Regular paper
Journal of the KIECS. pp. 1095-1102, vol. 18, no. 6, Dec. 31. 2023, t. 122, pISSN 1975-8170 | eISSN 2288-2189
http://dx.doi.org/10.13067/JKIECS.2023.18.6.1095

CCN에서 실시간 콘텐츠 인기도 기반 캐시 정책

서민근* · 권태욱**

A Real-time Content Popularity-Based Cache Policy in Content Centric Network

Min-Keun Seo* · Tae-Wook Kwon**

요 약

CCN(: Content Centric Network)은 기존 네트워크 구조를 개선하기 위해 등장한 네트워크로, 주소 대신 콘텐츠 이름에 기반하여 통신한다. 캐시를 활용하여 트래픽을 분산시키고, 중간노드에서 콘텐츠를 전송함으로써 응답시간 감소 효과를 가져오고 있다. 본 논문에서는 CCN 환경에서 제한된 CS 공간을 효율적으로 활용할 수 있도록 인기도를 고려한 캐시 정책을 제안한다. 어떤 콘텐츠에 우선순위를 두어 저장하고 방출할지를 결정하는지에 따라 CCN의 성능이 크게 달라질 수 있다. 가장 효율적인 캐시 교체를 위해 생성자 인기도, 생성자 거리, 콘텐츠 히트수를 기반으로 콘텐츠 인기도를 계산해 우선순위를 정하는 실시간 콘텐츠 인기도 기반 효율적인 캐시 교체정책을 제안하였으며, 새로운 정책의 효율성을 실험을 통해 입증하였다.

ABSTRACT

Content Centric Network (CCN) is a network that emerged to improve the existing network structure and communicates based on content names instead of addresses. It utilises caches to distribute traffic and reduce response time by delivering content from intermediate nodes. In this paper, we propose a popularity-based caching policy to efficiently utilise the limited CS space in CCN environment. The performance of CCNs can vary significantly depending on which content is prioritised to be stored and released. To achieve the most efficient cache replacement, we propose a real-time content popularity-based efficient cache replacement policy that calculates and prioritises content popularity based on constructor popularity, constructor distance, and content hits, and demonstrate the effectiveness of the new policy through experiments.

키워드

Cahe Replacement Policy, Content Centric Network(CCN), Least Frequently Used(LFU)
캐시 교체 정책, 콘텐츠 중심 네트워크, LFU 방식

Ⅰ. 서 론

인터넷은 등장한 이후 지속적으로 발전하고 있으며, 그 발전의 속도는 나날이 빨라지고 있다. 현대의 인터넷에서는 사용자 수와 활용범위가 급격하게 늘어났고, 이에 따라 유통되는 콘텐츠의 수와 상호 연결되는 기기들의 수도 증가하고 있다. Cisco의 연례 보고서에 따르면, 인터넷을 사용하는 유저의 수는 2018년 39억명에서 2023년에는 전 세계 인구의 66% 수준인 53억명까지 증가할 것이며, 사용자 1인당 연결되는 디

* 국방대학교 관리대학원(rebron_n@naver.com)
** 교신저자 : 국방대학교 컴퓨터공학과
· 접 수 일 : 2023. 09. 18
· 수정완료일 : 2023. 10. 30
· 게재확정일 : 2023. 12. 27
· Received : Sep. 18, 2023, Revised : Oct. 30, 2023, Accepted : Dec. 27, 2023
· Corresponding Author : Tae-Wook Kwon
Dept. Computer engineering, Korea National Defense University
Email : rebron_n@naver.com

바이스의 수는 2018년 2.4개에서 2023년 3.6개로 증가할 것으로 예측하였다[1]. 콘텐츠의 양도 늘어남에 따라, 2017년에 비해 2022년에는 분야별로 약 3~7배 트래픽이 증가하였다. 네트워크 성능 또한 따라서 향상되어, 2017년 39Mbps였던 전 세계 유선 광대역 속도는 2022년에 75.4Mbps로 두 배 가까이 빨라졌다[2].

그러나 기존 IP주소 기반 네트워크에서는 Host to Host 연결을 통해 생성자가 요청받은 만큼 콘텐츠를 복사하여 전송해야 한다. 따라서 콘텐츠의 양과 사용자가 많아질수록 트래픽이 급격하게 증가하는 비효율성이 존재한다. 이에 따라 인기 콘텐츠의 경우에는, 서버 및 회선의 병목현상이 발생하여 응답시간을 증가시키고 패킷이 손실되는 등 통신 품질을 저하시키는 주요 원인이 되고 있다.

