NFT를 활용한 실시간 원본 증명 등록 및 비실시간 봉인 원본 증명 등록 방법

김창배[1], 김진욱[2]
[1]한국방송통신대학교 대학원 정보과학과 석사과정
[2]한국방송통신대학교 대학원 정보과학과 교수
khaebo@gmail.com, gnugi@knou.ac.kr

Real-time and Non-real-time Original Proof Registration Methods Using NFTs

ChangBae Kim[1], Jin Wook Kim[2]
[1]Dept. of Computer Science, Korea National Open University
[2]Dept. of Computer Science, Korea National Open University

요 약

NFT는 중앙화되어 있는 디지털 자산의 원본성과 진위성 증명을 블록체인 기술로 대체할 수 있는 혁신적인 방법으로 주목받고 있으나, NFT도 블록체인에 등록되기 전에 생성된 콘텐츠의 원본성과 진위성은 보장할 수 없다. 본 논문에서는 원본 콘텐츠 파일과 그에 대한 원본 증명 정보에 암호화 해시를 적용하여 실시간으로 NFT로 발행하고 추후 원본성과 진위성을 증명할 수 있는 방법을 제안한다. 또한 실시간으로 NFT 발행이 불가능할 경우 원본 콘텐츠 파일과 원본 증명 정보를 암호화하여 보관하였다가 나중에 NFT로 발행하여 원본성과 진위성을 증명할 수 있는 방법도 제안한다. 추가로, 발행된 NFT URL을 별도로 메모하지 않고 원본 콘텐츠 파일 자체만으로 NFT를 검색하여 원본을 증명할 수 있는 방법도 제안한다.

1. 서론

SNS 사용자가 많아지면서 사진을 이용한 가짜 뉴스가 이슈로 등장하고 있다. 2020년 코로나19 시기에 코로나19와의 사투에 지친 의료진의 모습이라고 소개되었던 사진은 중국 의료진 사진이 대구 의료진 사진으로 둔갑한 장소가 다른 가짜 뉴스였으며[1], 2023년 태풍 카눈 부산 피해 상황으로 알려진 사진은 2022년 태풍 힌남노 부산 피해 사진을 2023년 사진으로 속인 시간이 다른 가짜 뉴스였다[2]. 2023년 펜타콘 폭발 사진은 AI 생성 사진을 사용한 가짜 뉴스였다[3]. 만약 뉴스 제보 시점이나 게시 시점에 사진에 대한 원본 증명 정보를 함께 제공하는 전자적, 사회적 구조가 마련된다면 가짜 뉴스와 사실 뉴스, 원본이 증명된 뉴스들을 쉽게 구분할 수 있을 것이다.

지금까지의 디지털 콘텐츠의 원본 증명 방법은 주민등록등본 등 중앙기관이 발급하는 특정 서류로 제한되거나 위변조에 취약한 단점이 있다[4]. NFT(Non-Fungible Token, 대체불가토큰)는 블록체인 기술을 활용하여 디지털 자산의 원본성과 소유권을 증명할 수 있는 혁신적인 방법으로 주목받고 있으나, 블록체인은 블록체인에 저장된 데이터에 대해서는 무결성과 신뢰성을 보장하지만 블록체인에 저장되기 전에 생성되는 데이터에 대해서는 무결성과 신뢰성을 보장할 수 없다.

또 다른 측면에서 NFT의 대상이 되는 원본 콘텐츠가 저장되는 저장 공간에 대한 문제가 존재한다. 블록체인은 저장되는 데이터의 크기에 따라 비용이 발생하며, 비용과 시간 문제를 피하기 위하여 NFT는 블록체인(온체인)에 발행하고 해당 NFT의 콘텐츠 정보가 되는 메타데이터나 대상 콘텐츠는 블록체인 외부(오프체인)에 저장하는 것이 일반적이다. 이때 사용되는 오프체인 저장소는 온프레미스 파일 서버, S3와 같은 클라우드 저장소, IPFS로 대표되는 분산 저장소 등으로 구분할 수 있다[5].

