Journal of Digital Convergence
Vol. 16. No. 1, pp. 271-279, 2018

ISSN 1738-1916
https://doi.org/10.14400/JDC.2018.16.1.271

MHT 기반 콘텐츠 인증 기술의 전송량 개선

김대엽
수원대학교 정보보호학과

Network Overhead Improvement for MHT-based Content Authentication Scheme

KIM DAEYOUB
Dept. of Information Security, Suwon Univ.

요 약 인터넷을 이용하여 콘텐츠를 보다 효율적으로 공유하기 위하여 P2P, CDN과 같은 다양한 기술들이 개발되어왔다. 이러한 기술들은 콘텐츠 배포자에 집중되는 요청 패킷이 네트워크에 분산된 다수의 노드들에 의해 처리되도록 설계되어, 네트워크 병목 문제를 해결할 뿐만 아니라 콘텐츠 배포 시스템과 네트워크의 상태와 상관없이 지속적으로 콘테츠를 배포할 수 있다. 그러나 분산 노드로부터 콘텐츠를 전송 받는 경우, 사용자가 실제 콘텐츠 전송 노드를 식별/인증할 수 없기 때문에 공격자 개입 및 악의적인 콘텐츠 변경을 통한 다양한 해킹이 가능하다. 그러므로 분산 노드/호스트를 이용한 네트워킹 기술의 경우, 콘텐츠 인증 기술이 반드시 필요하다. 본 논문에서는 CCN에 적용된 콘텐츠 인증 기술인 MHT 기반의 콘텐츠 인증 기법을 소개하고, MTH의 인증 정보 중복 전송 문제를 해결하여 전송량을 개선할 수 있는 방안을 제안하고, 기존 기술과의 성능 비교를 통하여 개선안의 성능을 평가한다.

주제어 : P2P, CDN, CCN, 데이터 인증, MHT

Abstract Various technologies have been developed to more efficiently share content such as P2P and CDN. These technologies take a common approach that request packets are responded by distributed network nodes, not by a single distributor. Such approaches not only resolve network congestion around content distributors, but also make it possible to distribute content regardless of the system and network status of content distributors. However, when receiving content from distributed nodes/hosts, not from authenticated distributors, users cannot practically identify which node/host sent content to them. Due to this characteristic, various hacking caused by the malicious modification of content is possible. Therefore, to make such approaches more secure, a content authentication technique is required. In this paper, we propose a improved operation of MHT used in CCN for authenticating distributed content. Then we evaluate the proposed method by comparing its performance with the existing technology.

Key Words : P2P, CDN, CCN, Data Authentication, MHT

1. 서론

실시간 데이터 및 대용량 콘텐츠를 유/무선 네트워크 기술을 이용하여 사용자에게 전송하고 공유하는 다양한 서비스가 소개되고 있다. 특히, 클라우드 기반의 콘텐츠 서비스가 보급됨에 따라 이와 같은 콘텐츠 서비스를 이용하는 네트워크 전송양은 폭발적으로 증가할 것으로 예상된다[1]. 그러나 네트워크를 통해 전송되는 데이터의 증가에 대응하기 위하여 유/무선 통신 선로를 전체적으로 확장시키는 것은 서비스 제공자에게 매우 많은 투자

*The paper was supported by The research grant of the University of Suwon in 2016 (No. 2017-0061).
*Corresponding Author : Kim DaeYoub (daeyoub69@suwon.ac.kr)
Received October 19, 2017 Revised December 13, 2017
Accepted January 20, 2018 Published January 28, 2018

