디지털 포렌식에서 머클 트리 기반 무결성 검증의 적용 한계 분석 및 향후 연구 방향

이재욱*, 최상훈[1], 박기웅†
세종대학교 SysCore Lab. (대학원생, 연구 교수[1])
† 세종대학교 정보보호학과 (교수†)

Analysis of the Limitations of Merkle Tree-Based Integrity Verification in Digital Forensics and Future Research Directions

Jae-Wook Lee*, Sang-Hoon Choi[1], Ki-Woong Park†
Syscore Lab., Sejong University (Graduate Student, Research Professor[1])
† Dept. of Computer and Information Security, Sejong University
(Professor†)

요 약

디지털 포렌식에서 해시 기반 무결성 검증은 증거의 신뢰성을 보장하기 위한 필수 절차이며, 데이터 크기 n에 비례하는 $O(n)$의 시간복잡도를 갖는다. 이러한 시간복잡도 특성과 데이터 규모의 증가로 인해 기존 검증 방식은 분석 과정에서 병목 현상을 유발한다.

본 연구에서는 기존 해시 기반 검증 방식의 한계를 정리하고, 머클 트리 구조가 포렌식 분석 환경에서 활용되지 않는 원인을 기술적·구조적·법적 관점에서 분석한다. 또한 부분 무결성 검증의 특성과 법적 수용 가능성을 고려하여 향후 머클 트리의 적용 방안과 연구 방향을 제시한다.

1. 서론

디지털 포렌식(Digital Forensics)은 범죄 수사 및 법적 분쟁에서 디지털 증거를 수집, 보존, 분석하는 핵심 분야로, 증거의 무결성 확보는 수사의 신뢰성과 증거의 법적 효력을 결정짓는 중요한 요소이다[1]. 디지털 증거의 무결성은 수집 시점부터 법정 제출까지의 전 과정에서 데이터가 변조되지 않았음을 보장함으로써 확보되며, 이를 위해 SHA256, MD5 해시 검증 방식이 주로 사용되고 있다[2].

기존 해시 기반 검증 과정은 전체 데이터를 순차적으로 처리하는 구조로 인해 데이터 크기 n에 비례하는 $O(n)$의 시간복잡도를 갖는다. 오늘날 대용량 저장매체의 보급과 디지털 서비스의 확산으로 처리해야 할 데이터의 크기가 급격히 증가하였으며, 이에 따라 해시 검증 과정에서 병목 현상이 발생하고 있다[2][3][4].

한편, 이러한 데이터 무결성 검증 단계에서의 시간복잡도 문제를 해결하기 위한 대안으로 머클 트리(Merkle Tree) 구조의 검증이 연구되고 있다[5][6][7][8].

본 논문의 구성은 다음과 같다. 2장에서는 기존 해시 검증에서 발생하는 병목 문제를 시간복잡도 관점에서 분석하며, 3장에서는 머클 트리 기반 검증 방식의 구조적 특징과 적용 사례를 분석하고, 데이터 무결성 입증 측면에서의 효율성을 검토한다. 4장에서는 포렌식 환경에서의 머클 트리 검증 방식 적용 한계를 기술적 문제를 넘어 법적 기준, 구조적 제약 등 복합적인 관점으로 분석하며, 마지막 5장에는 디지털 포렌식 환경 내 머클 트리의 적용 가능성 및 향후 연구 방향을 제시한다.

2. 기존 해시 검증의 문제점

디지털 포렌식 실무에서는 저장매체 이미징 이후 전체 데이터에 대한 해시값을 재생성하고, 이를 기존 해시값과 비교하여 무결성을 검증한다. 이 과정

† 교신저자 박기웅 : (세종대학교 정보보호학과 교수)
본 연구는 본 연구은 과학기술정보통신부의 재원으로 정보통신기획평가원 (IITP)의 정보보호핵심원천기술개발(Project No. RS-2026-25519773, 30%; RS-2024-004 38551, 10%), 실감콘텐츠핵심기술개발(Project No. RS2023-00228996, 20%), ICT (ITRC) (Project 대학 연구센터 No. ,10% IITP-2026-RS-2021-II211816, 10%), 한국연구재단 핵심연구사업 (NRF) (Project No. RS-2026-25481431 30%) 의 지원을 받아 수행된 연구임

은 수집 시점뿐만 아니라 보관, 분석 및 법정 제출 단계 등 Chain of Custody를 유지하는 여러 시점에서 반복적으로 수행되어 무결성을 보장한다.

