템은 사용자가 구매한 아이템과 비슷한 이웃을 찾아 제품을 추천해준다. 유사도는 사용자  $i$ 와  $j$  사이에서 얼마나 공통된 구매를 가졌는지에 따라 결정된다. 공통 구매한

제품이 많을수록 비슷한 이웃(CommonNeighbors)이며, 가장 유사한 이웃을 기반으로 아직 구매하지 않은 제품을 추천한다.

$$\text{CommonNeighbors}(i, j) = |S_i \cap S_j|$$

이러한 방법만을 사용한 그래프 추천시스템은 기존 협력필터링과 차이점이 없다. 따라서, 본 연구에서는 다양성 구현을 위해 사용자-아이템 이분 그래프만이 아니라 사용자와 상품만으로 네트워크가 구성된 일분 그래프 혹은 네트워크(Unipartite Graph)를 만들었다. 일분 그래프는 이분 그래프로부터 투사(Projection)하여 얻어낼 수 있으나, 무게중심에 다양성 알고리즘을 적용하기 위해 투사하여 얻지 않고 새롭게 만들었다. 사용자 일분 그래프는  $G(U, E)$ 이고, 상품 일분 그래프는  $G(I, E)$ 이다.

사용자 일분 그래프와 상품 일분 그래프는 무방향이며 두 그래프가 동일한 원리로 만들어진다. 공통으로 상품을 구매한 사용자들 혹은 공통으로 사용자를 가지고 있는 상품들을 연결하되,



〈Figure 5〉 Users graph G(U, E) and items graph G(I, E)

〈Table 1〉 Graph weight according to categories

	카테고리	Jaccard Distance(JD)	$G(i, j)_{weight}$
사용자 A	{음료, 커피, 스낵, 초콜릿, 사탕}	JD(A, B) = 4/10, JD(A, C) = 1/10	$G(A, B)_{weight} = 2.5$ $G(A, C)_{weight} = 10$
사용자 B	{음료, 커피, 녹차, 초콜릿, 사탕}	JD(B, A) = 4/10, JD(B, C) = 2/10	$G(B, A)_{weight} = 2.5$ $G(B, C)_{weight} = 5$
사용자 C	{음료, 녹차, 신선우유, 꽃, 통조림}	JD(C, A) = 1/10 JC(C, B) = 2/10	$G(C, A)_{weight} = 10$ $G(C, B)_{weight} = 5$

카테고리가 다들수록 강하게 연결되는 그래프 G(U, E)와 G(I, E)를 구성한다. 두 일분 그래프에서 무게중심은 다양성을 구현하는 결정적인 역할을 하는데, 자카드 거리를 변형하여 값을 얻는다. 자카드 거리(Jaccard Distance)는 유사도를 계산하기 위해 많이 사용하는 방식이다. 자카드 거리는 0에서 1사이의 값을 가지며 1에 가까울수록 공통된 원소가 많다는 의미이다.

$$\text{Jaccard Distance}(i, j) = \frac{|i \cap j|}{|i \cup j|}$$

무게중심(w)은 User A와 User B가 구매한 상품 카테고리들의 자카드 거리(Jaccard Distance)의 역수를 취하여 공통된 카테고리가 작을수록

무게중심이 높게 매겨지도록 한다.

$$G(i, j)_{weight} = \frac{1}{\text{Jaccard Distance}(i, j)}$$

예를 들면, G(U, E)에서 사용자 A, B, C는 셋 모두 공통구매한 상품들이 있고 네트워크에서 연결 되어있다. <표 1>과 같이 상품에 대한 카테고리들을 가졌을 때, 자카드 거리가 작을수록 무게중심은 높게 갖는다. 이렇게 얻어진 무게중심 값은 유사도가 낮은 사용자를 찾는데 역할을 한다.

사용자 일분 그래프를 통해서 두 가지 결과를 얻을 수 있다. 첫 번째는 사용자의 상대적 다양



성 선호 수준 값을 얻을 수 있고, 두 번째는 다양한 추천결과를 위해 유사도가 낮은 사용자를 찾을 수 있다. 일분 그래프에서 노드의 연결이 얼마나 많은가에 따라 각 노드의 중요도를 값으로 얻을 수 있는데, 얻어진 값을 나열하면 사용자의 다양성 선호 수준을 나눌 수 있어 사용자의 수준에 따른 모델 적용 결과를 비교할 수 있게 된다.