콘텐츠 중심 네트워크(Content Centric Network, CCN)는 기존 네트워크가 가지는 이러한 비효율성을 해결하기 위해 등장했다. CCN에서는 트래픽을 분산시키기 위해 캐시를 활용한다. 중간노드의 제한된 CS 공간을 어떻게 효율적으로 활용하는지에 따라 CCN의 성능이 좌우되며, 인기도를 기반으로 한 효율적인 캐시 정책을 통해 네트워크의 성능을 향상시킬 수 있다.

본 논문에서 제안하는 캐시 정책은 LFU(: Least Frequently Used)를 기반으로 한 캐시 정책으로, 인기도를 통해 우선순위를 부여하기 위해 생성자 인기도 및 거리를 고려한 선행 연구에 각 콘텐츠의 개별 히트수를 함께 고려하여 적중률을 향상시키는 것을 목표로 했다.

본 논문의 구성은 2장에서 콘텐츠 중심 네트워크(CCN)의 개념과 캐시 정책에 대해 논하고, 3장에서는 제안 캐시정책인 실시간 콘텐츠 인기도 기반 효율적인 캐시정책을 소개한다. 4장에서 제안 캐시정책의 성능을 NS-3 기반 ndnSIM을 활용해 분석한 뒤 5장에서 결론과 함께 향후 연구과제를 논하였다.

II. 관련연구

2.1 CCN

CCN은 주소가 아닌 콘텐츠 이름을 기반으로 한 서비스를 제공하는 네트워크로, V.Jacobson에 의해 2006년 제안되었다. 기존 네트워크가 IP주소를 기반으로 목적지(Where)를 찾아가는 것이 목적이었다면, CCN에서는 이름을 기반으로 원하는 콘텐츠(What)를 찾아가는 것이 목적이 된다[3]. 기존 네트워크와 CCN에서의 데이터 흐름도는 그림 1과 같다.

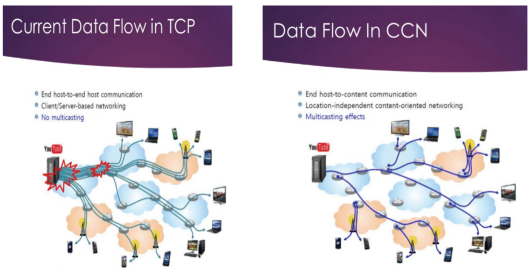


그림 1. 기존 네트워크와 CCN에서의 Data Flow
Fig. 1 Data Flow in TCP and CCN

CCN에서 사용되는 패킷의 종류로는 그림 2와 같이 Interest 패킷과 Data 패킷이 있다. Interest 패킷은 콘텐츠 요청을 위해 사용되는 패킷으로 콘텐츠 이름(Content Name), 선택자(Selector), 랜덤 임시변수(Nonce)로 구성되어 있다. Data 패킷은 콘텐츠 요청에 대해 콘텐츠 제공자가 응답하는 패킷으로 콘텐츠 이름(Content Name), 서명(Signature), 사인 정보(Signed Info), 데이터(Data)로 구성되어 있다.

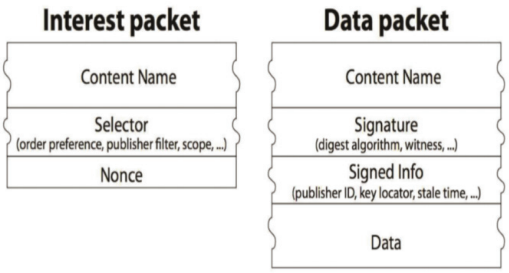


그림 2. CCN 패킷의 구성요소
Fig. 2 Components of a CCN packet

CCN은 CS(: Content Store), PIT(: Pending Interest Table), FIB(: Forwarding Information Base)로 포워딩 엔진이 구성되어 있다. PIT는 수신한 Interest packet에 대한 응답으로 콘텐츠를 전송하기 위해 기록하는 테이블로, Interest packet의 입·출력을

기록 관리한다. FIB는 Interest packet을 포워딩하는데 사용되는 일종의 포워딩 테이블이다. CS에는 데이터 콘텐츠와 콘텐츠 리스트가 저장 / 관리된다. 각 콘텐츠들은 유통되는 과정에서 중간노드에 캐시되는데, 위치에 관계없이 콘텐츠 자체를 찾아가기 때문에 반드시 생성자까지 Interest packet이 도달할 필요가 없다. 콘텐츠 생성자가 아니더라도 사용자가 원하는 콘텐츠를 요청자와 가까운 노드에서 제공할 수 있게 하여 병목현상을 방지하고 응답속도를 감소시킨다[4].