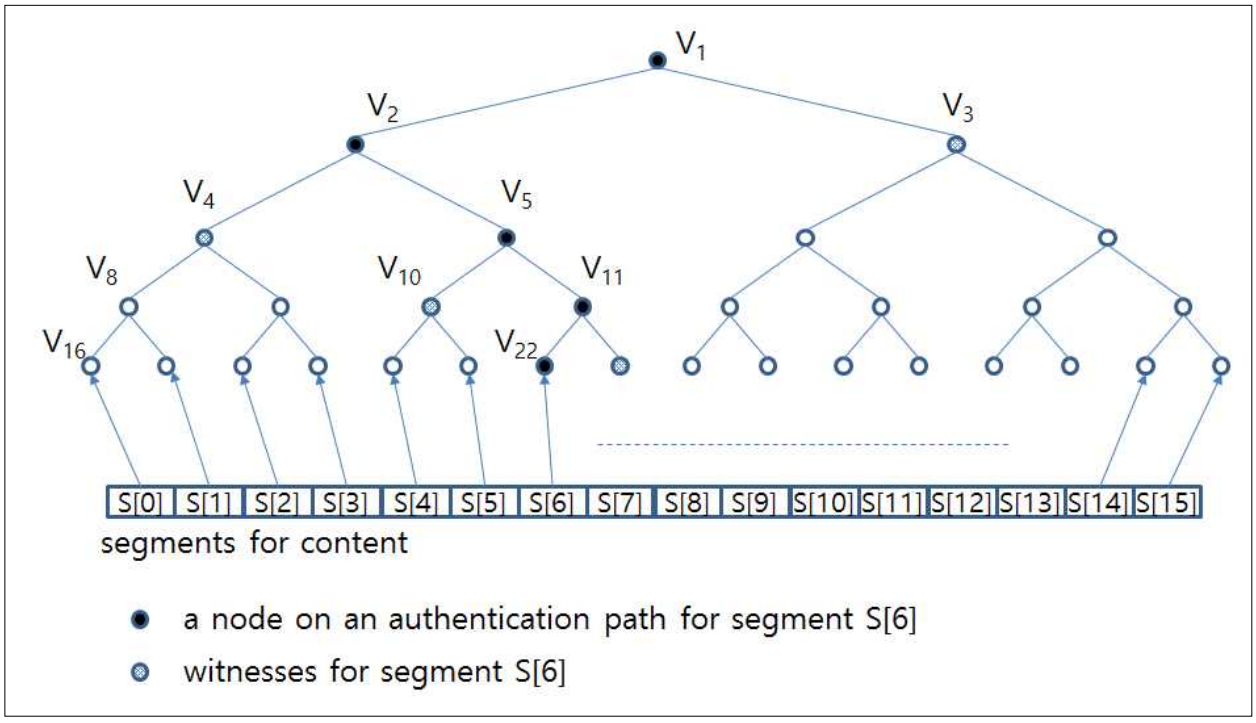
비용을 요구할 뿐만 아니라, 물리적 통신 선로 확장 속도에 비하여 네트워크를 통한 데이터 전송량 증가 속도가 훨씬 더 빠르게 증가하는 추세이기 때문에 물리적 해결책 외에 네트워크 효율성을 개선할 수 있는 기술적 대안이 요구 된다 [1, 2]. 이러한 기술적 대안 중 하나가 P2P (Peer-to-Peer) 네트워킹 기술과 CDN (Content Delivery Network) 서비스라 할 수 있다. P2P/CDN은 사용자가 요청하는 콘텐츠를 해당 콘텐츠의 원배포자 (Content Distributor/Provider, CP) 뿐만 아니라 동일한 콘텐츠를 이전에 다운로드 받은 사용자들이나 콘텐츠 프락시 서버(Content Proxy Server, CPS)와 같이 현재 요청된 콘텐츠를 저장하고 있는 다양한 소스들을 통하여 콘텐츠를 제공 받을 수 있도록 설계되었다. 이러한 기술들은 CP로 집중되는 콘텐츠 요청 메시지들을 네트워크 내에서 효과적으로 분산처리 함으로써 네트워크 효율성을 개선할 수 있을 뿐만 아니라, 원배포자의 시스템/네트워크 상태와 상관없이 지속적으로 콘텐츠를 배포할 수 있다 [3, 4].

인터넷의 초기 개발 목적은 원격 호스트들 사이의 네트워킹을 안전하게 연결하는 것이었기 때문에 현재 인터넷을 기반으로 하는 다양한 서비스들의 요구사항들을 고려하여 설계되지 않았다. 그러므로 급속한 데이터 전송량 증가로 인한 네트워크 병목현상과 호스트 간 인증 기술의 부재와 같은 취약한 보안 구조로 인한 보안 침해 사고뿐만 아니라, 모바일 기기 증가로 인한 디바이스의 빈번한 이동에 따른 비효율성 증가와 같은 다양한 문제점들이 드러나고 있다 [5]. 특히, 다양한 IT 융합 서비스의 증가와 모바일 기기 및 네트워크에 연결된 센서의 증가는 이와 같은 인터넷의 문제점들을 더욱 심화시킬 것으로 예상되기 때문에 인터넷이 갖고 있는 구조적인 문제들은 IT 융복합 서비스의 발전 및 저변 확대를 방해하는 주요 요인이 될 수 있다.

그러므로 인터넷이 갖고 있는 내재된 기술적 문제들을 해결하고, 효율적으로 보다 다양한 데이터 및 정보를 인터넷을 통하여 제공하기 위한 미래 인터넷 기술 및 아키텍처 연구가 활발하게 진행되고 있다 [6-8]. 정보 중심의 네트워킹 기술(Information Centric Networking, ICN)은 이와 같은 미래 인터넷 아키텍처 중 하나이다. ICN은 CP에게 집중되는 콘텐츠 요청 메시지를 네트워크 내에서 효율적으로 분산 처리하기 위하여 CPS나 네트워크 라우터와 같은 네트워크 장비/노드 (Network Node)에 콘텐츠를 저장한 후, CP를 대신하여 이들 CPS/노드들이 콘텐츠 요청 메시지에 직접 응답/처리하도록 설계 되었다. ICN 기술 중 하나인 콘텐츠 중심 네트워킹 (Content Centric, CCN)은 네트워크 노드에 구현된 콘텐츠 임시 저장 기능 (Caching)과 콘텐츠의 계층화된 고유 이름을 참조하여 패킷 라우팅 경로를 결정하는 기술을 이용하여 콘텐츠 요청 메시지를 네트워크 전송 경로 상의 중간 노드들이 직접 응답할 수 있도록 설계되었다 [7-10].