기존의 해시 검증 방식은 전체 데이터를 순차적으로 처리하는 구조를 가지므로, 데이터 크기를 n이라 할 때 $O(n)$의 시간복잡도를 가진다[2]. 이는 데이터 크기가 증가할수록 처리 시간이 선형적으로 증가함을 나타내며, 변조 여부를 확인하기 위해서도 전체 데이터를 다시 처리해야 하는 비효율성을 가진다. 또한, 디지털 포렌식 과정에서는 동일한 데이터에 대해 검증 절차가 반복적으로 수행되기 때문에 이러한 선형 시간 구조는 수사 및 분석 과정에서 주요 병목 요소로 작용하게 된다. 실제 포렌식 실무에서는 디스크 이미징 작업보다 해시 검증 과정에 더 많은 시간이 소요된다고 보고된 바 있다[3].

아울러, 포렌식 대상 데이터의 규모는 급격히 증가하고 있다. 최신 저장매체의 용량은 수 TB에서 수십 TB에 이르며, 클라우드 및 분산 환경에서는 그 규모가 더욱 확대된다. 이에 따라 디지털 증거의 해시 검증 과정에서 심각한 성능 문제가 발생하고 있다[3][4].

결과적으로 기존 해시 기반 검증 방식은 무결성 보장 측면에서는 높은 신뢰성을 제공하지만, 대용량 데이터 환경에서는 시간 비용이 과도하게 증가하는 한계를 가지며, 이는 포렌식 분석의 효율성을 저해하는 요인으로 작용한다.

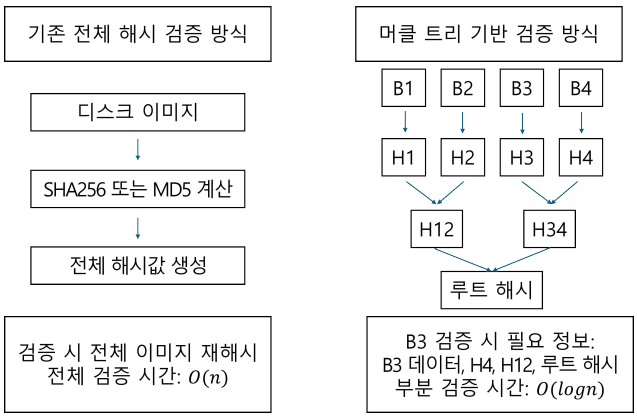
3. 개선 연구 동향

앞서 무결성 검증 단계의 병목 현상이 수사 및 분석 과정에서 제약 요인으로 작용함을 살펴보았다. 이에 많은 전문가는 효율적으로 검증 단계를 처리하기 위해 다양한 분야에서 적용되는 알고리즘을 분석 및 응용하는 추세이다[5][6][7][8].

본 장에서는 무결성 검증 단계에서의 시간복잡도를 감소시키는 효율적인 해시 검증 알고리즘인 머클 트리(Merkle Tree)를 소개하며, 해당 알고리즘이 사용 또는 연구된 분야에 관해 설명한다.

3.1 머클 트리

머클 트리는 데이터를 블록 단위로 나눈 후 트리 형태로 구성하고, 각 노드의 해시값을 결합하여 최종적으로 루트 해시를 생성함으로써 전체 데이터의 무결성을 나타내는 구조이다. 머클 트리는 일반적으로 각 노드가 두 개의 자식 노드를 가지는 이진 트리 구조로 구성되며, 잎 노드는 데이터 블록의 해시값으로 이루어지고, 내부 노드는 자식 노드들의 해시값을 결합하여 생성된다. 이러한 구조는 블록체인 시스템과 클라우드 내 데이터 검증 및 무결성 감사 등 다양한 분야에서 활용되고 있다[5][9].



