Maya Protocol
A decentralized protocol for digital media authenticity

Maya Labs

May 29, 2024

Abstract

Recent advancements in AI-generated digital media technologies, such as DALL·E 2, Midjourney, and Sora, have reached a level of realism where distinguishing between authentic and synthetic media has become increasingly challenging. While these innovations offer exciting possibilities, they also pose significant risks such as societal unrest, political polarization and also provide a fertile environment for widespread misinformation and potential damage.

Advanced cryptographic methods like Zero Knowledge technology offer a way to solve this problem. This technology binds verification proofs to media content, which can then be verified by anyone while also preserving the privacy of the original media. Furthermore, the use of blockchains and decentralized storage solutions creates an immutable audit trail of data that are publicly verifiable. This ensures a trustless system that is open and unrestricted, enabling easy verification of media authenticity.

By combining these cryptographic techniques with the economic mechanisms, it's possible to build a decentralized protocol that could serve as a global authenticity layer for all digital media content across the internet.

1 Introduction

The inception of the internet revolutionized how information was cataloged and accessed, with HTTP (Hypertext Transfer Protocol) spearheaded by Tim-Berners Lee serving as the backbone for data exchange on the World Wide Web. This innovation transformed the internet into a repository of knowledge accessible to the masses. However, as the internet landscape evolved from an academic haven to a vast digital expanse inhabited by a diverse array of users including advertisers, anonymous entities, and nefarious actors, the pressing need for a security layer became apparent. To address this, Netscape Communications introduced SSL (Secure Sockets Layer), later succeeded by TLS (Transport Layer Security), enhancing the web's integrity by encrypting communications and transitioning standard HTTP websites to more secure HTTPS counterparts.

Parallelly, the rise of camera phones and social media platforms catalyzed an unprecedented surge in digital media production, evidenced by staggering statistics: daily creation of 4.5 billion images and 250 hours of video content every minute. This deluge of digital content has not only become a cornerstone of modern entertainment and communication but has also blurred the lines between reality and artificiality in media. Projections indicate that by 2026, AI-generated content will constitute 90% of all online material, making it increasingly difficult to distinguish between what is real and what is fabricated. This looming reality underscores the urgent need for a protocol that can assure the authenticity of digital media akin to how HTTPS has become synonymous with providing authenticity of the messages exchanged between the server and client.

We present the Maya Protocol, a decentralized protocol designed to redefine the authenticity of digital media across the internet. Inspired by the transformative role of HTTPS in enhancing web security, the Maya Protocol aims to become the new cornerstone for ensuring the authenticity and integrity of digital media. The protocol seeks to make the authenticity of every digital media piece transparent and verifiable, addressing the challenges posed by the growing prevalence of AI-generated content in our digital ecosystem.

2 Current Solutions

In the evolving landscape of digital content verification, two primary approaches have emerged: AI detection and cryptographic provenance-based methods. Both have their merits but also significant limitations that hinder their effectiveness.

2.1 AI Detection

AI detection methods utilize artificial intelligence to identify manipulated or AI-generated content. While these systems have made strides in recognizing fake media, they are inherently flawed.

- **Detection is Unreliable**: AI detection systems often produce false positives, incorrectly flagging legitimate content as manipulated. This undermines trust in the detection process and can lead to significant credibility issues for content creators and platforms.
- **Generative AI Outpaces Detection**: The rapid advancement of generative AI technologies has outpaced the development of detection methods. As generative AI becomes more sophisticated, creating increasingly realistic fake content, detection systems struggle to keep up, rendering them less effective.

2.2 Cryptographic Attestations

Cryptographic attestation methods involve digital signatures or certificates to verify the authenticity of content. While cryptographic techniques provide a foundational level of security, they also come with notable drawbacks.

