Proof of Novelty

Daniel Severo
Independent Scientist
São Paulo, Brazil
danielsouzasevero@gmail.com

***Abstract*—We propose a design for securing novelty of archived content in distributed ledgers, called *Proof of Novelty*. What constitutes as novel is decided through a consensus mechanism together with a similarity function, which is selected according to the content type (e.g. full-motion videos, textual documents). Scalability is guaranteed by forming a validation committee with cryptographic sortition, which use statistical hypothesis testing to decide on the probability of a content being novel or not. The system can trade-off computational with statistical performance by manipulating parameters. We discuss the usage of this design to secure the novelty of full-motion videos and end with a proposal of future lines of research that can extend the systems capabilities.**

***Index Terms*—blockchain, patent, smart contracts, prior art, full-motion videos, proof of novelty**

I. Introduction

Guaranteeing novelty and authenticity of modern media content (e.g. full-motion videos, textual documents, audio files) is crucial for decision making in areas such as court trials and journalism. Failing to do so risks undermining the credibility of the decision maker. For example, journalists fact check their findings before publishing news articles [12]; patent clerks exhaustively gather evidence of absence of prior art before emitting patent certificates [7]; and media archives (e.g. YouTube) protect artists by taking down copyright-infringing content [23]. Scholarly peer review systems also share this characteristic, where a reviewer is tasked with evaluating the claims made in an academic paper during the process of submission to a conference or journal.

The central concepts underlying the aforementioned examples are *novelty* and *authenticity*. An item is deemed authentic when its provenance is indisputable. What constitutes as novel is often the cause of discourse due to its subjective nature and the trust bestowed upon a centralized agent to discriminate on what is, and what is is not, genuinely novel. More recently, a new issue has emerged due to the proliferation of digital content in the internet age; verifying novelty against large archives. Even if the definition of novelty can be agreed upon by interested parties, exhaustively comparing candidate and existing contents can be computationally intractable due to large data volumes (e.g. 500 hours of video images are uploaded to YouTube every minute [26]). Defining and securing authenticity is comparably more mature and has benefited considerably in recent years with the advent of Digital Signatures [25].

In the case of media archives, checking for novelty prior to acceptance can be seen as a way of preserving the moral integrity of *currently* archived content. A malicious agent can attack content by submitting a modified version and leverage social media as a vector for propagating false information. Considerable time may pass until the false content is debunked and may cause irreversible damage to the owner of the original content. Complexity of attacks range from audio manipulation and video context clipping to modern computer vision techniques (e.g. Deepfakes) [8].

Pior work has tackled the issue of detecting video tampering *from within* the archive by using content-sensitive identifiers (e.g. digests from cryptographic hash functions) at the moment of content ingestion and storing them in permanent ledgers (e.g. blockchains). The same techniques are also employed for combating spoofing, where content provenance is contested; closely relating to issues with authenticity [14]. To the best of our knowledge, little to no work has been done in preventing the ill-usage of created content.

In this paper, we contribute with

1) a consensus mechanism for securing genuity, called *Proof of Novelty*[1] (PoN);
2) an approach to combat false media content in digital archives with PoN.

Our design draws inspiration from *Patent Systems*. We discuss shortcomings of our approach as well as model details in the following sections.

II. Background

A. Patent Systems and Prior Art

A patent is a legal document that provides proof of ownership of intellectual property (IP) [7]. It is commonly issued by government run agencies to individuals or organizations. Its main function is to secure legal exclusivity regarding sales, production and distribution of the IP to the owner. The procedure of emission is initiated with a formal request by the party interested in obtaining the patent, called a *patentee*. A *patent clerk*, representing the emitting agency, is then attributed with the task of verifying patentability conditions such as (but not limited to) novelty and non-obviousness of the invention. This is done by collecting evidence of absence of previous

This short paper was written for the VIRTUAL DESIGN CHALLENGE FOR AUTHENTICATING AND PROTECTING FULL MOTION VIDEOS: https://blockchain.ubc.ca/news/virtual-design-challenge-authenticating-and-protecting-full-motion-videos

[1] https://github.com/dsevero/Proof-of-Novelty

similar work, called *Prior Art*. Searching is done against niche databases specialized in storing technical documents [7].

Granting a patent is an attestation of novelty, conditioned on trust in the emitting agency. An improper Prior Art search can lead to erroneous conclusions of novelty regarding the patentee's invention [4].

