

MEC 환경의 영상 스트리밍 효율화를 위한 강화학습 기반 적응형 캐싱 기법

최민지, 박소민, 곽희주, 유철우*
명지대학교

choiminji@mju.ac.kr, thalsi@mju.ac.kr, wopelt8@mju.ac.kr, *cwyou@mju.ac.kr

Reinforcement Learning-Based Adaptive Caching for Efficient Video Streaming in MEC environments

Minji Choi, Somin Park, Heeju Kwak, Cheolwoo You*
Myongji University

요 약

MEC(Multi-access Edge Computing, MEC)는 네트워크 엣지에 위치함으로써 지연 시간을 줄일 수 있는 이점을 가지지만, 저장 용량이 제한적이기 때문에 효율적인 서비스 제공을 위해서는 캐시 성능에 대한 연구가 필요하다. 본 논문에서는 대용량 데이터의 실시간 전송이 요구되는 360° 영상 스트리밍 시스템에 적합한 MEC 캐싱을 위해 적응형 최신성 마스킹 캐싱(Adaptive Recency Masking Caching, ARMC) 알고리즘을 제안한다. 제안 기법은 중요도 기반 최신성 마스킹과 FFLRU(Frequency-Filtered Least Recently Used) 삭제, 관찰 창 개념을 결합해 효율적이고 적응적인 캐시 교체를 실현한다. HMD(Head Mounted Display)를 착용한 사용자의 시야(Field of View, FoV) 데이터셋을 기반으로 실험한 결과, 사용자의 FoV에 해당하는 고화질의 타일을 MEC 캐시에 요구하는 시나리오에서, 기존 캐시 교체 알고리즘 대비 우수한 성능을 보였다. 또한, 최적의 관찰 창 크기가 캐시 크기와 시청 패턴에 따라 가변적일 수 있는 문제를 해결하기 위해, 강화 학습(Reinforcement Learning, RL) 기반의 ARMC-RL을 함께 제안한다. 실험 결과는 제안한 기법이 ARMC의 효율성과 성능을 동시에 개선하며 안정적 자동화를 실현할 수 있음을 보여준다.

I. 서론

최근 3 차원 및 360° 영상과 같은 고해상도 및 고몰입 콘텐츠의 수요 증가로 인해, 대규모 데이터를 실시간 처리할 수 있는 멀티-엑세스 엣지 컴퓨팅(Multi-access Edge Computing, MEC)이 B5G/6G 네트워크의 차세대 기술로 주목받고 있다[1]. 그러나 MEC 시스템은 대용량 데이터에 비해 사용 가능한 메모리 용량이 상대적으로 제한적이기 때문에 MEC 메모리 최적화 연구의 필요성이 제기되고 있다[2]. 본 논문에서는 MEC 기반 360° 영상 스트리밍 시스템에서의 캐시 효율을 높이기 위해 새로운 캐시 교체 알고리즘인 적응형 최신성 마스킹 캐싱(Adaptive Recency Masking Caching, ARMC)을 제안한다. ARMC는 관찰 창, FFLRU(Frequency-Filtered Least Recently Used) 삭제, 최신성 마스킹 기법을 결합하여, 요청 패턴에 따라 정책을 실시간으로 조정한다. 또한 주요 변수인 관찰 창을 강화 학습(Reinforcement Learning, RL)으로 추정하는 ARMC-RL을 함께 제안한다. 이는 실제 시스템에서의 ARMC의 실시간 적응성과 실용성을 입증할 것이다.

II. 본론

1. System Model

본 논문에서 고려하는 360° 영상 스트리밍을 제공하는 MEC 캐싱 시스템 시나리오는 그림 1과 같다. 서버는 원본 영상을 각 세그먼트 내에서 여러 타일로 조각내는 타일링(Tiling)을 한다. 360° 영상 스트리밍을 고려하기 때문에 클라이언트는 HMD(Head Mounted Display)이다. HMD는 영상의 일부분인 사용자의 시야(Field of View,

