

사용자 유사도 기반 Cooperative MEC 콘텐츠 캐싱

김희수*, 임유진^o

Content Caching in Cooperative MEC Based on User Similarity

Heesu Kim*, Yujin Lim^o

요약

Multi-access edge computing(MEC) 환경은 5G의 보급과 함께 가장 주목받는 기술 중 하나이다. MEC는 기존의 중앙 집중식 클라우드 방식에서 5G의 처리 속도 한계점을 보완하여 초지연성의 실현을 가능하게 한다. 본 논문은 MEC 내 캐싱 스토리지를 이용해 사용자에게 효율적으로 데이터를 제공하는 방법에 대해 제안한다. 특히 사용자의 콘텐츠 소비 패턴을 활용한 사용자 유사도를 계산하고, 이를 기반으로 한 MEC 콘텐츠 캐싱 방법을 제안하여, 효율적인 MEC 콘텐츠 캐싱을 가능하게 한다. 또한 Cooperative MEC 콘텐츠 캐싱을 통해 사용자가 사용할 수 있는 캐싱 공간을 확장시켜 사용자 요청 콘텐츠 데이터의 적중률(hit ratio)을 높이고, 지연(latency)을 줄여 서비스 제공의 질(Quality of Service, QoS)을 향상시키고자 한다.

키워드 : 콘텐츠 캐싱, 클라우드 컴퓨팅, Cooperative MEC, 사용자 유사도, 캐시 교체 정책

Key Words : Contents Caching, Cloud Computing, Cooperative MEC, User Similarity, Content Cache Replacement Policy

ABSTRACT

The Multi-access edge computing(MEC) environment is one of the most promising technologies for 5G network. MEC mitigates the processing bottlenecks that limit the provision of ultra low latency services in 5G centralized cloud environment. We focus on how to efficiently provide data to users using caching storage of MEC in 5G distributed cloud environment. In particular, we propose how to calculate user similarity using user's content request pattern for efficient MEC content caching. We also expand the caching space available to users by adopting Cooperative MEC content caching to increase the hit ratio of requested contents and to reduce the latency for improving the quality of service(QoS).

1. 서론

최근 디바이스들의 성능이 크게 향상되었고 스마트폰 사용 시간의 증가와 더불어 초고음질, 초고화질 스트리밍 서비스의 사용량과 사용시간이 폭발적으로 증가하였다. 또한 인터넷에 접속하는 모바일 디바이스의

양이 증가하고 있다. 이러한 통신 수요를 해결할 수 있는 기술로 5G가 주목받고 있다. 5G는 차세대 무선 통신 기술로 초고속, 초저지연, 초연결을 특징으로 한다. 5G 서비스에 대한 사용자의 기대가 커지면서 사용자들의 서비스 지연시간에 대한 요구사항은 엄격해졌다. 하지만 현재 5G 망에서 사용하고 있는

* 이 성과는 2018년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. NRF-2018R1A2B60025 05)

• First Author : Sookmyung Women's University Department of IT Engineering, ozo789@sookmyung.ac.kr, 학생(석사), 학생회원

^o Corresponding Author : Sookmyung Women's University Department of IT Engineering, yujin91@sookmyung.ac.kr, 정교수, 종신회원

논문번호 : KICS202008-194-B-RN, Received August 11, 2020; Revised September 29, 2020; Accepted October 6, 2020

Centralized Radio Access Network(C-RAN) 방식은 사용자의 요구사항을 충족시키기 어렵다. C-RAN 방식은 모든 트래픽이 반드시 중앙 서버에 집결된 후 처리되는 형태를 의미한다. 하지만 수많은 트래픽 요청들을 모두 중앙 서버에 집결시키는 과정에서 지연을 발생시켜 5G 서비스의 초저지연성 구현을 어렵게 한다.

이러한 이유로 Multi-access Edge Computing (MEC) 방식이 주목받게 되었다. MEC는 기지국에 설치되어 사용자와 보다 가까운 곳에서 서비스를 제공한다. 이는 중앙 서버보다 훨씬 가까운 뿐만 아니라 트래픽이 중앙 서버에 집결되는 과정 없이 기지국과의 무선 통신을 통해 서비스를 제공하므로 대역폭 절감, 보안 강화, 지연 저하, 에너지 절감 등의 효과를 가져올 수 있다¹¹.

