2022 International Conference for Advancement in Technology (ICONAT)
Goa, India. Jan 21-22, 2022

Blockchains to curb Fake News in an Online World

1st Abhishek Wahane
School of Computer Engineering and Technology
MIT World Peace University
Pune, India
abhishekwahane1@gmail.com

2nd Balaji Patil
School of Computer Engineering and Technology
MIT World Peace University
Pune, India
balaji.patil@mitwpu.edu.in

2022 International Conference for Advancement in Technology (ICONAT) | 978-1-6654-2577-3/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICONAT53423.2022.9725933

***Abstract*—Fake news has become a significant problem in today's online world. Misinformation is easily propagated through online social media platforms. It is essential to curb the spread of fake news. Blockchain has provided a framework for tamperproof and traceable data, and can therefore be used for developing a news platform. Applying blockchain technology to news tracing has formed the groundwork for supressing the spread of fake news. However, the scale of online platforms and the characteristics of blockchain add several new challenges that hinder the real world deployment of such solutions. This study reviews various blockchain based methods that combat the spread of fake news on online platforms and discusses the challenges faced in realizing such solutions.**

***Index Terms*—Blockchain, Blockchain scalability, Fake News, Fake-news detection in social media**

I. Introduction

The rise of the internet dramatically changed the way information is shared. Information would reach people across the world faster than ever before and almost effortlessly. Despite its numerous benefits and noble applications, the widespread adoption of the Internet paved way for a new problem - misinformation. Fake news is when users with malicious intents deliberately propagate misinformation by modifying contents of an authentic news feed and mislead other users into believing an unreal state of events [1].

The motivation behind posting fake news stories is usually to fetch people's attraction, to gather more page views on a particular site or indirectly harm another political party, company, or a community. Fake news influences the sentiments of citizens and can have real-world consequences. It can potentially influence a country's election results or lead people into believing unscientific pronouncements. It can change people's mindset towards a particular entity or community, which might lead to hatred, protests and other violent acts.

Traditional means of news consumption, such as newspapers and well-known television channels, are much less likely to contain fake news items, as to be worthy of their recognition, they have an obligation to maintain themselves as a reliable source. However, digital platforms, especially social media platforms, which also serve as news consumption sources, do not have similar incentives as traditional media. Digital media as a news source, has been rising for the past few years.

A. Fake News and Social media

Social media is rising day by day. It has witnessed rapid adoption in the past decade. Table I shows the number of active users on few leading social media platforms as of July 2021 [2].

TABLE I
NUMBER OF USERS ON LEADING SOCIAL MEDIA PLATFORMS AS OF JULY 2021 [2]

Platform	**Number of Users (in billions)**
Facebook	2.85
YouTube	2.29
WhatsApp	2.00
Instagram	1.38

Social media has been prevalent as the source of news for most people [3]. An opportunist organization or individual can create fake news items easily with catchy headlines. Such clickbait stories can spread exponentially on social media platforms. A news item may spread to millions of users around the world within days. This is largely due to people's tendency to share appealing news without checking its authenticity or understanding its consequences.

TABLE II
PROPORTION OF PEOPLE USING SOCIAL MEDIA FOR NEWS [3]

Platform	**Platform's users who get news there**
Facebook	73%
Twitter	71%
Reddit	62%
YouTube	38%
Instagram	36%

It is safe to say that Facebook is the epicentre of media attention. Facebook had implemented a feature to flag fake news on the site [4]. However, the methodology behind identifying fake news is still something to ponder upon. Flagging news items as fake may not prove to be useful if actual fake news items are never detected by the system. The effectivity of this feature is still questionable [5].

978-1-6654-2577-3/22/$31.00 ©2022 IEEE

II. Approaches to Mitigating Fake News

A. The Problem with Artificial Intelligence

Artificial Intelligence is a rising technology. It has immense potential to improve human life. AI might be the first consideration for developing a system to accurately detect fake news and prevent its spread. However, there are reasons why it may not be the best solution for the preventing fake news.

If AI can be used to detect fake news, it can also easily be used to create it. The adoption of AI has increased the ability to create fake content [6], [7]. With the rising popularity of synthetic content generation, it is only going to become more difficult to identify fake content using an AI [8].

While AI can help curb the propagation of fake news, it cannot guarantee complete elimination of fake news. A false negative in the system would lead to a fake news article being labelled as true, and a false positive would lead to a true news article being labelled as fake.

