콘텐츠 인기도 예측을 위한 연합학습 기반 개인화된 Variational Autoencoder

김유노, 최민석

경희대학교

rladbsh456@naver.com, choims@khu.ac.kr

Personalized Variational Autoencoder Based on Federated Learning for Content Popularity Prediction

Yunoh Kim, Minseok Choi

Kyung Hee Univ.

요 약

본 연구는 캐싱 네트워크에서 사용자의 프라이버시를 보호하는 캐싱 결정을 위해 연합학습 기반으로 인기도 예측을 수행한다. 이때, 실제 데이터셋으로 모델링된 콘텐츠 인기도 분포에 맞추어 Variational Autoencoder(VAE)를 학습시키고자 하며, 각 사용자와 지역별 특색을 반영한 개인화 VAE 모델을 위한 연합학습 방법을 제안한다. 또한, 현실적인 조건을 적용하여 많은 콘텐츠, 연합학습 불참여 사용자에 대한 캐시히트 성능을 분석한다.

Ⅰ. 서 론

각종 스마트 디바이스의 보급과 함께, 비디오 트래픽은 지속적으로 증가하고 있다. 비디오 트래픽은 중복되고 반복적인 콘텐츠 요청이 많다는 특성이 있는데, 이를 반영하여 에지 단에서의 스토리지를 활용하여 트래픽을 줄일 수 있는 캐싱 네트워크가 주목 받아왔다. 캐싱 네트워크에서는 사용자와 가까운 노드에서 콘텐츠를 캐싱하여, 클라우드 서버를 거치지 않고 사용자의 요청에 대해 빠르게 콘텐츠를 제공할 수 있다. 캐싱을 잘 수행하기 위해서는 어떤 콘텐츠를 캐싱하는 것이 좋을지 결정해야 하고, 이를 위해서는 사용자의 비디오 콘텐츠에 대한 인기도를 예측이 필수적이다. 딥러닝을 활용한 콘텐츠 인기도 분포 예측 연구가 많이 이루어졌지만, 여전히 다음과 같은 어려움이 존재한다: 1) 사용자의 콘텐츠 요청 데이터 프라이버시 보호, 2) 사용자 개인의 콘텐츠 요청 오픈 소스 데이터셋의 부재, 3) 사용자 및 지역적 선호도의 반영.

최근 EU에서 사용자 개인의 데이터 보호에 대한 법안을 제출하는 등 [1] 사용자의 데이터 프라이버시를 보호할 수 있는 학습 환경이 중요해졌다. 이러한 흐름에 맞추어 사용자가 데이터를 공유하지 않아도 되는 연합학습 기반의 인기도 분포 예측 연구가 이루어졌다. 그러나, 공개된 사용자 개인이 요청한 콘텐츠 데이터가 없어서 사용자 개인이 아닌 기지국 단위로 학습을 수행하거나 [2], 콘텐츠 평가 데이터셋인 MovieLens dataset [3] 등을 콘텐츠 요청 데이터라고 가정한 연합학습 기반 인기도 분포 예측만 이루어졌다 [4]. 또한, 콘텐츠 인기도 분포에는 지역적 선호도뿐만이 아니라 개인적 선호도 또한 존재한다. 기존 연합학습 방법은 지역적 선호도는 파악할 수 있어도, 개인적 선호도는 파악이 어려울 수 있다. 이를 해결하기 위해 연합학습 구조는 유지하면서도, 각 사용자에 대한 개인화를 수행하여 각 사용자의 개인적 선호도를 파악할 필요가 있다.

본 논문에서는 BBC dataset을 기반으로 여러 단계에 거쳐 콘텐츠 인기도 분포를 모델링한 콘텐츠 요청 데이터셋 생성기 [5]를 이용하여 최대한 실제 환경과 유사한 콘텐츠 요청 데이터를 바탕으로 사용자와 지역적 선호도를 반영하여 개인화된 인기도 분포를 예측할 수 있는 연합학습 기술을 제안한다.

