

# OurTube: 소셜 및 카테고리 선호도 기반 유튜브 채널 추천 시스템\*

김의석<sup>01</sup>, 박상근<sup>2</sup>

<sup>1</sup>경희대학교 산업경영공학과

<sup>2</sup>경희대학교 소프트웨어융합학과

uiseok227@gmail.com, sk.park@khu.ac.kr

## OurTube: Social and Category Based YouTube Video Recommendation System

UiSeok Kim<sup>01</sup>, Sangkuen Park<sup>2</sup>

<sup>1</sup>Department of Industrial and Management Systems Engineering, Kyung Hee University

<sup>2</sup>Department of Software Convergence, Kyung Hee University

### 요 약

방대한 양의 영상이 유튜브에 축적되면서 개인의 선호에 맞는 개인화된 추천 시스템을 제공하고 있다. 그러나 이러한 개인 맞춤 추천은 사용자의 다른 분야로의 접근을 방해하고 선호도에 기반한 제한된 분야의 영상만 접하게 한다는 단점이 있다. 사용자 만족도를 높이기 위한 다양한 개인화 영상 추천 연구가 진행되었지만, 영상 추천에서 개인의 소셜 관계의 중요성은 크게 주목받지 못했다. 본 논문에서는 개인의 소셜 관계에 기반한 영상 채널 추천 시스템을 개발하였으며, 사용성 평가를 위해 그 효과를 검증하였다.

## 1. 서 론

온라인에 영상을 업로드하고 타인과의 공유가 쉬워지면서, 전 세계적으로 1인 크리에이터들 및 인플루언서가 인기를 끌고 있고, 다양한 영상이 인기를 끌고 있다. 유튜브에서는 다양한 요소를 고려한 추천 알고리즘을 통해 사용자가 시청할만한 추천 영상을 제공하고 있다. 이러한 개인화 추천은 사용자가 관심 있어 하는 분야를 계속해서 볼 수 있다는 장점을 갖지만 보다 다양한 분야를 접할 기회를 잃는 필터버블 문제를 발생시킬 수 있다 [1,2]. 이 문제를 해결하기 위해, 본 연구에서는 개인의 소셜 관계를 활용한 영상 채널 추천을 제안한다.

본 연구는 SNS 속에서 추천 알고리즘의 개선점을 찾았다. 현대인들은 SNS 속에서 관계를 맺고 있는 사람들의 사진 혹은 글을 보며 그들의 모습과 특성을 살펴보거나 새로운 분야에 대해 흥미를 가지기도 한다. 이런 소셜 관계를 맺는 기능을 서비스에 추가했고, 다른 친구들의 구독 목록을 살펴볼 수 있게 하여 자연스럽게 사용자들이 구독 채널을 공유하고 새롭게 알아갈 수 있도록 했다. 그리고 소셜 및 사용자 선호도 기반의 추천과 새로운 분야의 확장을 돕는 추천을 제공하여 개인맞춤형 추천의 문제를 개선했다.

## 2. 기존 연구

### 2.1 추천 알고리즘

개인에게 맞춤형 정보를 제공하기 위한 추천 시스템의 연구가 활발하게 진행되어 왔다. 사용자의 과거 행동 기반으로 사용자와 아이템의 상호 관계를 분석하는 Content Based Filtering을 시작으로, 유사한 아이템에 대해 비슷한 선호도를 보이는 다른 사용자를 찾아서 아이템에 대한 사용자의 선호도를 예측하는 Memory-based Collaborative Filtering의 알고리즘이 개발되었다. 그리고 이를 혼합한 Hybrid Collaborative Filtering 방법이 개발되는 등, 2000년대에는 다양한 추천시스템의 연구가 눈에 띄게 증가하였다 [3]. 추천 시스템에 대한 연구가 많아지면서 피어슨 상관관계수(pearson correlation coefficient)나 코사인 벡터(cosine vector)를 주로 이용했던 Memory-based Collaborative Filtering의 연구[4, 5]와 다르게 새로운 유사도 개념을 접목시킨 추천 연구가 등장했다. 개선된 유사도를 통해 추천을 제공했던 [6]의 연구에서는 Personal Rank를 유사도에 접목시켜 추천의 정확도를 높이기도 했다. 본 논문의 연구 또한 보다 개선된 추천을 제공하기 위해 새로운 유사도의 개념을 제공하며 Hybrid 형식의 추천을 제공한다.