Fake news content can be generated using AI, using AI itself to fight against it is counter-intuitive [9]. The problem stems from the misuse of social media and the ability to propagate swiftly among users. The better solution should therefore leverage the common supervising strength of society. There is need to focus more on the source of the news and emphasize the publishers rather than the individual stories [1].

B. A Better Approach

Recent AI based solutions for tackling fake news usually involve analysing the content of the news item once it is published or in case of social media, has already been posted by an entity. The mitigating action relies on the mere strength of the AI software used to verify contents of news, hence cannot be perfect.

The better approach to curbing fake news should focus on 'who published the news' rather than 'what is the news'. This shifts the focus from analysis to accountability. Fake News is created by entities because they believe they cannot be identified and be held accountable. Since, blockchain is a consensus based system that creates immutable records, participants can see who created or modified a record.

The use of Blockchain is not to analyse contents of the news item, but to maintain a trace of its origination and therefore guarantee a source. Such systems can thus identify news as fake before being viewed by thousands of people on a platform.

To understand this further, It is important to establish the distinction between what is factual and what is the truth. To claim something as the truth requires examining individual facts mentioned. Consider the example of a statement made by a politician. The politician on day 1 claims that a particular fund money was distributed. On day 2, he/she claims that the money was not distributed. Here, determining whether the money was actually distributed or not is not of concern. What is of concern though is what the politician stated - that is factual news. AI based solutions will attempt to verify whether the content of the news is true or not. The use of blockchain technology is not to verify whether the news item contains the truth or not, but to verify whether it is factually correct.

III. News Tracing through Blockchain

A. Blockchain Basics

Blockchain is a set of blocks linked together thus forming a chain. It provides a decentralized database of any transaction involving value. It creates a record of whose authenticity can be verified by participants of the network. Once data is recorded on a blockchain, it is almost impossible to change it. The properties of blockchain have made it suitable for application such as cryptocurrencies, ownership verification, logistics tracking etc. The same properties can prove useful for building a news infrastructure that can verify news items.

Each block on a blockchain stores data, hash of the block, and hash of the previous block. Since hashes of previous blocks are also stored, it essentially links all blocks. If data is changed or tampered on a block, its hash changes, thus breaking the link and rendering following blocks invalid. In this case modification of data on the block can be identified.

To prevent recalculation of hashes of all following blocks, in an attempt to revalidate the blockchain, a consensus algorithm is used. Consensus algorithms are a set of computations that need to be performed before adding a block to the blockchain. This computation intentionally slows the process of adding blocks, thereby preventing it from tampering.

B. Tracing News Source

Wenqian Shang et al. describes a potential implementation of news tracing using blockchains [10]. They propose a block structure where the header contains, timestamps, hashes of the current block and previous blocks to help form a chain structure. When fake news is created i.e. a block is modified, its header hash changes, thus breaking the link. Since, blockchain is decentralized, change on one block must be replicated on all other blocks. To prove validity of the changed blocked,

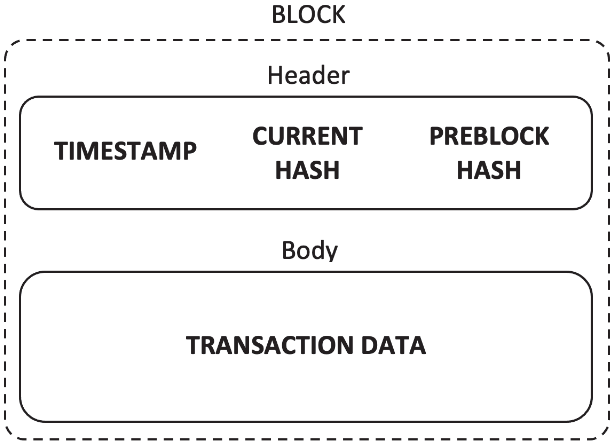

Fig. 1. Sample Block Structure

more than half the network must agree to the change, which is almost impossible to achieve.

Along with block designs, researchers [10], [11] also propose use of tuples in their frameworks. In [10] the authors mention an annotation based method which represents news data with a tuple consisting of $< s, d, i >$, where s represents the data source, d represents the target data, and i represents the intermediate data result. During transmission of news, these tuples can help form a path to trace. In [11] the authors propose an architecture, wherein news publishers can publish news onto the platform, which computes a tuple $((id, n, tn), H(n))$. Here, id is the unique transaction identifier, n is the news payload, tn is the timestamp of news generation, and $H(n)$ is the hash of the news payload.