유사도가 낮은 사용자를 얻기 위해서 일분 그래프  $G(U, E)$  각 사용자를 중심으로 하는 에고-그래프(Ego-Graph)를 거쳐야 한다. 이 때 무게중심 임계치 기준값은 다양성의 수준을 조절할 수 있고, 무게중심을 크게 잡으면 구매 카테고리 차이가 큰 사용자 집합을 얻을 수 있다. 에고-그래프의 목적은 유사성이 낮은 사용자들의 집합을 얻기 위함이므로,  $n$ 개 이상의 사용자들을 얻을 수 있다. 그래프 네트워크의 특징 중 하나는 중심성을 활용하여 노드들의 특징을 확인할 수 있다는 점이다. 중심성, 매개중심성, 근접중심성 등이 있는데 그 중 고유벡터 중심성(Eigenvector Centrality)을 사용하여 에고 네트워크 중에서 영향력 있는 사용자를 얻는다.  $C_i$ 는 A의 고유벡터이며  $\lambda$ 는 고유값으로, 인접행렬에 값에 대한 고유값을 계산하면 고유벡터 중심성을 얻을 수 있다.

$$\lambda C_i = AC_i$$

고유벡터 중심성은 연결 중심성에 비해 중요 노드에 대한 가중치를 더하기 때문에 단순히 연결된 노드 숫자만으로 중심성을 계산하지 않는다는 장점이 있다.  $G(U, E)$ 로부터 무게중심이 가장 큰 에고-그래프를 얻었을 때, 사용자들의 중요도에 따라 값을 차별화시키는 작업이 필요하다. 고유벡터 중심성 값이 가장 큰 사용자는 가장 낮은 유사도를 얻게 됨으로 추천의 다양성을

위한 사용자를 얻어 추천결과를 가질 수 있다.

상품 일분 그래프  $G(I, E)$ 는 다양성을 극대화시키는 역할을 할 수 있다. 유사도가 낮은 사용자를 사용해  $B(U, I, E)$  추천시스템에서 다양성이 포함된 추천상품을 얻을 수 있으나 추천 결과와  $G(I, E)$ 를 결합하면 매우 큰 다양성을 얻을 수 있다. 예를 들어,  $B(U, I, E)$ 를 통해 얻은 추천 결과가  $\{A, B, C, D, E\}$  라면, 이 중 A를  $G(I, E)$ 와 결합했을 때 다시 다양성이 극대화된  $\{K, L, M, O\}$ 와 같은 추천항목을 얻을 수 있다. 모델은  $G(I, E)$ 와 결합한 결과들의 비율을 조절하면서 다양성과 정확도의 트레이드-오프를 해결할 수 있는 지점을 찾게 된다.

## 4. 평가방법

추천시스템이 처음 개발된 이후로 많은 추천시스템이 설계되었고, 설계자의 목표에 맞게 추천시스템이 구축되었는지 다양한 평가지표들도 나와있다. 설계자의 필요에 따라 측정지표가 결정되지만 설계자는 좋은 성능의 추천시스템을 얻기 위해서 다양한 기준으로 평가할 필요가 있다(Shani, 2011). 본 논문에서는 가장 널리 사용되는 평가지표인 정확도와 다양성을 목표로 설계된 추천시스템 평가를 위한 다양성 지표로 평가하였다.

### 4.1 정확도

정확도 측정을 위해서 사용자의 평점을 예측하는 방식도 있지만, 새로운 상품을 추천할 때 사용자가 제품을 선택할지 여부를 예측하는 방식도 있다. 본 논문에서는 구매한 제품들 중에서

〈Table 2〉 Evaluation of recommender system

	추천상품	추천하지 않은 상품
실제 구매	True-Positive (TP)	False-Negative (FN)
실제 구매하지 않음	False-Positive (FP)	True-Negative (TN)



〈Figure 6〉 User's preference level and inflection point

추천한 제품들이 얼마나 포함되어 있는지를 확인하는 Hit Rate로 정확도를 평가하였다.

