Establishing Environment Setup for Preventing Deepfakes using Blockchain Technology

[1]Alok Chauhan,[2]Amit Kumar

[1]AsstProfessorDepartment of Information Technology, RGCER

[2,]AsstProfessorDepartment of Information Technology, PCE

Nagpur,India

Abstract –The Rapid growth of agile gadgets has led to a tremendous increase in digital media utilization, mostly for mobile videos and images at ease of marketing and for entertainment purpose. With the rise of artificial intelligence (AI) and deep learning techniques, fake videos and audio recording that look and sound just like the real thing. Blockchain is the technology which provides a way to store the permanent and tamper proof data where once the data is stored it becomes immutable. With the help of this technology, we are setting up a Blockchain environment on multiple nodes connected in private networkto avoid the increased amount of deepfakes. We are making the use of Ethereum smart contract and Interplanetary File System (IPFS) over those nodes in private network for data transaction. Ethereum smart contracts allows us to trace and track the history of digital content to be altered. The Interplanetary File System is a decentralized storage system where the digital content like videos and images can be stored and it generates hashes of the digital content for smart contracts. The work is underway for designing and implementing a fully functional and operational decentralized blockchain system for combating deepfake**.**

***Keywords*-**Deepfake, Blockchain, Ethereum, Smart contract, IPFS system.

INTRODUCTION

The "deepfake", a term which is the combination of two distinct terms 'deep learning' and 'fake', is a technique for human image synthesis based on Artificial Intelligence. It combines and superimposes existing digital media ontosource images or videos using various machine learning techniques. Because of these technological capabilities, deepfakes have been used to create fake news , malicious hoaxes, fake X-rated videos or financial frauds. The combination of such superimposed digital contents and source videos results in a fake video that shows an individual performing an action at an event that never occurred in reality.

Deepfakes can influence public opinions, political issues or intense the situations that can lead to armed conflicts. On the other side, Mostly deepfakes have been used to create fake videos of celebrities, videos which negatively impacted the reputation of that person once posted online, even if they weren't real.

"When it comes to the area of deepfake, eminent technologies like blockchain can come to the fore to provide some levels of security, approval and validation. Blockchain has typically been touted as a visibility and transparency play, where once something is

done, the "when" and "who" becomes clearly visible. When a user with a digital identity wants to do something, they could be prompted for proof of their identity before access to something (like funds, videos) can be permitted.

A deepfake of the president of India's ruling Bharatiya Janata Party (BJP), Manoj Tiwari, went viral on social media in the country earlier few months, ahead of legislative assembly elections in Delhi. It's the first time a political party anywhere has used a deepfake to offend opposition parties. In the original video Manoj Tiwari speaks in English, criticizing his political opponent Arvind Kejriwal and encouraging voters to vote for the BJP. The duplicate video has been manipulated using deepfake technology so his mouth moves convincingly as he speaks in Haryanvi, the Hindi dialect spoken by the target voters for the BJP.

The BJP has partnered with political communications firm The Ideaz Factory to create deepfakes that let it target voters across the India who used over 20 different languages. The party told Vice that the Manoj Tiwari deepfake reached approximately 15 million people in 5,800 WhatsApp groups. The big risk is that we reach a point where people can no longer trust what they see or hear. In that scenario, a video wouldn't even need to be digitally altered for people to blame it as fake.

As of today ,Technology has raised to the sky . Many centralized companies provide solutions to combat deepfakes and the person or organization making deepfakes can be caught easily. These centralized companies do not allow the users to have an access to a trusted data provenance of digital content. To consider the digital content to be real, users have to use search engines and check various blogs ,posts, news and so on relevant to the fake ones in order to identify the original sources and therefore able to avoid the deepfakes spreading disinformation.

This praperpresents a decentralized system called Blockchain for allowing the social media users to have an access to trusted data provenance of digital content .Blockchain technology offers an immutable and tamper-proof ledger of data and transactions as a shared database, validated by a wide community. Each record created forms a block, and as each block is confirmed by the community, it is paired up with the previous entry in the chain, creating a chain of blocks. Blockchain could be used to prove authenticity and originality of digital media in a way that is decentralized, trusted, and secure, with tamper-proof records, logs, and transactions.

We proposed Ethereum Blockchain based solution with the use of Ethereum Smart Contracts and Interplanetary File System(IPFS) .Ethereum smart contracts allows us to trace and track the history of digital content to be altered using cryptographic hashes. The Interplanetary File System (IPFS) is a decentralized storage system where the digital contents like videos and images can be stored and it generates hashes of the digital content for smart contracts. The Environment setup of the Blockchain system is done by creating a private network by initializing genesis block on multiple nodes and deploying ethereum Wallets.

RELATED WORK

In this section gives the description of related work found in literature on the originality of deepfakes content.