B. Similarity Measures

A similarity measure is any function that quantifies the degree of similarity between objects [18]. For example, the Jaccard index measures similarity between sets $\mathbb{A}$ and $\mathbb{B}$, by comparing the number of common elements. Formally, it is the ratio between the the intersection and union, and varies between 0 and 1; $J(\mathbb{A}, \mathbb{B}) = |\mathbb{A} \cap \mathbb{B}| / |\mathbb{A} \cup \mathbb{B}|$. These metrics are common in applications regarding information retrieval (e.g. search systems) and recommendation systems [20].

Distance functions d are common throughout literature and can be converted into a similarity measure by a simple transformation such as $d \to \frac{1}{1+d}$. Both distance and similarity measures are sensitive to their domain. For example, the Jaccard index previously mentioned can not be used directly on real valued data (what would $J(5, 2.4)$ mean?) [18].

There is a vast amount of literature addressing metrics for text, full-motion videos and audio applications [16]. The field of *Similarity Learning* is concerned with the task of learning similarity and distance functions through the usage of labeled examples and Machine Learning techniques (i.e. Supervised Machine Learning) [16].

C. Blockchains and Smart Contracts

A Blockchain is a growing list of ordered records that contain digitally signed transactions [2]. Appending a record to the chain is done by cryptographically hashing the list and adding it to the incoming record. Blockchains are considered immutable, since altering a transaction anywhere in the list would require recomputing the hash of all records that came after it [2]. This data structure often lives on a distributed public peer-to-peer network, where all peers hold a copy of the list and must reach consensus on which incoming blocks should be appended.

Multiple consensus mechanism have been proposed throughout literature. Examples are Proof of Work [10], Stake [15] and Authority [9] (PoW, PoS, and PoA; respectively). Some protocols require partial centralization to be secure, such as PoS and PoA. PoW can be fully distributed, but is inefficient in the time it takes to append a new block to the chain and requires a surplus amount of resources [10].

A *Smart Contract* usually refers to a computer program that is executed collaboratively on a blockchain [28]. Consensus on the state of execution of the program is reached through the consensus mechanisms previously mentioned. The name comes from its usage in enforcing terms of agreements between parties in a way that doesn't require trust in a centralized moderator, such as a government agency.

The *Ethereum Virtual Machine* is a distributed blockchain system that implements a set of machine-level instructions similar to a general purpose computer [29]. It supports Smart Contracts, which are usually written in languages such as Solidity, Vyper and LLL. Programs running on the EVM can be used as the back-end (i.e. server-side) for software applications called *Decentralized Apps* (Dapps).

D. Cryptographic Sortition

Sortition is the act of randomly selecting elements of a group to form a committee. It is usually associated with political systems as an alternative to elections, where representatives are selected by sampling from the population of eligible citizens. For element i, the probability p_i of being selected is proportional to some non-negative real valued weight ω_i, such as $p_i = \omega_i / \sum_i \omega_i$. Choosing equal weights (i.e. $\omega_i = \omega_j; \forall i, j$) results in a uniform distribution, where every element has the same chance of being selected.

Sortition can be used to scale blockchain systems, by forming a committee of peers to reach consensus on new blocks; possibly replacing PoW [11]. Although promising, a naive implementation can cause security concerns due to interactions between committee members and their exposure to malicious agents. Algorand [11] created a consensus protocol which forms committees without the need of interactions or exposure, using *Cryptographic Sortition*. Any peer holding a private key can verify and prove self-membership in the committee using a *Verifiable Random Function* [21].

E. Content-Addressable Storage

Location-addressable systems reference resources based on location. Examples are everyday systems such as bank accounts, housing addresses, traditional databases, and the internet. Content-addressable systems use the content itself (usually in the form of a hash) as a locator [3].

The *InterPlanetary File System* (IPFS) [5] is a distributed content-addressable storage system that is commonly used with blockchain technology and Smart Contracts. Referencing stored content by its cryptographic hash allows IPFS to remove duplicated files, as any two files that are exactly the same will have equal hashes.

III. Problem Statement

Let $\mathbb{C}$ represent a collection of content archived in a distributed content-addressable storage system such as IPFS. For each content $c \in \mathbb{C}$ there exists some owner $o \in \mathbb{O}$ such that the mapping $c \to o$ is one-to-one but the inverse $o \to c$ is one-to-many. In other words, one owner can have multiple contents, but each content has a unique owner.

