

웹 환경에 적합한 보관수명 기반 캐시 교체정책

한성민*, 박흥순°, 권태욱*

Shelf-Life Time Based Cache Replacement Policy Suitable for Web Environment

Sungmin Han*, Heungsoon Park°, Taewook Kwon*

요약

오랜기간 컴퓨터 분야의 연구주제였던 캐시 메커니즘은 네트워크 영역에서 웹 캐시로 응용되었다. 응답시간 감소, 네트워크 자원 절약 등의 다양한 이점을 갖는 웹 캐시는 교체정책에 의해 성능이 좌우되므로, 보다 나은 교체정책의 설계를 위해 웹 캐시가 운용되는 환경에 대한 분석과 고찰이 필수적이다. 따라서 과거에 비해 급속도로 다변화된 현재 웹 환경에서는 그러한 변화를 반영할 수 있는 교체정책이 요구된다. 따라서 본 논문에서는 현재 웹 환경의 특성을 규정하고, 이에 적합한 캐시 교체정책을 설계하고 평가한다.

Key Words : Shelf-life Time, Cache, Cache Replacement Policy

ABSTRACT

Cache mechanism, which has been a research subject for a significant period of time in computer science, has become realized in the form of web caching in network practice. Web caching has various advantages, such as saving of network resources and response time reduction, depends its performance on cache replacement policy, therefore, analysis and consideration of the environment in which a web cache operates is essential for designing better replacement policies. Thus, in the current web environment where is rapidly changing relative to the past, a new cache replacement policy is necessary to reflect those changes. In this paper we stipulate some characteristics of the web at present, propose a new cache replacement policy, and evaluate it.

1. 서론

캐시 메커니즘은 컴퓨터 분야에서 오랜 기간 연구되어 온 주제이다. 지역성 이론에 근거한 캐시 메커니즘은 다양한 분야에 응용되어졌는데, 네트워크 영역에서는 웹 캐시라는 형태로 발전하였다. 원천 서버와 사용자의 중간에서 콘텐츠를 보관했다가 사용자의 요청이 있을 때 원천서버를 대신하여 응답하는 웹 캐시를

통해 네트워크 관리자는 지연을 감소시켜 사용자의 서비스 만족도를 향상시키고 네트워크 자원도 절약할 수 있다. 웹 캐시의 성능은 교체 알고리즘에 의해 좌우되었기 때문에 많은 연구가 교체 알고리즘의 성능 향상에 초점을 두고 이루어져왔다. 하지만 보다 적절한 교체 알고리즘의 설계를 위해서는 캐시가 운용되는 환경이 갖는 특성을 도출하는 것이 우선적으로 요구된다. 근래에 이르러 인터넷의 보급과 이동성을 갖

* First Author : Dept. of Computer Science & Engineering, Korea National Defense University, eall99@naver.com, 정희원

° Corresponding Author : Dept. of Computer Science & Engineering, Korea National Defense University, heungsoon.park@gmail.com, 흥신희원

* Dept. of Computer Science & Engineering, Korea National Defense University, kwontw9042@naver.com, 흥신희원
논문번호 : KICS2015-03-064, Received March 23, 2015; Revised June 4, 2015; Accepted June 4, 2015

춘 통신기기의 대중화로 인해 사용자 누구나 손쉽게 콘텐츠를 생산할 수 있는 환경이 조성되었고, 생성된 수많은 콘텐츠들을 소셜 네트워크 서비스로 대표되는 Web 2.0 환경에서 손쉽게 접할 수 있게 됨에 따라 사용자들의 관심도 시시각각 급변하게 되었다. 콘텐츠가 빠르게 생성되고 사용자의 관심이 새롭게 생겨나는 콘텐츠에 일시적 집중되는 현재의 웹 환경에서 캐시 교체정책은 이러한 환경적 특성을 적절히 반영해야 한다. 따라서 본 논문에서는 현재의 웹 환경이 과거에 비해 어떠한 측면에서 변화하였는지 정의하고, 이에 적합한 캐시 교체정책을 설계하여 성능을 평가하는 것을 목표로 하였다. 이를 위해 논문은 다음과 같이 구성되었다. 2장에서는 기존의 캐시정책과 연구를 통해 밝혀진 웹 환경의 특성을 알아보고, 과거에 비해 현재의 웹 환경이 보이는 두드러진 특성을 살펴본 뒤, 현재의 환경에서 기존 캐시 교체정책이 갖는 문제점을 알아본다. 3장에서는 이러한 문제점을 해결하기 위해 보관수명의 개념을 도입한 캐시 교체정책을 제안하고, 4장에서 제안된 교체정책과 기존 교체정책의 성능 비교를 통해 의미하는 바를 분석한다.

II. 관련연구

2.1 웹 캐시(Web cache)와 교체정책

컴퓨터 분야의 많은 영역에서 캐시 메커니즘(Cache Mechanism)은 오랜 기간 연구되어온 주제이다. 계층적 기억장치 구조에서 하위 저장장치로의 접근은 필연적으로 지연시간의 증가를 발생시키게 되는데, 캐시 메커니즘을 통해 특정 콘텐츠(Content)¹⁾를 상위 저장장치와 근접한 별도 공간(캐시 메모리)에 보관함으로써 이러한 문제를 일부 해소하고 처리율을 향상시킬 있었다. 이러한 캐시 메커니즘은 네트워크, 데이터베이스, 포렌식 등에 응용되고 있으며, 수많은 연구자들에 의해 발전을 거듭해왔다.^[1]

네트워크 영역에서의 캐시 메커니즘의 응용은 그림 1에서 보는 바와 같이 사용자와 원천서버의 중간에 있는 캐시 서버(Cache server)를 이용하여 자주 요청되거나, 미래에 예측될 가능성이 높은 콘텐츠를 저장하였다가 사용자의 요청이 있을 경우 캐시 서버가 원천 서버를 대신하여 응답하는 웹 캐시의 형태로 이루어

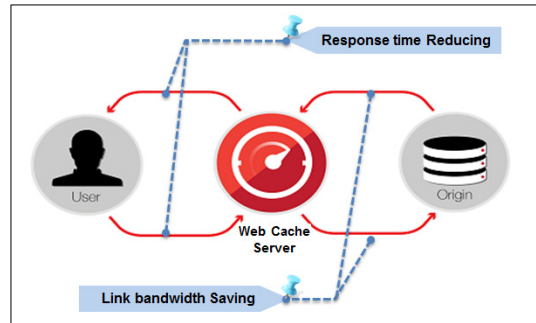


그림 1. 웹 캐시 동작 개요
Fig. 1. Overview of web cache

어졌다. 이를 통해 네트워크 관리자는 캐시 서버에서 원천 서버에 이르는 구간의 네트워크 자원을 절약하고, 응답시간을 감소시켜 사용자들의 서비스 만족도를 향상시킬 수 있었다.^[2] 때문에 어떤 콘텐츠를 캐시 메모리에 보관할 것인지 결정하는 교체 알고리즘의 중요성이 자연스레 관심의 대상이 되었는데, 이는 부적절한 교체 알고리즘이 적용될 경우 불필요한 콘텐츠가 메모리에 보관됨으로써 캐시 검색에 따른 부수적인 지연을 발생시켜 비용과 노력의 투입에도 불구하고 만족할 만한 효과를 얻을 수 없었기 때문이다.