온프레미스 또는 클라우드 저장소는 과도하게 중앙집중화되어 있으며, 서비스의 가용성을 신뢰하기 어렵고, 원본 콘텐츠가 변경/위변조/삭제되는 무결성의 문제가 발생할 수 있다. 분산 저장소들은 가용성, 신뢰성을 좀더 보장할 수 있으나 여전히 원본 콘텐츠가 변경/위변조/삭제되는 무결성의 문제가 발생할 수 있다.

따라서, 블록체인에 NFT로 발행되기 전, 그리고

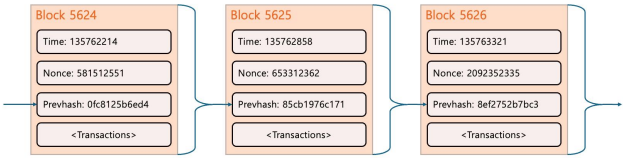
메타데이터와 원본 콘텐츠가 온체인/오프체인 어디에 저장되든 관계없이 원본 콘텐츠가 최초 생성된 시점부터 위변조되지 않고 블록체인으로 전송되었다고 확신할 수 있는 프로토콜이 필요하다. 또한, 시간이 흐른 뒤에도 블록체인을 통해 원본 콘텐츠가 위변조되지 않고 동일한 콘텐츠임을 증명할 수 있는 프로토콜도 필요하다.

본 연구에서는 앞에서 기술한 NFT의 문제를 해결하기 위해 원본 콘텐츠의 해시를 생성하고, 원본 콘텐츠 생성 시의 원본 증명 정보를 NFT 메타데이터에 포함하여 NFT로 발행함으로써 NFT 콘텐츠의 원본성과 진위성을 증명하고, 역으로 콘텐츠의 해시를 사용하여 NFT를 조회할 수 있도록 함으로써 원본성과 진위성을 재검증할 수 있는 방법을 제안하고자 한다.

2. 배경 지식 및 관련 연구

2.1 블록체인의 투명성과 불변성

블록체인은 분산 공개 원장으로서 모든 데이터가 공개된다. 모든 계정 활동이 블록체인에 기록되어 추적 가능하며, 모든 기록을 암호학적으로 검증할 수 있다. 또한 블록체인은 제네시스 블록을 제외한 모든 블록에는 이전 블록의 해시(prevhash)를 저장하고 있어서 특정 블록의 값을 변경하려면 특정 블록 이후의 모든 블록을 수정해야 하며 이는 실질적으로 불가능하다. 한 번 기록된 데이터가 변경되거나 삭제될 수 없다. 이는 보안과 데이터 무결성을 보장하는 핵심 요소이다[6].



(그림 1) 블록체인의 구조 [6]

2.2 오라클 문제

오라클 문제는 데이터를 블록체인으로 올리는 주체를 신뢰할 수 없고, 전송 중에 데이터가 변조될 수 있어 탈중앙화에 필요한 신뢰가 필요 없는(trustless) 환경을 훼손되는 것이다. 따라서 전송 데이터의 무결성을 입증할 수 있는 오프체인 방식의 검증 방법이 중요하다[7].

2.3 NFT

NFT는 이더리움의 ERC-721에 정의된 대체 불가능 토큰 표준으로 소유권 증명, 저작권 증명 등에 사용된다.

NFT에서 메타데이터(metadata)는 NFT의 추가적인 정보를 의미한다. metadata url은 스마트 컨트랙트의 tokenURI 메서드로 조회하며 metadata의 내용은 JSON 형식이다. 메타데이터의 저장소는 온체인 또는 오프체인 중에서 선택할 수 있다. 온체인은 수수료가 비싸기 때문에 주로 오프체인 저장 방식을 사용한다.

2.4 NFT와 자산 간 연결의 취약성

Wang et al.[5]은 조사 대상 6,234,141건의 NFT 중 실제 자산을 확인할 수 있는 것은 2,851,894건으로 45.75%에 불과하여 비효율적인 URL 문제가 예상보다 더 심각함을 확인하였다.

2.5 NFT 위변조 방지 및 보호

송효준[8]은 NFT 메타데이터의 description 속성에 파일의 해시 값을 저장하고 NFT에서 메타데이터를 조회하고 image 속성에 저장된 파일을 다운로드하여 해시를 비교하는 방식으로 NFT 데이터 무결성 검사를 설계하고 클레이튼 환경에서 구현하였다.