### 2.2 소셜 기반 추천 시스템

미디어 서비스에서의 추천 알고리즘에 문제를 제기하며 다른 각도로 개선시키려는 연구는 지속되어 왔다. 소셜 네트워크 사용자 기반의 클러스터링 추천 시스템[7]의 연구는 사용자의 과거 패턴뿐만 아니라 사용자의 온·오프라인에서 관계를 맺고 있는 다른 사용자의 행동 패턴을 반영하여 추천시스템의 성능을 높일 것을 제안했다. OtherTube [8]는 소셜 미디어 서비스의

\* “본연구는 과학기술정보통신부 및 정보통신기획평가원의 2023년도 SW중심대학사업의 결과로 수행되었음”(2023-0-00042)

추천 알고리즘이 사용자의 과거 기록을 기반으로 맞춤형 추천을 하기에 사용자가 보다 다양한 콘텐츠를 접할 기회를 잃고 제한된 콘텐츠만 소비하는 필터버블 현상[1,2]이 발생할 수 있음을 지적하며 대표적인 소셜 서비스인 YouTube의 필터버블 현상을 해결할 수 있는 새로운 방안을 제시했다. 무작위의 다른 사용자의 성별, 위치 정보, 취미를 확인할 수 있는 프로필과 함께 그 사용자에게 추천되는 유튜브 홈 화면의 영상을 제공하는 확장 프로그램을 만들어 추천의 다양성을 높였다. [9]의 연구 또한 유튜브 API를 통해 사용자의 구독 목록, 채널 카테고리를 가져와 소셜 기반의 추천 시스템을 제안했다.

그러나 기존의 연구는 실제 사용자가 관계를 맺고 있는 사람들의 구독 목록을 살펴볼 수 없다는 한계를 가졌다. 또한 직접 사용자들을 확보하지 못했기에 Memory-based Collaborative Filtering 추천 시스템의 사용자 만족도를 살펴볼 수 없었다. 본 논문에서는 이런 선행 연구를 바탕으로 소셜 기반의 친구 시스템을 제공해 오프라인 및 온라인에서 관계를 맺고 있는 친구들의 구독 정보를 볼 수 있고 채널 추천을 받을 수 있는 OurTube 시스템을 개발하고, 사용자 평가를 통해 그 효과를 검증한다.

### 3. 추천 알고리즘 구현

#### 3.1 채널 카테고리 설정

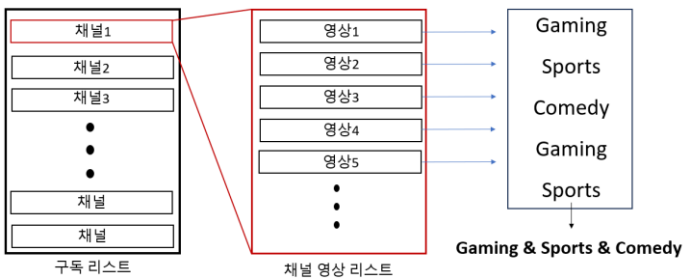


그림 1. 채널 카테고리 선정 알고리즘

사용자의 선호 카테고리를 기반으로 채널을 추천하기 위해 채널 별 카테고리 파악이 필요하다. YouTube Data API v3의 Subscription 리소스를 이용하여 사용자의 구글 고유 식별 ID를 통해 사용자 구독 채널 정보를 가져왔다. 하지만 YouTube Data API v3에서 채널 별 카테고리 정보를 직접 제공하고 있지는 않기 때문에 별도의 알고리즘으로 채널 별 카테고리를 지정했다.

Playlist 리소스를 활용해 사용자 구독 리스트에 있는 채널의 영상 리스트에 접근하고, Video 리소스를 활용하여 이 영상 리스트에 있는 영상들의 카테고리를 파악했다. 이때, 카테고리는 유튜브 공식 홈페이지에 나와있는 카테고리 분류를 기준으로 설정했다. 각 채널별로 가장 최근 영상 5개의 카테고리를 찾아 [그림 1]과 같이 채널의 카테고리를 설정해 주었다.

#### 3.2 추천 알고리즘

본 연구는 협업 필터링 방식과 콘텐츠 기반 필터링 방식을 모두 이용하면서, 보다 높은 정확도의 추천을 제공하기 위해 가중치 개념을 도입했다. 채널 추천을 위해 사용자가 구독하지 않는 DB에 있는 모든 채널 목록을 가져오고, 초기 추천 가중치를 모두 0으로 설정한다. 그리고 사용자의 구독 현황 파악을 위해 사용자의 구독 카테고리 비율을 가져온다. 이때 구독 카테고리

비율은 사용자가 구독하는 카테고리 별 채널 개수를 각각 전체 구독 채널 개수로 나누는 방식으로 계산한다.