$$\text{정확도} = \frac{n(TP)}{n(TP) + n(FN)}$$

#### 4.2 다양성

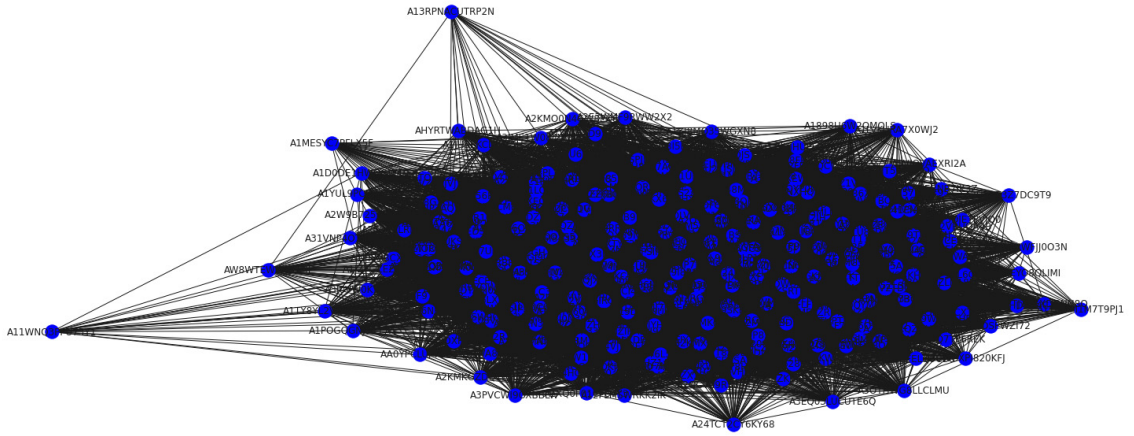
추천된 결과의 카테고리의 다양성을 확인하기 위해서 개인별 추천된 아이템들의 카테고리가 기존 구매한 카테고리보다 얼마나 더 다양해졌는지 확인해야 한다. 본 연구에서는 카테고리의 다양성 수준이 얼마나 넓어졌는지 알기 위해 다양성 지표를 개발하였다. 다양성 지표는 사용자  $i$ 가 구매한 총 카테고리  $w$ 에  $i$ 에게 추천된 카테고리와의 공통집합이 작을수록 카테고리 다양성이 늘어났음을 의미한다.

$$\text{다양성 지표} = 1 - \frac{\sum i_c \cap \sum i_{r,c}}{\sum i_c}$$

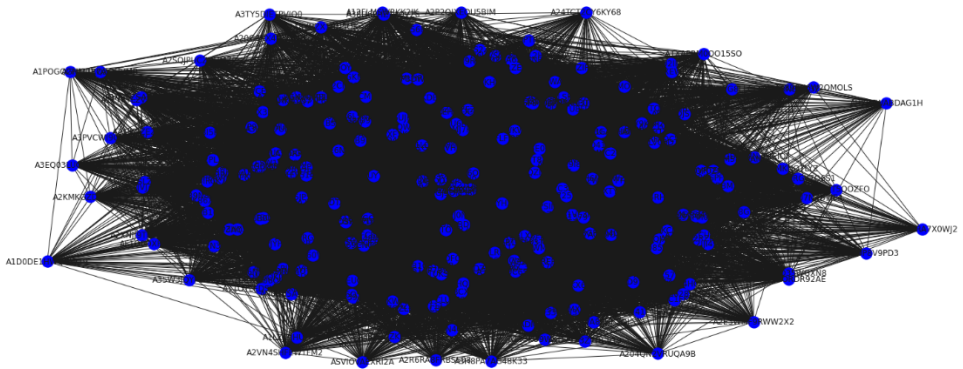
### 5. 연구 결과

본 연구에서 중요한 역할을 하는 사용자 일부 네트워크의 두 가지 기능은 사용자의 다양성 수준 수치화와 낮은 유사도의 사용자를 얻게 하는 것이다. 〈Figure 7〉는 사용자들 네트워크이다. 네트워크 사용자들의 정도 중심성을 구하고 값의 변화율을 바탕으로 변곡점이 되는 지점을 찾을 수 있다. 〈Figure 6〉는 다양성 수준이 변화되는 지점이 검은색 선으로 표시되어 있는데 해당 사용자부터 다양성 수준이 넓어지고 있음을 알 수 있다.

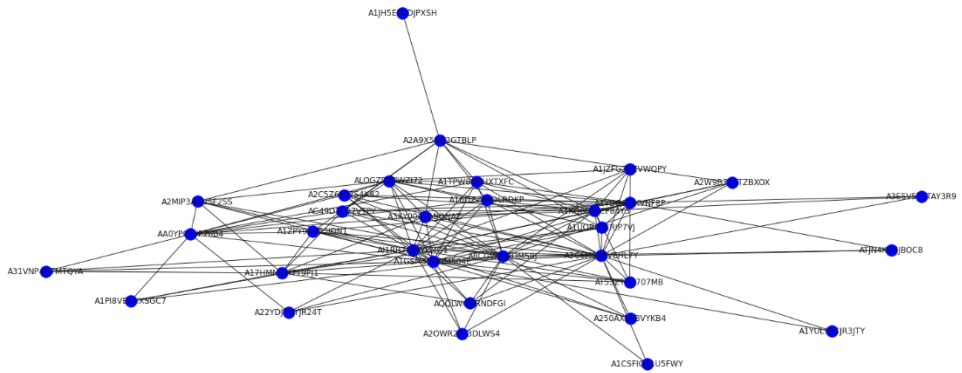
다양성 추천시스템 구축을 위해 사용자 일부 네트워크를 만들고 유사도가 낮은 사용자를 얻기까지의 네트워크 변형 과정을 시각적으로 확인할 수 있다. 중심으로 들어갈수록 구매가 비슷한 사용자들이 분포되어 있고, 외곽으로 나갈수록 구매한 상품이 다른 사용자들로 구성된다. 사용자 그래프에서는 한 사용자의 에고-그래프를