Visconti et al.[9]은 스마트 계약을 활용하여 고유 식별자와 디지털 자산 간의 연관성을 더 잘 보장하는 설계 메커니즘을 제안하였다. NFT 발행 과정에서 NFT 식별자로 일련번호 대신 디지털 콘텐츠의 충돌 저항 해시 함수(예: sha256)를 사용하여 식별자를 생성한다.

그러나, 관련 연구들에서 오라클 문제를 해결하고 NFT URL 등 NFT 정보 없이 파일만으로 원본 증명이 가능한 연구는 찾지 못했다.

3. 원본 증명 NFT 발행

3.1 실시간 원본 증명 NFT 발행

실시간 원본 증명 NFT 발행은 모바일 앱에서 촬영한 이미지가 위변조 없이 NFT로 발행되고, 나중에 NFT의 블록체인 정보를 모르더라도 원본임을 증명하는 데 목적을 두고 있다. 이는 모바일 앱의 사진 촬영 시각과 원본 증명 서버에서 해당 이미지를 NFT로 발행하는 시각의 차이를 최소화 함으로써 달성할 수 있다. 원본 증명 서버는 사진 촬영 시각과 NFT 발행 요청 도착 시각을 비교하여 그 차이(경과 시간)가 특정 시간(허용 경과 시간)을 초과하였다면 원본으로 인정하지 않고 NFT 발행을 거부한다.

본 연구에서는 사진 촬영 후 파일 해싱 소요 시간 2초 이상, 원본 증명 정보 생성 시간 2초 이상,

네트워크 전송 시간 5초 이상 등으로 가정하여 허용 경과 시간을 10초로 설정하였다. serverCurrTime - created_at <= 10초이다. 원본 증명 정보는 모바일 앱의 장치 식별자(device_id), 이미지 생성 시각(created_at), 위치 정보(location) 등을 사용한다.

본 연구에서 암호화 해시 알고리즘은 keccak256을 사용한다.

모바일 앱에서 사진 촬영 즉시 파일을 해싱하고 경로에 file_hash를 포함하는 image_url을 생성하고, image_url을 원본 증명 정보(continfo)와 함께 NFT metadata에 저장하고, 실시간으로 NFT로 발행하여 블록체인에 기록한다. metadata의 proofinfo 속성에 continfo와 continfo_hash도 함께 포함한다. metadata는 오프체인에 저장되며 metadata_hash 역시 metadata_url에 포함시킨다.

<표 1> 데이터 생성 규칙

데이터	규칙
file_data	**사진 파일의 바이너리 데이터**
file_hash	keccak256(**file_data**)
image_url	https://file.nftproof.kr/images/**[file_hash]**
continfo	원본 증명 정보(file_hash, device_id, created_at, location 등으로 구성)
continfo_hash	keccak256(JSON.stringify(**continfo**))
proofinfo	continfo와 continfo_hash
metadata	proofinfo를 포함한 NFT 메타데이터
metadata_hash	keccak256(JSON.stringify(**metadata**))
metadata_url	https://file.nftproof.kr/metadata/**[metadata_hash]**
token_id	**file_hash를 그대로 사용**
NFT URL	https://data.nftproof.kr/[smartcontract_address]/**[token_id]**

<표 2> metadata 예시


```
{
  name: "NFTProof",
  description: "Proof of Origin",
  image:
  "https://file.nftproof.kr/images/0x8bb0456d45a7964cf1066b59a2fa99d2fd97aaf7c15111bab1efb8fd5a6a6698",
  proofinfo: {
    continfo_hash:
  "0x0e7da179f3594321d2a0d2d043d5eb277b5b65213b2e820b22eda7f7ba95b2d0",
    continfo:
  {"created_at":"2025-03-30T00:47:22.366Z","device_id":"8d0a82e5-1cd1-4a81-aea8-803e9252a3f8","file_hash":"0x8bb0456d45a7964cf1066b59a2fa99d2fd97aaf7c15111bab1efb8fd5a6a6698","location":{"latitude":"37.45137901","longitude":"127.05716081"}}
  }
}
```


3.2 비실시간 원본 증명 NFT 발행

실시간 원본 증명 NFT 발행은 네트워크 연결이 불가능하거나 네트워크가 안정적이지 않아 허용 경과 시간을 초과하는 경우에는 사용하기 어렵다. 긴급한 재난/사건/사고 등에는 증거를 빠르게 확보하는 것이 우선이므로 촬영할 때마다 NFT를 발행하는 데 필요한 시간을 기다릴 수 없다.