2.2 CCN에서의 캐시 정책

CCN에서는 CS 내에 콘텐츠를 캐시하여 해당 콘텐츠에 대한 Interest packet이 도착했을 때 즉시 콘텐츠를 제공할 수 있도록 하고 있다. CS에 모든 콘텐츠를 저장할 수 있다면 좋겠지만, 저장할 수 있는 공간은 유한하고 한정되어 있기에 어떤 콘텐츠를 우선적으로 저장해둘지에 대한 판단이 요구된다. 다시 요청될 가능성이 높은 콘텐츠, 즉 인기도가 높은 콘텐츠가 많이 저장될수록 네트워크의 효율은 높아질 것이고, 거의 요청되지 않는 콘텐츠, 즉 인기도가 낮은 콘텐츠가 많이 저장될수록 네트워크의 효율은 떨어질 것이다. 따라서 어떤 콘텐츠를 저장하고 방출할 지에 대한 캐시 정책이 네트워크 성능에 있어 굉장히 중요한 영향을 주는 요소가 된다.

CCN의 캐시 정책은 표 1과 같이 독립정책(Independent Policy)과 협력정책(Cooperative Policy)으로 분류된다. 독립정책은 각 라우터에서 단독으로 콘텐츠를 저장 또는 방출하도록 결정하는 정책이며, 협력정책은 라우터 간의 통신을 통해 콘텐츠를 저장 또는 방출하도록 결정한다. 본 논문에서 제안할 정책은 독립정책으로, 대표적인 정책 몇 가지를 소개한다.

표 1. 캐시 교체정책 분류[5]
Table 1. Classification of cache policy

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

LRU(: Least Recently Used)는 최근에 참조된 콘텐츠가 우선순위를 가지며, 캐시 공간이 가득 차면 새로운 콘텐츠가 진입할 때 가장 참조 시간이 오래된 콘텐츠를 방출하는 정책이다. 구현이 쉽고 간단하지만 단 1번만 참조되고 재참조 되지 않는 인기가 없는 콘텐츠도 오랫동안 캐시 공간 내에 잔류하게 되어 효율성을 저하시킬 수 있다는 단점이 있다[6].

LFU는 가장 많이 참조된 콘텐츠가 우선순위를 가지며, 캐시 공간이 가득 차면 새로운 콘텐츠가 진입할 때 가장 적게 참조된 콘텐츠를 방출하는 정책이다. 참조 횟수가 많은 인기 있는 콘텐츠가 오래 캐시공간에 잔류할 수 있지만, 오래 전에 참조가 많이 되고 최근에는 참조되지 않는 콘텐츠가 계속해서 캐시 공간을 차지하고 있는 캐시 오염현상이 발생할 수 있다는 단점이 있다[7].

TTL-base(: Time To Live-base)는 콘텐츠에 생존 시간(TTL)을 부여하고, 부여된 생존시간이 모두 소요되는 동안 재참조가 없을 시 방출하는 정책이다. LRU와 LFU의 단점을 보완가능하다는 장점이 있지만 TTL을 일괄적으로 적용하기 때문에 모든 콘텐츠의 생존 시간이 동일하다는 한계점이 있다[8].

III. 실시간 콘텐츠 인기도 기반 캐시 정책

3.1 개요

기존의 LFU의 정책은 참조 횟수를 통해 콘텐츠의 인기도를 반영할 수 있으나, 캐시 오염 현상의 문제점을 지닌다. 이를 해결하기 위해 선행 연구에서는 TTL 변수를 적용하여 생존시간을 부여하고, TTL이 만료되면 방출하여 캐시 오염 현상을 제거하였다.

그러나 모든 콘텐츠에 동일하게 TTL을 적용하면 인기 있는 콘텐츠와 인기 없는 콘텐츠가 같은 생존시간을 부여받기 때문에 인기도와 무관하게 같은 시간동안 생존하는 문제점을 지닌다. 따라서 TTL을 차등 부여하고자 하는 연구가 진행되어 왔다. 선행 연구로서 생성자의 인기도를 변수로 활용하는 정책[9], 생성자의 거리를 변수로 활용하는 쟁책[10], 생성자의 인기도와 거리 변수를 모두 혼합하여 TTL을 차등 부여하는 정책이 최근 연구되었다[11].