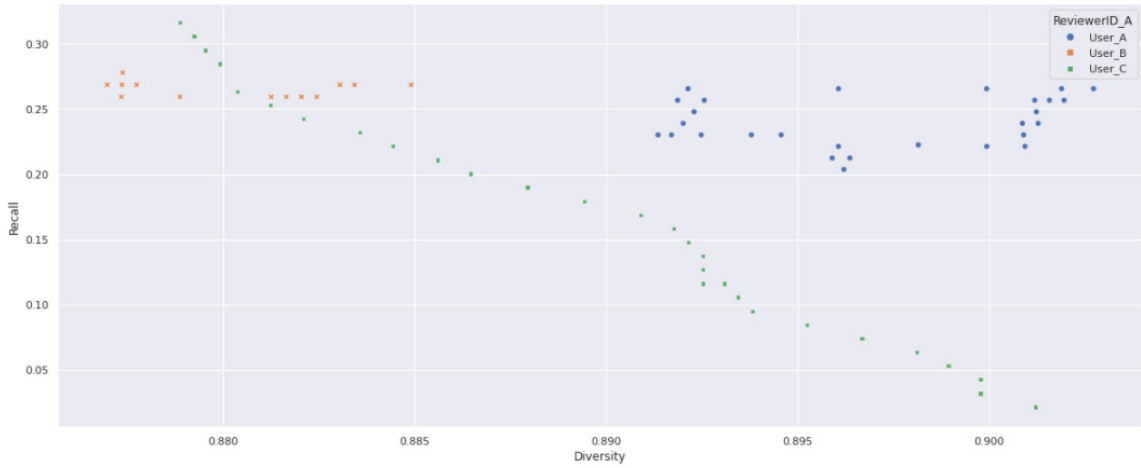
〈Figure 7〉 User's network



〈Figure 8〉 Ego-network of User A



〈Figure 9〉 Ego-network with maximum weight of User A



<Figure 10> The accuracy according to diversity level

얻어야 한다. <Figure 8>는 특정 사용자 A와 직접 연결된 네트워크로만 구성된 에고-그래프이며, <Figure 7>에 있는 가장 외곽의 사용자가 사라진 상태를 보여준다. 에고-네트워크에서 무계 중심의 기준값(threshold)을 변형하면 유사도가 낮은 사용자들을 얻을 수 있다. <Figure 9>은 최대의 무계중심 값을 가졌을 때 구성된 네트워크이다. 즉, 유사도가 매우 달라진 사용자들의 네트워크가 된 것이다.

최대의 무계중심 값을 가진 에고 네트워크로부터 유사도가 낮은 사용자를 얻어 B(U, I, E) 추천모델에 넣었을 때, 일반 모델과 달리 다양성이 큰 상품들을 추천하게 된다. 기본 모델과 다양성 모델의 평균 유사도 차이는 <표 3>과 같다.

<Table 3> Difference of similarity

	다양성 모델	일반 모델
유사도	0.0074	0.0434

다양성을 위한 그래프 기반의 추천시스템은

G(I, E)를 통해 다양성의 수준을 조절할 수 있다. 다양성 수준을 강하게 하고 싶을 때는 G(I, E)로부터 얻은 상품의 비율을 늘리는 것이 가능하다. 다양성 수준을 최대로 고정시켰을 때, 일반 모델과의 결과 비교는 <표 4>와 같다.

<Table 4> The highest diversity level and recommendation results

	Top-30		Top-20	
	정확도	다양성 지표	정확도	다양성 지표
그래프 기반 추천시스템 모델	0.2561	0.4175	0.1674	0.3890
그래프 기반 다양성 추천시스템 모델	0.2078	0.5969	0.1537	0.5451

<Figure 10>에서 확인할 수 있는 것처럼 G(I, E)의 추천항목의 비율에 따라 다양성과 정확도의 트레이드 오프 지점이 발생하는 것을 알 수 있다. 다양성 수준이 클 때, 정확도는 낮은 수준으로 떨어진다. 근소한 손실의 정확도 혹은 동일

<Table 5> Optimized diversity level and recommendation results

	Top-30		Top-20	
	정확도	다양성 지표	정확도	다양성 지표
그래프 기반 추천시스템 모델	0.2561	0.4175	0.1674	0.3890
그래프 기반 다양성 추천시스템 모델	0.2415	0.5111	0.1625	0.4362