비실시간 원본 증명 NFT 발행을 위한 중요 요소는 이미 촬영을 완료한 후 시간이 흐른 뒤에도 이미지와 원본 증명 정보들에 대해서 어떻게 하면 실시간 원본 증명 NFT 발행 시에 적용한 허용 경과 시간과 같은 신뢰성을 확보할 수 있는가이다.

비실시간 원본 증명 NFT 발행은 사진 촬영 직후 원본 증명 정보의 위변조를 방지하기 위해 continfo를 암호화하여 저장하고 있다가 NFT 발행이 가능할 때 NFT로 발행한다. 암호화는 ECIES(Elliptic Curve Integrated Encryption Scheme)를 사용한다. ECIES는 ECDH, AES-256-GCM, HMAC 등으로 구성된다[10-11]. 이 암호문은 원본 증명 서버의 개인키를 사용해야만 복호화할 수 있다. 시간 조작을 방지하기 위해서는 만료 시각을 가지는 JWT 인증키를 사용하여 인증키에 지정된 시간 범위에서 촬영한 원본 증명 정보만 인정하도록 한다. 인증키는 만료 시각이 있으므로 만료 시각 전에 주기적으로 재발급 받아야 한다. 암호화할 때 continfo에 encrypted_at(암호화 시각)이 추가된다.

<표 3> 비실시간 원본 증명 NFT 발행 시 조건

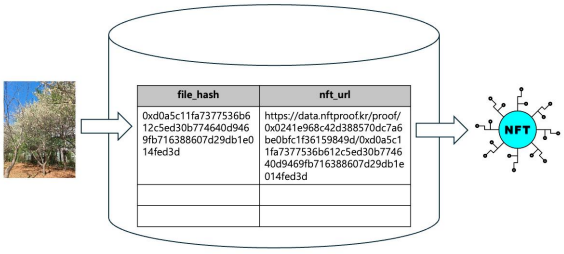
조건	규칙
인증키 무결성	인증키가 JWT 검증 유효해야 함
serverCurrTime - mobileCurrTime <= 10초	모바일 앱의 현재 시각과 원본 증명 서버의 현재 시각의 차이가 10초 이내이어야 함
created_at >= authKey.nbf	생성 시각은 인증키의 인증 유효 시작 시간 이상이어야 함
0초 < encrypted_at - created_at <= 2초	생성 시각과 암호화 시각의 차이는 해싱 소요 시간만을 고려하여 2초까지만 허용
encrypted_at <= serverCurrTime	암호화 시각은 원본 증명 서버의 현재 시각을 초과할 수 없음

3.3 파일 해시를 사용한 NFT 조회, 원본 증명

NFT는 NFT로부터 해당 콘텐츠를 조회하는 것은 가능하지만 콘텐츠로부터 그 콘텐츠의 NFT를 확인하기 어려워 중복된 NFT를 찾기가 어려우며, 관련 NFT 정보 없이 콘텐츠만으로는 원본성과 진위성을 검증하기 어렵다. 무단으로 복제 도용되는 디지털 저작물들이 관련 NFT와 함께 공유될 리는 없다.

NFT를 발행할 때 파일 해시와 NFT URL을 데이터베이스에 저장하면 사용자 개인이 NFT URL을 직접 저장하거나 원본 콘텐츠를 NFT URL과 함께 공유하지 않더라도 원본 콘텐츠만으로 원본 증명이

가능한 방법을 제공할 수 있다.

