

CCN에서 콘텐츠·생성자 인기도 기반 효율적 캐시 기법

이동건* · 권태욱**

Content and Producer Popularity-Based efficient Cache Policy in CCN

Dong-Geon Lee* · Tae-Wook Kwon**

요 약

현대의 네트워크와 인터넷은 전례 없이 많은 양의 데이터 패킷을 처리하기 때문에 효율적인 트래픽 관리가 점점 더 어려워지고 있다. 이 문제를 해결하기 위해 콘텐츠 중심 네트워크(CCN)와 같은 차세대 네트워크 체계가 등장했다. CCN은 불필요한 트래픽을 최소화하기 위해 콘텐츠 중심의 데이터 처리에 중점을 두어 기존 인터넷 아키텍처는 주로 end-to-end 접근 방식을 사용하므로 여러 사용자가 동시에 콘텐츠를 요청하면 트래픽이 급증할 수 있으나 이에 반해 CCN은 네트워크 내 캐시 기능을 활용하여 트래픽을 최적화함으로써 사용자에게 콘텐츠를 제공하는 데 필요한 시간을 단축한다. 따라서 CCN에서는 효과적인 콘텐츠 관리를 위해 어떤 캐싱 정책을 적용할지 결정하는 것이 중요하다. 기존 연구에서는 히트 수, 생성자 인기도 등의 요소를 기반으로 다양한 캐싱 방법을 제안했으나 이러한 방법들은 콘텐츠 유형을 고려하지 않아 인기가 낮은 동영상상이 이미지에 밀려 삭제되는 상황이 발생하고, 이에 따라 네트워크 트래픽이 증가할 가능성이 있다. 이 연구에서는 콘텐츠 유형과 제작자 인기도를 구분하여 네트워크 트래픽과 응답 시간을 더욱 줄일 수 있는 캐싱 전략을 제안한다.

ABSTRACT

Modern networks and the internet now handle an unprecedented volume of data packets, making efficient traffic management increasingly challenging. To address this issue, next-generation network architectures like Content-Centric Networks (CCN) have emerged. CCN focuses on content-centric data processing to minimize unnecessary traffic. Traditional internet architectures primarily use an end-to-end approach, where multiple users requesting content simultaneously can lead to a surge in traffic. In contrast, CCN optimizes traffic by utilizing in-network cache functions, reducing the time required to serve content to users. Therefore, in CCN, it is crucial to determine which caching policy to apply for effective content management. Existing research has proposed various caching methods based on factors like Hit count and producer popularity. However, these methods often fail to consider content type, resulting in situations where less popular videos are removed in favor of images, potentially increasing network traffic. This study proposes a caching strategy that distinguishes between content type and producer popularity to further reduce network traffic and response time.

키워드

CCN, Cache Replacement Policy, Data TTL(: Time To Live)
콘텐츠 중심 네트워크, 캐시 교체 정책, 데이터 생존 시간

* 국방대학교 관리대학원(swsunws@naver.com)

** 교신저자 : 국방대학교 컴퓨터공학과

• 접수 일 : 2024. 07. 12

• 수정완료일 : 2024. 08. 27

• 게재확정일 : 2024. 10. 12

• Received : Jul. 12, 2024, Revised : Aug. 27, 2024, Accepted : Oct. 12, 2024

• Corresponding Author : Tae-Wook Kwon

Dept. Computer engineering, Korea National Defense University

Email : swsunws@naver.com

1. 서 론

현대 모바일 기기의 발전과 SNS·OTT의 발달로 콘텐츠 양이 방대해짐에 따라 콘텐츠 수요는 기하급수적으로 증가하여 기존 IP네트워크 체계에서는 현재 발생하는 네트워크 트래픽을 처리하기에 한계에 이르게 되었다. '23년 에릭슨 모빌리티 자료에 따르면 총 데이터 트래픽은 '22년 매달 평균적으로 97EB에서 '23년 평균 130EB로 1년 동안 34% 증가하였고 최종적으로 '29년에는 403EB로 증가 될 것 이라고 예측하고 있어 네트워크 트래픽 과중화 문제는 더욱 가속화 될 것임이 분명하여 이를 해결하기 위해 콘텐츠 중심 네트워크(Content Centric Network, CCN)가 등장하였다[01]. 기존 legacy체계에서는 IP를 기반으로 client가 정보생성자에게 직접 데이터를 받는 End-to-End 방식으로 다수의 client가 데이터를 동시다발적으로 요청하게 되면 과도한 트래픽이 발생된다. 그래서 인기 콘텐츠가 등장할 경우 서버와 회선에 병목현상이 발생하여 응답시간을 증가시키고 패킷이 손실되는 현상을 초래한다. 이에 반해 CCN에서는 콘텐츠 이름을 중심으로 In-network Cache기능을 통해 CCN 포워딩 엔진에 있는 CS(Content Store)에서 콘텐츠를 캐시하고 관리한다. 이는 client가 정보생성자에게 데이터를 직접 요청하는 횟수를 최소화하여 전체 네트워크 트래픽과 응답시간 감소에 도움을 준다.