- **Overreliance on Signer**: Cryptographic attestations require significant trust in the signer or issuer of the certificate. If the entity providing the attestation is compromised or acts maliciously, the integrity of the attestation is nullified. This centralized trust model does not adequately ensure the authenticity and integrity of digital content in a decentralized ecosystem.

2.3 Limitations

Both AI detection and cryptographic attestation methods have critical limitations that reduce their effectiveness in ensuring the authenticity of digital media.

- **AI Detection**: Despite advances, AI detection remains susceptible to errors and the evolving capabilities of generative AI. For instance, deepfake technology continues to improve, often outpacing detection algorithms. This arms race between creation and detection means that detection methods are always a step behind.
- **Cryptographic Attestations**: The centralized nature of trust in cryptographic attestations poses a significant risk. If the trusted entity is compromised, all content verified by them becomes questionable. Additionally, this model does not scale well in a decentralized and trustless environment where multiple parties need to independently verify content without relying on a single authority.

These limitations underscore the need for a more robust and decentralized approach to digital content verification, one that Maya Protocol aims to provide through its innovative use of Zero-Knowledge proofs and decentralized infrastructure.

3 Maya Protocol

The Maya Protocol is a decentralized protocol to verify and ensure the authenticity of digital media content such as images, audios and videos. It allows users to cryptographically prove the authenticity of their images and videos, so that anyone can verify it and no-one can alter the integrity of the content.

3.1 Features

The Maya protocol provides multiple features such as authenticity, complete privacy, trustless verification, public verifiability, transparency, immutable provenance records, credible neutrality and composability.

3.1.1 Authenticity

The Maya Protocol guarantees the authenticity of digital media by ensuring that every piece of content, from images to videos, is immediately signed at the moment of capture, and any edits made to the original media are properly verified and recorded. It uses advanced cryptographic techniques like zero knowledge technology to certify the origin and integrity of media files, allowing users to distinguish genuine content from manipulated or counterfeit versions.

3.1.2 Privacy

Privacy is paramount in the Maya Protocol, allowing users to maintain the confidentiality of their original media while still providing proof of its authenticity. This is really crucial because raw images or videos are often edited before publication, sometimes to enhance their presentation or decrease file size, and at other times to omit sensitive details from the original content.

3.1.3 Trustless Verification

Trustless verification within the Maya Protocol eliminates the need for centralized parties to confirm the authenticity of digital media. The decentralized nature of the protocol means that the verification process is carried out on-chain using smart contracts, fostering a secure environment where trust is built into the system itself.

3.1.4 Public Verifiability

The Maya Protocol offers public verifiability, allowing anyone to independently verify the authenticity and integrity of digital media. This open verification process contributes to a transparent ecosystem where content can be trusted and verified by the wider community without compromising security.

3.1.5 Transparency

Transparency in the Maya Protocol is achieved through a transparent and open ledger of proofs of content authenticity. This feature ensures that the interactions within the protocol are visible to all participants, promoting accountability and trust in the system.

3.1.6 Immutable Provenance

Immutable provenance records are a key feature of the Maya Protocol, providing an immutable history of digital media from creation to its current state. This permanent record tracks ownership at source and all subsequent modifications, ensuring the integrity and origin of content are preserved over time.

3.1.7 Credible Neutrality

The Maya Protocol provides Credible neutrality by de-risking the system from centralized powers and ensures a widely accessible public good. The decentralized and crypto-economic nature of the protocol guarantees that all content is treated equally, fostering a balanced and censorship-resistant digital media environment.

3.1.8 Composability

Composability in the Maya Protocol refers to the ability to integrate and combine different features and services seamlessly. Composability is provided through integrations with multiple apps and services such as social networks and news platforms. This flexibility allows for the creation of an ecosystem of complex applications and services built on top of the protocol, enabling innovation and the development of new paradigms within the digital media landscape.

4 Technology

4.1 Zero-Knowledge (ZK)

Zero-Knowledge (ZK) technology is at the core of the Maya Protocol, providing unparalleled integrity for digital media.