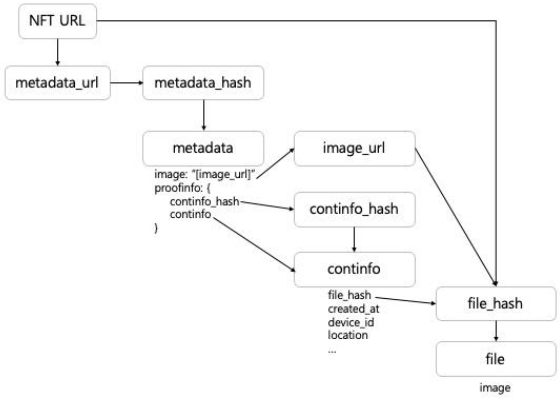


(그림 2) 파일 해시로 NFT를 조회하는 과정

4. 검증 - 무결성 검증 연결 고리

각 image_url, continfo, metadata_url 등 개별 정보들은 모두 자신을 해시 값을 포함하고 있어서 개별 요소들은 url과 파일 내용의 해시 값으로 스스로 검증이 가능하다.

전체적인 구조를 살펴본다면 file_hash가 원본 증명 NFT를 구성하는 각 요소에서 참조할 수 있도록 구성되어 있다. 무결성 검증을 위한 연결 고리를 구성하고 있다고 할 수 있다.



(그림 3) 무결성 검증 연결 고리

5. 결론 및 향후 과제

본 연구에서는 실시간 원본 증명 NFT 발행, 비실시간 원본 증명 NFT 발행, 파일 해시를 사용한 NFT 원본 증명 연구를 통해서 원본 증명의 오라클 문제를 해결하고자 하였다. 즉, NFT 발행 콘텐츠를 블록체인에 저장하기 전부터 원본으로 생성되었음을 증명하고, 이에 더해서 원본 증명 정보(사실 증명 정보)까지 함께 저장함으로써 NFT의 원본성과 진위성을 보장하는 방법을 제안하였다.

암호화 해시는 파일이 1비트만 변경되어도 다른 해시 값을 생성하므로 이미지 유사도 분석 등을 통하여 유사한 이미지를 판별하는 연구도 병행되면 원본 증명의 정확성을 높일 수 있으며, 모바일 앱의 이미지뿐만 아니라 여러 OS에서 다양한 파일을 위한 원본 증명 방법도 계속 연구될 필요가 있다.

참고문헌

[1] 박성호, KBS 뉴스 [팩트체크K] "이 사진, 진짜야?"…제보 사진 확인해보니, https://news.kbs.co.kr/news/pc/view/view.do?ncd=4394567, 2020년 3월 5일

[2] 연합뉴스, 부산 해안가 태풍 피해 '가짜 사진' 확산…상인들 고통 호소, https://www.youtube.com/watch?v=HaOjAliNk7g, 2023년 8월 11일

[3] 임선영, The JoongAng, "펜타곤 폭발" 사진 한장에 美발칵…AI발 가짜뉴스에 당했다, https://www.joongang.co.kr/article/25164636, 2023년 05월 23일

[4] Krawetz, Neal. "C2PA's Worst Case Scenario." The Hacker Factor Blog, https://www.hackerfactor.com/blog/index.php?%2Farchives%2F1013-C2PAs-Worst-Case-Scenario.html, 18 Dec. 2023

[5] Wang, Z., Gao, J., & Wei, X., "Do NFTs' owners really possess their assets? A first look at the NFT-to-asset connection fragility", In Proceedings of the ACM Web Conference 2023, pp. 2099-2109, April, 2023

[6] Buterin, V, "A next-generation smart contract and decentralized application platform.", white paper 3(37), 2-1, 2014

[7] Antonopoulos, A. M., & Wood, G., "Mastering ethereum: building smart contracts and dapps", O'reilly Media, 2018

[8] 송효준, "Hash 암호를 이용한 NFT속 파일 무결성 검증 및 위변조 탐지", 학위논문(석사)-전남대학교:정보보안협동과정, 2023년

[9] Visconti, I., Vitaletti, A., & Zecchini, M., "Preventing Content Cloning in NFT Collections.", In International Conference on Applied Cryptography and Network Security, (pp. 84-99). Cham: Springer Nature Switzerland, June, 2023

[10] Svetlin Nakov, Practical Cryptography for Developers, https://cryptobook.nakov.com, 2018

[11] Gayoso Martinez, V., Hernandez Encinas, L., Sanchez Avila, C., A Survey of the Elliptic Curve Integrated Encryption Scheme, Journal of Computer Science and Engineering, 2(2), pp. 7-13, August, 2010