CCN의 핵심은 제한적인 포워딩 엔진의 CS공간의 효율적인 활용성이다. 이것을 어떻게 활용하는지에 따라 CCN의 성능이 좌우되며, 적절한 캐시 교체정책을 통해 네트워크 성능을 향상 시킬 수 있다. 최근 캐시 교체정책은 참조횟수, 생성자의 인기도, 생성자와의 거리 등을 통해 TTL(: Time To Live)을 부여하는 캐시 교체정책을 제시하였다. 하지만 전 세계 인터넷 데이터 분석 기업인 Datareportal에 따르면 client가 구매하는 디지털 콘텐츠는 TV·영화 스트리밍 서비스(32%), 음악 스트리밍 서비스(24%), 음악 다운로드(20%) 순으로 집계되었다[02]. 해당 분석에서 알 수 있듯 현대 client들은 큰 데이터 용량을 가지는 콘텐츠를 희망하고 있고 이에 콘텐츠의 유형을 고려하지 않을 수 없는 현실이다.

따라서 본 논문에서 제안하는 캐시 정책은 생성자 인기도와 콘텐츠 유형별 인기도를 실시간·고정 인기

도로 분류하고 이를 통해 종합 TTL을 도출하여 네트워크 트래픽과 응답시간을 효율적으로 감소시키는 캐시정책을 제시한다.

II. 관련연구

2.1 CCN

CCN은 ICN(: Information Centric Network)에서 기초된 것으로 V.Jacobson에 의해 2006년 8월 처음 제기되었다[03]. CCN은 기존 IP 기반 라우팅 방식을 주소가 아닌 콘텐츠 이름을 통한 콘텐츠 중심 라우팅 방식으로 변경한 네트워크이다. 이는 legacy체계에서는 IP주소를 기반으로 client와 정보제공자를 찾아가는 목적지(Where)가 중요했다고 하면 CCN에서는 이름을 기반으로 콘텐츠(What)를 찾아가는 것이 중요하다. 그림 1과 같이 client1이 콘텐츠A를 Youtube에 요청하여 콘텐츠A가 전달 될 때 데이터가 지나가는 경로에 있는 노드에 콘텐츠를 저장함으로써 일정 시간 이후 동일한 콘텐츠를 client2가 요청하더라도 Youtube에 요청하지 않고 legacy노드인 D를 거쳐 노드B에서 제공받을 수 있다. 이를 통해 전체적인 네트워크 트래픽 감소효과를 얻을 수 있고 legacy체계의 노드들과 함께 운용도 가능하다.

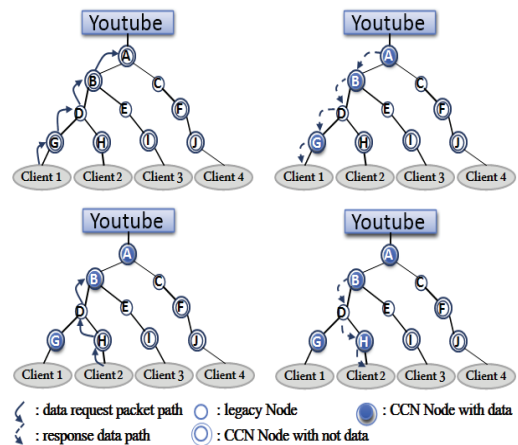


그림 1. CCN 프로토콜 수행 절차
Fig. 1 Procedure of CCN protocol

CCN 포워딩엔진은 그림 2와 같이 CS, PIT(: Pending Interest Table), FIB(: Forwarding Information Base)로 구성되며 각 노드에 설치되어 콘텐츠를 관리한다. CS는 캐시정책에 의해 수신된 콘텐츠를 저장하고 삭제하는 역할을 담당하고, PIT는 Interest Packet이 수신되었을 때 요구되는 콘텐츠가 CS에 없는 경우 해당 콘텐츠가 수신되어 client에게 Data Packet을 반환하기 전까지 콘텐츠의 이름과 요청 인터페이스 정보를 기억하는 역할을 담당하며, FIB는 CS와 PIT리스트에 요청받은 정보가 없을 때 Interest Packet을 인접 노드로 전파하기 위해 Face를 기록하여 목적지 정보를 저장하는 역할을 담당한다.

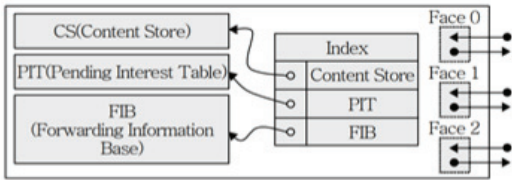


그림 2. CCN 포워딩 엔진
Fig. 2 CCN forwarding engine

2.2 CCN에서 캐시 정책

캐시정책은 프로세스들이 기억장치 내의 정보를 특정 시점에 시간·공간적 특정부분을 집중 참조하는 지역성 이론을 기초로 한다. 시간 지역성 이론은 가까운 미래에 최근 참조한 것이 다시 참조될 가능성이 높다는 것을 의미하고 공간 지역성 이론은 참조한 것과 가까운 것들이 계속 참조될 가능성이 높다는 것을 의미한다. CCN에서 캐시교체 정책이란 유한한 저장공간을 가지는 CS에 어떤 콘텐츠를 저장하고 삭제할지 결정하는 것으로 legacy 체계의 캐시 교체정책을 기초로 한다. 이에 다시 요청될 확률이 높은 콘텐츠가 CS에 오래 남아 있고, 다시 요청될 확률이 낮은 콘텐츠가 CS에 짧게 남아 있을수록 네트워크 효율은 높아지며 응답시간도 단축될 것이다. 따라서 네트워크에서 유통되는 데이터의 전체 양에 비해 각 라우터의 CS공간은 제한적이므로 어떤 캐시정책을 적용하는지에 따라 캐시의 효율성이 결정된다.