A *similarity measure* on $\mathbb{C}$ is a function $s \colon \mathbb{C}^2 \to \mathbb{R}$ that takes a pair of contents as input and returns a scalar that captures the notion of what constitutes as similar. s can be anything from an Artificial Neural Network to a simple comparison of descriptive statistics.

The system will accept or reject the insertion of a candidate $c' \notin \mathbb{C}$ based on a subset of similar content

$$\mathbb{S}(c', \mathbb{V}) = \{s(c, c') \mid \forall c \in \mathbb{V} \subseteq \mathbb{C}\}$$

with some rule R

$$\mathbb{S}(c', \mathbb{V}) \xrightarrow{R} r \in \{\text{accept, reject}\}$$

where (R, s) is defined through consensus.

Accepting c' can be viewed as the equivalent of emitting a *certificate of novelty*. The objective is to minimize the rate of false positives while maintaining a reasonable rate of acceptance. An owner o wishing to prove novelty of c can do so by showing that $c \in \mathbb{C}$ and $c \to o$ to any entity that trusts the system (R, s).

IV. Related Work

Previous work has tackled the problem of authenticity (i.e. proving provenance) by securing files on content-addressable storage (e.g. IPFS) with blockchain technology. [14] implements a Smart Contract on EVM where artists can register their content hosted on IPFS to be shared, modified, and purchased. The blockchain tracks a history of interactions and creates a hierarchy of parent-child relationships that makes it possible to trace content provenance. Bernstein [1] is a decentralized app that implements a Patent System. It can prove ownership, existence and integrity (i.e. tamper-proof) of stored files. ARCHANGEL [6] is a tamper-proof video archive equipped with a content hashing mechanism insensitive to most lossless compression algorithms.

None of these are capable of verifying novelty of files submitted to the system. A malicious agent can download content, modify it for ill-usage, and submit it back to the blockchain. Even though it is possible to prove that the modified content was inserted after the original (since blockchains are ordered lists), there is still the issue of locating the original content for comparison.

MediaChain was a peer-to-peer database for users to collaborate on open-media. It contained a metadata resolution protocol that was intended to map similar content to the same metadata. The project was halted when bought by Spotify in 2017 and the solver was never fully implemented [24].

V. Design Proposal

An overview of the system is presented through narrative. Details are addressed in later subsections.

A. Overview

Consider a credible archive $\mathbb{C}$ of similar content, with content-addressable hashes secured on a smart contract enabled blockchain, an associated similarity measure s, and an acceptance/rejection rule R. For example, arXiv articles stored on IPFS with community defined s and R secured by Ethereum. Note that credibility is a consensus amongst peers and not an intrinsic property of the archive (e.g. scholarly peer-reviewed venues). An owner o', wishing to prove novelty of content $c' \notin \mathbb{C}$, makes a transaction on the blockchain and receives a random subset of content-hashes of elements $c \in \mathbb{V} \subseteq \mathbb{C}$. The owner now uses s to calculate the similarity of c' with the elements of $\mathbb{V}$ (i.e. $\mathbb{S}(c', \mathbb{V})$) and submits the calculations back to the smart contract for evaluation. Using cryptographic sortition, the smart contract chooses a random committee that is tasked with off-chain verification of a subset of the results, $\mathbb{S}(c', \mathbb{F} \subseteq \mathbb{V})$. Consensus is reached by the committee regarding the legitimacy of $\mathbb{S}(c', \mathbb{V})$, and c' is accepted or rejected into the archive based on the rule R.

B. Choosing $\mathbb{V}$ and $\mathbb{F}$

The degree of novelty verified by the system is directly related to the cardinality (i.e. number of elements) of $\mathbb{V}$, represented by $|\mathbb{V}|$. Making $\mathbb{V} = \mathbb{C}$ will guarantee that all files in the archive are compared against the candidate content c', but will make the computation of $\mathbb{S}(c', \mathbb{V})$ resource intensive. Similarly, raising $|\mathbb{F}|$ decreases the probability that an owner can successfully manipulate the values of $\mathbb{S}(c', \mathbb{V})$, but requires more computational power during the verification process by network peers.