그래서 많은 연구자들은 다양한 기준과 고려요소를 함께 병합하여 다양한 웹 캐시 교체정책을 고안하였고, 이러한 다양한 교체정책을 각각의 특성에 따라 분류하고자 하는 연구도 지속적으로 이루어졌다.^{[3][4][5]} 이러한 연구들을 전반적 종합하여 봤을 때, 교체정책은 각각 최근요청(Recency), 빈도(Frequency), 함수(Function), 임의(Random), 혼합(Hybrid) 방식으로 표 1과 같이 분류될 수 있다.^[3]

표 1에서 보는 것처럼 캐시 교체정책은 다섯 가지 형태로 분류될 수 있으며, 특정 교체정책은 다른 분류의 특성을 일부 차용하기도 한다.^[3] 그런데 주목할 점은 다수의 교체정책들이 최근요청(Recency)과 빈도(Frequency)라는 두 가지 특성에 대체로 의존하는 형태를 취하거나, 혹은 이를 바탕으로 추가적인 고려요소나 판단 기준을 선정한다는 것이다. 예를 들어, LRU-T, Pitkow, SIZE, LRU-Min 등과 같은 정책은 순수한 최근요청 기반(LRU)에 콘텐츠의 크기나 교체 횟수라는 특정 요소를 추가적인 고려사항으로 판단한다. 이러한 사실은 최근요청과 빈도가 주요한 판단기준임에 틀림없으나, 그것 자체만으로는 상당한 단점이 있음을 반증하는 것이기도 하다. 왜냐하면 순수하게 최근 요청만을 고려한 정책(LRU)은 콘텐츠가 요청되고 있는 통계적 수치를 무시하게 되므로, 누적되는 요

1) 캐시 연구에서의 콘텐츠는 일반적으로 캐시의 단위가 되는 item, object, web object 등을 통칭하는데, 본 논문에서 주제와 연관되어 웹에서 유통되는 콘텐츠만을 의미할 때에도, 의미의 혼란이 발생하지 않을 범위 내에서는 콘텐츠로 통일하여 사용한다.

청횟수(빈도)의 통계적 계산을 통해 향후에 요청될 가능성을 판단할 수 있음에도 이를 등한시하기 때문이다. 반면 전적으로 빈도에 기반한 정책(LFU)은 근래에 이루어지고 있는 요청의 변화가 신속히 반영되지 못한다는 약점(최근의 변화가 캐시 우선순위 결정과정에 반영되려면 현재까지 누적된 요청횟수를 초과해야 하기 때문에)을 갖는다. 그럼에도 불구하고 최근 요청과 빈도라는 특성이 주요 판단기준이 될 수 있는 것은 이러한 두 가지 특성이 캐시 교체정책의 기반이 되는 참조 지역성 이론에 근거하기 때문이다. 참조 지역성 이론은 가상메모리 환경에서의 성능향상과 안정적인 시스템 구현을 위한 연구과정에서 파생되었다. 이는 프로그램들이 개발자에 의해 개발될 때 프로그래머의 의도에 관계없이 시·공간적 지역성(Temporal and Spatial Locality)을 갖는다는 것으로 캐시 교체정책에서 어떤 콘텐츠를 우선적으로 보관할 것인가에 대한 판단기준 선정의 이론적 근거가 되는 핵심 이론이다.^[11] 그러므로 일부 연구에서는 캐시 교체정책을 설계할 때, 캐시가 운용되는 환경에 적합한 교체정책을 제안하였는데^{[6][7]}, 이러한 연구의 기반에는 환경에 따라 상이한 지역성이 있으며, 이를 위해 우선순위 판단을 위한 기준과 고려요소가 달라져야 한다는 전제가 있다.

표 1. 웹 캐시 교체정책의 분류
Table 1. Classification of web cache replacement policy

Classification	Replacement Policy
Recency	LRU, LRU-T, Pitkow, SIZE, LRU-Min, HLRU, PSS, EXP1, Value-Aging, LRU-LSC
Frequency	LFU, LFU-aging, LFU-DA, α -Aging, swLFU
Hybrid	SLRU, GR, LRU*, LRU-Hot, HYPER-G, CSS, LRU-SP
Function	GD-Size, GDSF, GD*, Server-assisted cache replacement, TSP, Bolot/Hoschla's strategy, MIX, M-Metric, HYBRID, LNC-R-W3, LRV, LUV, LR-Model
Random	RAND, HARMONIC, LRU-C, LRU-S

2.2 현재 웹 환경에서의 콘텐츠 특성

초기의 웹은 전자메일의 송·수신이나 적은 수의 파일 전송을 위해 고안되었다. 사용자는 컴퓨터를 오랜 기간 연구해온 전문가이거나 어느 정도의 지식을 갖춘 사람들이었고 충분히 숙련되어야 했다. 당연히 콘

