

미디어 서비스 플랫폼에서의 사용자 행동 유형에 따른 CBF 추천 효과성 분석

박지수, 박지형, 이상범

SK 브로드밴드, SK 텔레콤, SK 브로드밴드

jstorys11@sk.com, jihyung.park@sk.com, sb.lee@sk.com

CBF recommendation effectiveness analysis according to user behavior type in media service platform

Park Jisoo, Park Jihyung, Lee Sangbeom

SK Broadband, SK Telecom, SK Broadband

요 약

본 논문은 미디어 서비스 플랫폼에서 점차 중요성이 커지고 있는 추천 시스템에 대한 효과성을 분석하였다. 대표적으로 사용되는 추천 알고리즘 중 CBF 기반의 콘텐츠-콘텐츠 간 유사성 추천 방식을 적용하여 연구를 진행하였다. 이때 해당 추천 방식에 반영되는 사용자의 행동 유형 2 가지(클릭, 재생)에 따른 추천 효과성을 살펴보았다. 추천 효과성은 3 가지 요인(콘텐츠의 다양성, 콜드 스타트 문제 해결 효과성, 콘텐츠 재생 전환율)을 기준으로 분석한 결과, 클릭 기반의 추천 방식을 사용할 경우 재생 기반의 추천 대비 콘텐츠의 다양성이 증가하고, 콜드 스타트 문제 해결에 효과적이라는 것을 알 수 있었다. 반면 재생 전환율 개선에는 효과가 미미하여 클릭 기반의 추천에서의 재생 전환율 개선 방향에 대한 후속 연구를 진행하고자 한다.

I. 서 론

미디어 서비스 플랫폼에서의 개인화 추천에 대한 중요성은 점차 커지고 있다. 대표적인 글로벌 미디어 서비스 플랫폼인 넷플릭스에서는 추천 시스템의 가치를 정량화하여 1 년에 1.2 조원 정도의 가치가 있다고 계산하기도 하였다[1]. 이와 같은 추천 시스템에서 사용되는 대표적인 두 가지 알고리즘에는 협업 필터링(Collaborative Filtering, CF)과 콘텐츠 기반 필터링(Content-Based Filtering, CBF)이 있다. 그 중 본 논문에서는 콘텐츠의 다양한 메타 정보를 기반으로 추천을 하는 방식인 CBF 알고리즘을 활용한 콘텐츠-콘텐츠 간 유사성 추천 방식에 대한 연구를 진행하였다. 넷플릭스에서는 BYW(Because You Watched)라고도 명기하는 해당 알고리즘[1]은 사용자가 시청하거나, 좋아요로 평가하거나, 찜한 콘텐츠에 대해 비슷한 콘텐츠를 추천한다. BYW 로 추천한 콘텐츠는 “OOO 와(과) 비슷한 콘텐츠”라는 타이틀로 묶여서 제공된다.

현재 SK 브로드밴드에서 운영하고 있는 IPTV 플랫폼의 A 월정액 상품은 약 20 만명의 가입자를 보유하고 있다. A 월정액 가입자에게 콘텐츠를 추천하는 방식 중 하나는 가입자가 재생한 콘텐츠와 비슷한 콘텐츠를 추천하는 콘텐츠-콘텐츠 간 유사성 추천 방식(이하 BYW)이다. 그러나 사용자가 월정액에 가입한 후 재생을 하기 전까지는 BYW 추천을 제공해줄 수 없는 콜드 스타트 문제가 발생한다. 이에 본 논문은 사용자의 행동 유형 중 재생 이벤트만을 기반으로 했던 BYW 추천

방식에 클릭 이벤트를 추가하여 사용자에게 얼마나 더 빨리 추천이 제공될 수 있는지에 대한 콜드 스타트 문제 해결 정도와 추천의 효과성에 대해 분석해보고자 한다.