<Table 6> Recommendation results according to user's diversity level

Top-30	강한 다양성 선호자		상대적으로 약한 다양성 선호자	
	정확도	다양성 지표	정확도	다양성 지표
그래프 기반 추천시스템 모델	0.2511	0.4068	0.2672	0.4408
그래프 기반 다양성 추천시스템 모델	0.2378	0.4410	0.2479	0.6371

한 정확도를 유지하면서 높은 다양성 지표가 나오는 지점이 최적화된 결과라 말할 수 있다. 최적화시킨 결과는 <표 5>와 같다. 다양성 값은 떨어지지만, 정확도가 상승하는 것을 알 수 있다. (Adomavicius and Kwon , 2011)에 따르면, 정확도가 0.1% 하락할 때 다양성이 20~25% 상승하면 좋은 결과치를 얻은 것이라 말한다. 다양성이 최적화된 모델은 정확도가 Top-30에서 0.06% 하락하였을 때, 22.42% 증가하였으므로 기본 모델에 비해 정확도와 다양성 모두 좋은 결과를 얻은 것이라 말할 수 있다.

G(U, E)로부터 사용자의 다양성 수준을 나눌 수 있었는데, <Figure 6>의 변곡점을 기준으로 오른쪽에 위치한 사용자들은 강한 다양성 선호자로, 변곡점 지점에 있는 사용자들은 상대적으로 약한 다양성 선호자로 분류하였다. <표 6>에 따르면, 다양성이 상대적으로 약한 집단이 정확도가 약 0.08% 하락할 때, 다양성 지표가 약 44.5% 상승하는 결과를 통해 다양성이 상대적으로 약한 사용자들에게 큰 성과를 나타내는 것으로 확인되었다.

## 6. 결론과 앞으로의 연구방향

사용자가 더욱 늘어나는 E-Commerce 환경에서 추천시스템은 사용자의 만족도를 좌우하는 중요한 서비스가 되고 있다. 추천시스템이 개발된 이래 만족도를 상승시키기 위한 연구가 꾸준히 하게 이뤄지고 있다.

본 연구에서는 그래프 기반의 추천시스템을 구축하여 추천의 다양성을 구현하였으며, 사용자의 다양성 선호 수준을 알 수 있었다. 사용자 일부 그래프를 통해 사용자의 다양성 선호가 상대적으로 다르며 이를 개인수준의 값으로 알 수 있었고, 본 연구에서 개발한 알고리즘을 통해 다양성을 위한 이질적인 사용자를 얻을 수 있었다. 본 연구 모델의 결과 정확도의 근소한 손실은 있었지만 높은 추천의 다양성을 얻을 수 있었고 연구 모델에서 다양성 컨트롤 지점을 활용하여 최적화를 시킬 수 있다는 결과를 알 수 있었다.

본 연구는 다음과 같은 이론적 공헌 세 가지를 갖는다. 첫째, 사용자 만족을 위한 추천시스템 연구를 확장시켰다. 인공지능 시대를 맞이하며

딥러닝, 강화학습 등 다양한 기술이 빠르게 발전하며 추천시스템으로의 적용도 이뤄지고 있다. 하지만, 많은 연구에서 기술 접목의 평가 기준으로 정확도를 선택 하고 있다. 추천시스템은 사용자로부터 평가받는 서비스인 만큼, 발전하는 기술의 접목 만이 아니라 사용자의 만족도를 향상시킬 수 있는 여러 지표 향상을 위한 연구가 필요하다. 그러한 점에서 본 연구는 추천시스템의 사용자 만족도를 상승시키기 위한 정확성 및 다양성 가치를 올리는 데 기여했다는 의의가 있다. 둘째, 그래프 방식의 추천시스템 구현 발전이다. 널리 알려진 추천시스템의 방식으로 협업필터링이 있다. 하지만, 최근 협업필터링의 약점이 지적되며 새로운 접근 방식을 필요로 했다. 그래프 기반의 추천시스템은 협업필터링의 개념과 유사하지만 다른 방법론과 결합이 유연하다는 점에서 발전 가능성이 크다. 또한, 사용자 혹은 상품과의 관계를 시각화시켜 결과를 확인할 수 있다는 장점을 가지고 있어 모델의 기능을 상승시키는 데 도움이 된다. 셋째, 다양성 평가 지표 개발이다. 추천시스템을 발전시키기 위한 다양한 알고리즘 변형이 연구되고 있다. 평가 지표는 추천시스템을 비교 평가할 수 있는 잣대로서 중요한 역할을 한다. 다양성 지표는 추천시스템의 다양성을 평가할 수 있는데, 다양성 정의와 어긋나거나 잘못 설정되어졌을 때 추천시스템을 평가하기 어렵다. 본 연구에서는 추천의 다양성을 카테고리 사용하며 추천시스템 알고리즘을 구현하였고 이에 맞는 지표를 개발하여 연구의 타당성을 높였다.