텐츠는 특정 소수의 주도로 생성되었고 일방향적으로 사용자들에게 제공되는 구조였다. 콘텐츠의 생산에는 오랜 시간과 많은 노력이 요구되고, 콘텐츠를 생산할 수 있는 사람은 한정적이었다. 하지만, 지속적인 변화가 이루어졌다. “Web 2.0”^[8]으로 일컬어지는 이러한 특성은 사용자의 참여와 공유로 대변되며 실시간 상호작용이 빈번하게 발생한다. Web 2.0의 환경에서 사용자는 누구라도 콘텐츠를 생성할 수 있고, 손쉽게 원하는 콘텐츠를 얻을 수 있게 되었다.^[8,10] 이동식 통신 기기의 보급과 대중화는 이러한 현상을 가속화시켰다. 현실에서 발생하는 사건들(Event)은 실시간으로 콘텐츠로 제작되어 전파되고 확산된다. YouTube, Facebook, Twitter 등과 같은 SNS는 사용자의 실시간 콘텐츠 접근과 전파, 그리고 확산에 기여하고 있는 대표적인 예이다. 때문에 현재의 웹 환경에서 콘텐츠는 대량 생산되고 소비되며, 수명은 짧아지고 있다. 네트워크 관련 기업인 Cisco는 백서(White Paper)를 통해 향후 2018년까지 연간 트래픽이 평균 21%씩 증가할 것이며, 그 증가폭이 매년 상승할 것으로 예측하고 있다. 뿐만 아니라, 콘텐츠 전달 네트워크(Content Delivery Networks)와 사용자의 콘텐츠 트래픽 역시 같은 추세가 될 것으로 전망하고 있다.^[9] 한편 콘텐츠에 대한 사용자의 관심은 일시적이고 급속히 변화하는 양상을 보이고 있는데, 일례로 콘텐츠 마케팅 전문가인 Ted Karczewski는 Bitly 및 Wisemetrics, Piqura 와 같은 기업들의 조사를 바탕으로 스카이워드(Skyword) 블로그를 통해 콘텐츠들이 생성초기 급격한 관심을 받는 기간(보관수명, Shelf-Life Time)을 갖는다고 발표하였다.^[11] 이러한 콘텐츠의 보관수명은 콘텐츠별로 Twitter는 2.8시간, Facebook은 5시간, 그리고 Pinterest Pins은 2일로 상이하나 초기에 집중되는 특성을 갖고 있다고 밝히고 있다. 학계에서도 콘텐츠의 인기예측에 대해 많은 연구가 이루어지고 있는데 가운데 콘텐츠 생성초기 사용자들의 집중적 요청과 시간에 따른 관심의 감소 추세가 다수의 연구를 통해서도 밝혀진바 있다. UGC(User Generated Content)를 대상으로 한 연구^[12]에서는 UGC 콘텐츠의 인기가 시간에 따라 감소하며, 생성초기에 일시적인 관심의 집중이 이루어질 뿐 아니라, 초기에 얻는 요청의 크기가 향후의 요청과도 상당한 연관성이 있다고 주장한다. 또한 트위터^[13], 북마킹 사이트^[14], 온라인 뉴스^[15] 등 많은 콘텐츠들에 관한 연구도 이와 유사한 생성초기 콘텐츠들에 대한 사용자의 집중적 요청이 발생함을 밝히고 있다. 이러한 연구들을 통해 밝혀진 현재의 웹 환경에서 콘텐츠가 갖는 두드러진 특성을 정리하

면 다음과 같다.

- 콘텐츠가 빠른 속도로 생성된다.
- 콘텐츠는 생성초기에 사용자들의 일시적인 관심을 받는다.
- 시간이 지날수록 콘텐츠에 대한 관심은 감소한다.

위와 같은 특성이 과거의 웹에서는 나타나지 않거나 더욱 강화되고 있는 특성이며, 또한 캐시의 성능을 좌우하는 교체정책에서 고려되어야 하는 지역성과 연관(요청, 집중도)이 있으므로 이러한 특성은 캐시 교체정책의 설계에 반영되어야 할 필요가 있다. 구체적으로 살펴보면, 캐시 교체정책의 성능을 평가하는 주요기준 중 하나인 적중률은 캐시 서버에 전체 요청되어진 횟수 중에 적중한 횟수의 비율로 정의되는데, 콘텐츠가 급속하게 많이 생성될 경우 사용자가 선택할 수 있는 전체 콘텐츠의 개수(사용자가 요청할 수 있는 콘텐츠의 선택폭)가 급속히 증가하게 됨을 의미한다. 또한 콘텐츠가 생성초기에 일시적으로 집중적인 관심을 받은 뒤, 시간에 따라 감소한다는 것은 각 콘텐츠의 가치(캐시될 필요가 있는지에 대한)가 생성 이후 시간에 따라 지속적으로 감소함을 의미한다. 그러므로 현재와 같이 위의 특성이 지속되거나 강화될 때 기존 교체정책의 성능 개선을 위해 위의 요소는 교체정책의 설계에 반영되어야 한다. 따라서 3장과 4장에서는 “콘텐츠의 시간에 따른 지속적 증가”, “생성된 콘텐츠에 대한 일시적이고 집중적 요청”, “오래된 콘텐츠에 대한 요청의 감소”라는 세 가지 특성을 반영하여 “보관수명(Shelf-Life Time)”을 정의하고, 이를 적용한 새로운 교체정책을 제안한 뒤 성능을 평가한다.

III. 보관수명 기반 웹 캐시 교체정책

일반적으로 캐시 교체정책의 성능을 평가하는 기준에는 (1) HR(적중률, Hit Rate), (2) BHR(byte 적중률, Byte Hit Rate), (3) DSR(지연절약비, Delay Saving Ratio)가 있는데²⁾, 본 논문에서는 (1)의 기준에서 성능을 향상시킬 수 있는 캐시 교체정책의 설계를 목표로 하였다. 적중률은 일반적으로 캐시 메모리의 크기에 영향을 받는다. 캐시 교체정책이 일관성있는 기준에 따라 콘텐츠의 우선순위를 판단할 때, 캐시 메모리의 증가는 향후 참조될 가능성이 높은 콘텐츠를 더욱 많이 보관할 수 있으므로 적중률이 메모리에 상당부분 의존적이라는 것은 직관적으로 이해된다. 또한 앞에서 언급했듯이 적중률은 캐시가 운영되는 환

경에 의존적인데, 이는 지역성과 관련이 있다. 왜냐하면 캐시가 운영되는 환경에서 어떠한 지역성이 만들어지느냐가 교체정책의 설계를 좌우하기 때문이다. 예를 들어 컴퓨터를 설계할 때, 컴퓨터의 하드웨어 환경에서 캐시의 대상은 프로그램 일부이다. 그런데 프로그램의 설계에서 사용되어지는 “배열(Array)”, “행렬(Matrix)”과 같은 자료구조는 공간적 지역성을 만들어내는 반면, “루프(Loop)”, “스택(Stack)” 등과 같은 구조는 시간적 지역성을 강하게 한다. 그런데 네트워크 환경에서의 지역성은 사용자들이 콘텐츠를 어떻게 요청하는지에 좌우되므로, 앞에서 살펴본 특성에 따라 콘텐츠 생성초기 사용자들의 일시적인 관심 집중과 시간에 따른 감소는 초기에 강했던 시간이 지날수록 지역성이 점차 약해진다는 것을 의미한다. 그런데 많은 콘텐츠가 지속적으로 생성되므로 매시간 콘텐츠 상호간 지역성의 우세는 그 정도를 달리할 것이다.

따라서 이러한 특성이 반영된 교체정책을 설계하기 위해 착안한 것이 생성시간이라는 요소를 반영하는 것인데, 순수하게 생성시간만을 고려했을 때 교체정책의 우선순위 판단은 불합리할 수 있다. 왜냐하면 신규 생성된 콘텐츠의 과급력²⁾이 경미할 경우에는 많이 참조되고 있던 기존의 콘텐츠들이 생성시간이 빠르다(오래되었다)는 이유로 캐시 저장공간에서 모두 방출되어버릴 것이기 때문이다. 따라서 위의 단점을 보완할 수 있도록 “보관수명(Shelf-Life Time)”이라는 용어를 정의하고 콘텐츠의 보관수명을 산출 또는 상호간 비교가 가능한 교체정책의 설계가 요구되어진다.