II. 본론

본 연구에서 분석하고자 하는 A 월정액의 사용 로그를 살펴보면, 콘텐츠를 클릭하여 줄거리, 출연 인물 등을 확인하는 클릭 이벤트는 콘텐츠를 시청하는 재생 이벤트보다 일평균 2 배의 규모로 발생하고 있으며, A 월정액 가입 1 일차로 기간에 제한을 두었을 때에는 클릭 이벤트가 재생 이벤트의 2.7 배의 규모로 발생하고 있다. 따라서 기존에 재생 이벤트를 기반으로 추천하던 BYW 추천 방식에서 클릭 이벤트를 기반으로 한 BYW 추천 방식을 추가함으로써 다음과 같은 2 가지 이점을 기대할 수 있다 :

(i) 다양성

재생 이벤트의 2 배 규모를 가지고 있는 클릭 이벤트를 기반으로 콘텐츠를 추천할 경우, BYW 추천에서 비슷한 콘텐츠를 추천하기 위해 기준이 되는 콘텐츠(“OOO 와(과) 비슷한 콘텐츠”에서의 OOO 에 들어가는 콘텐츠)가 더 다양해질 수 있다. 이에 따라 추천되는 콘텐츠의 다양성도 함께 증가하게 된다.

(ii) 콜드 스타트 문제 해결

월정액 가입 1 일차에는 재생 이벤트와 클릭 이벤트 간의 빈도 차가 더 벌어지는 현상을 볼 수 있다. 가입 초반에는 월정액 안에 포함된 콘텐츠들이 무엇이 있는지 탐색을 하기 위해 재생 보다는 클릭을 더 많이 하기 때문에 가입 1 일차에는 더 큰 격차가 벌어진다고 해석할 수 있다. 따라서 가입 초반에 더 빈번하고 빨리 발생하는 클릭을 기반으로 추천을 제공하게 될 경우 재생이 발생하기 까지 기다려야 하는 콜드 스타트 문제를 완화시킬 수 있다.

이는 인간-알고리즘 간의 인터랙션을 연구한 Luiz Agner 가 제안한 추천 시스템의 성공 요인 9 가지에 포함되는 2 가지의 요인(Diversity, Cold Start)으로서[2], 기존의 BYW 추천보다 더 향상된 추천 시스템을 제공할 수 있다는 가능성을 비추고 있다.

클릭 기반의 BYW 추천의 실제 효과성을 분석하기 위해 A 월정액의 신규 가입자를 대상으로 2 주의 기간 동안 실험을 진행하였다. 실험에서는 다음과 같은 3 가지 요인에 대한 변화를 분석해 보았다: (i) 콘텐츠의 다양성, (ii) 콜드 스타트 문제 해결 효과성, (iii) 콘텐츠 재생 전환율(클릭한 콘텐츠를 재생한 비율)

먼저 (i) 콘텐츠의 다양성을 분석한 결과, 클릭 기반 BYW 를 도입했을 경우 기존 콘텐츠의 종류가 13.4% 증가한 것으로 나왔다. 기존 콘텐츠 당 10 개 가량의 콘텐츠가 추천되는 것을 생각한다면, 콘텐츠 다양성에 대한 효과가 더 극대화 될 것이라 예상할 수 있다.

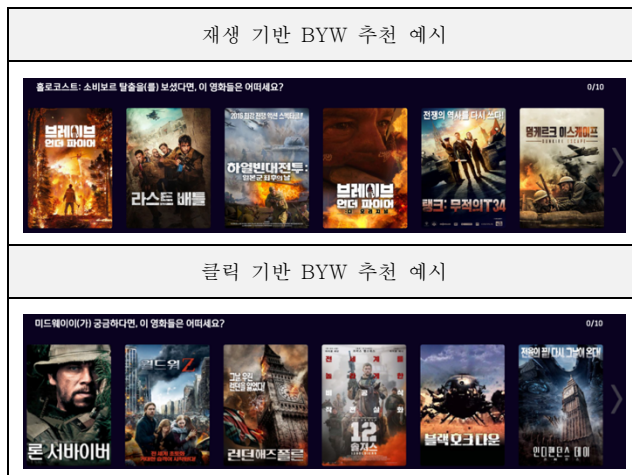


표 1. 재생/클릭 기반 BYW 추천 예시

두 번째로 (ii) 콜드 스타트 문제 해결 효과성은 클릭 기반 BYW 추천으로 인해 콜드 스타트 문제가 해결된 사용자의 수와 해당 사용자에게 대해 개인화 추천이 제공될 때까지 걸리는 리드 타임이 얼마나 단축되었는가를 살펴보았다. 먼저 콜드 스타트 문제가 해결된 사용자의 수는 신규 가입자의 9.4%에 해당하는 수치가 나왔다. 이는 2 주 안에 대부분의 사용자가 재생을 발생시키기 때문에 콜드 스타트 문제가 발생하는 사용자의 수 자체가 적다는 것을 의미하기도 한다. 후속 연구로 분석 대상 기간을 2 주, 1 주, 3 일으로 점차 줄여가며 콜드 스타트 해결 사용자수를 분석하는 것이 필요해 보인다.