하지만 생성자 인기도 및 거리로 TTL을 차등 부

여하는 선행 연구에도 제한사항은 있다. 인기가 많을 것으로 예상되는 콘텐츠에 보다 많은 TTL을 부여하는 것은, 예측에만 근거해 고정된 TTL을 산출하는 제한사항을 지닌다. 즉, 처음 부여된 TTL이 해당 콘텐츠의 고정 TTL값이 되는 것이다. 예상과 달리 인기가 많을 것으로 예상된 콘텐츠가 인기가 없거나, 인기가 적을 것으로 예상된 콘텐츠가 인기가 많을 경우 TTL값을 조정해줄 필요가 있다. 따라서 본 논문에서 제안하는 캐시 정책에서는 생성자 인기도 및 거리와 함께 콘텐츠 히트 수를 실시간으로 반영한다. TTL이 리셋 될 때 인기가 많은 콘텐츠의 TTL을 보다 더 길게 부여할 수 있도록 하는 방법을 제안하였다. 제안 정책의 Flowchart는 그림 3과 같다. 그리고 이 캐시 정책을 실시간 콘텐츠 인기도 기반 캐시 정책(Real-time Popularity-Based Efficient Cache Replacement Policy, RP-CRP)으로 정의하였다.

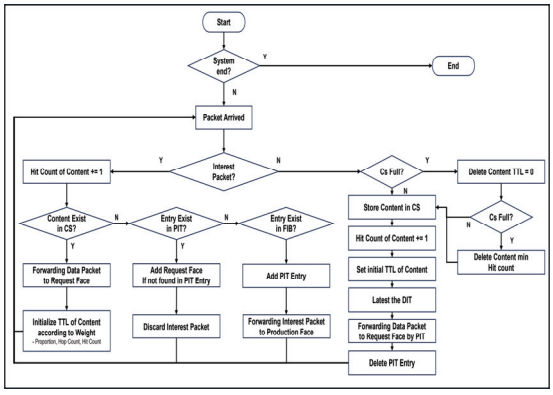


그림 3. 제안 캐시정책 Flowchart
Fig. 3 Flowchart of proposed cache policy

3.2 실시간 콘텐츠 인기도 기반 캐시 정책

생성자 인기도와 거리 변수를 적용해주기 위해 CS의 구조를 변경한다. 각각의 콘텐츠에 생성자 인기도 및 거리, 히트 수 기반의 TTL을 부여해주기 위해 그림 4와 같이 CS 구조를 변경하였다.

Name	Data	Hit Count
naver.com/Seo/A.ppt/	.....	4
naver.com/Seo/B.doc/	.....	3
youtube.com/Koo/C.mp4/	.....	2
google.com/Kim/D.mp4/	.....	2
google.com/Kim/E.Doc/	.....	1

(a) Content Information Table(CIT)

Name	Data	Hit Count	TTL
naver.com/Seo/A.ppt/	.....	4	5
naver.com/Seo/B.doc/	.....	3	5
youtube.com/Koo/C.mp4/	.....	2	4
google.com/Kim/D.mp4/	.....	2	3
google.com/Kim/E.Doc/	.....	1	3

(b) Detailed Information Table(DIT)

Name	Data	Number of Content	Proportion	#1 TTL	Hop Count	#2 TTL	#3 TTL
naver.com/Seo/A.ppt/	.....						6
naver.com/Seo/B.doc/	.....						5
youtube.com/Koo/C.mp4/	.....						3
google.com/Kim/D.mp4/	.....						2
google.com/Kim/E.Doc/	.....						1

그림 4. CS 구조변경
Fig. 4 Changing CS structure

CIT(Content Information Table, CIT)는 LFU 정책의 CS 테이블에서 TTL을 추가하여 Name, Data, Hit Count, TTL의 항목으로 구성되어 있다. Hit Count는 콘텐츠의 참조 횟수로, 최초 저장될 때 1로 설정되고 이후 재 참조 될 때마다 1씩 추가된다. TTL은 콘텐츠의 생존 시간이다. 콘텐츠가 CS에 최초 진입할 때는 기본 값인 2로 설정하고, 이후 생성자 인기도, 거리, 콘텐츠 참조 횟수에 따라 보정된다.