MEC에 캐시(cache)를 설치하면 사용자가 사용할 수 있는 저장 공간의 확장 효과도 볼 수 있다. MEC 캐시를 이용해 사용자가 요청하는 VR/AR, 스트리밍과 같은 큰 용량을 가진 콘텐츠를 MEC 단에서 처리함으로써 중앙 서버로의 요청 없이 바로 제공할 수 있게 된다. 그 중에서도 비디오 콘텐츠를 MEC에 캐싱해 두는 것은 비디오 콘텐츠가 용량이 크고, 일부 콘텐츠에 대한 반복적인 요청이 이루어진다는 점에서 좋은 효율을 낼 수 있을 것으로 기대되고 있다.

특히 각 MEC의 캐시 용량이 제한적인 것을 보완하기 위하여 등장한 이웃한 MEC와의 공동 작업을 진행하는 환경인 Cooperative MEC¹²에서의 비디오 캐싱을 통해 사용자가 속한 MEC 캐시에 존재하지 않는 콘텐츠를 전달받거나, 용량이 큰 비디오를 여러 MEC에 나누어 저장하여 가져오는 방법들이 함께 주목받고 있다.

MEC 캐시 적중률(hit ratio)을 극대화하기 위해서는 캐시 내 최적화된 콘텐츠 배치가 필수적이다. 본 논문은 사용자의 콘텐츠 요청에 대하여 MEC 캐시 적중률을 극대화하고, 지연시간(latency)을 최소화하기 위한 기법으로 사용자의 콘텐츠 소비 패턴을 분석하는 방법을 제안하고 이의 적용을 Cooperative MEC 환경으로 확장하여 사용자가 사용할 수 있는 캐시를 최대화하는 방법을 제안한다.

본 논문의 구성을 다음과 같다. 2장에서는 MEC 콘텐츠 캐싱에 대한 관련 연구를 설명하고, 3장에서는 제안된 기법에 대해 설명한다. 4장에서는 기존 기법들과의 성능 비교를 통해 제안된 기법의 향상된 효율성을 증명하고, 5장에서 결론을 맺는다.

II. 관련 연구

MEC 콘텐츠 캐싱은 사용자가 콘텐츠를 요청하였을 때 중앙 서버로의 연결 없이 사용자 기기와 MEC와의 무선 연결만으로 사용자에게 콘텐츠를 제공할 수 있는 기술이다. MEC 콘텐츠 캐싱은 적중률, 즉 사용자가 요청한 콘텐츠가 현재 캐시 내에 존재하는지 여부로 그 효율성을 판단한다. 적중률은 사용자들의 전체 요청 수에 대해 해당 콘텐츠가 캐시 내에 존재하였던 수를 백분율로 표현한 수치이다. 따라서 적중률을 높이기 위해서는 어떤 콘텐츠가 요청될 것인지를 예측하는 것이 주요 관건이다. 하지만 어떤 콘텐츠가 가장 많이 호출될지를 예측하는 것은 매우 어렵다. 적중률을 높이기 위하여 대부분의 연구는 가장 인기 있는 콘텐츠가 가장 많이 요청될 것이라 판단하고¹³, 가장 인기 있는 콘텐츠로 캐시를 구성한다^{4,5}.

이에 요청 빈도와 횟수를 기준으로 하는 기존의 캐시 교체 알고리즘 Least Recently Used(LRU)^{16,7}와 Least Frequently Used(LFU)¹⁷가 MEC 캐시 교체 알고리즘으로 사용되었다. LRU는 요청된 콘텐츠가 캐시에 존재하지 않을 경우 캐시에 삽입하고 캐시에 존재할 경우 우선순위를 높인다. 캐시 메모리 초과 시 가장 오랫동안 요청되지 않은 콘텐츠를 삭제한다. LFU는 콘텐츠의 요청 횟수를 저장하여, 요청 횟수가 높은 콘텐츠들을 캐시에 담는다. 캐시 메모리 초과 시 요청 횟수가 가장 작은 콘텐츠를 삭제한다. 하지만 인기 있는 콘텐츠로 캐시를 구성하는 것은 크기가 작아 저장과 삭제가 매우 자주 일어나는 구조인 MEC에서 치명적인 문제를 유발한다. LRU의 경우, 자주 요청되는 중요한 콘텐츠가 최근 요청되지 않아 삭제될 수 있다는 단점을 가지고 있다. LFU의 경우, 콘텐츠가 인기 있다고 판단되어 캐싱되기 전까지는 요청들에 효과적으로 대응할 수 없다는 점에서 콘텐츠 유형의 흐름을 반영하지 못한다는 단점이 존재한다.