이를 위해 본 논문에서는 “보관수명”을 “콘텐츠가 생성된 시점”에서부터 “사용자들의 충분한 관심을 받지 못해 캐시 할 가치를 잃는 특정 시점”까지로 정의하는데, 이 경우 크게 두 가지 방법에 의해 콘텐츠 상호간 보관수명 비교가 가능하다. 첫째는 각 콘텐츠의 보관수명을 산출하여 비교하는 직접적인 방식이고, 다른 하나는 콘텐츠의 보관수명을 산출하지는 않더라도 정황적 판단에 의해 콘텐츠 상호간 누가 우위에 있는지 비교가 가능토록 하는 간접적인 방식이다. 직접산출 방식은 관련연구를 통해 보았던 Ted Karczewski의 발표처럼 비디오 콘텐츠는 3시간, SNS는 1시간 등 콘텐츠 특성을 일반화하여 각각의 보관수명을 직접 산출함으로써 달성 가능하다. 이러한 방식을 수많은 다양성이 있는 웹 환경에 적용하기엔 일반화에 있어 많은 어려움이 따를 수 있으므로, 몇 가지 요소의 비

2) 본문에서 과급력은 사용자들로부터 콘텐츠가 생성초기 얻는 관심(요청)의 크기(횟수)로 정의하며, 과급력이 클수록 콘텐츠가 생성초기에 다수로부터 집중적인 요청을 받음을 의미한다.

교와 계산, 그리고 절차적 방식에 의해 이루어질 수 있는 간접적인 방법으로 문제를 해결하고자 한다.

3.1 LRU+SLT

기존의 LRU는 가장 최근에 요청된 것이 교체되지 않을 우선순위를 갖도록 하는 방식으로, 구현이 간단하고 요청 성향의 변화에 어느 정도 적응적이지만, 일시적인 관심이나 개별 사용자의 요청들에도 매우 민감하게 반응한다. 따라서 현재와 같이 웹 환경에서 콘텐츠가 지속적으로 생성될 때, LRU의 성능은 때때로 매우 낮다. 이는 많은 사용자의 관심이 제각각일 경우 특히 두드러진다. 다수의 콘텐츠에 대한 독립적인 요청은 캐시 우선순위에 대한 LRU정책에 따른 결정의 신뢰성을 손상시킨다. 때문에 이를 보완할 수 있는 의사 결정방법이 필요한데, 보관수명이라는 특성을 이용한다면, 간단하면서도 신뢰성있는 의사결정이 가능하다. LRU+SLT정책에서는 캐시 적중(hit)과 캐시 실패(miss)의 조건에 따라 동작을 달리한다. (1) 제한한 정책에서 캐시가 실패했을 때 캐시 서버는 원천 서버에 콘텐츠를 요청하게 되는데, 이를 통해 획득한 콘텐츠를 캐시 목록의 하위부터 탐색하여 자신보다 신규 콘텐츠인 콘텐츠를 찾아, 하위 엔트리에 저장한다. (2) 반면 캐시 적중시에 해당 콘텐츠는 LRU정책과 동일하게 다른 조건이나 변수와 관계없이 목록의 최상위에 위치시킨다. 이러한 과정의 반복을 통해 LRU+SLT정책은 캐시 메모리의 하위부분을 콘텐츠의 생성일자 순으로 정렬하고, 상위부분은 최근 반복해서 요청되고 있는 콘텐츠를 위치시킨다. 때문에, LRU+SLT정책을 통해 교체되어지는 콘텐츠는 “최근에 요청되지 않는 상대적으로 오래된 콘텐츠”이다. 다시 말해, 보관수명이 다한다는 의미를 앞 절의 보관수명 정의에서처럼 오래되어 상대적으로 요청되어지지 않는 콘텐츠라고 판단한다.

그림 2와 3을 통해 보듯이 캐시 메모리에 콘텐츠A~E가 저장되어 있을 때, 새로운 콘텐츠 N에 대한 사용자 요청은 적중하지 않으므로 (①), 캐시 서버는 원천 서버에서 콘텐츠 N을 획득하여 캐시 메모리를 하위 엔트리(D)부터 탐색 후, 콘텐츠 N 생성시간과 가장 근접한 생성시간을 가진 엔트리 하위에 보관한다. 이 때, 다시 한 번 사용자의 콘텐츠 N에 대한 요청이 있을 경우, 캐시가 적중하므로(②), N은 메모리의 최상위 엔트리로 이동하게 된다. 이러한 과정의 반복을 통해 앞서 언급했듯이 상위 엔트리에는 근래에 반복 요청되는 최신 콘텐츠를 보관하고, 하위 엔트리에는 근래에 요청되지 않은 오래된 콘텐츠들을 보관하게

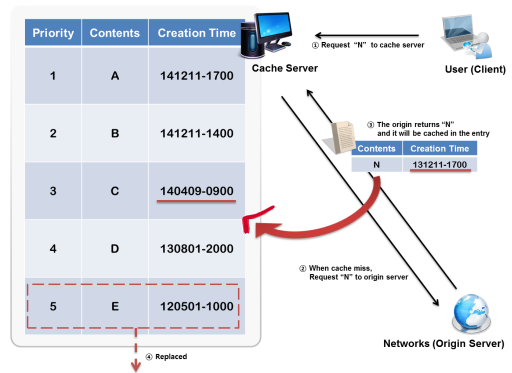


그림 2. 캐시 실패시 LRU+SLT의 동작
Fig. 2. Operation of LRU+SLT in the event of cache miss

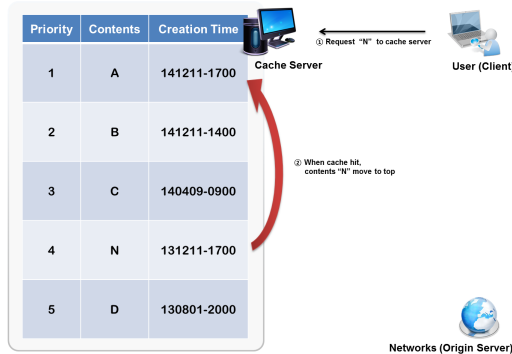


그림 3. 캐시 적중시 LRU+SLT의 동작
Fig. 3. Operation of LRU+SLT in the event of cache hit

된다. 즉, LRU+SLT에서는 생성시간과 최근 요청의 두 가지 요소를 이용하여 “신규 생성되었으면서 반복 요청되는 콘텐츠의 보관수명”이 “오래되었으면서 산발적으로 요청되는 콘텐츠의 보관수명”에 비해 더욱 오래남은 것으로 판단하고, 이를 통해 “상대적으로 보관수명이 다했다고 판단되는 콘텐츠”를 메모리에서 제거한다.