Ⅱ. 무선 캐싱 네트워크 모델

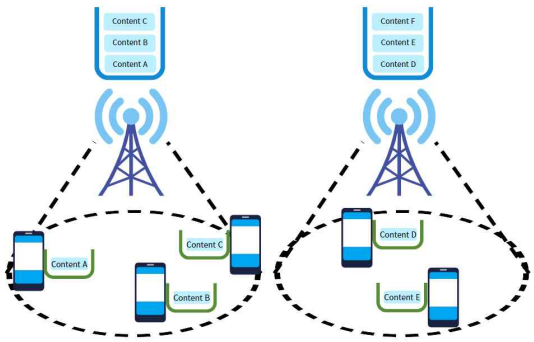

Fig.1 무선 캐싱 네트워크 구조

Fig.1과 같은 무선 캐싱 네트워크에는 클라우드 서버로부터 콘텐츠를 전달받아 사용자에게 전달해주는 Edge Server(ES)가 여럿 존재한다. 각 ES는 자신의 통신 영역 내에 분포된 여러 사용자들의 콘텐츠 요청에 따라 콘텐츠를 사용자에게 전송한다. 이때, ES 혹은 사용자의 디바이스는 사용자가 요청할 만한 콘텐츠를 미리 캐싱할 수 있다. 일반적으로 디바이스가 콘텐츠를 캐싱할 수 있는 저장용량은 작고, 그에 비해 ES가 콘텐츠를 캐싱할 수 있는 저장용량은 크다. 따라서, 본 연구에서는 캐싱하는 콘텐츠의 수를 바꿔가며 캐시 히트를 측정한다. 또한, ES는 보통 특정한 사용자 한 명에 치중하지 않고 지역적인 콘텐츠 인기도 분포를 예측하여 캐싱을 결정하는 경향성이 있다. 반대로 사용자 디바이스는 사용자 한 명의 개인적인 선호도를 예측하여 캐싱을 결정하는 경향성이 있다. 캐싱 네트워크에서 높은 캐시히트를 달성하기 위해서는 디바이스 캐싱에 유리한 개인적 선호도를 예측해야 하며, ES 캐싱에 유리한 지역적 선호도 또한 예측해야 한다. 따라서, 본 논문에서는 지역적 선호도와 개인적 선호도를 모두 반영한 [5]의 콘텐츠 인기도 분포 모델링을 사용한다. [5]에서 장르의 순위 분포는 지역적 선호도로서 반영된다. 또한, 사용자가 특정 순위의 장르를 얼마나 좋아하는지, 즉 장르 선호도 분포는 개인적 선호도로서 반영된다. 이때, 한 장르 내 콘텐츠들의 인기도 분포는 모든 사용자에 대해 동일하다고 가정하여 사용자들의 콘텐츠 인기도 분포 예측을 수행한다.

Ⅲ. 개인화된 VAE 연합학습

각 사용자는 자신의 콘텐츠 요청 데이터를 사용하여 자신의 VAE를 학습한다. 이에 따라 VAE는 해당 사용자의 콘텐츠 인기도 분포를 예측한다. 그러나, VAE를 사용해 모분포를 추정하기 위해서는 많은 요청 데이터가 필요하다. 이를 해결하기 위해, 연합학습을 통해 적은 콘텐츠 요청으로도 모분포 추정을 가능하도록 할 수 있다.

Algorithm 1 Personalized Federated Learning

procedure FEDERATED LEARNING
  N: Number of users
  E: Number of epochs
  for epoch $e = 1$ to E **do**
    for user $i = 1$ to N **do**
      $VAE_i \leftarrow TrainVAE(data_i)$
    end for
    $VAE_{global} \leftarrow Aggregate(VAE_1, VAE_2, \ldots, VAE_N)$
    for user $i = 1$ to N **do**
      $VAE_i \leftarrow VAE_{global}$
      $VAE_i \leftarrow Personalize(VAE_i)$
    end for
  end for
end procedure

Fig.2 개인화된 VAE 연합학습 과정

Fig. 2는 개인화된 VAE 연합학습 과정을 나타낸다. 각 사용자는 자신의 요청 데이터를 사용해 자신의 VAE를 학습한다. 모든 사용자의 학습이 끝난 후에 ES에서 모든 사용자의 VAE를 응집하여 하나의 글로벌 VAE를 생성한다. 그리고 글로벌 VAE를 각 사용자에게 다운로딩한다. 그 후에 각 사용자는 글로벌 VAE를 개인화하여 다음 연합학습 라운드에 시작 모델로 사용한다. 본 연구에서는 줄 11에서의 개인화 방법으로 두 가지 방법을 사용한다.

첫 번째는 글로벌 VAE와 자신의 VAE의 모델 가중치를 응집하는 방법(Weight Aggregation)이다. 기존 연합학습에서는 글로벌 모델을 사용자가 다운로딩 받아 새로운 학습 시작 모델로 바로 사용한다. 그러나 이 개인화 방법은 자신의 모델에 큰 가중치를 두고, 글로벌 모델에 적은 가중치를 두어 응집함으로써 자신의 모델을 완전히 덮어쓰는 과정을 없앤다. 이로 인해 자신의 요청 데이터로 지속적으로 학습하는 VAE가 어느정도 유지되기 때문에, 기존 연합학습보다 개인화된 성능을 보일 수 있다.

두 번째는 글로벌 VAE로 생성한 요청 데이터를 사용하여 학습하는 방법(Data Integration)이다. 각 사용자가 글로벌 VAE를 다운로딩 받았다면, 글로벌 분포를 생성할 수 있다. 해당 글로벌 분포를 통해 요청 데이터를 생성하여 자신의 VAE를 학습한다면, 자신의 VAE 특성이 어느정도 유지되면서 지역적 선호도를 학습할 수 있다.