그러나 ICN의 아키텍처들은 콘텐츠 전송의 효율성을 높이기 위해 CPS나 네트워크 노드에 캐싱 되어 있는 콘텐츠를 활용하므로, P2P/CDN처럼 콘텐츠를 사용자에게 제공하는 시스템/노드가 불특정하다. 그러므로 기존 호스트 중심 네트워킹과는 달리, 사용자가 콘텐츠를 수신했을 때, 실제 콘텐츠를 제공하는 시스템/노드를 사용자가 식별/인증할 수 없다. 이러한 취약점이 악의적으로 이용될 경우, 콘텐츠 위/변조가 가능하고, 위/변조된 콘텐츠를 이용한 다양한 공격이 가능할 수 있다. 그러므로 ICN을 이용하여 콘텐츠를 전송/배포하는 서비스를 구현할 때, 콘텐츠 인증 기술은 필수적으로 요구된다. 전송되는 콘텐츠의 인증을 위해 CCN은 콘텐츠에 해당 콘텐츠 최초 생성자 (Content Publisher)의 전자 서명이 첨부하도록 강제 규정하고 있다. 그러나 대용량 콘텐츠 전송을 지원하기 위해 CCN은 콘텐츠를 일정 크기 이하로 분해하여 세그먼트들로 단편화한 후, 각각의 세그먼트를 하나의 데이터로 간주하여 처리한다. 이 때, 콘텐츠를 구성하는 모든 세그먼트들을 인증하도록 했다. 그러나 개별 세그먼트 단위의 콘텐츠 인증 기술을 사용할 경우, 콘텐츠를 구성하는 전체 세그먼트들을 인증하는데 많은 시간이 소요되어 서비스 지연이 발생할 수 있다. 이와 같은 문제를 개선하기 위해서 CCN은 머클 해시 트리(Merkle Hash Tree, MHT)를 사용하여 콘텐츠 인증 및 개별 콘텐츠 세그먼트 인증을 동시에 수행할 수 있게 제안하였다. 그러나 [17]에서 MHT 운영 시에 발생되는 과도한 중복 연산이 지적되었고, 개선안이 함께 제안되었으나 인증을 위한 정보의 중복 전송은 해결하지 못했다.

본 논문에서는 MHT의 중복 연산 및 전송을 개선하는 새로운 운영 방법을 제안하고, 그 성능을 평가한다. 제안된 기법은 [17]에서 제안된 인증에 소요되는 시간 개선 기법을 바탕으로 인증을 위해 필요한 데이터의 전송 양을 감소시키는데 중점을 두었다.



[Fig. 1] MHT-based Content Authentication

2. MHT 기반 데이터 인증

2.1 CCN

네트워크 노드에 캐싱된 데이터를 활용하기 위하여, CCN은 다음과 같은 두 가지 차별화된 기술을 제안하고 있다:

(1) 데이터 이름 기반의 패킷 라우팅

(2) 전자 서명 기반의 패킷 인증

전자는 CCN 노드에 데이터 캐싱 기능을 구현할 때, 보다 효율적으로 캐싱 된 데이터를 탐색/관리할 수 있도록 한다. 후자는 캐싱 된 데이터가 사용자에게 전송되었을 때, 사용자가 수신된 데이터를 신뢰하고 이용할 수 있도록 한다 [7].

CCN은 콘텐츠 데이터 요청 메시지 (Interest) 전송에 의하여 네트워킹 프로세스가 진행되며, Interest에 대응되는 응답 메시지(Data)는 Interest가 전송된 경로의 역경로를 따라서 사용자에게 전송된다. 이 때, Interest 패킷 전송 경로 위에 있는 네트워크 노드는 Interest 패킷의 네트워크 계층 정보를 이용하여 요청된 데이터의 캐싱 여부를 확인해야 하므로, 네트워크 계층의 헤더 정보에 데이터의 식별자 정보가 포함되어야 한다. 특히, 특정 원격 호스트로부터 콘텐츠를 전송 받는 메커니즘이 아니기 때문에 특정 호스트와의 네트워크 접속을 위해 필요했던 호스트 식별자의 필요성이 매우 낮다. 그러나 네트워크 경로 상에 캐싱된 데이터가 없다면, 콘텐츠의 Publisher에게 Interest 패킷을 전송하기 위해 Publisher 식별자가 요구된다.

이와 같은 특성에 따라 CCN은 IP 주소와 같은 호스트 식별자 대신에 데이터 식별자와 Publisher 식별자 정보로 구성된 콘텐츠 이름(Content Name)을 네트워크 주소로 활용하고, 콘텐츠 이름을 기반으로 Interest/Data의 라우팅 경로를 결정한다. 또한, 콘텐츠 이름을 참조하여 Interest 패킷을 라우팅하기 위하여 콘텐츠 이름은 계층화하여 구성한다.

데이터의 위/변조 여부를 사용자가 검증 할 수 있도록 전송되는 각각의 Data는 Publisher의 전자 서명을 포함하고 있으며, 사용자는 Data에 첨부된 전자 서명 값을 검증하여 데이터의 Publisher를 인증하고, 동시에 데이터 위/변조 여부를 검증한다. CCN의 구성과 실제 운영 방안은 [17]에 자세하게 설명되어 있다.

2.2 MHT 기반 데이터 인증

CCN은 콘텐츠 Publisher 인증과 세그먼트 인증 기능을 제공한다. 콘텐츠 Publisher 인증은 수신된 콘텐츠의 Publisher를 식별하고 인증하는 것을 의미한다. 콘텐츠

Publisher 인증을 통해서 수신된 콘텐츠의 신뢰성을 일차적으로 검증할 수 있다.

전송 효율성과 안정성을 고려하여 CCN은 콘텐츠를 세그먼트들로 단편화한 후, 각각의 세그먼트를 독립된 Data로 처리한다. 만약 수신된 세그먼트가 요청한 콘텐츠의 세그먼트가 아니면, 콘텐츠 전체를 수신 한 후에도 해당 콘텐츠를 정상적으로 이용하지 못할 수도 있다. 그러므로 수신된 세그먼트가 사용자가 요청한 콘텐츠의 단편화된 세그먼트임을 검증하는 세그먼트 인증이 추가로 필요하다.

이와 같은 인증 요구사항을 만족시키기 위하여 CCN은 MHT 기반의 데이터 인증 기법을 제안하였다 [11-15].