(그림 1) 기존 해시 기반과 머클 트리 기반 검증 방식.

머클 트리는 루트 해시값과 머클 경로(Merkle Path)가 주어졌을 때, 특정 데이터 블록의 무결성을 $O(\log n)$의 시간복잡도로 검증할 수 있다. 데이터 블록의 무결성을 입증하기 위해서는 루트 해시값과 함께 해당 블록이 포함된 경로상의 형제 노드들의 해시값이 필요하다. 검증 과정에서는 대상 데이터 블록의 해시값을 계산한 뒤, 제공된 형제 노드의 해시값과 순차적으로 결합하여 상위 노드의 해시값을 재구성한다. 이후 최종적으로 계산된 해시값이 루트 해시값과 일치하는지를 확인함으로써 무결성을 검증한다. 이러한 방식은 전체 데이터에 대한 해시 재계산 과정을 거치지 않아도 특정 블록에 대한 무결성을 효율적으로 검증할 수 있게 한다[9].

3.2 비트코인 내 머클 트리

비트코인 시스템에서는 머클 트리를 활용하여 블록 내 모든 트랜잭션을 하나의 루트 해시로 요약하며, 루트 노드에서 잎 노드까지의 해시 경로로 구성된 머클 경로를 통해 특정 트랜잭션의 포함 여부를 효율적으로 검증할 수 있다. 또한 이러한 방식은 전체 블록 데이터를 저장하거나 전송하지 않고도 검증할 수 있게 한다[5].

3.3 머클 트리를 활용한 영상 무결성 검증 기법

영상 무결성 검증 분야에서는 대용량 데이터 대상 효율적인 변조 탐지를 위해 머클 트리를 활용한 다양한 기법이 제안되고 있다.

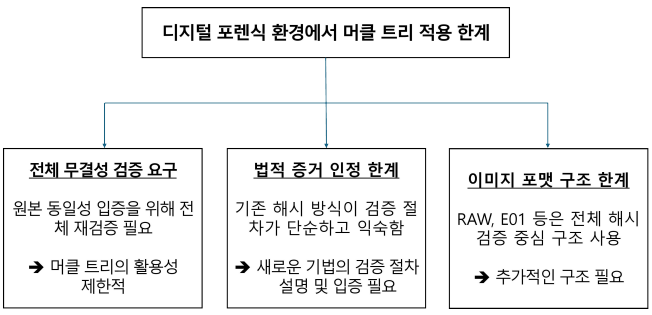
한 연구에서는 영상을 YUV420 형식으로 변환한 후 프레임별 해시값을 생성하여 경량 머클 트리를 구성하는 방식이 제안되었다. 해당 연구에서는 중복 해시 노드를 병합하여 네트워크 대역폭 부담을 줄임과 동시에 전체 영상을 재처리하지 않고 특정 프레임의 변조를 탐지할 수 있음을 도출하였다[6]. 다른 연구에서는 IoT 기기에서 산출된 세그먼트 단위 해시값을 허가형 블록체인의 분산 원장에 저장하고 스마트 컨트랙트를 통해 검증하는 포렌식 프레임워크도 제안되었다. 이를 통해 영상 데이터의 추적성과 신뢰성을 동시에 확보할 수 있음을 도출하였다[7].

3.4 네트워크 포렌식 내 머클 트리

네트워크 포렌식 분야에서는 로그 데이터의 무결성과 신뢰성을 보장하기 위해 머클 트리를 활용한 tamper-proof 기법이 제안되고 있다. 해당 방식은 각 로그 항목을 SHA256으로 해시화하고 이를 트리 구조로 재귀적으로 결합하여 루트 해시를 생성하며, 루트 해시의 변경 여부를 통해 로그 변조를 즉시 탐지할 수 있다. 특히 CNN 기반 AI 이상 탐지와 머클 트리 로깅을 결합한 프레임워크에서는 99.9%의 변조 탐지율과 260ms의 검증 시간을 달성하여 블록체인 기반 방식(1200ms) 대비 연산 효율성이 크게 향상하였으며, 사이버범죄 수사에서 법적 증거로서의 신뢰성을 확보할 수 있음을 도출하였다[10].