Scalability of this solution comes from the fact that $|\mathbb{F}| < |\mathbb{V}|$, which is possible only if $\mathbb{F}$ is hidden from the owner o'. This means that the candidate does not know which values of $\mathbb{V}$ will be verified by his peers, forcing him to compute $s(c, c')$ for all values of $c \in \mathbb{V}$. The probability of randomly guessing which elements are in $\mathbb{F}$ can be made as low as required.

C. Consensus of Peers Regarding $\mathbb{S}(c', \mathbb{V})$

Using only $\mathbb{S}(c', \mathbb{F} \subseteq \mathbb{V})$, the committee must infer if $\mathbb{S}(c', \mathbb{V})$ is valid. Inference through statistical hypothesis testing [22] can be leveraged to decide if candidate content is novel.

D. Certificate of Novelty

Showing that the content-hash of c is in the distributed archive (i.e. $c \in \mathbb{C}$) provides credibility with respect to the novelty of c. Agencies can request that candidates provide PoN on distinct (R, s) for different tasks. For example, peer-reviewed journals can use arXiv to build a blockchain where researchers submit their written work for tests against plagiarism before submission.

Notice that the burden of proof of novelty is bestowed upon the owner of the candidate content, and not the the network. This design can also provide varying levels of proof. For example, a content validated against a larger value of $|\mathbb{F}|$ will result in a higher confidence of novelty. This can be extended to the point where there are varying or even continuous levels of confidence, where the candidate owner can resubmit the same content to the blockchain when demanded by a third party that trusts the system.

VI. Application to Full-motion Video Archives

An advantage of PoN is that changing the content type (e.g. full-motion videos, audio tracks, images) only requires us to choose a different similarity function. Each content type has a unique property that usually implies different types of techniques for the same task. For example, although full-motion videos are a sequence of image frames, the similarity between sequential frames allows us to use specific compression algorithms that would not perform as well on a set of

random images [17]. Literature in Similarity Learning for full-motion videos as well as Near-Video Duplicate Detection [19] has evolved considerably with the use of Convolutional and Recursive Neural Network [27] and can be explored to find candidate similarity functions.

A single similarity function is not expected to perform well for all types of videos. The novelty certificate can be segmented by topic (e.g. famous speeches, car collisions) and the entity wishing to validate a candidate video can specify which certificate is needed.

Videos with long duration (i.e. tens of minutes) can be broken down with scene detection algorithms, such as in [6]. This framework can be extended to sample frames or scenes from a candidate video, and novelty can be secured stochastically.

VII. Conclusions and Future Work

In this paper we presented a design proposal for a Smart Contract compatible blockchain that can secure *novelty* of currently existing archives. Our work can scale by leveraging a cryptographic sortition algorithm to form a committee of validators during insertion of new content. It is easily applicable to full-motion videos, audio files, textual documents and any other type of content for which a similarity function can be defined. Through statistical hypothesis testing, guarantees can be made with respect to the degree of novelty (within the scope of the archive).

There are multiple points of improvement that future work can tackle, such as:

- implementing and running experiments on Smart Contract blockchains such as Ethereum;
- applying decentralized and collaborative machine learning to progressively update a similarity function [13];
- investigating statistical hypothesis tests and guarantees for different content and similarity functions.

Acknowledgment

We would like to thank Professor Chen Feng for the invitation to participate in the VIRTUAL DESIGN CHALLENGE FOR AUTHENTICATING AND PROTECTING FULL MOTION VIDEOS, hosted by Patriot One Technologies Inc. in collaboration with the Blockchain research cluster at The University of British Columbia.

References

[1] Digital ip protection.

[2] Andreas M Antonopoulos and Gavin Wood. *Mastering ethereum: building smart contracts and dapps*. O'Reilly Media, 2018.

[3] Benjamin Atkin, Grzegorz Calkowski, Cristian Ungureanu, and Cezary Dubnicki. System and method for content addressable storage, June 29 2010. US Patent 7,747,663.

[4] Shariq Bashir and Andreas Rauber. Improving retrievability of patents in prior-art search. In *European Conference on Information Retrieval*, pages 457–470. Springer, 2010.

[5] Juan Benet. Ipfs-content addressed, versioned, p2p file system. *arXiv preprint arXiv:1407.3561*, 2014.

[6] Tu Bui, Daniel Cooper, John Collomosse, Mark Bell, Alex Green, John Sheridan, Jez Higgins, Arindra Das, Jared Keller, Olivier Thereaux, et al. Archangel: Tamper-proofing video archives using temporal content hashes on the blockchain. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 0–0, 2019.