[Fig. 1]는 MHT 기반의 콘텐츠 인증 절차를 예를 들어 설명한 것이다:

(1) 콘텐츠가 일정 크기를 갖는 N ($\leq 2^n$) 개의 세그먼트들로 단편화 되었다면, 2^n개의 리프 노드(Leaf Node)로 구성된 이진 트리(Binary Tree)를 구성한 후, N개의 세그먼트를 순서에 따라 리프 노드에 할당한다. 리프 노드의 노드 값은 할당된 세그먼트의 해시 값으로 계산되며, 리프 노드를 제외한 모든 노드의 노드 값은 두 자식 노드 (Child Node)의 노드 값들을 연접하여 해시한 값으로 계산된다. 즉, 노드 N_k의 노드 값을 V_k라 하면, V_k는 다음과 같이 계산 한다:

$$V_k = H(V_{2k} \| V_{2k+1}).$$

(2) 콘텐츠 생성자는 전자 서명 키 ($priK$)를 이용하여 루트 노드 값 V_1에 서명하여 $sign = E_{priK}(V_1)$을 생성한 후, 콘텐츠의 세그먼트 $S[i]$와 $sign$을 함께 패키징 하여 Data를 생성한다.

(3) $S[i]$ 전송을 위한 Data가 수신되었을 때, 사용자가 Data에 패키징 된 $sign$을 검증하기 위해서는 V_1을 계산할 수 있어야 한다. 이를 위하여 각각의 세그먼트 마다 $sign$ 검증에 필요한 중간 노드들의 노드 값들(인증 정보, Witness)을 계산한 후, Data에 $sign$와 함께 패키징 한다. 즉, 해당 세그먼트에 대응하는 리프 노드부터 루트 노드까지의 경로(Path)에 포함된 노드들의 형제 노드(Sibling Node)의 노드 값을 세그먼트와 함께 전송하여 사용자가 $sign$을 검증할 수 있도록 한다.

MHT를 이용할 경우, 첫 번째 세그먼트 검증 과정에서 전자 서명이 검증되면, 나머지 세그먼트들은 해시 값 계산 및 비교만으로 검증을 완료할 수 있기 때문에 모든 세그먼트에 전자 서명을 첨부하고, 이를 검증하는 오버헤드를 줄 일 수 있다. 또한, 상위 노드의 노드 값(해시 값)을 계산할 때 하위 노드들의 모든 노드 값(해시 값)이 영향을 주기 때문에 일부 노드의 값이 충돌 되어도 최종 노드 값은 충돌 영향이 매우 낮기 때문에 안전성이 보장된다. 그러나 MHT를 대용량 콘텐츠 배포에 적용할 경우, 각각의 세그먼트 마다 인증 정보를 전송해야 하고, V_1 계산을 위하여 하위 노드 값들을 반복적으로 계산해야만 한다.

[17]에서는 데이터 해시 값의 중복 계산정도를 분석하고, 이러한 중복 요소로 인하여 발생되는 비효율성을 개선하기 위하여 계산된 해시 값을 저장한 후, 재사용하는 방법을 제안하였다. [Fig. 2]는 [17]에서 제안된 동적 프로그래밍 기법을 활용한 MHT 운영 개선 기법을 예를 들어 설명 한다: 이전 세그먼트 S[i] 인증 시, 인증 경로 상의 노드 값들을 노드의 계층에 따라 스택에 저장된다. 이후, 다음 세그먼트 S[k] (k>i) 를 수신한 사용자는 S[k]를 인증하기 위해 계산해야 되는 인증 경로 상의 노드들과 이전 인증 단계에서 스택에 저장된 노드들을 하위 계층부터 계층 별로 비교하여 동일 노드를 탐색한다. 탐색된 노드 중 최하 계층의 노드를 선택한다. [Fig. 2]에서 S[2]를 인증하기 위해서 인증 경로 상에 있는 노드 중에서 노드 N_2를 선택한다. 이 경우, 세그먼트 S[2]를 인증하기 위해서 사용자는 루트 노드 N_1의 노드 값까지 계산하지 않고 노드 N_2의 노드 값까지만 계산 한 후, 스택에 저장된 값과 비교하면 충분하다. 이와 같은 중복 계산 요인을 개선함으로써 대용량 콘텐츠 전송 시 두드러지는 개선 효과를 나타낼 수 있음을 보였다.

3. MHT 기반 데이터 인증 전송량 개선

CCN은 네트워크를 통해 전송되는 데이터의 최대 크기를 규정하고 있기 때문에, 실제 콘텐츠를 전송하기 위해서는 콘텐츠를 규정된 데이터 크기 이하로 단편화하여 처리한다. 그러므로 대용량 콘텐츠의 경우 단편화된 데이터(세그먼트)의 개수에 따라서 MHT의 크기도 증가하게 된다. 데이터의 크기 (ds)와 해시 값의 크기 (hs)가 둘 다 상수 값이므로, 콘테츠의 크기 (cs), 콘텐츠의 전송량 (ts)와 세그먼트의 개수 (2^n) 사이에는 다음과 같은 관계가 성립 된다:

$$cs = \sum_{i=1}^{2^n} ds, \qquad (2)$$
$$ts = \sum_{i=1}^{2^n} (ds + (n \times hs))$$

또한 각각의 단편화된 데이터에는 n개의 MHT 인증 정보가 포함되기 때문에, 콘텐츠의 크기에 비례해서 ts가 추가적으로 증가하게 된다.