CCN 캐시 교체정책의 유형에는 표 1과 같이 각 라우터가 독립적으로 매커니즘을 수행하는 독립유형과 주변 라우터들과 다양한 방법을 통해 협력하여 매

커니즘을 수행하는 협력유형이 있다.

대표적으로 LRU(: Least Recently Used)는 최근에 사용된 콘텐츠가 우선순위를 가지고 저장공간이 Full일 경우 새로운 콘텐츠를 저장하기 위해 가장 참조 시간이 오래된 콘텐츠를 삭제하는 정책이다[04]. LFU(: Least Frequently Used)는 가장 많이 사용된 콘텐츠가 우선순위를 가지고, 저장공간이 Full일 경우 새로운 콘텐츠를 저장하기 위해 가장 적게 사용된 콘텐츠를 삭제하는 정책이다[05]. TTL-based(: Time To Live-base)는 콘텐츠에 일정 시간의 생존시간을 부여하고, 재참조 시 해당 시간을 재부여, 재참조가 없을 시 콘텐츠를 방출하는 정책이다[06].

표 1. CCN 캐시정책 분류
Table 1. Classification of CCN cache policy

Classification		Cache Policy
Independent Policy	Legacy	LRU, HLRU, LFU, LFU-Aging, FIFO, Priority-FIFO, LNC-R-W3
	Novel	TTL-based, DA-LFU, SLT-based, PR-based, RTP-CRP, RRT-PD, RT-based
Cooperative Policy		LCE, LCD, MCD, Prob, WAVE, PBLRU, PBLFU

언급된 정책들을 기초로 최근 연구는 보완·발전하여 새로운 정책들이 제안되고 있고 대표적으로 LFU를 기반으로 생존시간을 생성자 인기도에 따라 차등 부여하여 캐싱하는 RT-PP-LFU, 생존시간은 생성자 거리(홉 카운트)에 따라 차등 부여하여 캐싱하는 RRT-PD, 효율성·Qos 모두 향상시키기 위해 생성자 거리와 인기도의 적절한 비율을 제시한 PD, 생성자 거리, 생성자 인기도, 참조횟수에 따라서 생존시간을 차등으로 부여하는 RP-CRP 등이 있다[07][08][09][10].

III. 콘텐츠·생성자 인기도 기반 캐시 정책

3.1 기본개념

앞서 언급한 최근 연구 동향은 생성자 인기도, 거리 등 모두 생성자 중심으로 연구되었고 콘텐츠의 유형은 고려되어 있지 않다. 이에 대중들이 현재 동영상, 음악 등 콘텐츠 용량이 상대적으로 큰 콘텐츠 유형들을 많이 요구하고 있어 데이터 트래픽이 많이 유

발되는 콘텐츠 유형들을 캐시정책 수립 간 반드시 고려해야 할 필요가 있다. 즉, 시간 지역성 이론과 공간 지역성 이론을 동시에 만족할 수 있게 인기 있는 생성자와 콘텐츠 유형에 따라 생존시간을 차등 부여하는 정책이 필요하다.

제안하는 정책은 콘텐츠의 인기도를 기반으로 생존 시간을 부여하여 캐시 교체정책을 수행한다. 인기도는 콘텐츠 유형 인기도와 생성자 인기도로 크게 구분하고 각각 고정(기본)인기도와 실시간 인기도의 종합값으로 산출한다. 생성자 고정 인기도는 인터넷 검색 시 이용 서비스 자료에서, 콘텐츠 유형 고정 인기도 콘텐츠 이용률 자료에서 연령대별 점수의 평균으로 세분화하였다[11][12]. 이는 생성자·콘텐츠 유형 인기도의 평균의 기준치를 제공하여 초기 적용 시에도 원활히 동작할 수 있는 기초가 된다. 실시간 인기도는 해당 시각에 얼마나 인기가 있는지를 판단하는 척도로 생성자·콘텐츠 유형을 CS 내 해당 생성자·콘텐츠 유형 수와 CS내 전체 콘텐츠 수의 비율로 인기도를 산출하여 실시간 변화 시에도 유연한 네트워크 환경 제공이 가능하다.

이에 최종적으로 종합 생존시간을 앞서 분류한 4가지 요소를 종합하여 최종 생존시간을 산출하여 CS에 저장되는 콘텐츠에 부여할 수 있도록 하는 방법을 제안한다. 제안 정책의 Flow-chart는 그림 3과 같으며 제안 정책을 콘텐츠·생성자 인기도 기반 캐시 정책(Content and Producer Popularity-Based Cache Replacement Policy, C&P-CRP)으로 정의하였다.