다음으로 개인화 추천의 리드 타임 단축에 대해서 살펴 보았을 때, 최소 5 일 정도의 리드 타임을

단축시켰다는 결과가 나왔다. 클릭 이벤트의 경우 발생하기까지 평균 7 일, 재생 이벤트의 경우 평균 12 일이 걸리기 때문에 재생 기반 BYW 추천을 제공하기 5 일전부터 클릭 기반 BYW 추천을 노출시킴으로써 추천 제공의 리드타임을 단축시켰다고 볼 수 있다.

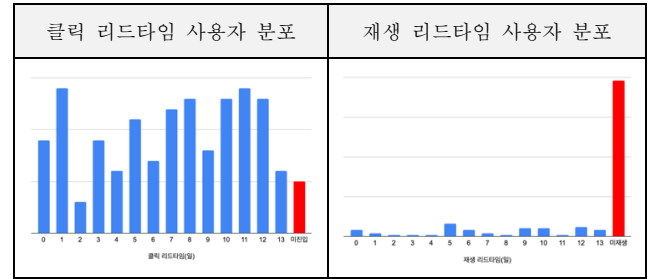


표 2. 리드타임(일) 별 사용자 수 분포

마지막으로 (iii) 콘텐츠 재생 전환율을 분석한 결과, 재생 기반 BYW 와 비교했을 때 클릭 기반 BYW 의 재생 전환율이 약 10% 정도 하락하는 것을 볼 수 있었다. 이는 예상 가능한 결과로, 클릭한 콘텐츠 보다는 재생한 콘텐츠에 대한 사용자의 선호도가 더 높기 때문에 그와 비슷한 추천 콘텐츠를 제공했을 때 콘텐츠를 재생할 확률이 더 높다. 재생 기반 BYW 추천의 재생전환율과 비교했을 때에는 낮은 수치이나, 해당 서비스에서 제공하는 추천에서의 평균 재생전환율과 비교했을 때, 평균 수치에 해당하는 성과를 보였다.

III. 결론

본 논문에서는 재생 기반 BYW 추천에서의 콜드 스타트 문제를 해결하고자 새로운 사용자 행동 유형인 클릭 이벤트를 기반으로 BYW 추천을 제공하고 그 효과를 살펴보고자 했다. 클릭 기반 BYW 를 도입해 사용자에게 제공하는 콘텐츠의 다양성을 13.4%를 향상시켰으며, 신규 가입자의 9.4%에 해당하는 사용자의 콜드 스타트 문제를 해결하고 추천 제공의 리드타임을 5 일 단축했다. 재생전환율의 측면에서는 기존에 제공했던 재생 기반 BYW 추천에 비해 10% 정도의 효과성이 떨어지는 것을 볼 수 있었는데, 후속 연구로 클릭 기반의 BYW 추천 방식에서 재생전환율을 높일 수 있는 알고리즘을 연구할 필요가 있다. 또한, 재생, 클릭 뿐만 아니라 좋아요, 찜과 같은 다양한 사용자 행동 유형으로 확장하여 추천 효과성을 분석하는 연구를 후속으로 진행하고자 한다.

참 고 문 헌

- [1] Carlos A. Gomez-Urbe and Neil Hunt, "The Netflix Recommender System: Algorithms, Business Value, and Innovation, ACM Transactions on Management Information Systems, Vol. 6, No. 4, Article 13, Dec. 2105.
- [2] Luiz Agner, Barbara Necyk, and Adriano Renzi, "Recommendation Systems and Machine Learning: Mapping the User Experience", International Conference on Human-Computer Interaction, pp3-17, July. 2020.