<Table 1> The duplicated transmission of MHT witness: $N = 2^n$

Level	$n=8$	$n=12$	$n=16$
L_{n-1}	$2^7 = 128$	$2^{11} = 2{,}048$	$2^{15} = 32{,}768$
L_{n-2}	$2^6 = 64$	$2^{10} = 1{,}024$	$2^{14} = 16{,}384$
L_{n-4}	$2^4 = 16$	$2^8 = 526$	$2^{12} = 4{,}096$
L_{n-8}	$2^0 = 1$	$2^4 = 16$	$2^8 = 526$

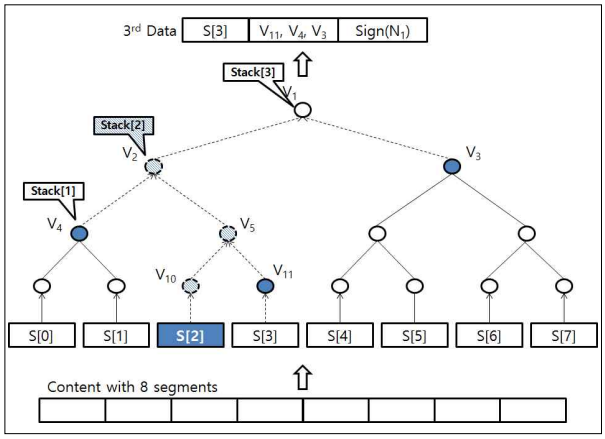
MHT의 운영에 있어서 해시 값의 중복 계산에 의한 비효율성뿐만 아니라 인증 정보도 중복해서 전송되는 비효율성이 발생된다. <Table 1>는 단편화된 개수에 따른 데이터의 인증 정보의 중복성을 나타낸다. L_i는 노드의 계층이 i인 노드 중 하나를 의미한다. 예를 들어 L_{n-1}은 계층이 $n-1$인 노드, N_2 또는 N_3을 의미한다. 그러므로 전체 리프 노드의 개수가 2^8개인 MHT의 경우, 전체 N_2의 노드 값인 V_2를 인증 정보 값으로 128번 전송하게 된다. 이와 같이 동일한 값을 중복해서 전송하는 것은 전체 전송 오버헤드뿐만 아니라, 전송된 정보를 처리하기 위한 계산 오버헤드까지 증가시키는 요인이 된다. 그러므로 인증 정보의 중복 전송 개선은 데이터 전송량 개선뿐만 아니라 해시 값 계산량 개선을 동시에 이룰 수 있다. 특히 [17]에서 제안된 해시 값 계산량 개선 기법과 함께 적용할 경우, MHT 운영의 최적화된 솔루션이 될 수 있다. 본 논문에서 제안하는 개선안은 [17]에서와 같이 동적 프로그래밍 기법을 활용하여, 인증 정보를 한 번씩만 전송/수신하고, 수신된 인증 정보를 저장한 후, 다음과 같이 재사용한다 [16]: MHT의 리프노드의 개수를 $N = 2^n$개라고 할 때, 각각의 세그먼트에는 구성 순서에 따라 세그먼트 번호(sn)가 0부터 $N-1$까지 순서대로 할당된다고 가정하자.

(1) 수신자는 n개의 인증 정보를 저장할 캐시 $\{C_{w[i]} | 0 \le i \le n-1\}$와 n개의 인증 경로를 저장할 캐시 $\{C_{v[i]} | 0 \le i \le n-1\}$를 준비한다. $C_{w[i]}$와 $C_{v[i]}$에는 계층 i인 인증정보 $w[i]$와 인증경로의 노드 값이 저장된다.

(2) 콘텐츠의 i번째 Data를 수신하면, 첨부된 인증정보를 순서에 맞게 $C_{w[i]}$에 저장한다. 이 후, 해당 인증정보를 재사용해야 할 경우, 캐시에 저장된 값을 사용한다.

(3) 인증경로의 노드 값들을 [17]에서 제안한 방법에 따라 순차적으로 계산한 후, $C_{v[i]}$에 저장 한다: k번째 Data 인증을 위하여, 앞서 인증된 $k-1$번째 Data에 할당된 리프 노드부터 루트 노드에 이르는 인증경로와 k번째 세그먼트의 인증 경로를 비교하여, 두 경로에 동일하게 포함된 노드들 중에서 가장 하위 노드의 노드 값까지만 계산하여 저장된 노드 값과 비교한다. 특히, 이 최하위 노드를 인증 노드 (Authentication Node)라고 부른다.

(4) Data에 할당된 리프노드가 $C_{v[i]}$를 인증경로로 하는 마지막 노드인 경우, 저장된 $C_{w[i]}$와 $C_{v[i]}$을 서로 교환하여 저장한다.



[Fig. 2] Improved MHT-based Authentication Process

<Table 2> Authentication Level Calculation

```
Input: Segment Number (sn)
Output: Authentication Level (level)
SET mask := 1;
FOR level := 0 to n-1 :
    IF (i & mask) is not zero THEN
      STOP this routine
    ELSE
      SET mask := mask << 1
    END_IF
END_FOR
OUTPUT level;
```

<Table 3> Segments Authentication Pseudo Code

```
Input: data[i], 0 ≤ i ≤ 2^n − 1
Output: Authentication Result
FOR segment number i = 0 to 2^n − 1 :
  RECEIVE data[i];
  SET level := Authentication Level for i;                    //(1)
  FOR witness number k := 0 to level − 1:
    SET C_w[j] := data[i].w[j] ;
  END_FOR
  IF i = 0 THEN
    SET C_w[n−1] := data[i].w[n−1] ;
  END_IF                                                      //(2)
  s[0] := data[i].seg; //i-th segment
  FOR j := 0 to level :
    CALCULATE h = H(C_w[j], s[j]);                            //(3)
    SET s[j+1] := h;
    IF j = level THEN                                         //(4)
      IF level = n − 1 AND i = 0 THEN                         //(5)
        VERIFY sign USING h;
        IF sign is valid THEN
          SET C_v[n] := h;
        ELSE
          REPORT error;
          STOP;
        END_IF
      ELSE
        IF C_v[level+1] = h THEN                              //(6)
          SET C_w[0] := H(data[i].seg);                       //(7)
        ELSE
          REPORT an error
        END_IF
      END_IF
      FOR k := 0 to level:
        SET C_v[k] := s[k];                                   //(8)
      END_FOR
    END_IF
  END_FOR
END_FOR
```