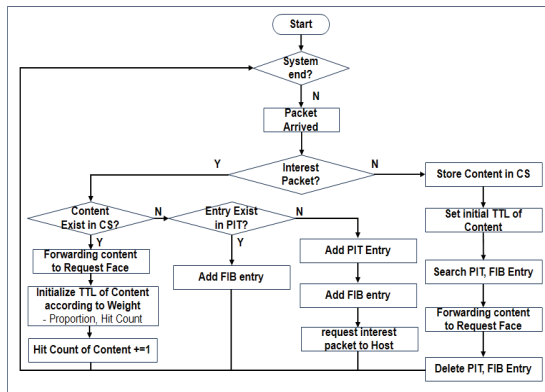


그림 3. C&P-CRP 흐름도
Fig. 3 C&P-CRP flow chart

3.2 세부 수행절차

C&P-CRP는 Interest 패킷이 CCN 포워딩 엔진에 도착할 때 CS내 콘텐츠의 유무에 따라 분류할 수 있다. CS에 콘텐츠가 있을 경우에는 ① Data packet을 요청face로 전송하고 ② 해당 콘텐츠에 종합 생존시간을 재부여한 후 ③ 콘텐츠의 Hit count를 가산하는 절차가 수행된다. CS에 콘텐츠가 저장되어 있지 않을 때에는 PIT 목록 기재 유무에 따라 나뉘게 된다. PIT 목록에 있을 때에는 face를 FIB 리스트에 추가하는 것으로 절차가 종료되지만, PIT 목록에 없을 때에는 ① PIT, FIB에 리스트 추가하고 ② Interest 패킷을 호스트(상위 노드)에 요청하는 절차가 이루어진다.

다음으로 Data 패킷이 CCN 포워딩 엔진에 도착할 때에는 ① CS에 콘텐츠를 저장하고 해당 콘텐츠의 종합 생존시간을 부여하고, ②PIT와 FIB에 기재되어 있는 해당 콘텐츠를 각 FIB의 Face로 전송한 후, ③ PIT와 FIB 목록에서 삭제하는 절차가 이루어진다.

C&P-CRP의 고정인기도는 사용하는 국가·지역에 따라 상이해 질 수 있고 본 논문은 통계청 자료 선호도를 통해 표2, 3과 같이 고정 인기도를 선정하였다.

표 2. 생성자 고정 인기도 분류

Table 2. Classification of producer static popularity

popularity weighting	0.4	0.3	0.2	0.1
services	Naver ...	YouTube ...	Google Daum Facebook	Yahoo others
Notes(%)	$x > 70$	$70 > x > 50$	$50 > x > 30$	$30 > x$

표 3. 콘텐츠 타입 고정 인기도 분류

Table 3. Classification of content type static popularity

popularity weighting	0.4	0.3	0.2	0.1
type	-	media	music picture document	others
Notes(avr)	$x > 5.5$	$5.5 > x > 5$	$5 > x > 3.5$	$3.5 > x$

실시간 인기도 변수의 적용과 종합 생존시간을 산출하기위해 CS의 구조를 변경해야 한다. 콘텐츠에 생

생성자 실시간 인기도, 콘텐츠 타입 인기도 기반의 생존 시간(TTL)을 부여해 주기 위해 그림 4.(a),(b)와 같이 CS구조를 변경하였고, 고정 인기도를 포함한 종합 생존시간을 산출하기 위한 CS의 변경은 그림 4.(c)와 같으며 종합 생존시간 산출식(1)은 다음과 같다.

$$total\ TTL = basic\ TTL + (1 + total\ P) \quad \dots (1)$$

종합인기도(totalP)의 산출식(2)은 다음과 같다.

$$total\ P = Producer\ Static\ P \times a\% + Producer\ Real\ time\ P \times b\% + Content\ Static\ P \times c\% + Content\ Realtime\ P \times d\%$$

basic CS Table

Name	Data	Hit count
youtube.com/Lee/A.mp4/seq...	...	4
naver.com/Seo/A.ppt/seq...	...	2
naver.com/Do/A.jpeg/seq...	...	5
google.com/Lee/A.avi/seq...	...	1



(a)

Name	Number of Content	total Content	Popularity
youtube.com	4	10	0.4
naver.com	3	10	0.3
google.com	2	10	0.2
...

(b)

Name	Number of Content	total Content	Popularity
.avi	4	10	0.4
.mp4	2	10	0.2
.ppt	3	10	0.3
...

(c)

Name	Producer Static Pop	Producer Realtime Pop	Content Static Pop	Content Realtime Pop	total TTL
youtube.com/Lee/A.mp4/...	0.3	0.4	0.3	0.2	...
naver.com/Seo/A.ppt/...	0.4	0.3	0.2	0.3	...
google.com/Lee/A.avi/...	0.2	0.2	0.3	0.4	...
...