[7] Julie Callaert, Bart Van Looy, Arnold Verbeek, Koenraad Debackere, and Bart Thijs. Traces of prior art: An analysis of non-patent references found in patent documents. *Scientometrics*, 69(1):3–20, 2006.

[8] Robert Chesney and Danielle Keats Citron. Deep fakes: a looming challenge for privacy, democracy, and national security. 2018.

[9] Stefano De Angelis, Leonardo Aniello, Roberto Baldoni, Federico Lombardi, Andrea Margheri, and Vladimiro Sassone. Pbft vs proof-of-authority: applying the cap theorem to permissioned blockchain. 2018.

[10] Cynthia Dwork and Moni Naor. Pricing via processing or combatting junk mail. In *Annual International Cryptology Conference*, pages 139–147. Springer, 1992.

[11] Yossi Gilad, Rotem Hemo, Silvio Micali, Georgios Vlachos, and Nickolai Zeldovich. Algorand: Scaling byzantine agreements for cryptocurrencies. In *Proceedings of the 26th Symposium on Operating Systems Principles*, pages 51–68. ACM, 2017.

[12] Lucas Graves. *Deciding what's true: The rise of political fact-checking in American journalism*. Columbia University Press, 2016.

[13] Justin D Harris and Bo Waggoner. Decentralized & collaborative ai on blockchain. *arXiv preprint arXiv:1907.07247*, 2019.

[14] Haya R Hasan and Khaled Salah. Combating deepfake videos using blockchain and smart contracts. *IEEE Access*, 7:41596–41606, 2019.

[15] Sunny King and Scott Nadal. Ppcoin: Peer-to-peer crypto-currency with proof-of-stake. *self-published paper, August*, 19, 2012.

[16] Brian Kulis et al. Metric learning: A survey. *Foundations and Trends® in Machine Learning*, 5(4):287–364, 2013.

[17] Didier Le Gall. Mpeg: A video compression standard for multimedia applications. *Communications of the ACM*, 34(4):46–59, 1991.

[18] Marie-Jeanne Lesot, Maria Rifqi, and Hamid Benhadda. Similarity measures for binary and numerical data: a survey. *International Journal of Knowledge Engineering and Soft Data Paradigms*, 1(1):63, 2009.

[19] Yeguang Li, Liang Hu, Ke Xia, and Jie Luo. Fast distributed video deduplication via locality-sensitive hashing with similarity ranking. *EURASIP Journal on Image and Video Processing*, 2019(1):51, 2019.

[20] Donald Metzler, Susan Dumais, and Christopher Meek. Similarity measures for short segments of text. In *European conference on information retrieval*, pages 16–27. Springer, 2007.

[21] Silvio Micali, Michael Rabin, and Salil Vadhan. Verifiable random functions. In *40th Annual Symposium on Foundations of Computer Science (Cat. No. 99CB37039)*, pages 120–130. IEEE, 1999.

[22] Whitney K Newey and Daniel McFadden. Large sample estimation and hypothesis testing. *Handbook of econometrics*, 4:2111–2245, 1994.

[23] Damien S O'Brien and Brian F Fitzgerald. Digital copyright law in a youtube world. *Internet Law Bulletin*, 9(6 & 7):71–74, 2006.

[24] Sarah Perez. Spotify acquires blockchain startup mediachain to solve music's attribution problem. *TechCrunch. com. Available at https://techcrunch. com/2017/04/26/spotifyacquires-blockchain-startup-mediachain-to-solve-musics-attribution-problem*, 2017.

[25] David Pointcheval and Jacques Stern. Security arguments for digital signatures and blind signatures. *Journal of cryptology*, 13(3):361–396, 2000.

[26] Mark R Robertson. 500 hours of video uploaded to youtube every minute [forecast]. *Tubular Insights: Video Marketing Insights*, 2015.

[27] Parsa Saadatpanah, Ali Shafahi, and Tom Goldstein. Adversarial attacks on copyright detection systems. *arXiv preprint arXiv:1906.07153*, 2019.

[28] Nick Szabo. The idea of smart contracts. *Nick Szabo's Papers and Concise Tutorials*, 6, 1997.

[29] Gavin Wood et al. Ethereum: A secure decentralised generalised transaction ledger. *Ethereum project yellow paper*, 151(2014):1–32, 2014.