<Table 2>와 <Table 3>은 제안하는 콘텐츠 인증 절차의 유사 코드(pseudo code) 이다. <Table 2>는 i번째 세그먼트 $S[i]$를 인증할 때, 어느 계층까지의 노드 값을 계산해야 되는지를 결정한다. 이 때, 실제 인증을 위해 계산되어야할 최상위 노드의 계층은 <Table 2>의 결과 값(Output)인 $level$에 1을 더한 값인 $level+1$이 된다. 실제로 i번째 Data가 수신되었을 때, 수신된 Data에 할당된 리프 노드부터 루트 노드까지의 계산된 인증 경로와 $i-1$번째 Data 인증 절차에 저장된 $\{C_{v[i]}\}_{i=0,\ldots,n}$을 하위 계층부터 비교하여 같은 노드가 있는지 확인한다. i번째 Data의 경우, $k=0$부터 순차적으로 증가 시키면서 i가 2^k의 배수인지 확인한다. 만약 i가 2^k의 배수이면, 인증 경로 상의 $k+1$번째 계층의 노드와 $C_{v[k+1]}$에 저장된 노드가 동일한 노드가 된다. 즉, 이 $k+1$ 계층의 노드가 인증 노드가 된다. 그러므로 <Table 2>의 절차에 따라 최종 $level$ 값은 k가 출력된다.

Data를 수신하였을 때, 수신된 Data의 인증을 위해 <Table 3>은 다음과 같이 동작 한다:

(1) Data를 수신한 사용자는 수신된 Data의 세그먼트 번호를 참조하여, <Table 2>의 절차를 따라서 인증에 사용될 최상위 노드의 계층 수 ($level$)를 계산한다.

(2) 앞서 계산된 $level$ 값에 따라서 Data 인증에 필요한 인증 정보 $\{w[0],\ldots,w[level-1]\}$를 Data에서 읽어 와서 $C_{w[i]}$에 순서대로 저장한다. 단, 첫 번째 세그먼트의 경우, $level$의 값이 $n-1$이지만 $w[n-1]$을 $C_{w[level]}$에 추가로 저장한다.

(3) $level$ 값에 따라서 MHT의 인증 경로 위의 노드 값을 Data에 대응되는 리프 노드부터 차례로 계산한다. 이 때 인증 경로 위의 노드의 계층들은 0부터 $level+1$까지의 값을 가진다.

(4) 인증 경로의 최상위 계층($level+1$)의 노드 값까지 계산되었다면, 해당 노드 값을 이용하여 Data를 최종 인증한다. 인증 절차는 세그먼트 번호에 따라서 두 가지 경우, (5)와 (6),로 나눠서 수행된다.

(5) 세그먼트 번호가 0인 Data의 경우, 계산된 MHT의 루트노드 값을 이용하여 Data의 서명 값을 검증한다. 검증 결과 유효한 서명으로 판정되면, 루트 노드 값을 검증경로 캐시 $C_{v[n]}$에 저장한 후, (9)단계를 수행한다.

(6) 세그먼트 번호가 0보다 큰 Data의 경우, 계산된 $level+1$ 계층의 인증 경로 위의 노드 값과 $C_{v[level+1]}$ 값을 비교한다. 만약 두 값이 다르면 인증 오류를 사용자에게 알린다. 두 값이 같으면, Data가 정상적으로 인증된 것으로 간주한다.

(7) 다음 순서의 Data 인증에 사용될 $w[0]$의 값을 미리 계산하여 $C_{w[0]}$에 저장한다.

(8) Data 인증이 완료되면, 인증 경로의 노드 값들을 순서에 따라서 $\{C_{v[0]},\ldots,C_{v[level]}\}$에 저장한 후, 다음 순서의 Data에 대한 인증을 계속한다.

<Table 4>는 제안된 Data 패키징 절차를 설명한다. Data는 단편화된 세그먼트와, 인증 정보, 그리고 전자 서명으로 구성된다. 특히 2^n개로 단편화된 콘텐츠에 대한 인증을 위해 CCN은 모든 Data에 n개의 인증정보를 저장하도록 설계 되었다. 본 논문에서는 Data 마다 인증경

로 위에 있는 특정 중간 노드까지만 노드 값을 계산하도록 설계되었기 때문에, 첫 번째 Data를 제외한 모든 Data가 서로 다른 개수의 인증정보를 저장하도록 설계되었다. 특히, 첫 번째 Data를 제외한 Data의 경우, 인증경로 상의 인증노드의 계층이 $level+1$이면, 인증정보는 순서에 따라 $level-1$개면 충분하다.