- (a). PRT 구조
- (b). CRT 구조
- (c). DPIT 구조

그림 4. CS 구조변경

- (a). PRT structure
- (b). CRT structure
- (c). DPIT structure

Fig. 4 Changing CS structure

PRT(: Producer Real-time Table)는 해당 생성자 이름을 가진 콘텐츠를 관리하기 위한 테이블이다. 테이블에는 생성자 이름(producer Name)과 DPIT 내부에 저장되어 있는 해당 콘텐츠 생성자 이름을 가진 콘텐츠의 수(Number of Content), 전체 콘텐츠의 수(total Content), 전체 콘텐츠 대비 해당 생성자 비율로 산정된 인기도(Popularity)의 정보를 유지하고 있다. CS에 존재하는 콘텐츠가 다시 요청될 경우 콘텐츠 수의 변경은 없으므로 PRT 정보를 수정할 필요가 없다.

CRT(: Content Real-time Table)는 해당 콘텐츠 유형별로 콘텐츠를 관리하기 위한 테이블이다. 테이블에는 생성자 이름(producer Name)과 DPIT 내부에 저장되어 있는 해당 콘텐츠 유형과 같은 콘텐츠의 수(Number of Content), 전체 콘텐츠의 수(total Content), 전체 콘텐츠 대비 해당 콘텐츠 유형별 비율로 산정된 인기도(Popularity)의 정보를 유지하고 있다. PRT와 마찬가지로 CS에 존재하는 콘텐츠의 재요청 발생시 CRT 정보를 수정할 필요는 없다.

DPIT(: Detail Popularity Information Table)는 콘텐츠를 저장할 때 생성자·콘텐츠 유형별 고정·실시간 인기도를 관리하여 종합 생존시간을 부여한다. CS 공간이 Full이 되어 더 이상의 저장공간이 존재하지 않을 경우 최소 종합 생존시간을 가지고 있는 콘텐츠부터 순차적으로 삭제하며 이때, PRT와 CRT의 정보를 수정해야 한다. 콘텐츠 재참조 요청 발생시 콘텐츠 수의 변화는 발생하지 않으므로 PRT와 CRT의 정보를 재수신하여 종합 생존시간을 재부여한다.

IV. 실험 및 분석

4.1 실험환경

CCN 환경 구현을 위해 NS-3 기반의 ndnSIM을 사용하여 실험을 진행하였고 토폴로지 구성은 그림 5와 같다.

콘텐츠를 생성하는 Producer를 4개 배치하고, Producer들이 연결되어 있는 노드와 client들이 연결되어 있는 노드 2개로 구성하였으며 Target Node에서 CHR(: Cache Hit Ratio)을 측정하였다.

client의 요청은 Zipf - Mandelbrot(: MZipf) 분포를 적용하는 경우와 랜덤으로 요청하는 경우로 분류하였다. MZipf 분포는 콘텐츠 인기도를 나타내는데 가장 적합한 모델로 N 중 i 순위의 객체에 액세스 할 확률

은 식 (3)과 같이 산출한다.

$$p(i) = \frac{K}{(i+q)^\alpha} \quad \dots (3)$$

K는 식 (4)와 같이 산출할 수 있다.

$$K = \sum_{c=1}^N \frac{1}{(c+q)^\alpha} \quad \dots (4)$$

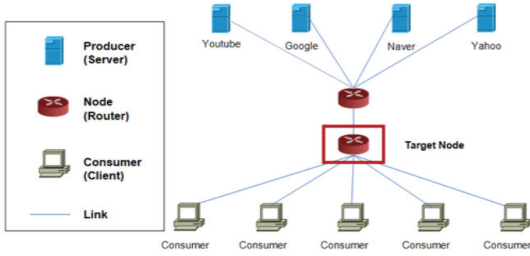


그림 5. 실험 토폴로지 구성
Fig. 5 Experimental Topology

skeness factor(α)는 데이터 분포의 비대칭성을 나타내는 것으로 MZipf 분포 곡선의 기울기를 제어할 때 사용되는데 α 가 높을수록 CHR이 높아지고, $q \geq 0$ 은 곡선의 평탄도를 나타낸다. 본 논문에서 skeness factor(α)는 0.7로 실험하였다.

실험결과는 CHR을 통해 네트워크 성능을 향상유무를 판단하며 CHR은 식 5와 같이 산출한다.

$$Cache Hit Ratio = \frac{\sum_{n=1}^N hit_j}{\sum_{n=1}^N (hit_j + miss_j)} \quad \dots (5)$$

표 4. 세부 실험환경
Table 4. Detailed experimental environment

Topology(nodes)	NDN Tested(2)
Simulator	ndnSIM
Simulation time	600 sec
Number of Content	10,000(2,500 Contents each)
Number of Content type	5(500 Contents each type)
Contents request method	Random, MZipf(Zipf $\alpha = 0.7$)
Contents Request rate	100 contents/s

Forwarding strategy	Best Route
Cache Size	5%
Basic TTL	15sec

세부 실험환경은 표 4와 같이 구성하였다. Target Node의 CS 캐시 용량은 전체 콘텐츠 용량의 5%이며 전체 콘텐츠 총 개수는 10,000개로 설정했다. 각 생성자 별로 2,500개의 콘텐츠를 분배하였으며 비디오, 음악, 이미지, 문서, 기타별로 각 500개씩 보유하고 있다. 랜덤으로 콘텐츠를 요구할 때는 생성자의 인기도를 별도 설정하지 않았지만 MZipf분포를 활용하여 콘텐츠를 요구할 때는 생성자에게 요청비율을 네이버 40%, 구글 20%, 유튜브 30%, 야후 10%로 각각 설정하였고 콘텐츠 유형별 요청은 MZipf분포에 의해 설정되었다.