3.2 LFU+SLT

한편, LFU는 사용자의 콘텐츠에 대한 관심이 변화할 때, 누적된 요청횟수로 인한 우위가 사라질 때까지 이러한 변화 양상의 반영이 어렵다. 이러한 문제점을 해결하기 위해 보정(Aging)과 같은 기법들이 도입되었지만, 무엇을 기준으로 보정할 것인지에 대한 판단이 중요하다. 본 논문에서 제안하는 LFU+SLT는 생성시간을 이용하여 콘텐츠의 요청횟수를 보정하는 방식으로 기존의 LFU에 보관수명의 개념을 도입한다. 즉, 어떤 콘텐츠가 상대적으로 생성된 시간에 비해 총

분한 요청을 받지 못하고 있는가를 판단한다. 이를 수식으로 표현하면 다음과 같다.

$$P_n = \frac{R_n}{C^{(T-T_0)}} \quad (1)$$

수식에서 T는 현재시간, T₀는 콘텐츠의 생성시간이다. 콘텐츠 n의 보정된 참조횟수 P_n은 콘텐츠의 전체 참조횟수 R_n을 보정상수 C에 콘텐츠가 존재한 시간 (T - T₀) 지수승으로 보정한다. 따라서 콘텐츠의 생성시간(T₀)이 최근일수록(T-T₀가 작을수록), 참조횟수 R_n의 값은 P_n에 그대로 반영되어지며, 계산되어진 보정 참조횟수 P_n이 큰 콘텐츠일수록 캐시 우선순위를 갖는다. 다시 말해, 보정된 참조횟수 P_n은 콘텐츠가 생성되어진 이후 콘텐츠가 얻은 참조의 정도를 의미하며, 존재했던 시간에 비해 많은 요청을 얻지 못한 콘텐츠는 캐시에서 우선순위를 얻지 못하고 방출될 것이다.

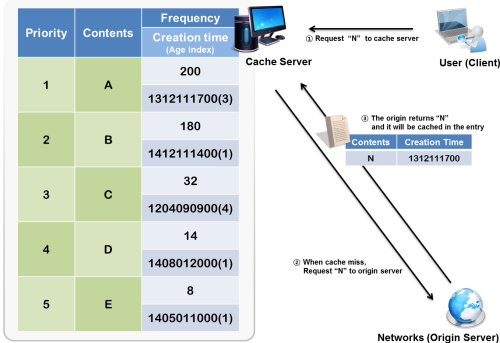


그림 4. LFU+SLT의 캐시 엔트리 구조
Fig. 4. Cache entry structure for LFU+SLT

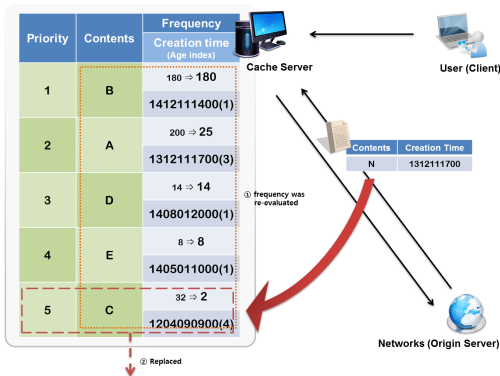


그림 5. LFU+SLT의 동작
Fig. 5. Operation of LFU+SLT

그림 4와 5에서 보는 것과 같이 캐시 메모리에 콘텐츠 A~E가 저장되어 있을 때, 새로운 콘텐츠 N이 요청되면 기존 LFU 교체정책에서는 단순히 요청횟수가 가장 적은 콘텐츠 E를 방출했지만, LFU+SLT에서는 위 함수에 의해 보정된 참조횟수를 판단하여 “오래 전에 생성되었지만, 충분한 요청을 얻지 못하고 있는 콘텐츠(C)”가 “최근에 생성되어 충분한 요청을 얻지 못하는 콘텐츠(D, E)”보다 상대적으로 보관수명의 도래가 가까워졌다고 판단하고, 이를 우선적으로 방출한다.

IV. 실험 및 분석

4.1 실험방법 및 환경 구성

앞서 언급했듯 현재의 웹 환경은 콘텐츠 생성 후 요청 감소, 파급력 증가, 빠른 콘텐츠 생성과 같은 특성을 갖는데, 실제 환경에서 이러한 특성은 무수한 다양성을 갖기 때문에 일반화가 어려운 요소이다. 이러한 어려움을 해결하고 제안한 LRU+SLT와 LFU+SLT의 성능을 기존의 캐시 교체정책과 비교하기 위해 본 장에서는 (1) 요청횟수 감소, (2) 파급력, (3) 콘텐츠 생성량이라는 세 가지 특성을 낮은 수준에서 높은 수준으로 변화시키면서 제안한 교체정책의 성능과 기존의 교체정책을 비교 및 해석하고, 특정구간에서 세부적 관찰을 통해 교체정책의 신뢰성을 비교하고자 한다.

표 2는 이를 위한 실험환경을 도표화한 것으로 (1) 콘텐츠 생성은 초당 10~60개 (2) 오래된 콘텐츠 전체에 대한 요청 감소는 5~25% (3) 새롭게 생성되는 콘텐츠가 갖는 파급력은 초당 1~600회 범위에서 변화시키면서 특성이 거의 나타나지 않는 환경에서부터 높은 환경에 이르기까지 순차적으로 모의한다. 개별 요소가 성능에 미치는 영향을 분석하는 2절에서는 분석을 위한 요소를 제외한 나머지를 변인 통제하여 결과를 산출하고 세부적인 의미를 분석한다. 난수 생성기는 기반 시뮬레이터가 제공하는 생성기를 사용하였으며, 신뢰도 상승을 위해 반복수행하여 평균값을 도출하였고, 대조실험에는 난수모의가 동일한 값을 얻도록 하여 대조실험이 동일조건에서 공정히 비교될 수 있도록 구성하였다. 또한 콘텐츠 전체에 대한 요청의 분포는 zipf분포를 따르도록 하였다.^[16]

표 2. 시뮬레이션 환경
Table 2. Simulation variables configuration

Variable Configuration	Value
Simulator	NS-3
Number of Total Contents	100000
zipf- α Index (Degree of Popularity Distribution)	0.8
Ratio of Cache Size	0.015 ~ 0.040%
Request for Continuously Lasting Contents	1000 / sec.
1) Request Reduction Rate for Old Contents	5 ~ 25%
2) Request for Newly Created Contents	1 ~ 600 / sec.
3) The Number of Content reation	10 ~ 60 / sec.