이러한 단점들을 보완하기 위하여 캐시를 나누어 관리하는 방법이 제안되었다. 해당 방법은 MEC 클러스터를 여러 클러스터로 나누고, 각 클러스터에 캐시를 할당하여 관리하는 방법이다. 이러한 방법 중 하나로 Segmented Least Recently Used(SLRU)¹⁸가 제안되었다. SLRU를 활용한 MEC 캐싱¹⁹은 캐시를 N개의 클러스터로 나누어 요청된 콘텐츠가 캐시에 존재하지 않을 경우 클러스터 0에 삽입하고, 콘텐츠가 클러스터 n-1에 존재할 경우 자주 요청되는 콘텐츠로 판단, 상위 클러스터 n에 교체하여 저장한다 (0, 1, 2, ..., n < N). 클러스터 메모리 초과 시 LRU 기법과 같이

최하위 클러스터에서 가장 오랫동안 요청되지 않은 콘텐츠를 삭제하거나, 상위 클러스터에서 아래 하위 클러스터로 이동하여 저장한다. 해당 기법은 기존 LRU가 중요 콘텐츠를 쉽게 삭제할 수 있다는 단점을 보완한다. 하지만 콘텐츠를 기반으로 캐시를 나누어 관리하는 기법들은 사용자들의 요청 패턴과 같은 콘텐츠 요청 형태를 파악할 수 없어, 콘텐츠 유형의 흐름을 캐시에 반영하기가 어렵다는 단점이 여전히 존재한다. 또한 MEC 캐시 크기를 캐시 개수 N 으로 나누어 고정된 크기의 캐시를 사용함으로써, 효율적인 캐시 사용이 어렵다는 단점이 존재한다.

이러한 기법의 문제점을 해결할 수 있는 방안으로 제안된 것이 사용자 기반 분배기법이다¹⁰⁾. 사용자 기반 분배기법은 MEC가 담당하고 있는 지역 내에 있는 사용자들을 클러스터링하고, 각 클러스터에 캐시를 분배하는 방법이다. 사용자들의 콘텐츠 요청 패턴 정보를 기반으로 클러스터링을 진행하여 사용하기 때문에, 콘텐츠 유형의 흐름을 캐시에 반영할 수 있고, 클러스터의 크기를 클러스터에 속한 사용자 수와 비례하게 설정할 수 있어 캐시의 크기를 유동적으로 활용할 수 있다. 뿐만 아니라 같은 클러스터에 속한 사용자들은 콘텐츠 요청에 있어 유사점을 가지고 있기 때문에, 클러스터 별 콘텐츠 요청 유형에 따른 개별 관리가 가능해진다.

그러나 사용자 기반 분배기법에서 가장 중요한 문제는 어떻게 효과적으로 사용자를 분배하는가이다. 본 논문에서는 사용자를 콘텐츠 요청 유형 별로 분배하여 클러스터를 구성하고, 이를 기반으로 하여 캐시의 최고 효율을 이끌어내는 방법을 제안하고자 한다.

III. 제안 기법

본 논문은 MEC 캐시의 효율성을 높이기 위한 방법으로 사용자 기반 분배기법을 사용하고자 한다. 그리고 사용자 기반 분배기법의 분배기준으로 사용자가 콘텐츠를 소비하는 패턴인 사용자의 콘텐츠 요청 다양성과 요청 반복도를 사용한다.

사용자가 콘텐츠를 소비하는 패턴이 유사하다고 판단되는 사용자 그룹을 ‘사용자 유사도’가 높은 그룹이라고 정의하였고, 이를 활용한 분배기법을 ‘사용자 유사도 기반 분배기법’이라고 칭하였다. 사용자 유사도 기반 분배기법은 캐시를 나누어 사용하여 중요 콘텐츠가 쉽게 삭제되는 것을 방지하고, 콘텐츠 유형의 흐름을 캐시에 반영할 수 있으며 사용자의 콘텐츠 소비 형태를 모두 가지고 있기 때문에 캐시의 효율적인 관

표 1. 알고리즘 파라미터
Table 1. Algorithm parameter

Variables	Description
M_k	M_k indicates the MEC which is responsible for area k
CM	Cooperative MEC which is consisted of nearby MECs ($CM = \{M_1, M_2, \dots, M_k\}$)
C_k	C_k is the cache of M_k C_k is divided into J subclusters $C_k^j, j \in J$
Limit	Limit of mismatch rate between requested MEC and MEC that provided contents

리가 가능해진다. 자세한 내용은 아래와 같다.