DIT(Detailed Information Table, DIT)는 콘텐츠 생성자 등 세부적인 정보를 포함한다. Number of Content는 내 캐시 공간에 저장된 콘텐츠 중 해당 생성자가 생성한 콘텐츠 수이다. Proportion은 전체 콘텐츠 중 해당 생성자가 차지하는 비중으로, 생성자 인기도의 척도가 된다. #1 TTL은 Proportion을 이용해 계산하며, TTL 기본 값인 2초에 (1+proportion)을 곱하여 산출한다.

DIT에는 생성자 거리에 대한 정보도 있는데, Hop Count는 CS 내 캐시 공간에 저장된 콘텐츠가 생성자로부터 라우터로 도달하기 까지 거친 중간노드 수를 의미한다. #2 TTL은 TTL 기본 값인 2초에 Hop Count를 곱하여 산출한다.

#3 TTL은 CIT의 Hit Count를 기반으로 계산된다. #3 TTL은 CS 내 캐시 공간에 저장된 콘텐츠들 중 실제로 가장 많이 참조된 인기 있는 콘텐츠일수록 길게 부여된다. #3 TTL은 캐시 공간 내 데이터의 전체 수를 기준으로 부여한다. 캐시 공간에 5개의 콘텐츠가 있다면, 이중 가장 많은 참조 횟수를 가진 콘텐츠의 #3 TTL은 5가 되고, 그 다음 콘텐츠는 4가 된다. 즉 n개의 콘텐츠가 있다면 가장 많이 참조된 콘텐츠의 #3 TTL은 n이 되고, 그 다음 많이 참조된 콘텐츠의 TTL은 n-1이 된다.

결론적으로, 생성자 인기도가 높은 콘텐츠는 상대적으로 긴 생존시간(#1 TTL)을 부여한다. 생성자 거리가 먼 콘텐츠는 다른 콘텐츠들보다 상대적으로 긴 생존시간(#2 TTL)을 부여한다. 그리고 실제 참조 횟수가 많은 인기 있는 콘텐츠일수록 긴 생존시간(#3 TTL)을 부여한다. 기존 연구[11]에서 #1 과 #2 TTL의 비율이 5:5일 때 가장 높은 효율을 얻을 수 있음을 확인 했기 때문에, #1과 #2, #3 TTL의 비율은 1:1:1로 고정하였다. 각 콘텐츠가 가지는 TTL은 #1, #2, #3 TTL의 평균으로 산출되게 된다.

Ⅳ. 실험 및 성능분석

4.1 실험환경

CCN 환경 구현을 위해 NS-3 기반의 ndnSIM을 활용하였다. 토폴로지 구성은 그림 5와 같다. 콘텐츠를 생성하는 Producer를 5개 배치하고, 중간 노드를 5개 배치하여 Hop 거리를 1부터 5까지 다양하게 구성하였다.

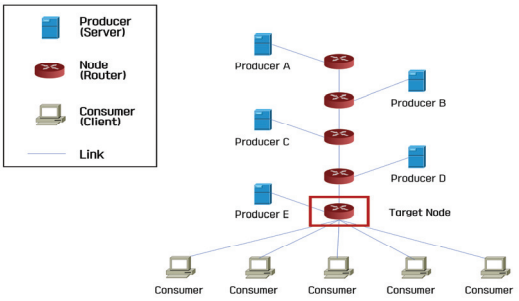


그림 5. 실험 토폴로지
Fig. 5 Experimental Topology

세부 실험환경은 표 2와 같이 구성하였다. Target Node의 CS 캐시 용량은 전체 콘텐츠 용량의 1%이며 시뮬레이션 시간은 총 300초로 하였다. 총 콘텐츠 개수는 10,000개이며 각 Producer가 2,000개 씩 콘텐츠를 생성한다. 콘텐츠 생성자의 인기도를 조절하기 위해 Producer A에 40%, B에 30%, C에 15%, D에 10%, E에 5%의 비중으로 요청하도록 할당하였다.

표 2. 세부 실험환경
Table 2. Detailed experimental environment

Parameter	Value
Simulator	ndnSIM
Simulation Time	300s
Number of Content	2,000 Contens each
Cache Size	1%
Chunk Size	1,024 Bytes
Zipf-α	0.7
Contents Request	100/s

4.2 실험 결과 및 분석

실험 대조군으로는 기본 LFU 정책과, 단점 보완을 위해 TTL 개념을 도입한 LFU-TTL, 생성자 인기도 및 거리 기반의 캐시정책(EQ-CRP)과 적중률을 비교하였다. 실험 결과는 그림 6과 같다.