4.2 요청횟수 감소, 파급력, 생성량 증가에 따른 교체정책 성능 비교

4.2.1 생성 후 요청횟수 감소에 따른 교체정책 성능 비교

그림 6은 캐시 크기와 신규 콘텐츠의 파급력을 동일하게 유지한 상태에서 오래된 콘텐츠에 대한 요청의 감소정도가 변화할 때 교체정책 성능변화를 나타낸 도표이다. 전반적으로 보았을 때 모든 정책의 성능이 감소비율 증가에 따라 약간 상승하는 것으로 보여지나, 세부적으로 살펴보면 차이가 나타난다. 그림 7은 그림 6의 실험 중 감소율 20%일 때 각 교체정책의 성능변화를 시간에 따라 추적한 그래프이다. 그림 7에서 보는 바와 같이 LRU와 LRU+SLT정책은 최근참조 기반 정책의 특성상 요청횟수의 변화(감소)가 지속 되더라도 시간에 따른 적응률은 소폭 상승하는 모습

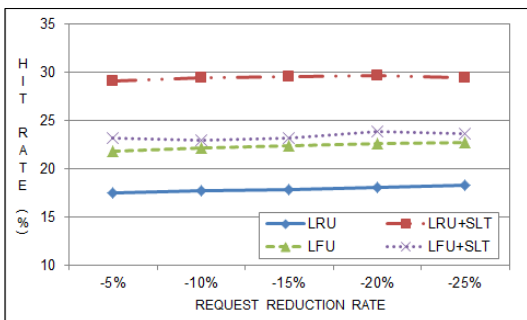


그림 6. 콘텐츠 요청횟수 감소폭에 따른 교체정책 성능 변화
Fig. 6. Performance change between policies in decreasing requests for long-lasting contents

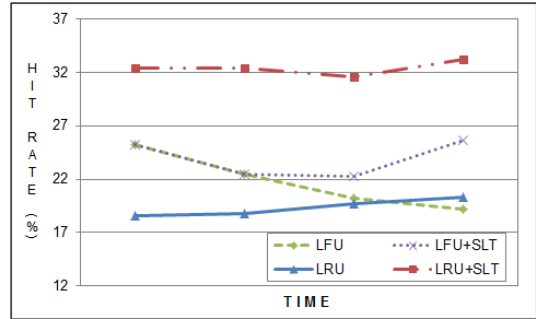


그림 7. 콘텐츠 요청횟수 감소시 시간에 따른 교체정책 성능 변화
Fig. 7. Chronological performance change between policies in decreasing requests

이다. 반면, LFU와 LFU+SLT정책은 초기에는 성능이 유사하게 나타나지만, 점차 시간이 흐를수록 LFU+SLT의 성능이 LFU와는 다른 양상으로 전개되는 모습으로 나타나는데, 모의 초기에는 모든 콘텐츠의 생성시간이 유사하고 누적되는 요청이 적으므로 성능차이가 나타나지 않다가, 콘텐츠가 누적되고 생성시간이 다양화 될수록 LFU가 변화의 양상을 따라가지 못하는데 반해 LFU+SLT는 적응적인 모습을 보여준다.

4.2.2 파급력 증가에 따른 교체정책 성능 비교

그림 8은 신규 콘텐츠가 생성초기 얻는 요청의 크기에 따른 교체 정책의 성능변화이다. 앞선 실험과 마찬가지로 이를 측정하기 위해 캐시 용량과 오래된 콘텐츠에 대한 요청의 감소비율을 동일하게 통제된 상태에서 콘텐츠의 생성초기 집중되는 요청크기를 변화시켰다.

그림 8에서 보여지는 것과 같이 모든 정책이 새로운 콘텐츠의 파급력이 클수록 성능이 향상하는 모습을 보였다. 새로운 콘텐츠가 얻는 요청이 크다는 것은

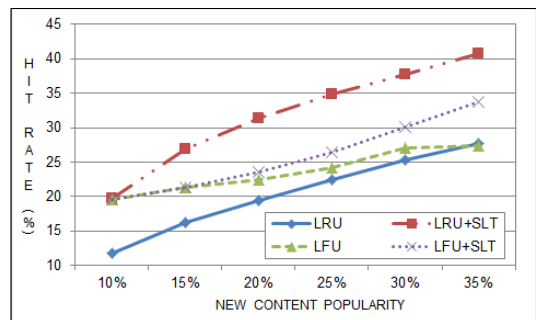


그림 8. 신규 콘텐츠 파급력 증가에 따른 교체정책 성능 변화
Fig. 8. Performance change between policies in increasing popularity with newly created contents

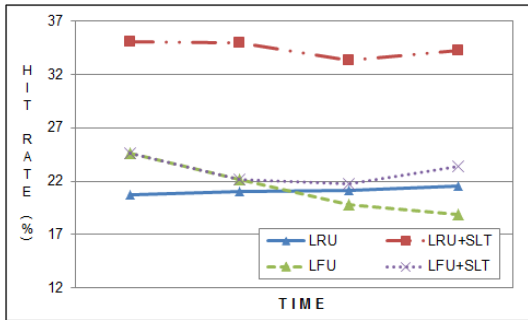


그림 9. 신규 콘텐츠 파급력 증가시 시간에 따른 교체정책 성능 변화
 Fig. 9. Chronological performance change between policies in increasing popularity

기존에 캐시되어 있던 콘텐츠보다 잦은 빈도로 많이 요청됨을 의미하므로 이는 직관적으로도 예측가능하다. 또한 관련연구에서 요청횟수 기반 정책의 문제점으로 지적했던 사용자의 관심이 변화할 때 성능문제가 위의 실험에서도 확연히 보여지는데, 신규 콘텐츠에 대한 일시적 요청이 증가하자 LRU의 성능이 LFU를 점차 초과하는 모습이 그것이다. 논문을 통해 제안했던 보관수명 기반의 LRU+SLT와 LFU+SLT정책은 모두 기존의 정책에 비해 향상된 성능을 보여준다. LRU+SLT은 시간에 따라 적중률이 점차 특정값으로 수렴하는 것으로 보여지는데, 이는 캐시용량의 한계로 인한 것으로 판단된다. 또한 LRU와 LRU+SLT의 성능이 유사하게 증가하는데 비해, LFU와 LFU+SLT의 성능은 파급력이 클수록 성능의 차이 심화된다. 이를 통해 볼 때 제안한 LFU+SLT 정책이 변화의 양상을 어느 정도 실시간으로 수용하며 콘텐츠에 대한 캐시 우선순위를 판단할 수 있음을 보여주는 부분이다. 그림 9은 동일한 실험에서 파급력이 초당 200이하일 때의 성능의 변화를 시간대별로 추적한 도표인데 앞의 실험과 동일한 이유로 LFU와 LFU+SLT가 초기에 유사한 성능을 보이나 시간이 지날수록 성능차이가 심화되어 나타난다.

4.2.3 콘텐츠 생성량 증가에 따른 교체정책 성능 비교

그림 10은 신규 콘텐츠의 생성량에 따른 성능변화이다. 예측한바와 같이 콘텐츠 생성량이 많을수록 기존의 정책과 제안한 교체정책 모두 성능은 저하된다. 하지만, 그림에서 보여지는 바와 같이 각 정책이 특정 값으로 수렴하는 양상을 보이고 있으며 성능에 있어서 제안한 교체정책이 약간 더 우위에 있음을 볼 수 있다.