Creating deepfakes of original images, videos or audios leads to the problem of altering the truth and spreading false information around the world . The authors of [2] aimed a system to detect fake videos using Blockchain and smart contracts and to prove the truthfulness of

the videos. The solution provided by the authors is to prove the ethnicity of digital videos in which a safe and trustworthy tracing can be created to the origin. Their solution makes use of a time stamping technique which ensures unique recording of critical situation without hampering it's actual situation.

Original media is often edited for creative content preparation or tampered with to fabricate false propaganda over social media. The authors in [4] proposed a novel watermarking based Multimedia Blockchain framework that can address such issues. The unique watermark information contains two pieces of information: a) a cryptographic hash that contains transaction histories (blockchain transactions log) and b) an image hash that preserves retrievable original media content. Once the watermark is extracted, first part of the watermark is passed to distributed ledger to retrieve the historical transaction trail and the latter part is used to identify the edited / tampered regions. The author has further outlined the requirements, the challenges and demonstrated the proof of this concept in this paper[4].

METHODOLOGY

Firstly, we start to establish private ethereum network using Geth to develop and deploy our smart contracts. Then we initialize the very first block in our Blockchain,the genesis block - genesis.json. After that IPC endpoint will be opened, which is used to process connections to Geth with following programs:

1.Metamask - It is a Ethereum wallet and browser extension that can be on Chrome.It serves as an interface for ethereum based dApps.

2.Ethereum - It is an open-source, blockchain-based, software platform used for its own cryptocurrency i.e. ether. It allows Smart Contracts and decentralised Applications (DApps) to be built and run without any downtime and interference from a third party.

3. Mist - It is ethereum interface which allows user to access dApps available in ethereumnetwork .It manages the contracts of individual in ethereum Blockchain.

4. Remix - It is an open source compiler and IDE that enables user to build ethereum smart contracts with solidity language and to debug to it .

5.IPFS - protocol and Peer to peer decentralized network for sharing and storing in distributed file system .It generates hashes of digital content for smart contracts.

A. System overview:

In this section , we give a brief overview of ethereumblockchain System and entities participating in the system.

Video: A video has important information other than the video frame such as date and time of video and device information on which the video is captured. This information is stored as metadata in an EXIF format (Exchangeable Image File). Every video will be associated with ethereum smart contract which can be created by a new artist. The ethereum address of artist as well as the address of smart contract are essential part of metadata.

IPFS storage: The video and its associated metadata are stored on decentralized, peer to peer file system such as Interplanetary File System (IPFS). IPFS is used to generate a unique hash which is the address of video content and its metadata. This unique hash address is used to locate and access the files stored on IPFS network. Files on IPFS may include terms and conditions agreement of editing and copying the video in case the video is being copied to create different video content by another artists. The unique hash generated by IPFS would be used by smart contract.

B. Implementation Details

The original artist of a video create a smart contract where other artist request for a permission to alter, share, edit according to terms and conditions of an agreement which is saved on IPFS server and it hash is available in smart contract.

The secondary artist (artist who copies the digital content) request for permission to share, edit or alter. A request sent by the secondary artist is also for confirming the terms and conditions of an agreement. The original artist assess this request and then the result is announced. The original contract can handle multiple requests at the same time by the same or different artists.

Once a secondary artist gets an approval to their request, they create a child contract which is same as original contract. Then the secondary artist request for the verification of newly created contract from the original artist. The original artist then approves and grants the attestation after verifying newly created smart contract. After the attestation of smart contract , it is added as child in original contract. Hence, both the contracts point to each other as each of them has the Ethereum address of the other contract.