기업에서 추천시스템을 적극적으로 활용하고 있는 만큼 본 연구의 실무적 공헌을 꼽을 수 있다. 우선, 기업의 장기적 이익을 높이는 데 기여할 수 있다. 추천시스템의 다양성은 사용자가 평

소에 관심을 갖지 않은 범위로까지 클릭 혹은 구매를 확장시킬 수 있다. 따라서, 단기적으로 봤을 때는 정확도가 낮은 서비스로 보일 수 있지만, 장기적으로 봤을 때 베스트 셀러 만이 아니라 롱테일 지점의 상품들도 가치를 올리게 하여 기업의 이윤을 높일 수 있다. 또한, 사용자의 다양성 수준에 맞춰 서비스를 제공할 수 있다. 추천의 다양성을 시도할 때, 사용자가 추천의 다양성을 어떻게 받아들일지에 대한 의문이 있다. 추천의 다양성이 효과적으로 적용되기 위해서는 사용자가 추천의 다양성을 선호해야 한다. 본 연구를 통해 사용자의 다양성 선호 수준을 알 수 있고, 이에 맞춰 추천 서비스를 선택적으로 제공할 수 있기 때문에 서비스를 처음 시도할 때 리스크를 줄일 수 있다. 서비스를 처음 적용할 때 기업에서 A/B테스트 등 다양한 실험을 하지만 그럼에도 어느 정도의 리스크를 갖기 마련이다. 본 연구에서 사용자 다양성 수준을 분류한 것은 이러한 초기 리스크 비용을 감소시킬 수 있다는 점에서 실무적 공헌이 있다고 할 수 있다. 마지막으로, 본 연구의 활용이 저관여 제품을 다루는 분야에서 효과를 크게 볼 수 있을 것이라 본다. 다양성 발견이 제품의 성질에 영향을 미칠 것이라 판단하여 본 연구에서 식품 데이터를 사용하였다. 사용자가 냉장고, TV, 자동차 등 고관여 제품보다는 생활상품에서 다양한 시도를 쉽게 할 수 있는 만큼 본 연구를 해당 분야에 적절하게 적용했을 때 나은 결과를 얻을 수 있을 것이라 여긴다.

본 연구는 정확도에서 근소한 손실의 트레이드 오프였으나, 완벽히 극복하지 못했다는 한계점 역시 가지고 있다. 최적화 및 모델의 변형을 통해 현재보다 더 높은 정확도를 얻는 지점을 발견한다면 트레이드 오프를 극복한 중요한 연구

가 될 것이라 생각한다. 또한, 이번 연구 결과를 바탕으로 다양성을 강하게 선호하는 사람들에게 대해서는 더 나은 결과를 위한 알고리즘 적용이 필요해 보인다. 따라서, 향후 연구에서 저관여 제품만이 아니라 여러 분야의 데이터를 바탕으로 정확도와 다양성의 트레이드 오프를 극복하고 각기 사용자 수준에 맞는 알고리즘들을 발전시키는 연구를 하고자 한다.