또한 지속적으로 콘텐츠가 생성되는 환경에서 시간

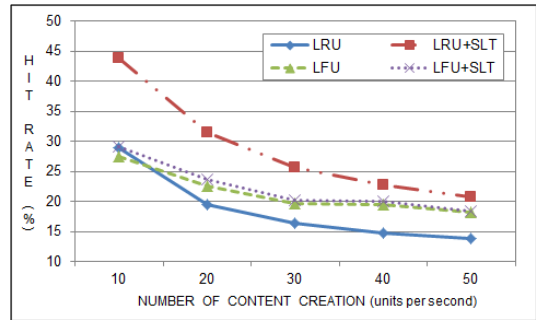


그림 10. 콘텐츠 생성량 증가에 따른 교체정책의 성능 변화
 Fig. 10. Performance change between policies in increasing content creation

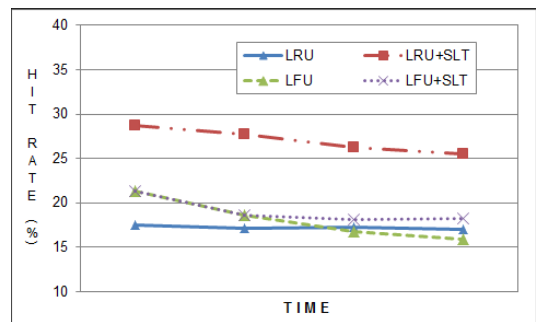


그림 11. 콘텐츠 생성량 증가시 시간에 따른 교체정책 성능 변화
 Fig. 11. Chronological performance change between policies in increasing popularity

에 따른 캐시율 변화를 세부적으로 살펴보면 그림 11에서 보는 바와 같이 시간이 지날수록 콘텐츠가 증가하게 되고 LFU는 적중률이 지속적으로 낮아지는 반면 LFU+SLT는 특정 수준을 유지함으로써 장기간 캐시 운영시 성능의 차이를 보이게 될 것으로 예상된다. 반면, LRU가 특정 수준의 적중률을 지속적으로 유지하는 반면, LRU+SLT가 조금 낮아지는 양상을 보이는 것으로 볼 수 있으나, 전체적으로 볼 때, LRU+SLT가 LRU에 비해 지속적으로 높은 수준의 성능을 보이면서 LFU+SLT와 동일하게 특정 수준에서 수렴하며 할 것으로 예상된다.

4.3 저장공간 크기 증가에 따른 캐시 교체정책 성능 변화 비교

2절에서 살펴본 바와 같이 (1) 요청횟수 감소, (2) 파급력, (3) 콘텐츠 생성량이라는 세 가지 특성이 존재하는 환경에서는 일반적으로 제한한 캐시 교체정책의 성능이 우위에 있음을 알 수 있다. 또한 그러한 양상이 복합적으로 나타나고 두드러질수록 제한한 캐시 교체정책의 성능이 기존의 정책에 비해 우수할 것임

을 판단할 수 있다. 하지만, 이러한 성능이 특정 캐시 용량에서만 나타나는 것이라면 교체정책의 신뢰성이 확보될 수 없을 것이다. 왜냐하면 일반적으로 낮은 기준에 의해 우선순위가 판단되거나, 적절하지 못한 교체정책은 캐시 용량이 증가하더라도 적중률이 증가하지 않거나 오히려 감소하는 기형적인 모습을 나타낼 것이기 때문이다. 때문에 본 절에서는 위의 실험 중 임의의 구간(초당 10~20개의 신규 콘텐츠가 생성되고, 생성초기에 초당 200개 이하의 요청을 받으며, 모든 콘텐츠에 대한 요청이 20%씩 감소)을 선정하여 세 가지 특성을 유지한 상태에서 캐시 저장공간의 크기를 증가시키면서 교체정책의 성능변화를 측정하였다.

그림 12는 선정된 구간에서 기존의 교체정책과 제안된 교체정책의 성능을 비교한 도표이다. 그림에서 보는 바와 같이 LRU와 LFU는 각각 약 18%와 22%의 성능을 보인 것에 반하여, LRU+SLT와 LFU+SLT는 약 30%와 24%의 적중률을 각각 보임으로써 기존의 교체정책의 성능에 비해 제안한 정책이 약 2~12%가량 향상된 적중률을 보였다. 이는 기존의 정책에 비해 제안한 정책이 환경적 특성을 적절히 반영한 결과로 보여진다. 특이한 점은 LRU에 SLT를 적용했을 때의 성능 증가폭이, LFU에 SLT를 적용했을 때의 성능 증가폭보다 높음에 따라 LRU+SLT가 LFU+SLT에 비해 성능면에서 우위를 보인다는 점이다. 이는 시간적 지역성의 반영이라는 LRU의 강점을 SLT를 통해서 보완함으로써 상승효과가 발생하는 것으로 보여진다. LFU+SLT 도 시간적 지역성을 통해 보완함에 따라 LFU에 비해 향상된 성능을 보여준다. 더불어 LFU+SLT가 시간이 어느 정도 지나서 LFU와 성능 격차가 나타나는 것을 보여주는 앞 절의 그림 7, 9, 11을 함께 고려해본다면, LFU+SLT의 성능은 조금 더 상승될 여지가 있는 것으로 예측해 볼 수 있다.

해당 구간에서 캐시 용량을 증가시키면서 적중률

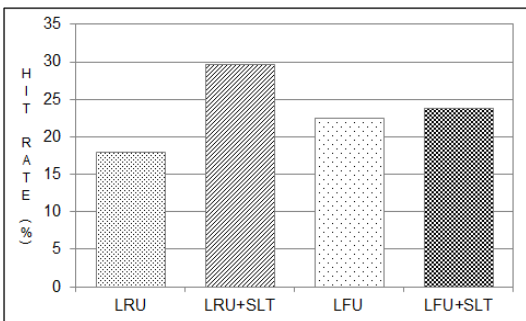


그림 12. 기존 정책과 제안된 정책의 적중률 비교
Fig. 12. Hit rate comparison between policies

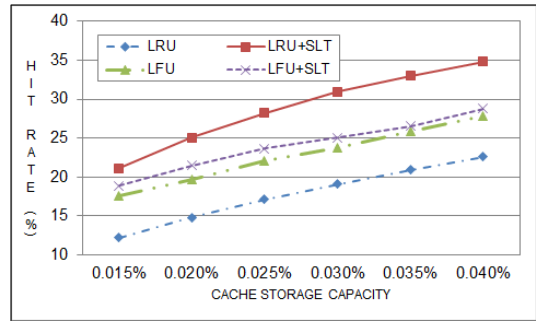


그림 13. 캐시 용량에 따른 적중률 변화
Fig. 13. Hit rate change in increasing Cache Capacity

변화를 측정하였을 때에는 기존 정책과 제안한 정책 모두 그림 13과 같은 적중률의 지속적인 상승을 보였다. 세부적으로 볼 때, 캐시 용량증가시 LFU+SLT는 LFU에 비해 지속적으로 약 2~3%가량 우수한 성능을 보이며, LRU+SLT의 경우 LRU에 비해 8~12%의 적중률 향상을 관찰할 수 있다. 이처럼 캐시 용량 증가에 따른 제안한 정책의 성능 향상은 캐시 공간 전체가 타당하고 일관적 기준에 의해 정렬되고 있다는 것을 반증한다.