사용자 유사도 기반 분배기법은 콘텐츠 클러스터링과 사용자 클러스터링이라는 두 번의 클러스터링 과정을 거친다. 먼저 콘텐츠 클러스터링을 위하여 콘텐츠 간 장르에 대한 Cosine similarity^[11] 연산을 통해 ‘콘텐츠 간 유사도’를 계산한다. 그리고 콘텐츠 간 유사도를 활용한 Spherical k-means 클러스터링^[12]을 통하여 콘텐츠들의 클러스터링이 이루어진다. Spherical k-means 클러스터링의 최적의 클러스터 개수 k 를 계산하는 것은 성능에 큰 영향을 끼친다. 본 논문은 적절한 클러스터의 개수를 결정하기 위하여, 게임 이론 등에서 발생 가능한 손실을 최소화하고자 하는 개념으로 주로 사용되는 Minimax 방법을 활용한다^[13]. 이는 적용 시 초기 클러스터의 개수를 지정하지 않을 수 있으며, 가장 효율적인 클러스터의 개수를 유동적으로 지정할 수 있다는 장점이 있다. 본 논문에서는 위 방법을 통해 결정된 클러스터의 개수를 k-means 클러스터링의 최종 클러스터 개수로 지정한다 (표 2, Line 1).

다음으로 사용자 클러스터링은 사용자의 콘텐츠 요청 목록들을 기반으로 사용자를 분류하기 위해 사용되는 분배기법이며 Hierarchical 클러스터링^[14]을 사용한다. 이 클러스터링은 앞서 이루어진 콘텐츠 클러스터링의 결과를 활용한 ‘클러스터 다양성’, ‘클러스터 요청 반복도’를 이용한다. 콘텐츠 다양성은 사용자가 요청하는 콘텐츠의 클러스터의 다양성을 뜻하며, 콘텐츠 요청 반복도는 해당 사용자가 한 클러스터를 여러 번 반복하여 요청했던 정도를 의미한다. 클러스터 다양성과 클러스터 요청 반복도를 통해 사용자의 콘텐츠 소비 패턴을 파악할 수 있다(표 2, Line 2-4). Hierarchical 클러스터링의 과정은 다음과 같다.

- 1) 모든 사용자는 요청 콘텐츠 클러스터 종류, 요청

표 2. Cooperative MEC 내 사용자 유사도 기반 분배기법
Table 2. Content caching algorithm based on user similarity in Cooperative MEC

Algorithm	
1:	Cluster contents based on contents genre and production year
2:	Calculate repeatability and diversity of contents cluster in user's requests
3:	Calculate user's similarity based on the repeatability and diversity
4:	Cluster users based on the similarity
5:	C_k^j is divided by the number of user clusters in the area and the size of each cluster is proportion to the number of users in the cluster
6:	for each contents request r :
7:	/* when a user who requested r is in coverage of M_k^* /
8:	if r is in C_k^j
9:	rank r at the top of C_k^j and provide to the user
10:	elif r is in $C_{k'}^j$ ($k' \neq k, C_{k'}^j \in CM$)
11:	rank r at the top of $C_{k'}^j$ and provide to the user
12:	calculate mismatch rate m_{rate} between requested MEC and MEC that provided contents
13:	if $m_{rate} \geq Limit$:
14:	go back to line 1
15:	/* for reclustering */
16:	else :
17:	/*when requested content is not located in CM cache */
18:	provide r to the user from Datacenter
19:	put r to top of C_k^j

반복도에 대한 값을 가지고 있는 하나의 클러스터로 지정된다.

2) 각각의 클러스터가 가지고 있는 요청 콘텐츠 클러스터 종류, 요청 반복도 값에 대한 거리 계산을 통해 가장 가깝다고 판단되는 두 개의 클러스터가 병합하는 과정을 반복한다.

3) Minimax 방법을 통해 결정된 최적 군집의 수가 되면 클러스터 병합을 중지한다.

이와 같은 과정을 거쳐 최종 클러스터의 개수, 각 클러스터에 속한 사용자의 수가 정해진다. 본 논문에서는 각 클러스터의 크기에 비례하게 캐시를 나누는 것으로 계산한다.(표 2, Line 5)

이렇게 만들어진 MEC 캐시를 사용하여 사용자의 콘텐츠 요청을 처리한다. 사용자가 요청한 콘텐츠가

MEC 캐시 내에 존재하지 않을 경우 중앙 데이터 센터에서 데이터를 가져와 사용자에게 제공하고, 요청이 발생한 MEC 캐시에 해당 콘텐츠를 저장한다(표 2, Line 6-16).