#### 3.2.1 협업 필터링

협업 필터링 방식은 서비스를 이용하는 모든 사용자의 구독 채널 목록과 구독 카테고리 비율을 가져와서 추천을 받을 사용자와 전체 사용자의 구독 카테고리 비율의 유사도를 통해 가중치를 더해준다.

이때 사용자와 다른 사용자 간의 유사도 측정을 위해서는 공통적으로 구독하는 카테고리가 얼마나 있는지 그리고 그 카테고리 별 비율이 수치적으로 얼마나 비슷한지 반영이 필요하다. 이를 위해 크기와 방향을 모두 고려할 수 있도록 Jaccard유사도와 MSD유사도를 합친  $(Jaccard(x, y) * (1 - msd(x, y)))$ 의 유사도를 활용했다 [10]. 이 유사도를 통해 사용자와 다른 사용자의 유사도를 구해줬고 이를 가중치에 더해줬다. 그리고 Min-Max Scaling을 진행해 1차적으로 채널별로 가중치를 부여한다.

#### 3.2.2 콘텐츠 기반 필터링

이후 사용자의 선호 카테고리를 추천에 반영하기 위해 콘텐츠 기반 필터링을 이용한다. 사용자의 구독 카테고리 비율을 통해 사용자의 카테고리 선호도를 파악한다. 이후 추천 채널 리스트에서 카테고리 별 사용자의 구독 비율을 더해줄게 된다. 예를 들어 채널의 카테고리가 Sports이고 사용자가 Sports를 50% 구독한다면 0.5의 가중치가 해당 채널에의 가중치에 더해진다. 이 방식을 통해 사용자의 선호 비율을 더해줄 수 있게 된다. 그리고 [0-1]로 정규화하기 위해 Min-Max Scaling을 진행한다.

3.2.1과 3.2.2의 추천 방식을 거치게 되면 전체 채널 별로 추천 가중치가 만들어진다. 이 가중치는 사용자가 얼마나 해당 채널을 선호하는지를 예상하는 수치이다. 본 연구에서 제안하는 추천 시스템은 총 3개의 채널 추천을 제공한다. 첫 번째는 사용자에게 가장 관련된 채널을 추천해 주기 위해 가중치를 오름차순 했을 때 상위 5%에서 무작위로 뽑아서 제공한다. 그리고 두 번째는 너무 선호 비율로 편향될 것을 방지하여 상위 10%에서 무작위로 뽑아서 제공한다. 마지막 세 번째 추천의 경우 필터버블 현상 [1,2]을 방지하기 위하여 완전 무작위 추천을 제공하여 사용자가 다른 카테고리의 채널을 접할 수 있도록 도와준다.

### 4. 추천 시스템 개발

#### 4.1 추천 화면

위에서 설명한 추천 알고리즘을 활용해서, 소셜 기반 유튜브 채널 서비스 OurTube를 개발하였다 [그림 2]. OurTube는 한번에 3가지 채널을 추천한다. 추천이 마음에 들지 않는다면 Refresh 버튼을 클릭해서 새로고침 할 수 있다. 각 추천은 썸네일, 채널 명, 카테고리를 제공하며 구독하러 가기 버튼으로 채널 공식 URL로 접속할 수 있다.

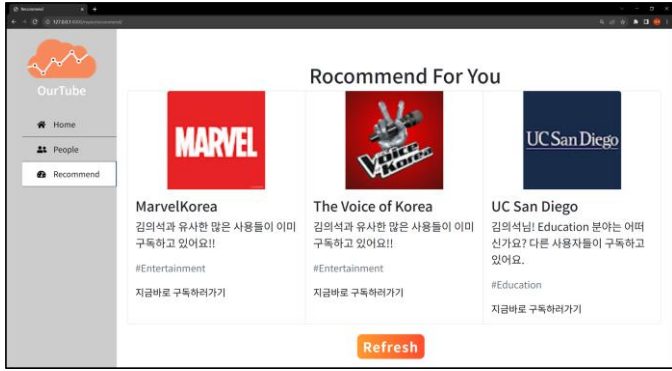


그림 2. OurTube 채널 추천 화면

## 4.2 사용자 평가

OurTube평가를 위해 2023년 9월, 3명의 사용자를 모집해 사용성 평가를 진행했다. 참가자들은 15분 동안 20번의 Refresh 버튼을 누르며 채널 추천을 받았고 그중 관심 있는 채널을 구독했다. 이때 3개의 추천 방식 중 어떤 채널에서 추천이 이어졌는지 그리고 해당 카테고리를 기록했다.