V. 결 론

지금까지 현재의 웹 환경이 갖는 특성을 정의하고, 이러한 특성을 보다 적절히 반영할 수 있는 새로운 교체정책이 필요함을 밝히고, 새롭게 제안한 캐시 교체정책의 성능을 평가하였다. 최초로 캐시 메커니즘이 고안되었을 때 캐시는 운용되는 환경적 특성에서 출발했다. 환경적 특성을 기반으로 메커니즘의 성능을 최대한 발휘할 수 있는 교체정책을 설계하였다. 오랜 기간 캐시 메커니즘은 다양한 분야에 응용되어왔고, 현재에도 상당한 영향력을 발휘한다. 하지만, 캐시 메커니즘이 응용되어진 각각의 분야 역시 그 동안 수많은 변화를 통해 환경적인 변화가 있었다. 따라서 적절한 캐시 교체정책의 설계를 위해서는 환경적 변화를 통해 새롭게 나타난 특성이 반영되어야 할 필요가 있다. 웹 환경은 초기에 비해 사용자들의 선택폭이 넓어졌고, 시시각각으로 사용자의 관심이 변화하며, 사용자는 주도적이 되었다. 이러한 특성을 반영하기 위해 본 논문에서는 보관수명의 개념을 도입하여 현재 웹 환경에 적합한 캐시 교체정책을 제안하고 성능을 평가하였다. 앞서 이야기한 바와 같이 새로운 교체정책은 기존에 비해 조금 더 나아진 성능을 보이며, 이는 오래된 콘텐츠에 대한 사용자의 요청 감소정도가 클

수록, 그리고 새로운 콘텐츠의 파급력이 클수록 좋은 성능을 보인다. 물론, 논문에서 제안한 방식이 변화하는 양상에서 교체정책을 위한 최선의 방법이 될 수는 없다. 보다 세밀하게 환경적 요소가 검토되고, 정확히 특성화되어진다면, 더욱 우수하고 발전적인 교체정책의 설계가 가능할 것이다. 따라서 본 논문을 통해 캐시가 운용되는 환경에 대한 세밀한 고찰이 이루어지고, 각각의 환경에 더욱 적합한 교체정책이 제안될 것으로 기대한다.

References

[1] P. J. Denning. "The locality principle," *Commun. ACM*, vol. 48, no. 7, pp. 19-24, Jul. 2005.

[2] B. D. Davison, "A web caching primer," *IEEE Internet Computing*, vol. 5, no. 4, pp. 38-45, 2001.

[3] S. Podling and L. Böszörményi, "A survey of web cache replacement strategies," *ACM Computing Surveys(CSUR)*, vol. 35, no. 4, pp. 374-398, 2003.

[4] C. Aggarwal, J. L. Wolf, and P. S. Yu, "Caching on the world wide web," *IEEE Trans. Knowledge and Data Eng.*, vol. 11, no. 1, pp. 94-107, 1999.

[5] S. Jin and A. Bestavros, "GreedyDual* Web caching algorithm: exploiting the two sources of temporal locality in Web request streams," *J. Comput. Commun.*, vol. 24, no. 2, pp. 174-183, 2001.

[6] J. Chen and L. Subramanian, "Interactive web caching for slow or intermittent networks," in *Proc. 4th Annu. Symp. Comput. for Development (ACM DEV-4 '13)*, no. 5, 2013.

[7] Y. Li, H. Xie, Y. Wen, and Z. Zhang, "Coordinating in-network caching in content-centric networks: model and analysis," *IEEE ICDCS*, pp. 62-72, Philadelphia, Jul. 2013.

[8] T. o'Reilly, *What is web 2.0*, O'Reilly Media, Inc., 2009.

[9] Cisco, *Cisco visual networking index: Forecast and methodology, 2013-2018*, Retrieved Nov., 2, 2015, from "http://www.cisco.com/c/en/us/solutions/collateral/service-provider/ip-ngn-ip-n

ext-generation-network/white_paper_c11-481360.html

[10] K. C. Laudon and J. P. Laudon, *Essentials of management information systems*, Upper Saddle River : Pearson, 2011.

[11] T. Karczewski, *How long does your content last online?*, Retrieved Nov., 10, 2014, from <http://www.skyword.com/contentstandard/enterprise-marketing/how-long-does-your-content-last-online-data/>

[12] M. Cha, H. Kwak, P. Rodriguez, Y. Ahn, and S. Moon. "Analyzing the video popularity characteristics of large-scale user generated content system," *IEEE/ACM Trans. Networking (TON)*, vol. 17, no. 5, pp. 1357-1370, 2009.

[13] T. Zaman, E. B. Fox, and E. T. Bradlow, "A bayesian approach for predicting the popularity of tweets," *The annals of Applied Statistics*, vol. 8, no. 3, pp. 1583-1611, 2014.

[14] G. Szabo and B. A. Huberman, "Predicting the popularity of online content," *Commun. ACM*, vol. 53, no. 8, pp. 80-88, 2010.

[15] C. Castillo, M. El-Haddad, J. Pfeffer, and M. Stempeck, "Characterizing the life cycle of online news stories using social media reactions," in *Proc. 17th ACM Computer supported cooperative work & social computing*, pp. 211-223, 2014.

[16] L. Breslau, P. cao. L. fan, G. Philips, and S. Shenker, "Web caching and zipf-like distributions : Evidence and implications," *IEEE INFOCOM'99*, vol. 1, 1999.

한 성 민 (Sungmin Han)



2007년 : 육군사관학교 전산학과 졸업
 2014년~현재 : 국방대학교 컴퓨터공학과 석사과정
 <관심분야> Operation System, Software Defined Networking, Cyber Warfare

박 흥 순 (Heungsoon Park)



2002년 : 육군사관학교 전산학과 졸업

2007년 : Air Force Institute of Technology 컴퓨터공학과 석사

2012년~현재 : 국방대학교 컴퓨터공학과 박사과정

<관심분야> Military Communications, MANET, Content-Centric Networking, Internet of Things, Network Security, Cyber Warfare, Network Modeling & Simulations

권 태 옥 (Taewook Kwon)



1986년 : 육군사관학교 전산학과 졸업

1995년 : 미 해군대학원 컴퓨터공학과 석사

2001년 : 연세대학교 컴퓨터공학과 박사

2007년~현재 : 국방대학교 컴퓨터공학과 교수

<관심분야> Next Generation Networking, Sensor Networking, Cloud Computing, RFID/USN, Virtual Reality, Content-Centric Networking, Software Defined Networking