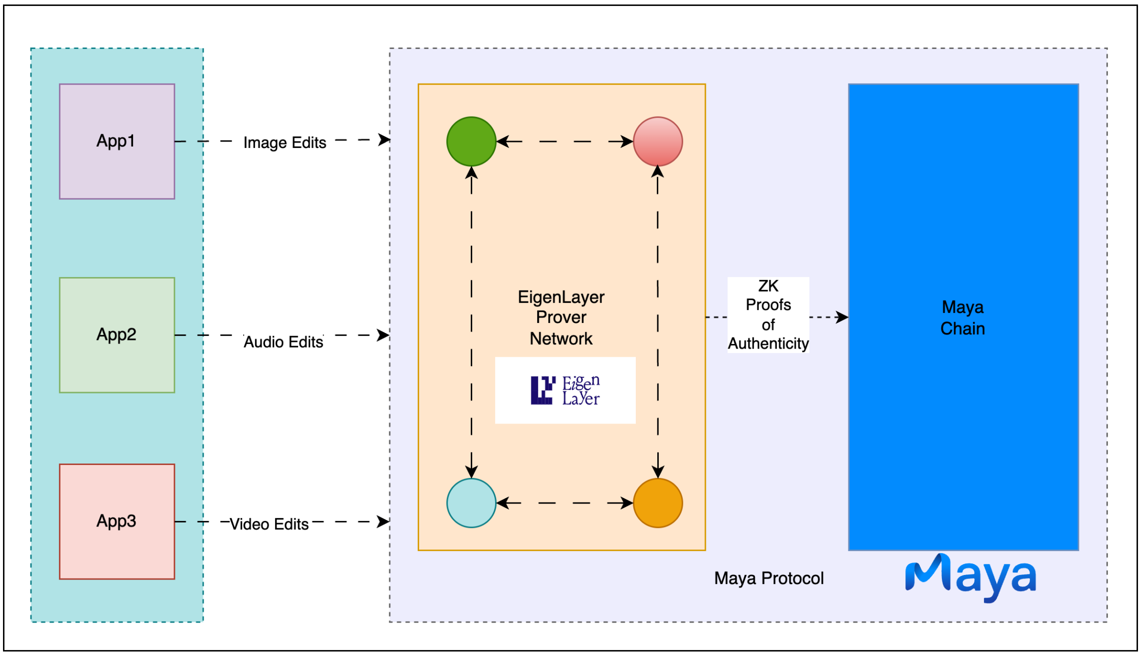

Figure 1: High-level architecture diagram of the Maya Protocol.

- **Ultimate Integrity**: ZK provides ultimate integrity, also known as glass-to-glass security, ensuring that the authenticity of digital content is maintained from the moment of capture to its display on a visual display.
- **Verifiable Media Backbone**: As ZK technology becomes more scalable, it creates the backbone for verifiable media across the internet. It secures all digital media interactions by generating cryptographic proofs that verify the integrity and authenticity of content without revealing any additional information.

Zero-Knowledge technology is the endgame for media authenticity because it offers a robust, scalable, and privacy-preserving solution to verify the provenance and integrity of digital media. By leveraging ZK proofs, Maya Protocol ensures that all interactions with digital media are verifiable and trustworthy, providing a solid foundation for combating AI-deepfakes.

4.2 Blockchain

Blockchain technology underpins the **Maya Chain**, offering a transparent and immutable audit trail for digital media.

- **Immutable Audit Trail**: The Maya Chain functions as a transparent and immutable audit trail, recording all provenance and authenticity information about digital media. This ensures that the history of any piece of content can be traced and verified, enhancing trust and accountability.
- **Provenance**: By maintaining detailed records of provenance and authenticity, the blockchain ensures that any alterations or edits to digital media are accurately documented and verifiable. This capability is crucial for maintaining the integrity of content in a decentralized environment.
- **Decentralizing Trust**: Blockchain technology is essential for decentralizing trust and reputation, moving away from reliance on centralized institutions. By distributing trust across a decentralized network, the Maya Protocol enhances the security and reliability of digital media verification, making it resilient to tampering and fraud.