이러한 사용자 기반 콘텐츠 캐싱은 Cooperative MEC 환경으로 확장하여 적용된다. Cooperative MEC 환경은 사용자가 위치한 지역의 MEC뿐만 아니라, 주변의 MEC를 함께 사용할 수 있게 되는 것으로 캐싱 공간을 확장하여 적중률 향상 및 지연시간 저하가 매우 뛰어나다. 본 논문은 사용자가 요청한 콘텐츠가 위치한 MEC 캐시 내에 존재하지 않을 경우, 같은 Cooperative MEC에 속한 다른 MEC 캐시를 먼저 확인한다. 콘텐츠가 Cooperative MEC에 속한 다른 MEC 캐시에 저장되어 있을 경우 해당 콘텐츠를 사용자에게 제공한다. 콘텐츠 요청이 발생한 MEC와 콘텐츠가 제공된 MEC가 다른 경우가 일정 횟수 이상 발생할 경우 재 클러스터링을 진행하였다(표 2, Line 12-13).

또한 Cooperative MEC 내 사용자는 사용자들의 MEC 유입, 유출에 있어 자유롭다. 사용자 기반 콘텐츠 캐싱의 경우 사용자의 요청 콘텐츠는 계속해서 변화하나 사용자가 콘텐츠를 소비하는 패턴은 쉽게 변하지 않기 때문에 다른 방법에 비하여 클러스터링 재구성 필요성이 낮아 효율성이 좋다. 하지만 클러스터 간 크기의 편차가 심해지거나, Cooperative MEC에 대한 요청 의존도가 높아질 경우 클러스터를 재구성하였다.

IV. 성능평가

이번 장에서는 본 논문에서 제안한 Cooperative MEC 환경을 적용한 사용자 유사도 기반 분배기법의 성능을 평가한다.

본 논문은 MEC 콘텐츠 캐싱 환경에서 캐시 교체 정책의 변화에 따른 요청 데이터 적중률과 지연시간을 측정하기 위하여 Python 기반의 시뮬레이터를 구성하여 활용하였다. 실험에 사용된 데이터셋은 MovieLens 데이터셋^[15]으로, 해당 데이터셋은 138,493명의 사용자의 27,278편의 영화에 대한 약 20년간의 시청 기록으로 이루어져있다. 본 논문은 해당 데이터셋의 사용자 중 7,800명의 사용자의 영화 요청 기록을 사용하여 분석하였다. 각 MEC의 서비스 범위 내 사용자의 수는 평균 1,300명으로 6개의 MEC에 골고루 분포되어있다고 가정하였다. 사용자의 콘텐츠 요청은 MovieLens 데이터셋에서 제공하는 사용자의 콘

텐츠 요청 기록을 사용하였다. MEC 캐시의 크기는 사용자의 전체 콘텐츠 요청 대비 2%-10%이다. 콘텐츠는 모두 같은 크기로 나누어져있다고 가정한다.

본 논문의 실험 환경은 표 3과 같이 설정하였다. MEC를 소형 기지국에 설치한다고 가정하였고, MEC가 담당하는 구역은 소형 기지국이 담당하고 있는 구역과 같다고 가정한다. 6개의 소형 기지국은 Cooperative MEC로 서로 연결되어 서로의 캐시 내역에 대해 알고 있다고 가정했다. MEC 캐시에는 사용자가 요청한 영화 데이터가 같은 크기로 분할되어 저장되며 실험은 본 논문에서 제시한 사용자 유사도 기반 분배기법과 Cooperative MEC 내 사용자 중심 네트워크 분배기법, 그리고 LFU, SLRU으로 진행되었다. 실험 내 적중률은 사용자의 콘텐츠 요청 중 데이터센터로의 접근 없이 콘텐츠를 제공할 수 있었던 비율을 뜻한다. 그리고 지연시간은 사용자의 콘텐츠 요청이 발생한 후부터 사용자에게 콘텐츠가 제공되는 시간을 뜻한다.

그림 1은 MEC 캐시 크기가 사용자 요청량에 따라 변화할 때의 적중률과 지연시간이다. Cooperative MEC 내 사용자 유사도 기반 분배기법은 사용자 유사도 기반 분배기법에 비하여 최대 약 26% 높은 적중률과 약 20% 이상 감소된 지연 시간을 보였다. LFU에 비하여 최대 약 38% 높은 적중률과 약 30% 이상 감소된 지연시간을 보였으며, SLRU에 비하여 적중률 약 33% 향상과 지연시간을 약 26%가량 감소시켰음을 알 수 있다. 이는 사용자 기반 분배기법이 사용자가 콘텐츠 소비 형태로 MEC 내 클러스터를 나누어 사용하기 때문에, 클러스터 내 중요 콘텐츠가 쉽게 삭제되지 않으며, 클러스터 내 사용자의 콘텐츠 유행이