<Table 4> Data Generation Pseudo Code

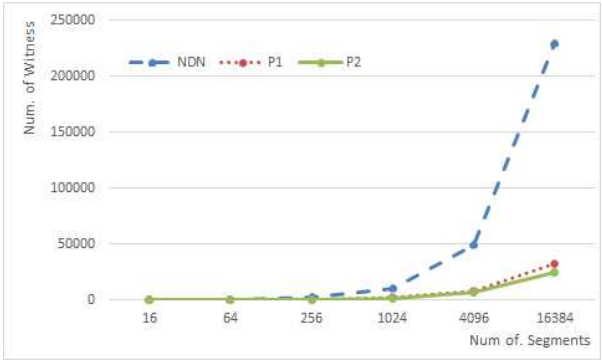
Input: V_0, $priK$, *and* $seg[i]$ *and* $w_i[k]$, $0 \le i \le 2^n-1$, $0 \le k \le n-1$
Output: $data[i]$, $0 \le i \le 2^n-1$

SET $data[0].seg := seg[0]$;
FOR $k:=0$ *to* $n-1$:
  SET $data[0].w[k] := w_0[k]$;
END_FOR
SET $data[0].sign := E_{priK}(V_0)$;
FOR segment number $i:=1$ *to* 2^n-1 :
  SET $data[i].seg := seg[i]$;
  IF i *is EVEN THEN*
  SET $level$:=*Authentication Level for* i;
  FOR $k:=0$ *to* $level-1$:
  $data[i].w[k] := w_i[k]$;
  END_FOR
  END_IF
END_FOR

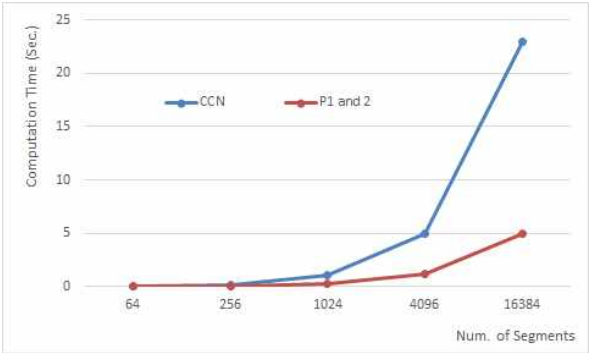
4. 성능 분석

제안된 운영 방안의 성능을 분석하기 위하여 Data 인증 시 필요한 인증정보의 전송량과 해시 계산 횟수를 분석한다. 콘텐츠를 구성하는 세그먼트의 수를 $N=2^n$이라 하면, CCN의 MHT 운영 방법을 적용할 때, 각각의 Data는 n개의 인증 정보를 포함한다. 그러므로 콘텐츠를 구성하는 전체 세그먼트들을 고려할 때 $n \times 2^n$개의 인증 정보가 전송되어야 한다.

개선 안의 경우 leve 값이 n-1인 노드 중에 1개 노드의 노드 값, level 값이 n-2부터 1까지에 속한 모든 내부 노드들의 노드 값, 그리고 리프 노드들 중에서 절반의 노드 값들이 각각 한 번씩 전송된다. 그러므로 세그먼트의 수를 $N=2^n$이라 하면, $1+(\sum_{i=2}^{n-1}2^i)+2^{n-1}$ 개인 인증 정보가 전송된다. 즉, 2^n개로 단편화된 콘텐츠 전체를 인증하기 위해서는 $2^n+2^{n-1}-3$개의 인증 정보 전송이 필요하다. 이는 n이 증가할수록 성능 개선 효과 역시 증가한다. 예를 들어 $n=14$인 경우, CCN과 [17]에 비해 각각 90%와 30%의 성능 개선 효과를 보인다.



[Fig. 3] Transmission Overheads Comparison



[Fig. 4] Computation Overheads Comparison

[Fig. 3]은 인증 정보 전송량을 비교한 것이다. 전송량 변화를 관찰하기 위하여 세그먼트의 크기를 모두 동일하다고 가정하였고, 세그먼트의 수를 16부터 16,384까지 다양하게 적용하여 계산하였다. [CCN]은 MHT의 기본적인 운영 방법을 구현한 결과이며, [P1]은 [17]에서 제안된 개선안의 결과, 그리고 [P2]는 본 논문의 개선된 운영 방법을 구현한 결과이다. 세그먼트의 수가 많은 수록 개선안의 전송량 개선 효과가 두드러짐을 알 수 있다. 세그먼트의 수가 1만개 이상인 경우, [P1]은 CCN 대비 전송량을 85% 이상 개선할 수 있으나 [P2]의 경우는 91% 이상 개선할 수 있다.

해시 계산 횟수의 경우, CCN의 기본 구현에 따라서 수신된 세그먼트를 모두 인증하기 위해서는 $n \times 2^n$번의 해시 값 계산이 요구된다. 개선 안을 적용할 경우, 루트 노드부터 리프 노드의 부모 노드들까지 모든 내부 노드들의 노드 값들이 두 번씩 계산된다. 또한, 리프 노드의 노드 값은 한 번씩 계산된다. 그러므로 전체 해시 값 계

산 회수는 $3\times2^n-2$이다. [Fig. 4]는 인증 정보를 활용하여 세그먼트들을 인증할 때 계산량을 시간으로 비교한 결과이다. 각각의 세그먼트는 CCN의 기본 설정 크기인 4K 바이트의 크기를 갖는다고 가정하였고, SHA-512를 사용하여 해시 값을 계산한 경우를 가정하였다.

그러나 N $(=2^n)$개의 세그먼트로 구성된 콘텐츠를 인증하기 위해서는 루트 노드의 노드 값 외에 추가로 최대 n개의 인증 경로 정보와 $n-1$개의 인증 정보를 저장한 저장 공간이 필요하기 때문에 스토리지 오버헤드가 일부 증가할 수 있으나, 일반적인 세그먼트의 수를 고려할 때 최대 $2n$개의 스토리지 오버헤드는 수용 가능할 것으로 판단된다.