4. 디지털 포렌식 내 머클 트리 적용 한계점

위에서 살펴본 분야들은 공통적으로 대규모 데이터 처리, 부분 데이터에 대한 신속한 검증이 우선시되는 환경이다. 반면 디지털 포렌식은 증거 데이터의 동일성 보존과 법적 신뢰성 확보가 우선시되는 분야로, 무결성 검증의 목적과 운영 방식에서 차이가 있다. 본 장에서는 머클 트리가 적용할 수 있음에도 불구하고, 디지털 포렌식 실무에서 널리 사용되지 않은 주요 원인을 분석한다.



(그림 2) 디지털 포렌식 환경에서 머클 트리 적용 한계 요인.

4.1 전체 무결성 검증 요구

디지털 포렌식에서는 일부 데이터가 아닌 전체 디지털 증거의 무결성 검증이 핵심 요구사항이다. 저장매체로부터 획득한 이미지는 보관, 분석 및 법정 제출 과정 전반에 걸쳐 원본과 동일한 상태를 유지해야 하며, 이를 입증하기 위해 전체 데이터에 대한 해시값을 계산하고 비교하는 방식이 사용된다.

부분 데이터 검증 과정에서 시간적 이점이 있는 머클 트리 구조도 전체 무결성 검증 과정에서는 기존 방식과 유사한 수준의 연산이 요구된다. 또한 기존 포렌식 도구에서는 저장매체 전체뿐만 아니라 개별 파일 단위의 해시값도 함께 생성 및 검증하고 있어 추가적인 구조 없이도 실무적으로 충분한 무결성 검증이 가능한 상황이다. 따라서 머클 트리 기반 검증 방식은 일반적인 디지털 포렌식 환경에서는 그 활용성이 제한적이다.

4.2 법적 신뢰성 및 설명의 복잡성

디지털 포렌식에서 사용되는 무결성 검증 방식은 기술적 정확성뿐만 아니라 법적 신뢰성을 확보해야 한다. 현재 법정에서는 SHA256 기반 해시 검증 방식이 널리 사용되고 있으며, 계산 과정과 검증 절차가 비교적 단순하여 설명이 쉽다는 장점이 있다.

반면 머클 트리 기반 검증 방식은 트리 구조와 머클 경로를 포함하는 복잡한 검증 절차가 필요하며, 이러한 구조적 특성은 법적 증거로서의 설명 및 입증 과정에서 추가적인 복잡성을 유발할 수 있다.

4.3 포렌식 이미지 포맷의 구조적 한계

현재 널리 사용되는 포렌식 이미지 포맷인 RAW, E01 등은 원본 저장매체의 데이터를 복제하고 전체 이미지에 대한 해시값 검증을 통해 무결성을 확인하는 방식으로 널리 활용되고 있다. 일부 포맷은 압축, 분할 저장, 내부 오류 검출 등 추가 기능을 제공하지만, 블록 단위 해시 구조를 기본 검증 체계로 채택하고 있지 않다.

머클 트리를 적용하기 위해서는 데이터 블록별 해시값 생성 및 저장, 상위 노드 해시 계산, 루트 해시 관리, 검증 경로 제공, 관련 메타데이터 기록 등 추가적인 구조가 필요하다. 또한 기존 분석 도구와의 호환성, 이미지 포맷 표준, 검증 절차의 변경까지 함께 고려되어야 한다.

5. 결론 및 향후 연구 방향

디지털 포렌식 환경에서는 전체 저장매체의 무결성 검증이 요구되므로, 머클 트리 구조로 데이터를 관리하더라도 초기 모든 데이터 블록에 대한 해시 계산이 필요하다. 따라서 기존 해시 방식과 동일하게 $O(n)$의 시간복잡도를 가지며, 이러한 이유로 머클 트리의 이점은 제한적으로 작용한다. 이에 따라 일반적인 무결성 검증 방식으로는 널리 적용되지 않는다.