## 참고문헌(References)

- Adomavicius, G., Y. Kwon, "Improving Aggregate Recommendation Diversity Using Ranking-Based Techniques", *IEEE Transactions on Knowledge and Data Engineering*, VOL.24, NO.5(2011)
- Andrea Barraza-Urbina, Benjamin Heitmann, Conor Hayes, and Angela Carrillo-Ramos, "XPLODIV: An exploitation-exploration aware diversification approach for recommender systems" *the 28th International Florida Artificial Intelligence Research Society Conference*, Florida, 2015
- Barbara E Kahn, "Consumer variety-seeking among goods and services", *Journal of Retailing and Consumer Services*, Vol. 2, No 3(1995), 139~148
- Barry Smyth and Paul McClave. "Case-Based Reasoning Research and Development", *Springer*, vol 2080(2001), 347 - 361
- S. Miyamoto, T. Zamami and H. Yamana, "Improving Recommendation Diversity Across Users by Reducing Frequently Recommended Items," *IEEE International Conference on Big Data (Big Data)*, Seattle, WA, USA, 2018, 5392~5394
- Charu C. Aggarwal, *Recommender Systems*, Springer International Publishing Switzerland, 2016
- Cai-Nicolas Ziegler, Sean M. McNee, Joseph A. Konstan, and Georg Lausen, "Improving recommendation lists through topic diversification", *In Proceedings of the 14th International Conference on World Wide Web*, (2005), 22~32
- Daniel Billsus and Michael J. Pazzani. *User Modeling and User-Adapted Interaction 10*, Kluwer Academic, (2000), 147~180.
- Dokyun Lee, Kartik Hosanagar, "How Do Recommender Systems Affect Sales Diversity? A Cross-Category Investigation via Randomized Field Experiment", *Information Systems Research*, Vol. 30, No. 1(2019), 239~259.
- Fleder D, Hosanagar K, "Blockbuster culture's next rise or fall: The impact of recommender systems on sales diversity" *Information Systems Research*, Vol 55, No. 5, 697~712.
- Goldberg, K., Roeder, T., Gupta, D., and Perkins, C, "Eigentaste : A constant time collaborative filtering algorithm", *Information Retrieval*, (2001), 133-151.
- Itamar Simonson, "The effect of purchasing Quantity and Timing on Variety-Seeking Behavior", *Journal of Marketing Research*, Vol 27, No. 2(1990)
- Jieun Son, Seoung Bum Kim, Hyunjoong Kim, Sungzoon Cho, "Review and Analysis of Recommender Systems", *Journal of the Korean Institute of Industrial Engineers*, Vol. 41, No. 2(2015), 185~208
- John Paul Kelly and Derek Bridge. 2006. Enhancing the diversity of conversational collaborative recommendations: A comparison. *Artificial Intelligence Review* 25, (2006), 79~95.

- Lei shi, "Trading-off Among Accuracy, Similarity, Diversity, and Long-tail: A Graph-based Recommendation Approach", *Proceedings of the 7th ACM conference on Recommender systems*, 2013, 57~64
- Mouzhi Ge, Carla Amor Dvino Moreira Delgado, Dietmar Jannach, "Beyond accuracy: Evaluating recommender systems by coverage and serendipity", *Proceedings of the ACM Conference on Recommender Systems*, Barcelona, Spain, 2010
- McNee, S.M., J. Riedl, J.A. Konstan. "Being Accurate is Not Enough: How Accuracy Metrics have hurt Recommender Systems". *Conf. on Human Factors in Computing Systems*, 2006, 1097~1101.
- Maslow, A. H., "A theory of human motivation. Psychological Review", *Psychological Review*, Vol 50, No. 4(1943), 370~396.
- Marius Kaminskas and Derek Bridge. "Diversity, serendipity, novelty, and coverage: A survey and empirical analysis of beyond-accuracy objectives in recommender systems". *ACM Transactions on Interactive Intelligent Systems*, Vol 7, No. 1(2016)
- Moshe Givon, "Variety Seeking Through Brand Switching", *Marketing Science*, Vol 3. No. 1(1984)
- Mooney RJ, Roy L, "Content-based book recommending using learning for text categorization", *Proceedings of the fifth ACM conference on Digital libraries*, (2001), 195~204
- Mi Zhang and Neil Hurley, "Avoiding monotony: Improving the diversity of recommendation lists", *In Proceedings of the ACM Conference on Recommender Systems*, (2008).
- Paul Rsnick, Neophytos Iacovou, Mitesch Suchak, Peter Bergstrom, John Riedl, "GroupLens: an open architecture for collaborative filtering of netnews", *Proceedings of the 1994 ACM conference on Computer supported cooperative work*, (1994), 175~186
- P Adamopoulos, A Tuzhilin, "On unexpectedness in recommender systems: Or how to better expect the unexpected", *ACM Transactions on Intelligent Systems and Technology (TIST)*, Vol 5, No. 4(2014)
- Radder, Laetitia; Huang, Wei, "High-involvement and low-involvement products : a comparison of brand awareness among students at a South African university", *Journal of Fashion Marketing and Management*, Vol. 12, No. 2(2008), 232~243.
- Sa' ul Vargas and Pablo Castells, "Rank and relevance in novelty and diversity metrics for recommender systems", *In Proceedings of the 5th ACM Conference on Recommender Systems*. (2011), 109~116.
- Schafer JB, Konstan J, Riedl J, "Recommender systems in ecommerce". *1st ACM Conf. Electronic Commerce*, (1999), 158~166.
- Shani G., Gunawardana A, "Evaluating Recommendation Systems", Springer, Boston, 2011.
- 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". *Proceedings of the National Academy of Sciences*, Vol 107, No. 10(2010).
- Will Hill, Larry Stead, Mark Rosenstein, George Furnas, "Recommending and evaluating choices in a virtual community of use", *Proceedings*



*of the SIGCHI Conference on Human Factors in Computing Systems*, (1995), 194~201

Yuan Cao Zhang, Diarmuid O' Se'aghdha, Daniele Quercia, and Tamas Jambor. "Auralist: Introducing serendipity into music recommendation", *Proceedings of the 5th ACM International Conference on Web Search and Data Mining*, (2012).