5 Applications

The Maya Protocol has a wide range of applications wherever the integrity and authenticity of media are paramount. These applications span across various sectors, including legal and financial institutions, professional content creators, AI platforms, media companies, and news platforms. In this section, we discuss three key use cases.

5.1 Professional Content Creators

For professional content creators, especially social media influencers, maintaining the authenticity and integrity of their content is crucial.

- **Importance**: The reputation of their personal brand and the trust of their significant communities are at stake.
- **Maya's Benefits**:
  - Protects creator brand IP and reputation by ensuring all content is verified and authentic.
  - Enhances audience trust and integrity, leading to stronger community engagement.
  - Provides legal safeguards for content disputes by maintaining a verifiable history of content authenticity.
  - Ensures cross-platform authenticity, allowing creators to maintain their integrity across different social media platforms.

5.2 AI Inference

AI inference platforms must comply with regulations and maintain transparency in their training data to build credibility and trust.

- **Importance**: Compliance with regulations, transparency of training data, and maintaining credibility are critical.
- **Maya's Benefits**:
  - Ensures that training data is ethically and legally sourced, providing verifiable proof of its origins and authenticity.
  - Offers verifiable proof of origins and authenticity for AI-generated content (AIGC), ensuring that models are not used to create malicious deepfakes.

5.3 Legal

In the legal sector, the authenticity of digital evidence is essential to prevent false accusations and ensure the integrity of testimonies and the chain of custody.

- **Importance**: The authenticity of digital evidence, the ability to prevent plausible deniability, and the integrity of testimonies and the chain of custody are at stake.
- **Maya's Benefits**:
  - Provides tamper-proof proof of authenticity for digital evidence, ensuring its integrity from capture to courtroom presentation.
  - Maintains a verifiable chain of custody for all digital evidence, preventing any alterations or tampering.
  - Ensures the authenticity of testimonies, safeguarding the legal process from false claims and accusations.

6 Conclusion

The Maya Protocol represents a transformative approach to ensuring the authenticity and integrity of digital media in an increasingly digital world. By leveraging advanced cryptographic techniques, particularly Zero-Knowledge (ZK) proofs, and the transparency of blockchain technology, the protocol addresses critical issues related to media manipulation and misinformation.

Maya Protocol's features, including authenticity, privacy, trustless verification, public verifiability, transparency, immutable provenance records, credible neutrality, and composability, create a robust framework for verifying digital content. These features cater to various applications, from protecting the reputations of professional content creators to ensuring compliance and transparency for AI inference platforms and maintaining the integrity of digital evidence in legal contexts.

As the digital landscape continues to evolve, the need for reliable and decentralized solutions to verify digital media will only grow. The Maya Protocol stands at the forefront of this movement, providing a scalable and secure solution that empowers users to trust the digital content they encounter. By decentralizing trust and ensuring the verifiability of media, the Maya Protocol contributes to a more transparent and trustworthy digital ecosystem, paving the way for future innovations and applications.

7 Additional Resources

1. Dan Boneh's Talk: `https://www.youtube.com/watch?v=fF9VrtiwQ00`
2. CLI Tool: `https://github.com/0xmayalabs/maya-cli`
3. Benchmarks: `https://docs.mayalabs.tech/`
4. Blog: `https://blog.mayalabs.tech`
5. ZK-IMG: `https://arxiv.org/abs/2211.04775`
6. Article: `https://medium.com/@boneh/using-zk-proofs-to-fight-disinformation-17e7d57fe52f`
7. C2PA: `https://c2pa.org/`
8. Maya ZK Editor: `https://app.mayalabs.tech`