IV. Challenges for Implementation

A. Integration with Social Networks

As discussed earlier, social media platforms are major sources of news for the masses. Therefore whatever solution is deployed to curb fake news must also extend its benefits to online platforms. In other words, the blockchain based method must be scalable to social media platforms. Hence, researchers incorporate this consideration into their frameworks [11], [12].

However, social media is designed to be fast-paced. Users expect ease-of-use and quick response time for their actions. Adding transactions on the blockchain is computationally expensive and takes time. Integrating blockchain into a social network without hindering user experience is therefore a challenge.

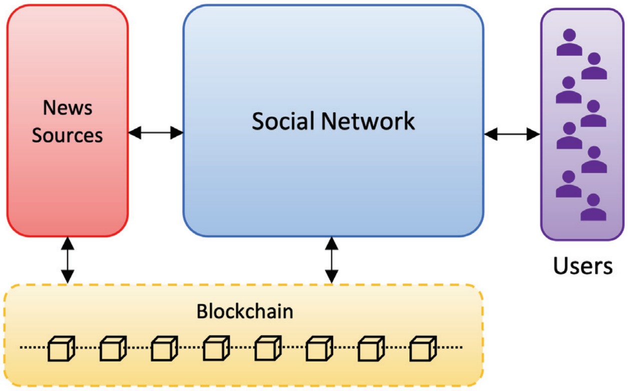

Fig. 2. Blockchain System integrated with a Social network

M. Saad et al. in [11] propose asynchronous execution of the backend blockchain based activities like verification, transaction generation along with frontend user based actions such as clicking on the 'share' or 'post' buttons. This might render a short period of time when the news is not verified. If the platform chooses to allow news to be shared during this time period i.e. before finishing the verification process, then it can be marked with certain visual indicators like tick marks or warning signs to caution users viewing the news.

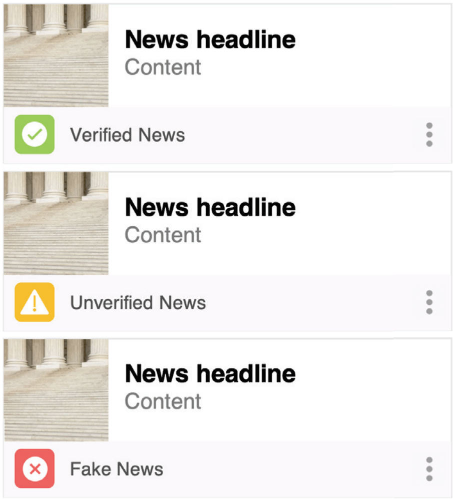

Fig. 3. Example of how News Visual Indicators can look

Websites and online platforms are conscious of storage requirements at both, their end and the user's end. This is mainly to reduce costs and ensure faster load times for the user. If a user wants to verify a news item, the blockchain can be queried to verify it. However, the size of a typical blockchain network is too large. Bitcoin is more than 360 GB in size [13] as of August 2021. Downloading such huge data for verification is not feasible. To overcome this problem, a Merkle Tree [14] is used. Each block on the blockchain is hashed, paired, then hashed again. This new hash is paired with the hash of another pair and so on. This forms a structure that resembles a tree. If a block needs to be verified, only the corresponding hashes at levels above are queried till the root hash. Thus, the block can be verified without traversing the entire chain of blocks.

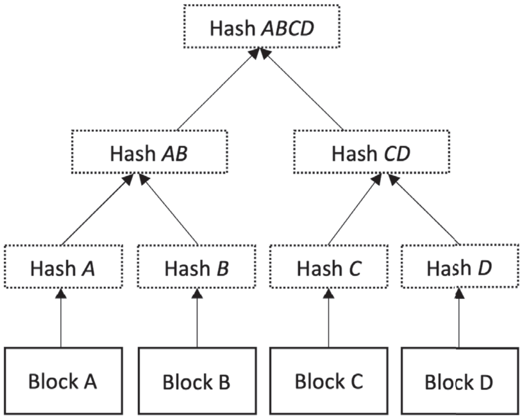

Fig. 4. Visualization of Merkle Tree

B. Blockchain Scalability

A block on the blockchain gets published only after certain amount of computation is performed. For Bitcoin, a nonce needs to be computed for a node to be appended. This process is also termed as mining. The time it takes to create a transaction on a blockchain depends on the consensus algorithm used. Various consensus algorithms are used by blockchain applications and can be broadly classified as Proof of Work (PoW) and Byzantine Fault Tolerant (BFT) protocols. Bitcoin uses the Proof of Work (PoW) algorithm. Each algorithm has its pros and cons [15]. Since social media users are extremely large in number, the consensus algorithm needs to be able to perform numerous transactions.