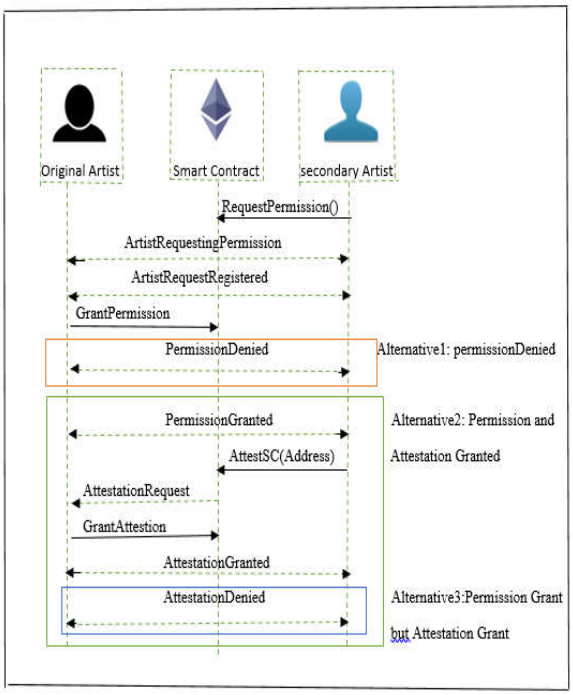

Fig a. Sequence diagram of peer to peer transaction

Figure a presents the sequence diagram that captures the interactions among the original artist, a secondary artist and the smart contract. The smart contract is owned by the original artist and the secondary artist is interested in requesting a permission to edit, alter and have distribution rights. Therefore , as shown in Figure a the secondary artist calls the RequestPermission() function which indicates that they have also read the terms and conditions agreementavailable on IPFS server. This creates two successful events announcing the registration of a secondary artist request. Then the original artist would reply back with the result of whether to grant the permission or deny it. Based on the result from the original artist three different scenarios take place. Either the permission is denied which is shown in Alternative 1 in Figure a or the permission is granted as shown in Alternatives 2 and 3. A granted permission allows the secondary artist to create a child smart contract which is an exact copy of the main contract in terms of function names and attributes. The child contract should have the Ethereum address

(EA) of the parent contract. The secondary artist then asks for an attestation using the AttestSC() function available in the original video's smart contract. Then the original artist will check the newly created child contract and grant the attestation as shown in Alternative 2 or deny it as shown in Alternative 3 in Figure a.

CONCLUSION

The free access to create and share information which has no fact behind it on social media platforms like WhatsApp and other digital platforms has popped out a new problem of fake information, which created rumors around the world. In this paper, we have presented a Environment set up of Ethereum blockchain-based solution for proof of authenticity of digital videos in which a secure and trusted traceability to the original video creator or source can be established, in a completely decentralized manner. Our solution allows the social media users to have an access to trusted data provenance of digital content so that they can trace the data and can trust that the data is real. The solution makes the use of ethereum smart contracts and IPFS decentralized storage system. Ethereum wallet deploys smart contracts for videos and IPFS is used to store the metadata of videos and also it generates unique hash of the videos to locate the files on IPFS. Our proposed solution framework, system design, sequence diagrams and implementation details can be applied to any digital content like videos and images. This Smart Contract based solution provides a trusted way for secondary artists o request permission form the original artist to copy, alter, share and edit videos. The work is underway for designing and implementing a fully functional and operational decentralized reputation system. Currently, the set up for ethereum framework, smart contracts, and private chain is successfully created.

REFERENCES

[1] P.Treleaven, R. G. Brown, and D. Yang, "Blockchain technology in finance," Computer, no. 9, pp. 14–17, 2017.

[2] Swapanali N. Tambe and A. B. Pawar, "Detecting Fake Video using Blockchain and Smart Contract", IJRTE, ISSN: 2277-3878, Volume-8 Issue-4S5, December 2019.

[3] Rasika Arsade,Ritika Charde, Alok Chauhan. " Design and Implementationof IOT based smart irrigation system",InternationalJournal of Advance Research, Ideas and innovationsin Technology,Vol 4 issue 2 2018www.IJRIIT.com

[4] Rajiv V Dharaskare, MM Goswami" Intelligent multipath routing protocol for mobile ad hoc network " , International Journal of computer science and Applications,Pg 135-1452009/11

[5]BHOWMIKDeepayan and FENG, Tian(2017), "The Multimedia Blockchain : a Distributed and Tamper-proof media Transaction Framework. In: Digital Singalprocessing(DSP),2017 22nd International Conference on,London, IEEE.

[6] B. Gipp, J. Kosti, and C. Breitinger, "Securing video integrity using decentralized trusted timestamping on the bitcoin blockchain." in MCIS,2016, p. 51.

[7]Aredeepfakes the newfakenews? [Online]. https://www.mediaupdate.co.za/media/144611/are-deepfakes-the-new-fakenews.

[8] Ipfs is the distributed web. [Online]. Available: https://ipfs.io/

[9]Deepfake-busting apps can spot even a single pixel http://www.technologyreview.com/s/612357/deepfake -busting-app-can-spot-even-a-single-pixel-out-of-

Place/
[10]https://bitcoin.stackexchange.com/questions/7330/whats-the-process-of-creating-a-block-on-the-blockchain
[11]Originalmy products. [Online] Available: https://originalmy.com/Products
[12]Ethereum nameservice. [Online]. Available: https://ens.domains/
[13]Ether gas station. [Online]. Available: https://ethgasstation.info/
[14]https://mlsdev.com/blog/156-how-to-build-your-own-blockchain-architecture
[15]https://blog.goodaudience.com/blockchain-for-beginners-what-is-blockchain-519db8c6677a7gt
[16]https://www.leewayhertz.com/blockch15ain-fake-News
[17]https://codeburst.io/build-your-first-ethereum-smartcontract-with-solidity-tutorial-94171d6b1c4b