4.2 랜덤 환경 실험 결과 및 분석

실험 대조군으로는 기본 정책인 LRU 정책과 LFU 정책을 선정하여 적중률을 비교하였다. 종합 생존시간을 위한 각 인기도의 가중치는 10% 단위로 모든 경우 측정을 완료하였고 가장 효율이 좋은 2개의 가중치를 선정하였으며 실험 결과는 그림 6, 데이터는 표 5와 같다.

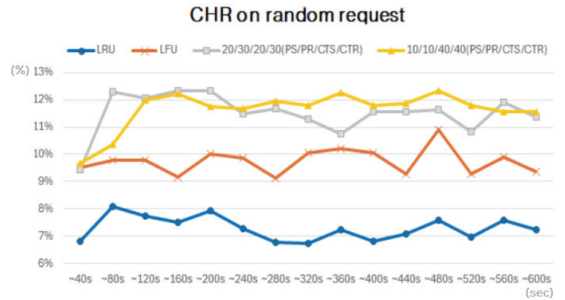


그림 6. 랜덤 요청에서 CHR 비교
Fig. 6 CHR comparison on random request

표 5. 랜덤 요청에서 세부 CHR
Table 5. detailed CHR on random request

Time(s)	LRU	LFU	C&P-CRP(PS, PR, CTS, CTR순)	
			20/30/20/30	10/10/40/40
~40	6.8	9.525	9.425	9.675
~80	8.075	9.775	12.275	10.375
~120	7.75	9.8	12.05	12

~160	7.525	9.175	12.325	12.225
~200	7.95	10.025	12.325	11.75
~240	7.275	9.85	11.5	11.675
~280	6.775	9.125	11.675	11.95
~320	6.75	10.05	11.3	11.775
~360	7.25	10.225	10.75	12.25
~400	6.825	10.05	11.55	11.8
~440	7.1	9.3	11.575	11.85
~480	7.575	10.9	11.65	12.325
~520	6.95	9.3	10.825	11.775
~560	7.575	9.9	11.9	11.55
~600	7.25	9.375	11.375	11.55

가장 CHR이 낮은 캐시 정책은 LRU로, 시물레이션 기간 평균 CHR은 7.3%가 나왔다. 제안한 정책인 C&P-CRP에서 생성자 고정 인기도(PS), 생성자 실시간 인기도(PR), 콘텐츠 타입 고정 인기도(CTS), 콘텐츠 타입 실시간 인기도(CTR)의 반영비율을 각각 10%, 10%, 40%, 40%로 설정하였을 때 평균 CHR은 11.64%로 가장 높았으며 LFU의 평균 CHR은 9.76%로 산출되었다. C&P-CRP는 LRU 대비 약 59% 가량 높고 LFU 대비 약 20% 높은 성능을 보임을 확인할 수 있었다.

시물레이션 초기 20초에는 LFU와 적중률이 비슷하다 랜덤요청으로 LFU는 많이 참조되었던 콘텐츠가 삭제되지 않는 데이터 오염현상으로 CHR이 낮아졌지만 C&P-CRP는 종합 인기도에 기반한 종합 생존시간을 부여하여 성능이 유지될 수 있었다고 생각된다.

4.3 MZip 분포 환경 실험 결과 및 분석

랜덤 요청과 동일하게 실험 대조군을 선정하였고 인기도의 가중치도 모든 경우 측정을 완료하였다.

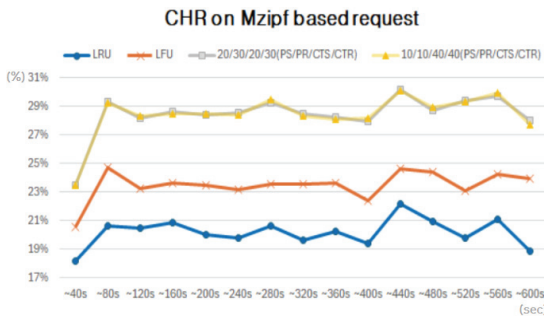


그림 7. MZip 요청에서 CHR 비교
Fig. 7. CHR comparison on MZip request

가장 효율이 좋은 2개의 가중치 비율을 선정하였는데 이는 랜덤 요청에서 실험 한 선정된 2개의 가중치 모델과 동일하다. 실험결과는 그림 7, 데이터는 표 6과 같다. 가장 CHR이 낮은 캐시 정책은 LRU로, 시물레이션 기간 평균 CHR은 20.18%가 나왔다. 제안한 정책인 C&P-CRP에서 각 인기도 가중치는 랜덤 요청 실험에서 선정된 각 인기도 반영비율인 10%, 10%, 40%, 40%로 설정하였을 때 평균 CHR이 28.45%로 가장 높았고 LFU의 평균 CHR은 23.49%를 기록하였다. 이는 C&P-CRP가 LRU 대비 약 40.1%가량 높고, LFU 대비 약 21% 높은 성능을 보임을 확인할 수 있다.