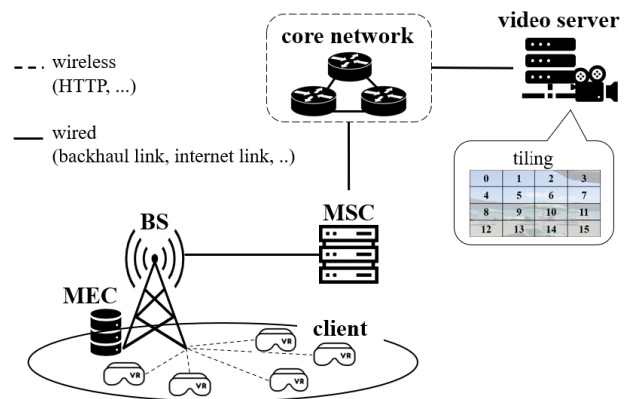


그림 1. MEC 환경에서의 영상 스트리밍 시나리오

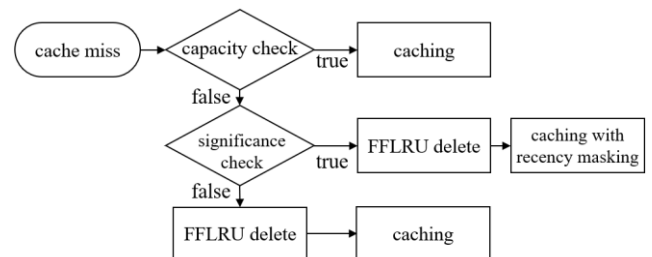


그림 2. ARMC 동작 흐름도

FoV)가 포함된 타일을 MEC 또는 서버에 요청한다. MEC는 캐시 교체 알고리즘에 따라 선택적으로 타일을 캐싱한다. 따라서 HMD가 필요로 하는 영상 타일이

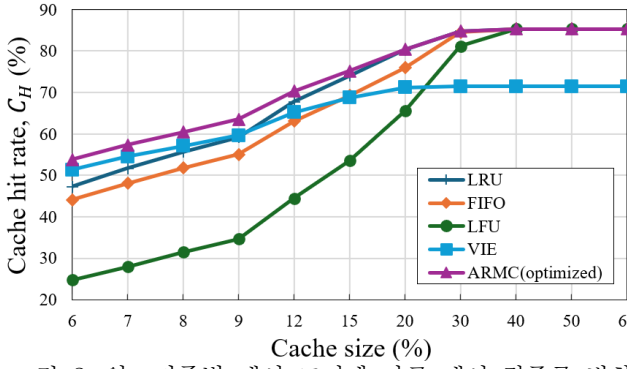


그림 3. 알고리즘별 캐시 크기에 따른 캐시 적중률 변화

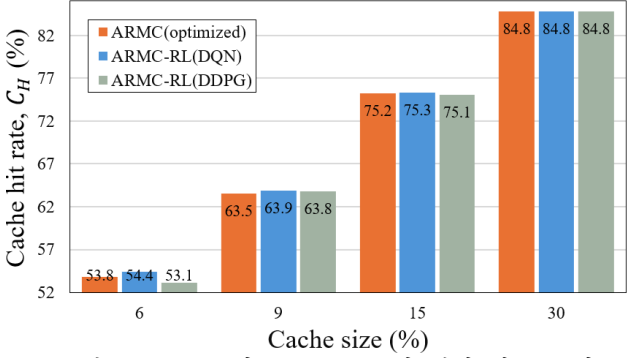


그림 4. ARMC 및 ARMC-RL의 캐시 적중률 비교

MEC에 존재하면 서버로부터 해당 타일을 전송받지 않아도 되므로 전송 지연을 감소시킬 수 있다.

그림 2는 ARMC의 동작 흐름도를 보여준다. MEC에서 캐시 미스가 발생했을 경우, 현재 캐시의 여유 공간이 충분하다면 미스된 타일을 서버로부터 수신하여 캐싱한다. 그렇지 않다면 캐시된 타일들의 중요도를 판단한다. 캐시된 데이터들이 중요하지 않다면 일부 타일을 삭제 후 미스된 타일을 캐싱하고, 그렇지 않다면 타일을 삭제 후 미스된 타일을 최신성 마스킹 기법에 따라 캐싱한다. 이때 중요도 여부와 상관없이 삭제 과정은 FFLRU를 따라 삭제한다. 최신성 마스킹의 경우에는 중요한 타일들을 삭제하면서 캐시 교체를 하는 것이기 때문에 새로 캐싱하는 타일에게는 패널티를 부여한다. 이는 추후에 해당 타일이 FFLRU 삭제가 고려될 때, 우선적으로 삭제되도록 캐싱된 시기를 오래된 것처럼 조작한다. FFLRU 삭제는 접근 빈도로 우선 삭제할 타일을 필터링하고, 동일한 접근 빈도를 가지는 데이터에 대해서 LRU를 적용하는 방식이다. 따라서 최신성 마스킹이 FFLRU 과정에서 삭제 우선순위에 패널티를 주는 설계 요소가 된다. 또한 스트리밍 추세의 최근성을 반영할 수 있도록 관찰 창 개념을 도입한다. 관찰 창은 최근 일정 기간을 의미하며 그 기간동안에 발생한 접근에 대해서만 접근 빈도, 최신성을 고려한다. 관찰 창의 최적 값을 매번 직접 설정해야 한다는 문제를 해결하기 위해서 RL을 활용하여 최적의 관찰 창 크기를 추정하고 업데이트하는 ARMC-RL도 함께 제안한다. RL 학습 모델로는 DQN(Deep Q-Network), DDPG(Deep Deterministic Policy Gradient)를 활용하며, 현재의 캐시 상태, 관찰 창 크기, 캐시 적중률을 상태 값으로 받아 캐시 적중률 향상을 목표로 학습한다.