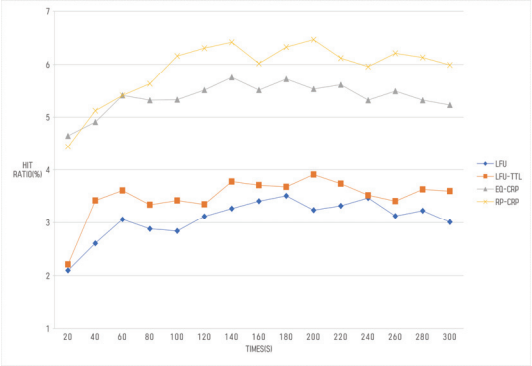


그림 6. 캐시 정책 별 적중률 비교
Fig. 6 Cache hit ratio comparison

가장 적중률이 낮은 캐시 정책은 LFU로, 시뮬레이션 기간 평균 캐시 적중률은 3.08%로 나타났다. 제안한 정책인 RP-CRP의 캐시 적중률은 5.91%로, 기존 LFU대비 92% 높은 적중률을 보였으며, 생산자 인기도

와 거리를 고려한 캐시정책인 EQ-CRP(5.37%) 대비 약 10% 가량 높은 적중률을 보였다.

시뮬레이션 초기 20초 구간까지는 기존 정책대비 가시적인 적중률 향상은 보이지 않다가 시간이 흐를수록 높은 적중률이 나타난 것을 볼 수 있다. 이는 추가된 변수인 콘텐츠별 참조 횟수가 실험 초기에는 영향이 미미하다가, 실험이 지속될수록 참조횟수가 누적되어 콘텐츠 인기도를 더 확실하게 반영할 수 있었기 때문으로 생각된다.

Ⅴ. 결론

이번 연구에서는 CCN 환경에서 콘텐츠 실시간 인기도를 고려한 캐시 정책을 제안하고 우수성을 확인하였다. LFU와 LFU-TTL 등 기본 캐시정책의 제한점은 보완하고 장점은 부각시킴으로써 캐시 적중률이 향상됨을 확인할 수 있었다.

하지만 이번 연구에서도 보완사항은 있다. 콘텐츠별로 부여되는 TTL의 기본 값에 따라 성능이 달라질 수 있다. 또한 생성자 인기도, 거리, 콘텐츠 히트수의 비율을 1:1:1이 아닌 다르게 설정할 경우 캐시 적중률에 있어 차이가 날 수 있다. 따라서 가장 높은 성능을 낼 수 있는 TTL 산출 방법에 대한 연구가 추가적으로 필요할 것으로 생각된다. 향후 TTL을 산출하는 더 효과적인 변수를 찾거나, 어떻게 비율을 설정할 것인지에 대한 추가 연구가 필요하다.

References

[1] Cisco, "Cisco Annual Internet Report(2018-2023) White paper," *report,* 2020.

[2] Cisco, "Cisco Visual Networking Index: Forecast and Trends, 2017–2022 White paper," *report,* 2019.

[3] V. Jacobson, D. K. Smetters, J. D. Tornton, 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.

[4] J. Kim, H. Jung, and W. Park, "Content Centric Networking Technology," *Journal of the Electronics and Telecommunications Trends,* vol. 25, no. 6, 2010, pp. 136-143.

[5] S. Podlipnig and L. B. S. Rmenyi, "A Survey of Web Cache Replacement Strategies," *ACM Computing Survey,* vol. 35, no 4, 2003.

[6] 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.

[7] A. K. Y. Wong, "Web cache replacement policies: a pragmatic approach," *IEEE Network magazine,* vol. 20, no. 1, 2006, pp. 28-34.

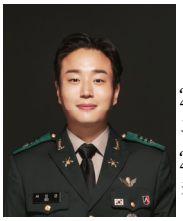
[8] N. C. Fofack, P. Nain, G. Neglia, and D. Towsley, "Analysis of TTL-based cache networks," *Proc. of 6th International ICST Conference on Performance Evaluation Methodologies and Tools,* Cargese, Freance, 2012, pp.1-10.

[9] J. Choi and T. Kwon, "Real-time Producer Popularity-based LFU in Content Centric Network," *J. of the Korean Institute of Electronic Communications Science,* vol. 16, no. 6, Dec. 2021, pp. 1113-1120.

[10] 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.

[11] 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.

저자 소개



서민근(Min-Keun Seo)

2016년 육군사관학교 토목환경공학과 졸업(공학사)
2022년 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, Software Defined Networking, Network Function Virtualization, U-Sensor Networking, VR, RFID