TABLE III
COMPARISON OF POW AND BFT CONSENSUS ALGORITHMS WITH IDEAL CHARACTERISTICS IN BOLD

	PoW	**BFT**
Scalability	**Excellent**	Limited
Throughput	Limited	**Excellent**
Latency	High	**Low**

Certain protocols based on Byzantine Fault Tolerance have high transaction throughput, but low network scalability [15]. Assuming that only news publishers are the ones publishing news and users only share the news items, then the system would need to scale only to transactions performed by news publishers on the social media platform. The users of the platform would be isolated from performing transactions directly. Hence BFT based consensus algorithms are relatively more suitable.

A. D. Dwivedi et al. proposed methods in [12] that address the scalability problem of blockchain using bloXroute [16]. BloXroute uses the blockchain distributed server concept to solve the scalability bottleneck at the network layer and allows up to 100 times faster data propagation.

Private blockchains [17] are another solution to improve performance of a blockchain. Private blockchain is a solution for corporates and organisations who want to reap the benefits of blockchain, but also want to confine the blockchain network to just their enterprise. Therefore, they are not completely decentralized. As concluded by studies in [18], [19], Hyperledger Fabric [20], a modular blockchain framework aimed for use at private enterprises, was observed to have significantly faster execution times, higher throughput, low latency and reduced computational load. When users share a news item, a smart contract runs to record the event. Chaincode [20] is a smart contract that can be executed with minimal overhead. These characteristics make private blockchains suitable for online platforms. The downside here is that immutability and decentralization are comprised.

C. Handling Multimedia

Newspapers were the most widely adopted news source and still are in many regions. However, with increasing use of social media and digital devices, news can now be in forms such as images, audio or video rather than just text. Any fake news prevention measure must also take this into consideration. Several researchers who have worked on blockchain solutions have also proposed methods to handle news in multimedia formats.

S. Huckle and M. White present their prototype distributed application, Provenator [21]. It is based on the Ethereum framework and can be used to verify media files. The application utilises the concept of metadata. It uses preservation metadata implementation strategies (PREMIS) [22] metadata definitions to store media on the blockchain. To verify ownership of a particular digital media, the file needs to be uploaded onto Provenator, which then looks for the generated hash on its blockchain network.

A similar video based implementation has been presented in [18] wherein the system generates MD5 hashes of a video and stores it on a blockchain. The hash of a particular video is also stored on neighbouring nodes. This helps to identify if the video was modified by comparing the hash of videos.

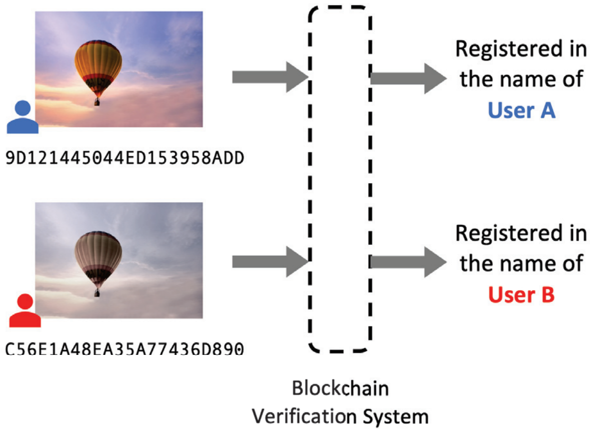
9D121445044ED153958ADD
C56E1A48EA35A77436D890
Registered in the name of User A
Registered in the name of User B
Blockchain Verification System

Fig. 5. Scenario where similar image can be registered by another entity

Data on social media platforms is humongous and classifies as big data. To reduce storage costs, most platforms use some form of compression to treat media files. For instance, when a user attempts to uploads images, they get compressed to reduce file size and then get uploaded to the servers. This is a problem since, compressing a file will change its hash and will therefore be treated as a new file. Moreover, anyone can alter the media, thereby changing the hash and publish it on the blockchain by their name. For example, an artist might create a digital painting and register it on the blockchain by their name. Another person can crop the image or rotate it a little, then register on the blockchain by their name. The blockchain system will allow this since, it treats the modified image as a new image due to the generated hash being different. Traditional hash based techniques would therefore not be effective in such cases.