다만 향후 법적 기준의 변화에 따라 파일 및 경로 단위에 대한 무결성 검증이 증거능력으로 인정되는 경우, 머클 트리 구조의 활용 가능성이 높아질 것으로 기대된다. 현재 디지털 포렌식 실무에서는 전체 데이터에 대한 무결성이 요구되나, 부분 무결성이 인정되는 환경에서는 특정 증거가 포함된 파일 경로 또는 저장 위치에 대한 머클 경로를 제시함으로써 해당 데이터의 무결성을 효율적으로 입증할 수 있다. 이 경우 초기 이미징 단계에서 머클 트리 구조로 데이터를 블록화하고, 이후 재검증 시에는 문제가 되는 경로의 해시값만을 선택적으로 활용하는 방식이 가능할 것이다.

또한 목적 기반 부분 수집 결과물이 증거능력으로 인정될 경우, 머클 트리는 수집 단계에서도 활용될 수 있다. 사건과 관련된 파일, 디렉터리, 사용자 영역 또는 특정 시점의 데이터 블록만을 선별적으로 수집하고, 해당 데이터의 머클 경로를 함께 확보함으로써 저장매체 전체를 이미징하지 않더라도 무결성을 입증할 수 있다. 이는 대용량 저장장치, 클라우드 환경 등 전체 이미징이 어려운 환경에서 현실적이며 효율적인 대안이 될 것으로 기대된다.

향후 연구에서는 전체 무결성 검증 요구를 유지하면서도 부분 검증을 함께 지원할 수 있는 하이브리드 검증 구조에 대한 연구가 필요하다. 또한 이미징이 어려운 환경을 고려하여, 선별 수집된 데이터의 무결성을 효과적으로 입증할 수 있는 부분 수집 검증 모델에 대한 검토가 요구된다.

참고문헌

[1] Karen Kent, Suzanne Chevalier, Tim Grance, Hung Dang, Guide to Integrating Forensic Techniques into Incident Response, NIST, 2006.

[2] Anwar, M. R. ., Apriani, D. ., & Adianita, I. R. (2021). Hash Algorithm In Verification Of Certificate Data Integrity And Security. Aptisi Transactions on Technopreneurship (ATT), 3(2), 67 - 74.

[3] Scott Richards, Rethinking digital forensic evidence, OpenText, 2026.

[4] Sans.org, Advanced Evidence Collection: DFIR's 2024 Mobile and Cloud Shift, Sans.org, 2024.

[5] S. Jing, X. Zheng and Z. Chen, "Review and Investigation of Merkle Tree's Technical Principles and Related Application Fields," 2021 International Conference on Artificial Intelligence, Big Data and Algorithms (CAIBDA), Xi'an, China, 2021, pp. 86-90, 2021.

[6] YunHee Kang, Eum-young Chang and Taeun Kwonk. (2022). Video Integrity Checking Scheme by Using Merkle Tree. Journal of Platform Technology, 10(4), 39-46.

[7] Mercan, Suat; Cebe, Mumin; Aygun, Ramazan S.; Akkaya, Kemal; Toussaint, Elijah; and Danko, Dominik, Blockchain-based Video Forensics and Integrity Verification Framework for Wireless Internet-of-Things Devices (2021). Computer Science Faculty Research and Publications. 53.

[8] Jaehyeok Han, Mee Lan Han, Sangjin Lee, Jungheum Park, ECo-Bag: An elastic container based on merkle tree as a universal digital evidence bag, Forensic Science International: Digital Investigation, Volume 49, 2024.

[9] Merkle RC, A Digital Signature Based on a Conventional Encryption Function, Santa Barbara, University of California, 1987.

[10] B. Abisha, V. S. Kumari, J. Premalatha, N. Jothy and S. K. V, AI-Driven Network Forensics: Ensuring Evidence Integrity with Merkle Tree-Based Tamper-Proof Logging, 2025 International Conference on Computing Technologies (ICOCT), Bengaluru, India, 2025.