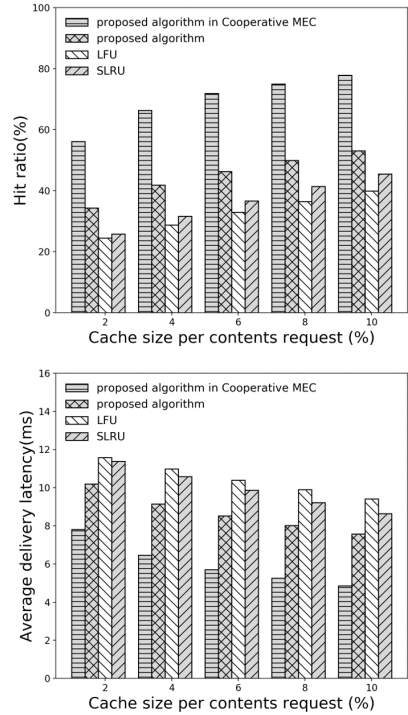


그림 1. Nc=20,000 , Nr= 30,000에 대한 적중률과 지연시간
Fig. 1. Hit ratio and Latency when Nc=20,000 and Nr= 30,000

캐시에 반영되어 콘텐츠 호출 횟수로 캐시를 구성하는 LFU, SLRU에 비하여 높은 효율을 보였음을 할 수 있다.

그림 1의 환경에서 진행된 실험인 그림 2는 콘텐츠 요청 대비 캐시의 사이즈를 콘텐츠 요청 대비 최대 100%까지 증가시킨 실험이다. 캐시 크기가 작은 MEC 환경에 맞게 구성된 사용자 유사도 기반 분배기

표 3. 실험 파라미터 값

Table 3. List of Simulation Parameters

Parameter	Description	Value
Nm	Number of Cooperative MEC	6
Nc	Number of Contents	
Nr	Number of Contents Requests	
St	Size of Test Case	25% of Nr
T1	Latency for data delivery from MEC to User	0.5-1.5ms
T2	Latency for data delivery from Cooperative MEC to User	2-4ms
T3	Latency for data delivery from Datacenter to User	10-20ms

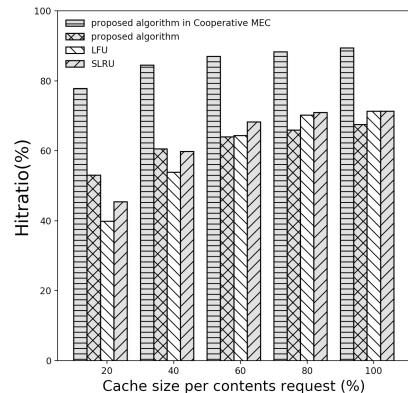


그림 2. Nc=20,000 , Nr= 30,000 의 경우 캐시 크기를 100%까지 증가시킬 때 적중률
Fig. 2. Hit ratio when increasing the cache size by 100% when Nc=20,000 and Nr= 30,000

법은 콘텐츠 요청 대비 캐시의 사이즈가 40% 이상으로 커질 경우 타 기법에 비하여 성능 수준이 낮아짐을 알 수 있다. 또한 실제 MEC 캐시의 크기는 총 콘텐츠 요청의 10% 정도에 불과한 수준이나, 이를 100%까지 증가시키면서도 Cooperative MEC 내 사용자 유사도 기반 분배기법의 적중률은 90%를 넘지 못하며, 사용자 유사도 기반 분배기법, LFU, SLRU의 적중률은 최대 약 72%를 넘지 못하는 포화 상태를 보인다.

그림 3에서 캐시 사이즈가 작을 경우, 고정되어 나누어진 캐시 클러스터를 가지고 있어 클러스터 저장 공간이 더욱 작아져 상대적으로 낮은 효율을 보이는 SLRU에 비하여 클러스터의 크기가 유동적으로 사용자 집단의 크기에 따라 변하는 사용자 유사도 기반 분배기법이 콘텐츠 삭제가 자주 발생하지 않아 높은 효율을 보였다. 또한 Cooperative MEC 내 사용자 유사도 기반 분배기법은 사용자의 요청 콘텐츠가 Cooperative MEC 내에 존재할 경우 사용자에게 콘텐츠를 제공할 수 있기 때문에 높은 적중률과 지연시간 저하를 보였음을 알 수 있다. 그림1과 그림3의 비교를 통해 사용자의 요청이 많아질수록, 실험 내 전체 기법들이 약 10% 이상 향상된 적중률과 지연 시간을 보임