2. Simulation Results

실험은 4x4 형식으로 타일링한 영상에 대해서 HMD가 매 세그먼트마다 FoV에 해당되는 타일 1개를 요청한다고 가정한다. 실제 오픈 데이터셋[3]을

활용하여 HMD를 통해 영상을 시청한 사용자의 FoV 기반으로 실험을 진행한다. 실험 결과, 기존 캐시 교체 알고리즘(LRU, LFU, FIFO, VIE[4])과 비교하여 다양한 캐시 크기에 대해서 모두 향상된 캐시 적중률을 가졌으며, 최대 29% 향상된 캐시 적중률을 보였다. ARMC-RL의 경우, DQN 모델을 사용했을 때 모든 캐시 크기에 대해서 ARMC보다 향상된 캐시 적중률을 보였다. 이는 ARMC-RL을 사용하면 빠르고 정확하게 최적의 관찰 창을 찾을 수 있음을 나타내고, 따라서 제안하는 기법이 실시간으로 데이터 요청이 변화하는 실제 시스템에서의 활용이 가능할 것으로 판단된다.

III. 결론

본 논문에서는 영상 스트리밍을 제공하는 MEC 시스템에서 효율적인 캐시 교체를 위해 새로운 캐싱 기법 ARMC를 제안하였다. ARMC는 캐시 상태와 데이터 요청 패턴을 기반으로 캐시 교체 정책을 실시간으로 조정하며, 관찰 창, FFLRU 삭제, 최신성 마스킹 기법을 결합해 구성된다. FFLRU 삭제는 관찰 창 내 요청 빈도가 낮은 데이터를 우선적으로 삭제하며, 동일한 요청 빈도에서는 LRU 기준을 따른다. 최신성 마스킹은 새로운 데이터를 캐싱할 때 해당 데이터의 삭제 우선순위를 낮추는 패널티를 부여하는 방식으로, 데이터가 캐싱된 시점을 오래된 것처럼 인위적으로 조작한다. 따라서 최신성 마스킹은 추후에 있을 FFLRU 삭제 과정에서 해당 타일에 대해 패널티를 주기 위한 설계 요소이다. 실험에서는 사용자의 FoV에 해당되는 360° 영상 타일을 MEC에 요청하는 시나리오를 가정하였고, 그 결과 ARMC는 다양한 캐시 크기에서 기존 캐시 교체 알고리즘 대비 향상된 캐시 적중률을 보였다. 또한, 관찰 창의 최적 크기를 실시간으로 추정할 수 있도록 RL을 접목시킨 ARMC-RL을 함께 제안하였다. ARMC-RL은 ARMC보다 미세하지만 향상된 캐시 적중률을 보였으며, 제안하는 기법의 실시간 적응성과 실용성을 입증하였다.

ACKNOWLEDGMENT

이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No. RS-2024-00335012). 또한, 이 성과는 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (No. RS-2024-00351187).

참고 문헌

- [1] M. Banafaa, I. Shaye, J. Din, M. H. Azmi, A. Alashbi, Y. I. Daradkeh, and A. Alhammadi, "6G mobile communication technology: Requirements, targets, applications, challenges, advantages, and opportunities," *Alexandria Eng. J.*, vol. 64, pp. 245–274, Feb. 2023.
- [2] J. Yao, T. Han, and N. Ansari, "On mobile edge caching," *IEEE Commun. Surveys Tuts.*, vol. 21, pp. 2525–2553, 3rd Quart., 2019.
- [3] C. Wu, Z. Tan, Z. Wang, and S. Yang, "A dataset for exploring user behaviors in VR spherical video streaming," in *Proc. 8th ACM Multimedia Syst. Conf.*, Jun. 2017, pp. 193–198.
- [4] Q. Liu, H. Chen, Z. Li, Y. Bai, D. Wu, and Y. Zhou, "Online caching algorithm for VR video streaming in mobile edge caching system," *Mobile Netw. Appl.*, 2024. [Online]. Available: <https://doi.org/10.1007/s11036-024-02291-2>.