시물레이터 실험 한 콘텐츠 유형에 따른 MZip 분포에 의한 데이터 요청으로 시물레이션 시간이 지속될수록 누적 요청 수가 많은 콘텐츠가 지속 저장되어 있는 LFU는 일정 수준을 유지하는 것을 볼 수 있으나 LRU는 인기도와 관계없이 캐시를 교체하므로 CHR이 낮아졌다고 생각된다.

표 6. MZip 요청에서 세부 CHR
Table 6. detailed CHR on MZip request

Time(s)	LRU	LFU	C&P-CRP(PS, PR, CTS, CTR순)	
			20/30/20/30	10/10/40/40
~40	18.2	20.525	23.5	23.45
~80	20.6	24.75	29.325	29.3
~120	20.45	23.275	28.2	28.35
~160	20.9	23.625	28.625	28.525
~200	20.05	23.45	28.45	28.5
~240	19.775	23.2	28.6	28.45
~280	20.625	23.575	29.3	29.475
~320	19.625	23.55	28.475	28.325
~360	20.25	23.625	28.3	28.075
~400	19.425	22.425	27.975	28.15
~440	22.2	24.65	30.175	30.15
~480	20.95	24.375	28.75	28.975
~520	19.75	23.125	29.4	29.35
~560	21.075	24.275	29.7	29.95
~600	18.825	23.975	28.05	27.725

client가 원하는 생성자는 항상 변화하였다. 그래서 대중이 선호하는 생성자별 요청비율을 일부 수정하여 100초마다 네이버의 요청비율을 5%씩 하향시키고 야후의 요청비율을 5%씩 상향시켜보았다. 실험결과는 그림 8, 데이터는 표 7과 같다.

가장 CHR이 낮은 캐시 정책은 LFU로, 시물레이션 기간 평균 CHR은 19.27%가 나왔다. C&P-CRP에서 인기도 가중치는 이전과 다르게 각 인기도 반영비율

이 10%, 40%, 10%, 40%로 설정하였을 때 평균 CHR이 27.56%로 가장 높았다. 하지만 앞선 실험들에서 가장 높은 효율을 보였던 10%, 10%, 40%, 40%로 설정된 가중치와 비교를 해보았을 때 0.03% 높은 수치로 큰 차이를 보이지 않는다.

시뮬레이터 실행 간 Client의 생성자별 요청이 변경되며 실시간 인기도 변화에 대응이 되지 않는 LFU는 시간이 지날수록 CHR이 낮아지고, 실시간 생성자 인기도(PR)를 반영하는 C&P-CRP는 CHR을 지속 유지할 수 있었다고 생각된다.

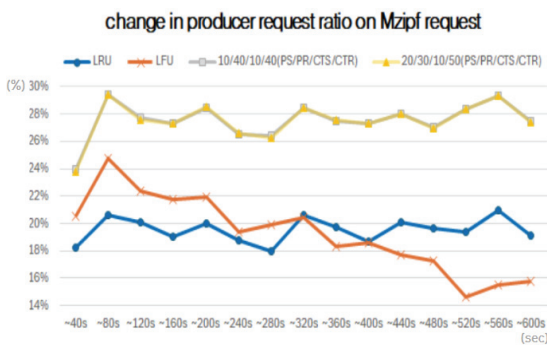


그림 8. MZipf 요청에서 생성자 요청비율 변화에 따른 CHR 비교

Fig. 8. Comparison of CHR based on changes in producer request ratio in MZipf request

표 7. MZipf 요청에서 생성자 요청비율 변화에 따른 세부 CHR

Table 7. detailed CHR based on changes in producer request ratio in MZipf request

Time(s)	LRU	LFU	C&P-CRP(PS, PR, CTS, CTR순)	
			10/40/10/40	20/30/10/50
~40	18.2	20.525	23.925	23.8
~80	20.6	24.75	29.4	29.4
~120	20.1	22.4	27.725	27.6
~160	19.025	21.725	27.325	27.275
~200	20.025	21.975	28.5	28.575
~240	18.775	19.4	26.55	26.625
~280	17.95	19.875	26.425	26.275
~320	20.65	20.45	28.5	28.475
~360	19.7	18.3	27.45	27.55
~400	18.65	18.6	27.275	27.35
~440	20.075	17.7	28.05	28.05
~480	19.625	17.3	27.05	26.975
~520	19.35	14.65	28.4	28.35
~560	21	15.55	29.35	29.35
~600	19.15	15.775	27.475	27.4

V. 결론

이번 연구에서 CCN 환경에서 콘텐츠·생성자 인기도 기반 캐시 정책을 제안하였고 생성자 고정 인기도(PS), 생성자 실시간 인기도(PR), 콘텐츠 타입 고정 인기도(CTS), 콘텐츠 타입 실시간 인기도(CTR)의 반영비율을 각각 10%, 10%, 40%, 40%로 설정하였을 때 최고의 성능을 보임과 기존 캐시 교체 정책 대비 우수함을 확인하였다. client의 관심은 생성자 중심에서 콘텐츠 유형 중심으로 이동하였고 이에 콘텐츠 유형을 고려함으로써 CHR을 향상시켜 네트워크 효율을 높이고 응답시간을 최소화할 수 있다.