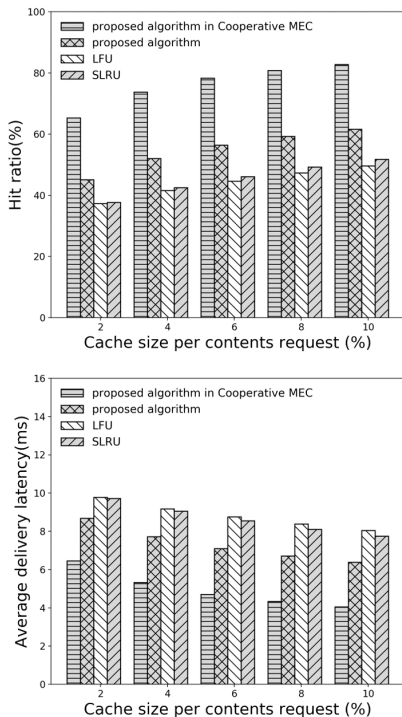


그림 3. Nc=20,000 , Nr= 50,000에 대한 적중률과 지연시간 Fig. 3. Hit ratio and latency when Nc=20,000 and Nr=50,000

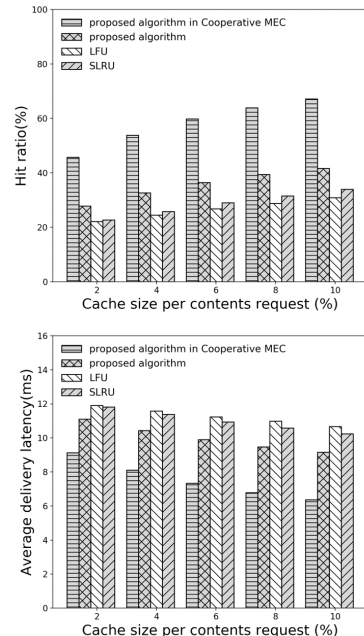


그림 4. Nc=10,000 , Nr=30,000에 대한 적중률과 지연시간 Fig. 4. Hit ratio and latency when Nc=10,000 and Nr=30,000

을 알 수 있었다.

그림 1과 그림 4의 비교를 통하여 콘텐츠 수(N_c)의 변화에 대한 성능 차이를 평가할 수 있다. Cooperative MEC 내 사용자 유사도 기반 분배기법은 콘텐츠 수가 많을 때 약 13% 높은 적중률과 약 10%의 지연시간 저하를 보였으며, 사용자 유사도 기반 분배기법은 약 9% 높은 적중률과 약 8%의 지연시간 저하 효과를 보였다. LFU와 SLRU는 콘텐츠 수의 변화에도 적중률과 지연시간 모두 유사했다. 이는 콘텐츠의 개수가 많아질수록 사용자 유사도 기반 분배기법이 재요청되지 않을 콘텐츠를 효과적으로 예측하고 있음을 알 수 있다. 이를 통해 Cooperative MEC 환경을 적용한 사용자 유사도 기반 분배기법이 콘텐츠 수가 다양하고, 사용자의 요청이 많은 MEC 환경에 효율적인 기법임을 확인할 수 있다.

그러나 Cooperative MEC를 효과적으로 유지하기 위해서는 Cooperative MEC로 포함되는 MEC의 범위를 효율적으로 지정해야 한다. 또한 각자의 MEC에 존재하는 콘텐츠에 대한 정보를 공유할 수 있는 백홀 환경이 존재해야 하며, 캐시 내의 콘텐츠 정보를 주고 받는 것에 대한 네트워크 오버헤드가 발생할 수 있다.

V. 결론

본 논문에서는 MEC 캐시를 가장 효율적으로 사용하기 위하여 사용자 유사도 기반 분배기법을 제안하였다. 사용자 유사도 기반 분배기법은 클러스터를 나누어 구성할 때 그 기준을 사용자 간 유사도를 활용하는 방법으로 제안 기법에서는 사용자 유사도를 계산하기 위하여 K-means 클러스터링을 통한 콘텐츠 분류를 진행하였다. 이의 결과 값을 활용하여 유사도와 거리를 계산하여 Hierarchical 클러스터링을 활용하였고 결과적으로 사용자들이 적절한 개수의 클러스터로 나누어지면 이를 기반으로 캐시를 구성하는 알고리즘을 제안하였다. 그리고 이러한 클러스터의 구성 과정을 클러스터의 크기 편차에 따라 재진행되게 하여, 캐시의 효율성을 높였다. 실험을 통하여 제안기법이 타 기법에 비해 적중률과 지연시간 감소 측면에서 최대 약 38% 이상 향상된 성능을 보임을 확인하였다. 그리고 콘텐츠의 수가 많고, 사용자 요청이 많은 경우 더욱 효과적인 실험 결과를 보여주었다.