4. 결론

CCN은 네트워크 노드에 캐싱 된 데이터를 활용하여 네트워크의 효율성을 높이기 위해 제안되었다. 이와 같은 목적을 달성하기 위하여 CCN은 콘텐츠 생성자에게 집중되는 요청 메시지를 효과적으로 분산 처리할 수 있게 네트워크 노드에 캐싱 기능을 구현하고, 캐싱되어 있는 콘텐츠에 대한 요청 메시지를 수신한 네트워크 노드가 콘텐츠 생성자를 대신하여 해당 요청 메시지에 응답할 수 있게 설계되었다. 그러나 이와 같은 중간 노드에 의한 응답 처리는 사용자가 실제 콘텐츠를 전송한 노드를 식별할 수 없기 때문에 호스트 인증 기반의 보안 체계를 적용할 수 없다는 문제점과 함께 데이터의 위/변조가 가능하다는 취약점을 갖고 있다. 이러한 문제를 해결하기 위해 CCN은 머클 해시 트리를 사용한 콘텐츠 인증 기법을 제안하고 있다. 그러나 머클 해시 트리를 이용한 콘텐츠 인증은 해시 값을 중복해서 계산해야하고, 루트 노드의 해시 값 계산에 필요한 인증 정보를 중복으로 전송해야만 한다.

본 논문에서는 이러한 문제점을 개선하기 위하여 동적 알고리즘 기법을 적용하여 해시 값의 중복 계산 및 전송 문제를 개선함으로써 CCN의 인증 비효율성을 개선하였다. 특히, 본 논문은 인증 경로 정보만 저장 후, 재사용하는 [17]의 기법을 개선하여 인증 경로와 인증 정보를 저장하고 재사용함으로써 성능을 개선하였다. 본 논문에서 제안된 기법을 적용할 경우, 대용량 콘텐츠의 인증 정보 인증 시간을 60~85%까지 개선하는 동시에 전송량을 92%까지 개선할 수 있다.

REFERENCES

[1] “Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2015 - 2020,” Cisco Public, February 3, 2016.

[2] “Cisco Visual Networking Index: Forecast and Methodology, 2015 - 2020,” Cisco Public, February 3, 2016.

[3] A. K. Pathan, and R. Buyya, “A Taxonomy and Survey of Content Delivery Networks,” Tech Report, Univ. of Melbourne, 2007.

[4] E. Meshkova, J. Riihijarvi, M. Petrova, and P. Mahonen, “A survey on resource discovery mechanisms, peer-to-peer and service discovery frameworks,” Computer Networks J., vol. 52, no. 11, pp. 2097-2128, 2008.

[5] D. Clark, “The Design Philosophy of the DARPA Internet Protocols,” ACM Sigcomm Comp. Comm. Review, Vol. 18, No. 1, pp. 106-114, Aug. 1988.

[6] B. Ahlgren, C. Dannewitz, C. Imbrenda, D. Kutscher and B. Ohlmann, “A Survey of Information-Centric Networking,” IEEE Communications Magazine, Vol. 50, No. 7, pp. 26-36, 2012.

[7] V. Jacobson, D. Smetters, J. Thornton, M. Plass, N. Briggs and R. Braynard, “Networking Named Content,” 5th International Conference on Emerging Networking Experiments and Technologies, pp. 1-12, 2009.

[8] D. Kim, “Content Centric Networking Naming Scheme for Efficient Data Sharing,” Journal of Korea Multimedia Society, Vol. 15, No. 9, pp. 1126-1132, 2012.

[9] D. Y. Kim, “Trend and Improvement for Privacy Protection of Future Internet,” Journal of Digital Convergence, Vol. 14, No. 6, pp. 405-413, 2016.

[10] D. Y. Kim, “A Comparison Study on Data Caching Policies of CCN,” Journal of Digital Convergence, Vol. 15, No. 1, pp. 327-334, 2017.

[11] R. Merkle, “Protocol for public key cryptosystems,” IEEE Sympo. Research in Security and Privacy, Apr. 1980.

[12] D. Y. Kim, J. S. Park, ″Efficient Contents Verification Scheme for Contents-Centric-Networking,″ The Journal of Korean Institute of Comm. and Inform. Sciences, Vol. 39, No. 4, pp. 234-241, 2014.

[13] D. Kim, “A Efficient Content Verification Scheme for Distributed Networking/Data Store,” Journal of The Korea Institue of Information Security & Cryptology, Vol. 25, No. 4, pp. 839-847, 2015.

[14] D. Kim, “Group-Interest-based Verifiable CCN,” Mobile

Information Systems, Volume 2016, Article ID 9202151
[15] B. Georg "Merkle Signature Schemes, Merkle Trees and Their Cryptanalysis,". Ruhr-Universität Bochum. Retrieved 2013-11-20.
[16] T. Cormen,"Introduction to Algorithm," The MIT Press, pp. 301-328, 1992
[17] D. Y. Kim, "Improvement of the Data Authentication of CCN," Journal of Digital Convergence, Vol. 15, No. 8, pp. 341-349, 2017.

김 대 엽(Kim, Dae Youb) [정회원]

- 1994년 2월 : 고려대학교 수학과 (이학사)
- 1997년 2월 : 고려대학교 수학과 (이학석사)
- 2000년 2월 : 고려대학교 수학과 (이학박사)
- 2000년 2월 ~ 2002년 8월 : 시큐아이 정보보호연구소 차장
- 2002년 9월 ~ 2012년 2월 : 삼정전자 종합기술원 수석 연구원
- 2012년 3월 ~ 현재 : 수원대학교 IT 대학 정보통신학부 학부장, 정보보호학과 조교수, IT 연구소장
- 관심분야 : 악성 코드 분석, 웹 보안, 콘텐츠 보안, 미래 인터넷 보안
- E-Mail : daeyoub69@suwon.ac.kr