하지만 이번 연구에서도 보완사항은 존재한다. 콘텐츠별로 부여되는 생성자·콘텐츠 유형 고정 인기도를 변화하는 환경에 맞춰 지속적으로 업데이트가 필요하고, CS공간이 하나의 생성자와 데이터 유형으로 점령될 수도 있다. 따라서 고정 인기도를 자동으로 업데이트 하거나 Interest Packet 기반의 요청을 count하여 순위를 매겨 변동시키는 등의 연구가 추가적으로 필요할 것으로 판단되며, 이를 보완한 후 하나의 생성자와 데이터 유형으로 점령되는 것을 방지하기 위한 최선의 제한책 도출이 필요하다.

References

[01] Ericsson, "Ericsson Mobility Report, November," report, 2023.

[02] DATAREPORTAL, "DIGITAL 2023: GLOBAL OVERVIEW REPORT," report, 2023.

[03] V. Jacobson, D. K. Smetters, J. D. Thornton, M. F. Plass, N. H. Briggs, and R. L. Braynard, "Networking Named Content," Proc. the 5th international conference on Emerging networking experiments and technologies, Rome, Italy, Dec. 2009, pp. 1-12. <https://doi.org/10.1145/1658939.1658941>

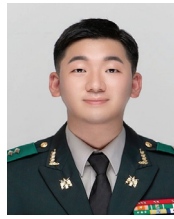
[04] M. K. Qureshi, A. Jaleel, Y. N. Patt, S. C. Steely, and J. Emer, "Adaptive Insertion Policies for High Performance Caching," Proc. of the 34th International Symposium on Computer Architecture (ISCA 2007), San Diego, CA, USA, 2007, pp.

381-391.

<https://doi.org/10.1145/1250662.1250709>

- [05] A. K. Y. Wong, "Web cache replacement policies: a pragmatic approach," *IEEE Network magazine*, vol. 20, no. 1, 2006, pp. 28-34.
<https://doi.org/10.1109/mnet.2006.1580916>
- [06] N. C. Fofack, P. Nain, G. Neglia, and D. Towsley, "Analysis of TTL-based cache networks," *Proc. of 6th International ICST Conference on Performance Evaluation Methodologies and Tools, Cargese, France*, 2012, pp. 1-10.
<https://doi.org/10.4108/valuertools.2012.250250>
- [07] J. Choi and T. Kwon, "A LFU based on Real-time Producer Popularity in Content Centric Networks," *J. of the Korean Institute of Electronic Communications Science*, vol. 16, no. 6, Dec. 2021, pp. 1113-1120.
<https://doi.org/10.13067/JKIECS.2021.16.6.1113>
- [08] G. Kim and T. Kwon, "Cache Policy based on Producer Distance to Reduce Response Time in CCN," *J. of the Korean Institute of Electronic Communications Science*, vol. 16, no. 6, Dec. 2021, pp. 1121-1132.
<https://doi.org/10.13067/JKIECS.2021.16.6.1121>
- [09] J. Min and T. Kwon, "A Cache Policy on Producer Popularity-Distance in CCN," *J. of the Korean Institute of Electronic Communications Science*, vol. 17, no. 5, Oct. 2022, pp. 791-799.
<https://doi.org/10.13067/JKIECS.2022.17.5.791>
- [10] M. Seo and T. Kwon, "A Real-time Content Popularity-Based Cache Policy in Content Centric Network," *J. of the Korean Institute of Electronic Communications Science*, vol. 18, no. 6, Dec. 2023, pp. 1095-1102.
<https://doi.org/10.13067/JKIECS.2023.18.6.1095>
- [11] *Korea Communications Commission*, "「Survey on Intelligent Information Society Users Panel」, Services used during search_1+2+3 ranking," *Report*, June 2022.
- [12] *Ministry of Science and ICT*, "「Survey on smartphone overdependence」 content usage level," *Report*, June 2022.

저자 소개

**이동건(Dong-Geon Lee)**

2016년 동아대학교 전자공학과 졸업(공학사)

2023년~현재 국방대학교 대학원 컴퓨터공학과

※ 관심분야 : Next Generation Networking, Content Centric Networking

**권태욱(Tae-Wook Kwon)**

1986년 육군사관학교 전자공학과 졸업(공학사)

1995년 美 해군대학원 컴퓨터공학과 졸업(공학석사)

2001년 연세대학교 대학원 컴퓨터공학과 졸업(공학박사)

2007년~현재 국방대학교 컴퓨터공학과 교수

※ 관심분야 : Next Generation Networking, Content Centric Networking, Network function Virtualization, Software Defined Networking, U-Sensor Networking, VR, RFID