향후 본 연구는 미래에 요청될 가능성이 높은 콘텐츠를 예측하기 위한 방법으로 추천 시스템을 도입하고, 콘텐츠의 인기도를 분석 및 예측하기 위한 방법으로 강화학습을 도입하여 연구를 발전시켜나가고자 한다.

References

- [1] D. Huang, X. Wang, and D. Wang, "Caching scheme based on user clustering and user requests prediction in small cells," in *2017 IEEE ICCT*, pp. 970-974, Chengdu, China, Nov. 2017.
- [2] Z. Chen, J. Lee, T. Q. S. Quek, and M. Kountouris, "Cooperative caching and transmission design in cluster-centric small cell networks," in *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 3401-3415, May 2017.
- [3] N. M. Markovich and U. R. Krieger, "A caching policy driven by clusters of high popularity," in *2016 IWCWC*, pp. 363-368, Paphos, Cyprus, Sep. 2016.
- [4] X. Wang, M. Chen, T. Taleb, and A. Ksentini, "Cache in the Air: Exploiting content caching and delivery techniques for 5G systems," in *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 131-139, Feb. 2014.
- [5] C. Bernardini, T. Silverston, and O. Festor, "Mpc: Popularity-based caching strategy for content centric networks," in *2013 IEEE ICC*, pp. 3619-3623,

- Budapest, Hungary, Jun. 2013.
- [6] J. Liu, G. Wang, T. Huang, J. Chen, and Y. Liu, "Modeling the sojourn time of items for in-network cache based on LRU policy," in *China Commun.*, vol. 11, no. 10, pp. 88-95, Nov. 2014.
- [7] M. Zhang, H. Luo, and H. Zhang, "A survey of caching mechanisms in information-centric networking," in *IEEE Commun. Surv. & Tuts.*, vol. 17, no. 3, pp. 1473-1499, Apr. 2015.
- [8] R. Karedla, J. S. Love, and B. G. Wherry, "Caching strategies to improve disk system performance," in *Computer*, vol. 27, no. 3, pp. 38-46, Mar. 1994.
- [9] L. Saino, "On the design of efficient caching systems," in *UCL*, 2015.
- [10] H. Fan, T. Zhang, J. Loo, and D. Liu, "Caching deployment algorithm based on user preference in device-to-device networks," in *2017 IEEE Global Commun. Conf.*, pp. 1-6, Singapore, Dec. 2017.
- [11] M. S. ElBamby, M. Bennis, W. Saad, and M. Latva-Aho, "Content-aware user clustering and caching in wireless small cell networks," in *11th ISWCS*, pp. 945-949, Barcelona, Aug. 2014.
- [12] K. Hornik, I. Feinerer, M. Kober, and C. Buchta, "Spherical k-Means Clustering," in *J. Statistical Softw.*, vol. 50, no. 10, Sep. 2012.
- [13] H. Wang, C. Weng, and J. Yuan, "Multi-feature spectral clustering with minimax optimization," in *2014 IEEE Conf. Comput. Vision and Pattern Recognition*, pp. 4106-4113, Columbus, OH, Jun. 2014.
- [14] S. C. Johnson, "Hierarchical clustering schemes," in *Psychometrika*, vol. 32, no. 3, pp. 241-254, Sep. 1967.
- [15] F. M. Harper and J. A. Konstan, "The MovieLens datasets: History and context," in *ACM Trans. Interactive Intell. Syst.*, vol. 5, no. 4, pp. 19:1-19:19, Dec. 2015.

김 희 수 (Heesu Kim)



2019년 8월 : 숙명여자대학교 IT
공학과 졸업
2019년 9월~현재 : 숙명여자대
학교 IT공학과 석사
<관심분야> Edge Computing,
캐싱, 강화학습
[ORCID:0000-0001-9783-9547]

임 유 진 (Yujin Lim)



2000년 2월 : 숙명여자대학교 전
산학과 박사
2013년 3월 : 일본 Tohoku Uni-
versity, Department of Infor-
mation Sciences 박사
2004년~2015년 : 수원대학교 정
보미디어학과 부교수
2016년~현재 : 숙명여자대학교 IT공학과 교수
<관심분야> 지능형 시스템, IoT, Edge Computing
[ORCID:0000-0002-3076-8040]