This limitation of plain hashing strategies is addressed by S. u. Rehman et al. in [24] by using perceptual hashing [25]. Using perceptual mean hash technique they were able to identify alteration of images by comparing their hash of similar images through hamming distance. Their proposed system can identify alerted images even if objects are added/removed in the image, the image is rotated or the colour scheme is changed.

Authors of [12] propose keyed watermarking as an approach for verifying integrity of data. Watermarking is similar to digital signatures. Digital signatures can be used for verification of documents and images, whereas, watermarking can be applied to digital contents such as audio/video as well. The authors of the paper propose robust watermarking which can also handle media being manipulated by rotation, scaling, compression etc.

V. Research Constraints

The idea of using blockchain for mitigating the fake news problem has been widely explored by researchers. However, most of these works are at a theoretical level and have little proof of concept. As discussed in previous sections, there are still many challenges in deploying a system that uses blockchains to curb fake news on online platforms:

- Various theoretical frameworks for a blockchain based news system, have been proposed. However, a deployment is yet to be seen as its practical implementation faces numerous challenges.
- A major challenge is the computational overhead brought about by the use of blockchain. This might make companies unwilling to adopt blockchain technologies without a strong incentive.
- Further study in consensus algorithms is needed to address the performance and scalability of blockchain for use on online platforms.
- The time it takes to add a transaction to the blockchain might hinder the user experience on media platforms. Therefore, platforms will have to undergo user experience research and work on designs to incorporate these changes.
- The prevalence of multimedia such as, images, audio and video, further complicate the matter. Generating hashes for such type of data is more computationally intensive than for text.
- Many social media platforms utilize compression techniques to reduce file size. This in itself changes the hash generated for the file. Slight alterations on the media file should not be treated as a new file by the system. Therefore, hash algorithms solely cannot be relied upon for such as use case.

Further research is needed to advance development of blockchain based system architectures. Studies should focus on development of novel consensus algorithms, hashing strategies and other innovative techniques to address scalability of blockchains.

VI. Conclusion

Blockchain has provided a solid foundation for building decentralized, consensus based applications without the need of a third party. Current research has mainly involved theoretical frameworks, methodologies and prototypes for the use of blockchain technology to solve the fake news problem. There is wide gap between characteristics of current blockchain based systems and the requirements of companies, organisations and online platforms. Further study is needed in areas of consensus algorithms, hashing techniques, and scalability of blockchains. Consensus algorithms need to accommodate as many participants on a network while maintaining good transaction speeds. Hashing techniques should be able to calculate hashes fast and be strategically implemented such that alterations in data does not affect ownership records. It is only with these advancements that a blockchain based news system can be realized. Bringing such systems to fruition will contribute to a society with a healthy public opinion environment.

References

[1] D.M.Lazer et al., “The science of fake news,” Science, vol. 359,Mar. 2018, pp. 1094-1096.

[2] S. Kemp. “Digital 2021 July Global Statshot Report” datareportal.com. https://datareportal.com/reports/digital-2021-july-global-statshot (accessed Aug. 14, 2021).

[3] E. Shearer and E. Grieco. “Americans Are Wary of the Role Social Media Sites Play in Delivering the News” pewresearch.org. https://www.pewresearch.org/journalism/2019/10/02/americans-are-wary-of-the-role-social-media-sites-play-in-delivering-the-news/ (accessed Aug. 14, 2021).

[4] A. Mosseri. “Addressing Hoaxes and Fake News” fb.com. https://about.fb.com/news/2016/12/news-feed-fyi-addressing-hoaxes-and-fake-news/ (accessed Aug. 14, 2021).

[5] M. Wilson. “Study: Facebook’s fake news labels have a fatal flaw” fastcompany.com. https://www.fastcompany.com/90471349/study-facebooks-fake-news-labels-have-a-fatal-flaw (accessed Aug. 14, 2021).

[6] D. Coldewey. “To detect fake news, this AI first learned to write it” techcrunch.com. https://techcrunch.com/2019/06/10/to-detect-fake-news-this-ai-first-learned-to-write-it/ (accessed Aug. 14, 2021).