Yehuda Koren, "Collaborative Filtering with

Temporal Dynamics", *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, (2009), 447~456.

Zan Huang, Wingyan Chung, Thian-Huat Ong, Hsinchun Chen, "A Graph-based Recommender System for Digital Library", *Proceedings of the Second ACM/IEEE-CS Joint Conference on Digital Libraries*, (2002), 65-73,

Abstract

## Application of diversity of recommender system according to user preference change

Hyeyeon Na\* · Kihwan Nam\*\*

Recommender Systems have been huge influence users and business more and more. Recently the importance of E-commerce has been reached rapid growth greatly in world-wide COVID-19 pandemic. Recommender system is the center of E-commerce lively. Top ranked E-commerce managers mentioned that recommender systems have a major influence on customer's purchase such as about 50% of Netflix, Amazon sales from their recommender systems. Most algorithms have been focused on improving accuracy of recommender system regardless of novelty, diversity, serendipity etc. Recommender systems with only high accuracy cannot satisfy business long-term profit because of generating sales polarization. In addition, customers do not experience enjoyment of shopping from only focusing accuracy recommender system because customer's preference is changed constantly. Therefore, recommender systems with various values need to be developed for user's high satisfaction.

Reranking is the most useful methodology to realize diversity of recommender system. In this paper, diversity of recommender system is represented through constructing high similarity with users who have different preference using each user's purchased item's category algorithm. It is distinguished from past research approach which is changing the algorithm of recommender system without user's diversity preference level. We tried to discover user's diversity preference level and observed the results how the effect was different according to user's diversity preference level. In addition, graph-based recommender system was used to show diversity through user's network, not collaborative filtering.

In this paper, Amazon Grocery and Gourmet Food data was used because the low-involvement product, such as habitual product, foods, low-priced goods etc., had high probability to show customer's diversity. First, a bipartite graph with users and items simultaneously is constructed to make graph-based recommender system. However, each users and items unipartite graph also need to be established to show

---

\* Graduate School of Interdisciplinary Management, Ulsan National Institute Science and Technology (UNIST)

\*\* Corresponding author: Kihwan Nam

Management Information Systems, Business School, Dongguk University

30, Pildong-ro, 1-gil, Jung-gu, Seoul, Korea

Tel: +82-2-2260-3684, Fax: +82-2-2260-3684, E-mail: namkh@dongguk.edu

diversity of recommender system. The weight of each unipartite graph has played crucial role changing Jaccard Distance of item's category. We can observe two important results from the user's unipartite network. First, the user's diversity preference level is observed from the network and second, dissimilar users can be discovered in the user's network. Through the research process, diversity of recommender system is presented highly with small accuracy loss and optimalization for higher accuracy is possible controlling diversity ratio.

This paper has three important theoretical points. First, this research expands recommender system research for user's satisfaction with various values. Second, the graph-based recommender system is developed newly. Third, the evaluation indicator of diversity is made for diversity. In addition, recommender systems are useful for corporate profit practically and this paper has contribution on business closely. Above all, business long-term profit can be improved using recommender system with diversity and the recommender system can provide right service according to user's diversity level. Lastly, the corporate selling low-involvement products have great effect based on the results.

**Key Words** : Recommender System, Decision Making Support System, Diversity, E-Business

Received : November 13, 2020 Revised : December 10, 2020 Accepted : December 16, 2020

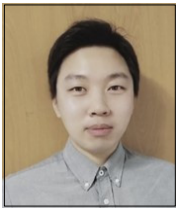
Corresponding Author : Kihwan Nam

## 저 자 소개



**나혜연**

UNIST 융합경영대학원 비즈니스 분석 석사과정에 있다. 주요 연구 관심 분야는 추천시스템 개발과 행동경제학, 빅데이터 분석, 딥러닝, 인공지능이다.



**남기환**

KAIST 경영대학원 경영공학부에서 MIS 박사학위를 취득하였다. 현재 동국대학교 경영대학 경영정보학과 조교수로 재직 중이다. 주요 관심분야는 Business Analytics & Business Intelligence, Big Data Analytics, Data Mining, Statistical Analysis, Recommender Systems, Econometrics Models, Machine Learning, Deep Learning 등 이다. 관련 연구들은 Decision Support Systems, Data Mining and Knowledge Discovery 등에 논문이 게재되었다. 학문적인 연구뿐만 아니라 이론을 바탕으로 실제 기업에서 다양한 프로젝트를 성공적으로 진행함으로써 학계와 산업계 모두에 실증적인 기여를 하고 있다.