위 두 가지 개인화 방법을 사용하여 각 사용자는 최종적으로 자신에게 특화된 VAE를 학습할 수 있게 된다. 이렇게 학습된 VAE를 사용하여 캐싱을 결정한다. 본 연구에서는 개인화 VAE를 가진 사용자에 대해서는 자신의 VAE에서 가장 인기있는 k개의 콘텐츠를 캐싱하여 캐시히트를 측정한다. VAE가 없는 연합학습 불참여 사용자에 대해서는, 연합학습 참여 사용자의 개인화된 VAE로 가장 인기가 높은 k개의 콘텐츠를 선택하여 ES에 전송하고, ES는 전송받은 콘텐츠들 중에서 가장 많은 사용자에게 선택된 콘텐츠 k개를 캐싱하여 캐시히트를 측정한다.

Ⅳ. 실험결과

제안한 개인화된 연합학습 기반 인기도 예측 방법의 성능 분석을 위해, 전체 2000개의 콘텐츠에 대하여 100명의 사용자가 각자 1000개의 과거 콘텐츠 요청 데이터를 바탕으로 VAE를 연합학습하는 시뮬레이션을 진행하였다. 이때, 모든 사용자가 연합학습에 참여하는 것을 동의하지 않을 환경을 고려하여 100명의 연합학습에 참여하지 않은 사용자도 존재한다고 가정한다. ES의 캐시 크기가 5, 30, 100일 때 예측한 인기도 분포를 바탕으로 가장 인기 있는 콘텐츠들을 캐싱했을 때의 캐시 히트 성능을 살펴보았다.

Top-K	5	30	100
Fed-AVG	68	297	555
Weight Aggregation	99	330	585
Data Integration	104	359	630

Table. 1 연합학습 참여 사용자에 대한 캐시히트

Top-K	5	30	100
Fed-AVG	64	276	528
Weight Aggregation	63	274	520
Data Integration	63	274	523

Table. 2 연합학습 불참여 사용자에 대한 캐시히트

Table. 1은 연합학습에 참여하여 개인화된 VAE로 캐싱한 사용자에 대한 캐시히트이다. 사용자의 콘텐츠 요청 1000번에 대해 콘텐츠를 5개, 30개, 100개 캐싱했을 때 모두 개인화된 방법에서의 캐시히트가 기존 연합학습방법의 캐시히트보다 좋은 것을 확인할 수 있다.

Table. 2는 연합학습에 참여하지 않아, VAE를 가지지 않은 사용자에 대한 캐시히트이다. 연합학습에 참여한 사용자들의 개인화된 모델을 집계하여 캐싱을 결정하였음에도 불구하고, 기존 연합학습의 글로벌 VAE로 캐싱을 결정한 방법에 근사한 캐시히트 성능을 보이고 있다. 이를 통해, 개인화 연합학습으로도 지역적 선호도를 파악하여 연합학습에 참여하지 않은 사용자에 대해 준수한 캐시히트 성능을 보일 수 있음을 알 수 있다.

Ⅴ. 결론

본 연구에서는 캐싱 네트워크에서 필요로 하는 콘텐츠 인기도 예측에 대해, 연합학습 기반 개인화된 VAE를 통해 보다 현실적인 조건에서도 많은 캐시히트를 발생할 수 있음을 확인하였다.

ACKNOWLEDGMENT

이 논문은 2024년도 정보(과학기술정보통신부)의 재원으로 한국연구재단과 정보통신기획평가원의 지원을 받아 수행된 연구임 (NRF-2022R1C1C1010766, No. 2022R1A4A3033401, No.2021-0-02201, 사용자 프라이버시를 보존하는 비디오 캐싱을 위한 연합 학습 시스템)

참 고 문 헌

[1] B. Clusters, et al. *EU Personal Data Protection in Policy and Practice*. Hague, The Netherlands: TMC Asser Press, 2019.

[2] Yongmoon Park, et al. "Proactive Content Caching via Interplay Between Deep Learning and Stochastic Optimization" IEEE MASS, Seoul, South Korea, 2024.

[3] MovieLens, available at: https://grouplens.org/datasets/movielens/.

[4] M. Ahn and M. Choi, "Federated Learning with Variational Autoencoder for Popularity Profile Prediction," 2023 14th International Conference on Information and Communication Technology Convergence (ICTC), Jeju Island, Korea, Republic of, 2023, pp. 1027-1032.

[5] Lee, Ming-Chun, et al. "Individual preference probability modeling and parameterization for video content in wireless caching networks." IEEE/ACM transactions on networking 27.2 (2019): 676-690.