[7] D. Robitzski. “This Site Uses AI to Generate Fake News Articles” futurism.com. https://futurism.com/site-ai-generate-fake-news-articles (accessed Aug. 14, 2021).

[8] A. Jaiman. “Deepfake detection is super hard!!!” towardsdatascience.com. https://towardsdatascience.com/deepfake-detection-is-super-hard-38f98241ee49 (accessed Aug. 14, 2021).

[9] S. Woolley. “We’re fighting fake news AI bots by using more AI. That’s a mistake.” technologyreview.com. https://www.technologyreview.com/2020/01/08/130983/were-fighting-fake-news-ai-bots-by-using-more-ai-thats-a-mistake/ (accessed Aug. 14, 2021).

[10] W. Shang, M. Liu, W. Lin and M. Jia, ”Tracing the Source of News Based on Blockchain,” 2018 IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS), 2018, pp. 377-381, doi: 10.1109/ICIS.2018.8466516.

[11] M. Saad, A. Ahmad and A. Mohaisen, ”Fighting Fake News Propagation with Blockchains,” 2019 IEEE Conference on Communications and Network Security (CNS), 2019, pp. 1-4, doi: 10.1109/CNS.2019.8802670.

[12] A. D. Dwivedi, R. Singh, S. Dhall, G. Srivastava and S. K. Pal, ”Tracing the Source of Fake News using a Scalable Blockchain Distributed Network,” 2020 IEEE 17th International Conference on Mobile Ad Hoc and Sensor Systems (MASS), 2020, pp. 38-43, doi: 10.1109/MASS50613.2020.00015.

[13] YCharts. ”Bitcoin Blockchain Size” ycharts.com. https://ycharts.com/indicators/bitcoin_blockchain_size (accessed Aug. 14, 2021).

[14] R. C. Merkle, ”A Digital Signature Based on a Conventional Encryption Function” in Proc. Advances in Cryptology — CRYPTO ’87, Aug. 1987, pp. 369-378, doi: 10.1007/3-540-48184-2_32.

[15] M. Vukolić, “The Quest for Scalable Blockchain Fabric: Proof-of-Work vs. BFT Replication” in Proc. Open Problems in Network Security, Oct. 2015, pp. 112-125 doi: 10.1007/978-3-319-39028-4_9.

[16] K. Uri, B. Soumya, K. Aleksandar, and S. E. Gun, “bloxroute: A scalable trustless blockchain distribution network,” Available at https://bloxroute.com/, 2019.

[17] S. Seth. “Public, Private, Permissioned Blockchains Compared” investopedia.com. https://www.investopedia.com/news/public-private-permissioned-blockchains-compared/ (accessed Aug. 14, 2021).

[18] I. R. Fedorov, A. V. Pimenov, G. A. Panin and S. V. Bezzateev, ”Blockchain in 5G Networks: Perfomance Evaluation of Private Blockchain,” 2021 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF), 2021, pp. 1-4, doi: 10.1109/WECONF51603.2021.9470519.

[19] S. Pongnumkul, C. Siripanpornchana and S. Thajchayapong, ”Performance Analysis of Private Blockchain Platforms in Varying Workloads,” 2017 26th International Conference on Computer Communication and Networks (ICCCN), 2017, pp. 1-6, doi: 10.1109/ICCCN.2017.8038517.

[20] E. Androulaki et al. ”Hyperledger fabric: a distributed operating system for permissioned blockchains,” in Proc. 13th EuroSys Conf., Apr. 2018, pp. 1–15, doi: 10.1145/3190508.3190538.

[21] S. Huckle and M. White, “Fake news: a technological approach to proving the origins of content, using blockchains,” Big data, vol. 5, no. 4, pp. 356–371, 2017.

[22] P.Caplan, *Understanding PREMIS*. The Library of Congress, 2009. [Online]. Available: www.loc.gov/standards/premis/understanding-premis.pdf. Accessed: Aug 15, 2021.

[23] A. Dhiran, D. Kumar, Abhishek and A. Arora, ”Video Fraud Detection using Blockchain,” 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA), 2020, pp. 102-107, doi: 10.1109/ICIRCA48905.2020.9182963.

[24] S. u. Rehman, M. U. S. Khan and M. Ali, ”Blockchain-Based Approach for Proving the Source of Digital Media,” 2020 3rd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), 2020, pp. 1-6, doi: 10.1109/iCoMET48670.2020.9073820.

[25] C. Zauner, “Implementation and benchmarking of perceptual image hash functions,” 2010.