표1 사용자 평가의 결과 카테고리 와 구독 수

	사용자1	사용자2	사용자3
평소 선호 카테고리	Entertainment, People & Blogs, Pets & Animals	Education, People & Blogs, Gaming	Entertainment, People & Blogs, Gaming
추천방식1	Entertainment (2)	People & Blogs (2)	Entertainment (2), Gaming (1)
추천방식2	People & Blogs (1)	Gaming (1)	People & Blogs (1)
추천방식3	Pet (1), Education (1)	Entertainment (1), Travel & Event (1)	Sports (2)

평가 결과는 표 1과 같이 나타났다. 실험 진행 동안 참가자들은 평균 5개(5,5,6)의 채널을 구독했다. 그 중 약 60%의 채널이 사용자가 평소 선호하던 카테고리였고 가중치를 더해준 추천방식 1,2에서 구독이 이루어졌다. 이때 가중치가 5%였던 방식1이 10%였던 방식2보다 약 2배 더 많이 구독으로 이루어졌다. 그 외 40%의 채널은 평소 선호와 다른 새롭게 구독한 카테고리였고 가중치를 고려하지 않고 모든 카테고리의 추천을 해주는 추천 방식3에서 이루어졌다. 해당 실험을 통해 평소 사용자가 자신이 선호하는 분야에 관심을 보이지만 새로운 분야의 추천에도 열려 있다는 것을 알 수 있었다.

## 5. 결론

본 연구는 다른 사용자들과 친구를 맺고 그들의 구독목록을 볼 수 있도록 만들었다. Memory Based Collaborative Filtering과 Content Based Filtering의 혼합된 추천방식에 새로운 유사도[10]를 도입하여 사용자가 관심 있어 하는 채널을 추천했고 추천의 다양성을 위해 랜덤의 추천을 도입했다. 해당 서비스를 통해 채널 추천을 받으며 관심 분야의 채널을 새롭게 알아가고 평소 보지 않았던 분야의 채널에 관심을 가질 수 있을 것으로 기대한다. 또한 온·오프라인에서 관계를 맺은 사람들과 친구추가를 하여 다른 사용자들이 어떤 채널을 구독하고 어떤 분야를 좋아하는지 살펴보고 취향 및 관심사를 공유할 수 있을 것이다.

## 참고 문헌

- [1] Qin, Peng Fei “개인화 추천 서비스, 필터버블, 광고 침입성이 이용의도에 미치는 영향에 관한 Youtube와 Tiktok 비교연구”(2022)
- [2] 신유진, “유튜브 추천 알고리즘으로 인한 필터버블 현상 연구”(2020)
- [3] Deuk Hee Park et al. “A Literature Review and Classification of Recommender Systems on Academic Journals”(2011)
- [4] Breese, J. S., Heckerman, D., and Kadie, C., “Empirical Analysis of Predictive Algorithms for Collaborative Filtering,” In Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence (UAI-98), San Francisco, California, pp. 43-52, 1998.
- [5] Delgado, J. and Ishii, N., “Memory-based Weighted-majority Prediction for Recommender Systems,” In ACM SIGIR'99 Workshop on Recommender Systems : Algorithms and Evaluation, 1999.
- [6] SHAN, ZIYAO “개선된 유사도 기반의 협업필터링 하이브리드 추천시스템”(2022)
- [7] 신사임 외 3명 “다중 미디어 융합 환경 및 소셜 네트워크 서비스 기반의 소셜 사용자 클러스터를 활용한 콘텐츠 추천 엔진 개발”(2011)
- [8] Carlos Augusto Bautista Isaza et al, “OtherTube: Facilitating Content Discovery and Reflection by Exchanging YouTube Recommendations with Strangers”(2022)
- [9] 유소엽 외 3명. “사용자의 소셜 카테고리를 이용한 유튜브 동영상 추천 알고리즘”(2015)
- [10] J.Babadilla et al “A new collaborative filtering metric that improves the behavior of recommender systems”(2010)